Volterra kernels and effective connectivity

Karl J Friston

The Wellcome Dept. of Cognitive Neurology,
University College London
Queen Square, London, UK WC1N 3BG
Tel (44) 020 7833 7456
Fax (44) 020 7813 1445
email k.friston@fil.ion.ucl.ac.uk

Contents

I INTRODUCTION

The purpose of this brief chapter is to establish the Volterra formulation of effective connectivity from a conceptual and neurobiological point of view. Recall from the previous chapter that the generalised convolution representation of dynamic systems, afforded by Volterra expansions, is just anther way of describing the input-output behaviour of systems that have an equivalent state-space representation. In the next chapter we will deal with

state-space representations in more detail. Before proceeding to explicit input-state-output models we will look at why the Volterra formulation can be so useful.

A The brain and dynamic systems

The brain can be regarded as an ensemble of connected dynamical systems and as such conforms to some simple principles relating the inputs and outputs of its constituent parts. The ensuing implications, for the way we think about, and measure, neuronal interactions can be quite profound. These range from implications for which aspects of neuronal activity are important to measure and how to characterise coupling among neuronal populations, to implications pertaining to dynamic instability and complexity, that is necessary for adaptive self-organisation.

This chapter focuses on the first issue by looking at neuronal interactions, coupling and implicit neuronal codes from a dynamical perspective. By considering the brain in this light one can show that a sufficient description of neuronal activity must comprise activity at the current time *and its recent history*. This history constitutes a neuronal transient. Such transients represent an essential metric of neuronal interactions and, implicitly, a code employed in the functional integration of brain systems. The nature of transients, expressed conjointly in different neuronal populations, reflects the underlying coupling among brain systems. A complete description of this coupling, or *effective connectivity*, can be expressed in terms of generalised convolution kernels [Volterra kernels] that embody high-order or nonlinear interactions. This coupling may be *synchronous*, and possibly oscillatory, or *asynchronous*. A critical distinction between synchronous and asynchronous coupling is that the former is essentially linear and the latter is nonlinear. The nonlinear nature of asynchronous coupling enables context-sensitive interactions that characterise real brain dynamics, suggesting that it plays an important role in functional integration.

Brain states are inherently labile, with a complexity and transience that renders their invariant characteristics elusive. The position adopted in this chapter is that the best approach is to embrace these dynamical aspects. Its aim is to introduce the notion of neuronal transients and the underlying framework (Friston 2000). The central tenet is that the dynamics of neuronal systems can be viewed as a succession of transient spatiotemporal patterns of activity. These transients are shaped by the brain's anatomical infrastructure, principally connections, which have been selected to ensure the adaptive nature of the resulting dynamics. Although rather obvious, this formulation embodies one fundamental point; namely that any description of brain state should have an explicit temporal dimension.

In other words, measures of brain activity are only meaningful when specified over periods of time. This is particularly important in relation to fast dynamic interactions among neuronal populations that are characterised by synchrony. Synchronisation has become popular in the past years (*e.g.* Gray and Singer 1991, Eckhorn *et al.* 1992, Engel *et al* 1991) and yet represents only one possible sort of interaction.

This chapter is divided into four sections. In the first we review the conceptual basis of neuronal transients. This section uses an equivalence, between two mathematical formulations of nonlinear systems, to show that descriptions of brain dynamics, in terms of (i) neuronal transients and (ii) the coupling among interacting brain systems, is complete and sufficient. The second section uses this equivalence to motivate a taxonomy of neuronal codes and establish the relationship among neuronal transients, asynchronous coupling, *dynamic correlations* and nonlinear interactions. In the third section we illustrate nonlinear coupling using magnetoencephalography (MEG) data. The final section discusses some neurobiological mechanisms that might mediate nonlinear coupling.

II. NEURONAL TRANSIENTS

The assertion that meaningful measures of brain dynamics have a temporal domain is neither new nor contentious (*e.g.* von der Malsburg 1983, Optican and Richmond 1987, Engel *et al* 1991, Aertsen *et al* 1994, Freeman and Barrie 1994, Abeles *et al* 1995, deCharms and Merzenich 1996). A straightforward analysis demonstrates its veracity: Suppose that one wanted to posit some variables x that represented a complete and self-consistent description of brain activity. In short, everything needed to determine the evolution of the brain's state, at a particular place and time, was embodied in these measurements. Consider a component of the brain (*e.g.* a neuron or neuronal population). If such a set of variables existed for this component system they would satisfy some immensely complicated nonlinear state equation

$$\frac{\partial x(t)}{\partial t} = f(x(t),u(t)) \qquad 1$$

where x is a huge vector of state variables which range from depolarisation at every point in the dendritic tree to the phosphorylation status of every relevant enzyme, from the biochemical status of every glial cell compartment to every aspect of gene expression.. $u(t)$

represents a set of inputs conveyed by afferent from other regions. Eq(1) simply says that the changes in the state variables are nonlinear functions of the variables themselves and some inputs. The vast majority of these variables are hidden and not measurable directly. However, there are a small number of derived measurements *y* that can be made,

$$y(t) = \lambda(x(t)) \qquad 2$$

such as activities of whole cells or populations. These activities could be measured in many ways, for example firing at the initial segment of an axon or local field potentials. The problem is that a complete and sufficient description appears unattainable, given that the underlying state variables cannot be observed directly. This is not the case. The resolution of this apparent impasse rests upon two things. (i) Firstly, a fundamental mathematical equivalence relating the inputs and outputs of a dynamical system and (ii) the fact that these measurable outputs constitute the inputs to other cells or populations.

Box 1 about here

A Convolution and state-space representations

Assume that every neuron in the brain is modelled by a nonlinear dynamical system of the sort described by Eq(1). Under this assumption it can be shown that *the output is a function of the recent history of its inputs*.

$$y(t) = h(u(t-\sigma)) \qquad 3$$

where $u(t-\tau)$ represents the inputs in the recent past. Furthermore, this relationship can be expressed as a Volterra series of the inputs (see Box 1). The critical thing here is that we never need to know the underlying and 'hidden' variables that describe the details of each cell's electrochemical and biochemical status. We only need to know the history of its inputs, which, of course, are the outputs of other cells. Eq(3) is, in principle, a sufficient description of brain dynamics and involves the variables $u(t-\sigma) = y(t-\sigma)$ that represent activity at all times σ preceding the moment in question. These are simply neuronal transients. The degree of transience depends on how far back in time it is necessary to go to fully capture the brain's dynamics. The sensible nature of Eq(3) can be readily seen. For example, if we wanted to determine the behaviour of a cell in V1 (primary visual cortex) then

we would need to know the activity of all connected cells in the immediate vicinity over the last millisecond or so to account for propagation delays down afferent axons. We would also need to know the activity in distant sources, like the lateral geniculate nucleus and higher cortical areas, some ten or more milliseconds ago. In short we need the recent history of all inputs.

Transients can be expressed in terms of firing rates (*e.g.* chaotic oscillations, Freeman and Barrie 1994) or individual spikes (e.g. syn-fire chains, Abeles *et al.* 1994; 1995). Transients are not just a mathematical abstraction, they have real implications at a number of levels: For example, the emergence of fast oscillatory interactions among simulated neuronal populations depends upon the time delays implicit in axonal transmission and the time-constants of postsynaptic responses. Another slightly more subtle aspect of this formulation is that changes in synaptic efficacy, such as short-term potentiation or depression, take some time to be mediated by intracellular mechanisms. This means that the interaction between inputs at different times, that models these activity-dependent effects, again depends on the relevant history of activity

1 levels of description

The above arguments lead to a conceptual model of the brain that comprises a collection of dynamical systems (*e.g.* cells or populations of cells), each representing as an input-state-output model, where the state remains forever hidden. However the inputs and outputs are accessible and are causally related where the output of one system constitutes the input to another. A complete description therefore comprises the nature of these relationships (the Volterra series corresponding to the function h) and the neuronal transients. This constitutes a *mesoscopic* level of description that permits a degree of 'black-boxness' but with no loss of information.

The equivalence, in terms of specifying the behaviour of a neuronal system, between microscopic and mesoscopic levels of description is critical. In short, the equivalence means that all the information inherent in unobservable microscopic variables that determine the response of a neuronal system is embodied in the history of its observable inputs and outputs. Although the microscopic level of description may be more mechanistically informative, from the point of view of response prediction, neuronal transients are an entirely equivalent representation[1].

[1]We have focussed on the distinction between microscopic and mesoscopic levels of description. The *macroscopic* level is reserved for approaches, exemplified by synergistics (Haken 1983), that characterise the

B Effective connectivity and Volterra kernels

The first conclusion so far is that neuronal transients are necessary to specify brain dynamics. The second conclusion is that a complete model of the influence one neuronal population exerts over another should take the form of a Volterra series[2]. This implies that a complete characterisation of these influences (*i.e*. effective connectivity) comprises the Volterra kernels that are applied to the inputs to yield the outputs. Effective connectivity refers explicitly to "the influence that one neural system exerts over another, either at a synaptic (*i.e.* synaptic efficacy) or population level" (Friston 1995a). It has been proposed (Aertsen and Preißl 1991) that "the notion of effective connectivity should be understood as the experiment- and time-dependent, simplest possible circuit diagram that would replicate the observed timing relationships between the recorded neurons". See the previous chapter.

If effective connectivity is the influence that one neural system exerts over another it should be possible, given the effective connectivity and the afferent activity, to predict the response of a recipient population. This is precisely what Volterra kernels do. Any model of effective connectivity can be expressed as a Volterra series and any measure of effective connectivity can be reduced to a set of Volterra kernels (see Box 1). An important aspect of effective connectivity is its context-sensitivity. Effective connectivity is simply the 'effect' that input has on the output of a target system. This effect will be sensitive to other inputs, its own history and, of course, the microscopic state and causal architecture intrinsic to the target population. This intrinsic dynamical structure is embodied in the Volterra kernels. In short, Volterra kernels are synonymous with effective connectivity because they characterise the measurable effect that an input has on its target. An example of using Volterra kernels, to characterise context-sensitive changes in effective connectivity, was provided in the previous chapter (see Figure 16). This example used hemodynamic responses to changes in neuronal activity as measured with functional magnetic resonance imaging (fMRI).

---

spatiotemporal evolution of brain dynamics in terms of a small number of macroscopic order parameters [see Kelso (1995) for an engaging exposition]. Order parameters are created and determined by the co-operation of microscopic quantities and yet, at the same time, govern the behaviour of the whole system. See Jirsa *et al* (1995) for a nice example.

[2] An important qualification here is that each system is 'controllable'. Systems which are not 'controlled' have quasi-periodic or chaotic behaviours that are maintained by interactions among the states of the system. Although an important issue at the microscopic level, it is fairly easy to show that the mean field approximation to any ensemble of subsystems is controllable.

III. NEURONAL CODES

Functional integration refers to the concerted interactions among neuronal populations that mediate perceptual binding, sensorimotor integration and cognition. It pertains to the mechanisms of, and constraints under which, the state of one population influences that of another. It has been suggested by many that functional integration, among neuronal populations, uses transient dynamics that represent a temporal 'code'. A compelling proposal is that population responses, encoding a percept, become organised in time, through reciprocal interactions, to discharge in synchrony (von der Malsburg 1985, Singer 1994). The use of the term 'encoding' here speaks directly to the notion of codes. Here a neuronal code is taken to be a metric that reveals interactions among neuronal systems by enabling some prediction of the response in one population given the same sort of measure in another[3]. Clearly, from the previous section, neuronal transients represent the most generic form of code because, given the Volterra kernels, the output can, in principle, be predicted exactly. Neuronal transients have a number of attributes (*e.g.* inter-spike interval, duration, mean level of firing, predominant frequency *etc.*) and any of these could be contenders for a more parsimonious code. The problem of identifying possible codes can be reduced to identifying the form the Volterra kernels in Box 1 can take. If we know their form then we can say which aspects of the input will cause a response. Conversely, it follows that the different forms of kernels should specify the various codes that might be encountered. This is quite an important point and leads to a clear formulation of what can and cannot constitute a code. We will review different codes in terms of the different sorts of kernels that could mediate them.

A Instantaneous vs. temporal codes

The first kernel characteristic that engenders a coding taxonomy is kernel depth. The limiting case here is when the kernels support shrinks to a point in time. This means that the only relevant history is the immediate activity of inputs (all earlier activities are 'ignored' by the kernel). In this case the activity in any unit is simply a nonlinear function of current activities elsewhere. An example of this is instantaneous rate coding.

[3] Although the term code is not being used to denote anything that 'codes' for something in the environment, it could be used to define some aspect of an evoked transient that expresses a high mutual information with a stimulus parameter (e.g. Optican and Richmond 1987, Tovee *et al* 1993).

Rate coding considers spike-trains as *stochastic processes* whose first order moments (*i.e.* mean activity) describe neuronal interactions. These moments may be in terms of spikes themselves or other compound events (*e.g.* the average rate of bursting, Bair *et al* 1994). Interactions based on rate coding are usually assessed in terms of cross-correlations. From the dynamical perspective instantaneous rate codes are considered insufficient. This is because they predict nothing about a cell, or population, response unless one knows the microscopic state of that cell or population.

The distinction between rate and temporal coding (see Shadlen & Newsome 1995, de Ruyter van Steveninck 1997) centres on whether the precise timing of individual spikes is sufficient to facilitate meaningful neuronal interactions. In temporal coding the exact time at which an individual spike occurs is the important measure and the spike-train is considered as a *point process*. There are clear examples of temporal codes that have predictive validity; for example, the primary cortical representation of sounds by the co-ordination of action potential timing (deCharms and Merzenich 1996). These codes depend on the relative timing of action potentials and implicitly, an extended temporal frame of reference. They therefore fall into the class of transient codes, where selective responses to particular inter-spike intervals are modelled by temporally extended second order kernels. A nice example is provided by de Ruyter van Steveninck *et al* (1997) who show that the temporal patterning of spike trains, elicited in fly motion-sensitive neurons by natural stimuli, can carry twice the amount of information than an equivalent [Poisson] rate code.

B Transient codes: Synchronous vs. asynchronous

The second distinction, assuming the kernels have a non-trivial depth, is whether they comprise high order terms or not. Expansions that encompass just first-order terms are only capable of meditating linear or synchronous interactions. Higher-order kernels confer nonlinearity on the influence of an input that leads to asynchronous interactions. Mathematically, if there are only first-order terms then the Fourier transform of the Volterra kernel completely specifies the relationship (the transfer function) between the spectral density of input and output in a way that precludes interactions among frequencies, or indeed inputs. In other words, the expression of any frequency in a recipient cell is predicted exactly by the expression of the same frequency in the source (after some scaling by the transfer function).

1 Synchronous codes

The proposal most pertinent to these forms of code is that population responses, participating in the encoding of a percept, become organised in time through reciprocal interactions so that they come to discharge in synchrony (von der Malsburg 1985, Singer 1994) with regular periodic bursting. It should be noted that synchronisation does not necessarily imply oscillations. However, synchronised activity is usually inferred operationally by oscillations implied by the periodic modulation of cross-correlograms of separable spike trains (e.g. Gray and Singer 1991, Eckhorn *et al.* 1992) or measures of coherence in multichannel electrical and neuromagnetic time-series (e.g. Llinas *et al* 1994). The underlying mechanism of these frequency-specific interactions is usually attributed to phase-locking among neuronal populations (*e.g.* Sporns *et al* 1992; Aertsen and Preißl 1991). The key aspect of these measures is that they refer to the extended temporal structure of synchronised firing patterns, either in terms of spiking (*e.g.* syn-fire chains, Abeles *et al* 1994, Lumer *et al* 1997) or oscillations in the ensuing population dynamics (*e.g*. Singer 1994).

Many aspects of functional integration and feature linking in the brain are thought to be mediated by synchronised dynamics among neuronal populations (Singer 1994). Synchronisation reflects the direct, reciprocal exchange of signals between two populations, whereby the activity in one population influences the second, such that the dynamics become entrained and mutually reinforcing. In this way the binding of different features of an object may be accomplished, in the temporal domain, through the transient synchronisation of oscillatory responses. This 'dynamical linking' defines their short-lived functional association. Physiological evidence is compatible with this theory (e.g. Engel *et al* 1991, Fries *et al* 1997). Synchronisation of oscillatory responses occurs within as well as among visual areas, for example between homologous areas of the left and right hemispheres and between areas at different levels of the visuomotor pathway (Engel *et al* 1991, Roelfsema *et al* 1997). Synchronisation in the visual cortex appears to depend on stimulus properties, such as continuity, orientation and motion coherence.

The problem with synchronisation is that there is nothing essentially dynamic about synchronous interactions *per se*. As argued by Erb and Aertsen (1992) "the question might not be so much how the brain functions by virtue of oscillations, as most researchers working on cortical oscillations seem to assume, but rather how it manages to do so in spite of them". In order to establish dynamic cell assemblies it is necessary to create and destroy synchronised couplings. It is precisely these dynamic aspects that speak to changes in synchrony (*e.g.* Desmedt and Tomberg 1994) and the asynchronous transitions between synchronous states as the more pertinent phenomenon. In other words, it is the successive

reformulation of dynamic cell assemblies, through nonlinear or asynchronous interactions, that is at the heart of 'dynamical linking' (Singer 1994).

2 Asynchronous codes

An alternative perspective on neuronal codes is provided by *dynamic correlations* (Aertsen *et al* 1994) as exemplified in Vaadia *et al* (1995). A fundamental phenomenon, observed by Vaadia *et al* (1995), is that, following behaviourally salient events, the degree of coherent firing between two neurons can change profoundly and systematically over the ensuing second or so. One implication is that a complete model of neuronal interactions has to accommodate dynamic changes in correlations, modulated on time-scales of 100-1000ms. Neuronal transients provide a simple explanation for this temporally modulated coherence or dynamic correlation. Imagine that two neurons respond to an event with a similar transient. For example, if two neurons respond to an event with decreased firing for 400ms, and this decrease was correlated over epochs, then positive correlations between the two firing rates would be seen for the first 400 of the epoch, and then fade away, exhibiting a dynamic modulation of coherence. In other words, a transient modulation of covariance can be equivalently formulated as covariance in the expression of transients. The generality of this equivalence can be established using singular value decomposition (*SVD*) of the joint-peristimulus time histogram (J-PSTH) as described in Friston (1995b). This is simply a mathematical device to show that dynamic changes in coherence are equivalent to the coherent expression of neural transients. In itself it is not important, in the sense that dynamic correlations are just as valid a characterisation as neuronal transients and indeed may provide more intuitive insights into how this phenomenon is mediated (*e.g.* Riehle *et al* 1997). A more important observation is that J-PSTHs can be asymmetric about the leading diagonal. This suggests that coupled transients in two units can have a different patterning of activity. This can only be explained by asynchronous or nonlinear coupling.

C Summary

In summary, the critical distinction between synchronous and asynchronous coupling is the difference between linear and nonlinear interactions among units or populations[4]. This difference is reduces to the existence of high-order Volterra kernels in the mediating the

[4] The term 'generalised synchrony' has been introduced to include nonlinear inter-dependencies (see Schiff *et al* 1996). Generalised synchrony subsumes synchronous and asynchronous coupling. An elegant method for making inferences about generalised synchrony is described in Schiff *et al* (1996). This approach is

input-output behaviour of coupled cortical regions. There is a close connection between asynchronous/nonlinear coupling and the expression of distinct transients in two brain regions: Both would be expressed as dynamic correlations, or, in the EEG, as event-related changes in synchronisation (e.g.. induced oscillations). If the full transient model is correct then important transactions among cortical areas will be overlooked by techniques that are predicated on rate coding (*e.g.* correlations, covariance patterns, spatial modes *etc.*) or synchronisation models (*e.g.* coherence analysis and cross-correlograms). Clearly the critical issue here is whether there is direct evidence for nonlinear or asynchronous coupling that would render high-order Volterra kernels necessary.

IV. EVIDENCE FOR NONLINEAR COUPLING

Why is asynchronous coupling so important? The reason is that asynchronous interactions embody all the nonlinear interactions implicit in functional integration and it is these that mediate the context-sensitive nature of neuronal interactions. Nonlinear interactions among cortical areas render the effective connectivity among them inherently dynamic and contextual. Compelling examples of context-sensitive interactions include the attentional modulation of evoked responses in functionally specialised sensory areas (*e.g.* Treue and Maunsell 1996) and other contextually dependent dynamics [see Phillips and Singer (1997)]. Whole classes of empirical phenomena such as extra-classical receptive field effects rely on nonlinear or asynchronous interactions.

A Nonlinear coupling and asynchronous interactions

If the temporal structures of recurring transients in two parts of the brain are distinct, then the prevalence of certain frequencies in one cortical area should predict the expression of *different* frequencies in another. In contrast, synchronisation posits the expression of the *same* frequencies. Correlations among different frequencies therefore provide a basis for discriminating between synchronous and asynchronous coupling.

Consider time-series from two neuronal populations or cortical areas. Synchrony requires that the expression of a particular frequency (e.g. 40 Hz) in one time series will be coupled with the expression of the same frequency in the other. In other words, the modulation of

---

particularly interesting from our point of view because it calls upon the recent history of the dynamics through the use of temporal embedding to reconstruct the attractors analysed.

this frequency in one area can be explained or predicted by its modulation in the second. Conversely, asynchronous coupling suggests that the power at a reference frequency, say 40 Hz, can be predicted by the spectral density in the second time-series at frequencies other than 40 Hz. These predictions can be tested empirically using standard time-frequency and regression analyses as described in Friston (2000). Figure 1 shows an example of this sort of analysis, revealing the dynamic changes in spectral density between 8 and 64 Hz over 16 seconds. The cross-correlation matrix of the time-dependent expression of different frequencies in the parietal and prefrontal regions is shown in the lower left panel. There is anecdotal evidence here for both synchronous and asynchronous coupling. Synchronous coupling, based upon the co-modulation of the same frequencies, is manifest as hot-spots along, or near, the leading diagonal of the cross-correlation matrix (*e.g.* around 20Hz). More interesting are correlations between high frequencies in one time-series and low frequencies in another. In particular note that the frequency modulation at about 34Hz in the parietal region (second time-series) could be explained by several frequencies in the prefrontal region. The most profound correlations are with lower frequencies in the first time-series (26Hz). Using a simple regression framework, statistical inferences can be made about the coupling within and between different frequencies (see Friston 2000 for details). A regression analysis shows that coupling at 34Hz has significant synchronous and asynchronous components, where as the coupling at 48Hz is purely asynchronous (middle and right peaks in the graphs) *i.e.* a coupling between beta dynamics in the premotor region and gamma dynamics in the parietal region.

Figure 1 about here

V. THE NEURAL BASIS OF NONLINEAR COUPLING

In Friston (1997) it was suggested that, from a neurobiological perspective, the distinction between nonlinear [asynchronous] and linear [synchronous] interactions could be viewed in the following way. Synchronisation emerges from the reciprocal exchange of signals between two populations, where each *drives* the other, such that the dynamics become entrained and mutually reinforcing. In asynchronous coding the afferents from one population exert a *modulatory* influence, not on the activity of the second, but on the interactions within it (*e.g.* a modulation of effective connectivity or synaptic efficacies within

the target population) leading to changes in the dynamics intrinsic to the second population. In this model there is no necessary synchrony between the intrinsic dynamics that ensue and the temporal pattern of modulatory input. To test this hypothesis one would need to demonstrate that asynchronous coupling emerges when extrinsic connections are changed from driving connections to modulatory connections. Clearly this cannot be done in the real brain. However, we can use computational techniques to create a biologically realistic model of interacting populations and test this hypothesis directly.

A. Interactions between simulated populations

Two populations were simulated using the model described in Friston (2000). This model simulates entire neuronal populations in a deterministic fashion based on known neurophysiological mechanisms. In particular we modelled three sorts of synapse, fast inhibitory (GABA), fast excitatory (AMPA) and slower voltage-dependent synapses (NMDA). Connections intrinsic to each population used only GABA and AMPA-like synapses. Simulated glutaminergic extrinsic connections between the two populations used either driving AMPA-like synapses or modulatory NMDA-like synapses. Transmission delays for extrinsic connections were fixed at 8ms. By using realistic time constants the characteristic oscillatory dynamics of each population were expressed in the gamma range.

Figure 2 about here

The results of coupling two populations with unidirectional AMPA-like connections are shown in the top of Figure 2 in terms of the simulated local field potentials (LFP). Occasional transients in the driving population were evoked by injecting a depolarising current, of the same magnitude, at random intervals (dotted lines). The tight synchronised coupling that ensues is evident. This example highlights the point that near-linear coupling can arise even in the context of loosely coupled, highly nonlinear neuronal oscillators of the sort modelled here. Contrast these entrained dynamics under driving connections with those that emerge when the connection is modulatory or NMDA-like (lower panel in Figure 2). Here there is no synchrony and, as predicted, fast transients of an oscillatory nature are facilitated by the low frequency input from the first population that has a lower frequency (*c.f.* the MEG analyses above). This is a nice example of asynchronous coupling that is underpinned by nonlinear modulatory interactions between neuronal populations. The nature of the coupling can be characterised using the time-frequency analysis (identical in every

detail) applied to the neuromagnetic data of the previous section. The results for the NMDA simulation are presented in Figure 3. The cross-correlation matrix resembles that obtained with the MEG data in Figure 1. Both in terms of the variance, and inference, asynchronous coupling supervenes at most frequencies but, as in the real data, mixed coupling is also evident. These results can be taken as a heuristic conformation of the notion that modulatory, in this case voltage-dependent, interactions are sufficiently nonlinear to account for the emergence of asynchronous coupling.

Figure 3 about here

B. Modulatory interactions and nonlinear coupling

In summary, asynchronous coupling is synonymous with nonlinear coupling. Nonlinear coupling can be framed in terms of the modulation of intrinsic interactions, within a cortical area or neuronal population, by extrinsic input offered by afferents from other parts of the brain. This mechanism predicts that the modulation of fast (*e.g.* gamma) activity in one cortical area can be predicted by much slower changes in other areas. This form of coupling is very different from coherence or other measures of synchronous coupling and concerns the relationship between the first-order dynamics in one area and the second order dynamics (spectral density) expressed in another. In terms of the above NMDA simulation, transient depolarisation in the modulating population causes a short-lived increased input to the second. These afferents impinge on voltage-sensitive NMDA-like synapses with time constants (in the model) of about 100ms. These synapses open and slowly close again, remaining open long after an afferent volley. Because of their voltage-sensitive nature this input will have no effect on the dynamics intrinsic to the second population unless there is already a substantial degree of depolarisation. If there is then, through self-excitation and inhibition, the concomitant opening of fast excitatory and inhibitory channels will generally increase membrane conductance, decrease the effective membrane time constants and lead to fast oscillatory transients. This is what we observe in the lower panel of Figure 2. In relation to the MEG analyses, the implied modulatory mechanisms, that may underpin this effect, are entirely consistent with the anatomy, laminar specificity and functional role attributed to prefrontal efferents (Rockland and Pandya 1979, Selemon and Goldman-Rakic 1988).

VI CONCLUSION

In this chapter we have dealt with some interesting and interrelated aspects of effective connectivity, neuronal codes, nonlinear coupling, neuronal transients and dynamic correlations (*e.g.* induced oscillations). The key points can be summarised as follows

- Starting with the premise that the brain can be represented as an ensemble of connected input-state-output systems (e.g. cellular compartments, cells or populations of cells) there exists an equivalent input-output formulation in terms of a Volterra series. This is simply a functional expansion of each system's inputs that produces its outputs (where the outputs to one system constitute the inputs to another).

- The existence of this expansion suggests that the history of inputs, or neuronal transients, and the Volterra kernels are a complete and sufficient specification of brain dynamics (to the extent they are controllable). This is the primary motivation for framing dynamics in terms of neuronal transients and using Volterra kernels to model effective connectivity.

- The Volterra formulation provides constraints on the form that neuronal interactions and implicit codes must conform to. There are two limiting cases; (i) when the kernel decays very quickly and (ii) when high order kernels disappear. The first case corresponds to instantaneous codes (*e.g.* rate codes) and the second to synchronous interactions (*e.g.* synchrony codes).

- High order kernels in the Volterra formulation of effective connectivity speak to nonlinear interactions and implicitly to asynchronous coupling. Asynchronous coupling implies coupling among the expression of different frequencies.

- Coupling among different frequencies is easy to demonstrate using neuromagnetic measurements of real brain dynamics. This implies that nonlinear, asynchronous coupling is a prevalent component of functional integration.

- High order kernels correspond to modulatory interactions that can be construed as a nonlinear effect of inputs that interact bilinearly with the intrinsic states of the recipient system. This implies that driving connections may be linear and engender

synchronous interactions. Conversely, modulatory connections, being nonlinear, may be revealed by asynchronous coupling and be expressed in high order kernels.

Box 1: Dynamical systems and Volterra kernels

Input-state-output systems and Volterra series

Neuronal systems are inherently nonlinear and lend themselves to modelling with nonlinear dynamical systems. However due to the complexity of biological systems it is difficult to find analytic equations that describe them adequately. Even if these equations were known the state variables are often not observable. An alternative approach to identification is to adopt a very general model (Wray and Green 1994) and focus on the inputs and outputs. Consider the single input-single output (SISO) system

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= \lambda(x(t)) \end{aligned} \qquad \text{B.1}$$

The Fliess fundamental formula (Fliess *et al* 1983) describes the causal relationship between the outputs and the recent history of the inputs. This relationship can be expressed as a Volterra series which expresses the output $y(t)$ as a nonlinear convolution of the inputs $u(t)$, critically without reference to the state variables $x(t)$. This series is simply a functional Taylor expansion of $y(t)$ in Eq(3) (main text).

$$\begin{aligned} y(t) = h(u(t-\sigma)) &= \sum_{i=1}^{\infty} \int_0^t \cdots \int_0^t \kappa_i(\sigma_1, \ldots \sigma_i) u(t-\sigma_1) \ldots u(t-\sigma_i) d\sigma_1 \ldots d\sigma_i \\ \kappa_i(\sigma_1, \ldots \sigma_i) &= \frac{\partial^i y(t)}{\partial u(t-\sigma_1) \ldots \partial u(t-\sigma_i)} \end{aligned} \qquad \text{B.2}$$

were $\kappa_i(\sigma_1, \ldots \sigma_i)$ is the ith order kernel. Volterra series have been described as a 'power series with memory' and are generally thought of as a high-order or 'nonlinear convolution' of the inputs to provide an output. See Bendat (1990) for a fuller discussion.

Volterra kernels and effective connectivity

Volterra kernels are essential in characterising the effective connectivity or influences that one neuronal system exerts over another because they represent the causal input-output

characteristics of the system in question. Neurobiologically they have a simple and compelling interpretation – *they are synonymous with effective connectivity*: From B.2

$$\kappa_1(\sigma_1) = \frac{\partial y(t)}{\partial u(t-\sigma_1)}, \qquad \kappa_2(\sigma_1,\sigma_2) = \frac{\partial^2 y(t)}{\partial u(t-\sigma_1)\partial u(t-\sigma_2)}, \qquad \ldots$$

It is evident that the first-order kernel embodies the response evoked by a change in input at $t-\sigma_1$. In other words, it is a time-dependant measure of *driving* efficacy. Similarly the second order kernel reflects the *modulatory* influence of the input at $t-\sigma_1$ on the response evoked by input at $t-\sigma_2$. And so on for higher orders.

Figure Legends

Figure 1

Time-frequency and regression analysis of MEG time-series designed to characterise the relative contribution of synchronous and asynchronous coupling. Neuromagnetic data were acquired from a normal subject using a KENIKRON 37 channel MEG system at one-millisecond intervals for periods of up to two minutes. During this time the subject made volitional joystick movements to the left, every two seconds or so. Paired epochs were taken from a left prefrontal and left parietal region. Top panels: The two times series (plots) and their corresponding time-frequency profiles (images). The first time-series comes from the left prefrontal region. The second comes from the left superior parietal region. Lower left panel: This is a simple characterisation of the coupling among frequencies in the two regions and represents the [squared] cross-correlations of the time-varying expression of different frequencies from the upper panels. Lower right panels: These are the results of a linear regression analysis that partitions the variance in the second (parietal) time-series into components that can be attributed to synchronous (broken lines) and asynchronous (solid lines) contributions from the first (prefrontal) time series. The upper graph shows the relative contribution in terms of the proportion of variance explained and in terms of the significance using a semi-log plot of the corresponding p-values (lower graph). The dotted line in the latter corresponds to $p = 0.05$.

This example was chosen because it illustrates three sorts of coupling (synchronous, asynchronous and mixed). From inspection of the cross correlation matrix it is evident that power in the beta range (20Hz) in the second time-series is correlated with similar frequency modulation in the first, albeit at a slightly lower frequency. The resulting correlations appear just off the leading diagonal (broken line) on the upper left. The proportion of variance explained by synchronous and asynchronous coupling is roughly the same and, in terms of significance, synchrony supervenes (see upper graph). In contrast, the high correlations, between 48Hz in the second time-series and 26Hz in the first, are well away from the leading diagonal, with little evidence of correlations within either of these frequencies. The regression analysis confirms that, at this frequency, asynchronous coupling prevails. The variation at about 34Hz in the parietal region could be explained by several frequencies in the prefrontal region. A formal analysis shows that both synchronous and asynchronous coupling coexist at this frequency (*i.e.* the middle peak in the graphs).

Figure 2

Simulated local field potentials (LFP) of two coupled populations using two different sorts of postsynaptic responses (AMPA and NMDA-like) to connections from the first to the target population. The dotted line shows the depolarisation effected by sporadic injections of current into the first population. The key thing to note is that under AMPA-like or driving connections the second population is synchronously entrained by the first. When the connections are modulatory or voltage-dependent (NMDA), the effects are much more subtle and resemble a frequency modulation. These data were simulated using a biologically plausible model of excitatory and inhibitory subpopulations. The model was deterministic with variables pertaining to the collective, probabilistic, behaviour of the subpopulations (*c.f.* a mean field treatment). See Friston (2000) for details.

Figure 3

As for Figure 1, but here using the simulated data employing voltage-dependent NMDA-like connections. The coupling here includes some profoundly asynchronous [nonlinear] components involving frequencies in the gamma range implicated in the analyses of real (MEG) data shown in Figure 1. In particular note the asymmetrical cross-correlation matrix and the presence of asynchronous and mixed coupling implicit in the p-value plots on the lower right.

References

Abeles M, Bergman H, Gat I, Meilijson I, Seidmann E, Tishby N and Vaadia E (1995) Cortical activity flips among quasi-stationary states. *Proc. Natl. Acad. Sci*. **92:**8616-8620

Aertsen A, Erb M and Palm G (1994) Dynamics of functional coupling in the cerebral cortex: an attempt at a model-based interpretation, *Physica D* **75**: 103 - 128

Aertsen A & Preißl H (1991) Dynamics of activity and connectivity in physiological neuronal Networks *in Non Linear Dynamics and Neuronal Networks* Ed Schuster HG VCH publishers Inc., New York NY USA p281-302

Bair W, Koch C, Newsome W and Britten K (1994) Relating temporal properties of spike trains from area MT neurons to the behaviour of the monkey. In *Temporal Coding in the Brain* Eds. Buzsaki, Llinas R, Singer W, Berthoz A and Christen T. Springer Verlag, Berlin. p221-250

Bendat JS (1990) *Nonlinear System Analysis and Identification from Random Data*. John Wiley and Sons, New York USA

Büchel C and Friston KJ (1997) Modulation of connectivity in visual pathways by attention: Cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex* **7**:768-778

deCharms RC and Merzenich MM (1996) Primary cortical representation of sounds by the co-ordination of action potential timing. *Nature* **381**:610-613

de Ruyter van Steveninck RR, Lewen GD, Strong SP, Koberie R and Bialek W (1997) Reproducibility and variability in neural spike trains. *Science* **275**:1085-1088

Desmedt JE and Tomberg C (1994) Transient phase-locking of 40Hz electrical oscillations in prefrontal and parietal human cortex reflects the process of conscious somatic perception. *Neurosci. Lett.* **168**:126-129

Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitboeck HJ (1988) Coherent oscillations: a mechanism of feature linking in the visual cortex?. Multiple electrode and correlation analysis in the cat. *Biol. Cybern* **60**:121-130

Erb M and Aertsen A (1992) Dynamics of activity in biology-oriented neural network models: Stability analysis at low firing rates. In *Information processing in the cortex. Experiments and Theory.* Eds. Aertsen A and Braitenberg V. Springer-Verlag, Berlin p201-223

Engel A K, Konig P and Singer W (1991) Direct physiological evidence for scene segmentation by temporal coding. *Proc. Natl. Acad Sci* USA vol. **88**: 9136-40

Fliess M, Lamnabhi M and Lamnabhi-Lagarrigue F (1983) An algebraic approach to nonlinear functional expansions. *IEEE Transactions on Circuits and Systems* **30**:554-570

Freeman W and Barrie J (1994) Chaotic oscillations and the genesis of meaning in cerebral cortex. *In Temporal Coding in the Brain* Eds. Buzsaki, Llinas R, Singer W, Berthoz A and Christen T. Springer Verlag, Berlin. p13-38

Fries P, Roelfsema PR, Engel A, Konig P and Singer W, 1997, Synchronisation of oscillatory responses in visual cortex correlates with perception in inter-ocular rivalry. *Pro Natl Acad Sci* USA vol. **94**:12699-12704

Friston KJ. (1995a) Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapp.* **2**:56-78

Friston KJ. (1995b) Neuronal transients. *Proc. Roy. Soc. Series B*. **261**:401-405

Friston KJ (1997) Transients metastability and neuronal dynamics. *NeuroImage* **5**:164-171

Friston KJ (2000) The Labile Brain I: Neuronal Transients and nonlinear coupling. *Phil. Trans. R. Soc.* (Lond.) **355**:215-236

Friston KJ and Büchel C (2000) Attentional modulation of effective connectivity from V2 to V5 in humans *Pro Natl Acad Sci* USA vol. **97**:7591-7596

Gerstein GL & Perkel DH (1969) Simultaneously recorded trains of action potentials: Analysis and functional interpretation *Science* **164**:828-830

Gray CM & Singer W (1989) Stimulus specific neuronal oscillations in orientation columns of cat visual cortex *Proc. Natl. Acad. Sci* USA **86**:1698-1702

Haken H (1983) *Synergistics: An introduction.* Springer Berlin. 3rd Edition

Jirsa VK, Friedrich R and Haken H (1995) Reconstruction of the spatio-temporal dynamics of a human magnetoencephalogram. *Physica D* **89**:100-122

Kelso JAS (1995) *Dynamic Patterns: The self-organisation of brain and behaviour.* MIT Press, USA

Llinas R Ribary U Joliot M and Wang X-J (1994) Content and context in temporal thalamocortical binding. In *Temporal Coding in the Brain* Eds. Buzsaki, Llinas R, Singer W, Berthoz A and Christen T. Springer Verlag, Berlin. p251-272

Lumer ED, Edelman GM, Tononi G, (1997) Neural Dynamics in a model of the thalamocortical System II. The role of neural synchrony tested through perturbations of spike timing. *Cerebral Cortex* **7**:228 – 236

Phillips WA and Singer W (1997) In search of common foundations for cortical computation. *Behavioural and Brain Sciences*. **20**:57-83

Riehle A, Grun S, Diesmann M and Aertsen A (1997) Spike synchronisation and rate modulation differentially involved in motor cortical function. *Science* **278**:1950-1953

Rockland K. S., Pandya D. N. (1979) Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. *Brain-Res* **179**: 3-20

Schiff SJ, So P, Chang T, Burke RE and Sauer T (1996) Detecting dynamical interdependence and generalised synchrony though mutual prediction in a neuronal ensemble. *Physical Review E* **54**:6708-6724

Selemon L. D., Goldman Rakic P. S. (1988) Common cortical and subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: evidence for a distributed neural network subserving spatially guided behavior. *J Neurosci.* **8**:4049-68

Shadlen MN and Newsome WT (1995) Noise, neural codes and cortical organisation. *Current Opinion in Neurobiology*. **4**:569-579

Singer W (1994) Time as coding space in neocortical processing: A hypothesis. In *Temporal Coding in the Brain* Eds. Buzsaki, Llinas R, Singer W, Berthoz A and Christen T. Springer Verlag, Berlin. p51-80

Sporns O, Gally, JA, Reeke GN & Edelman GM (1989) Reentrant Signalling among simulated neuronal groups leads to coherence in their oscillatory activity *Proc. Natl. Acad. Sci* USA **86**:7265-7269

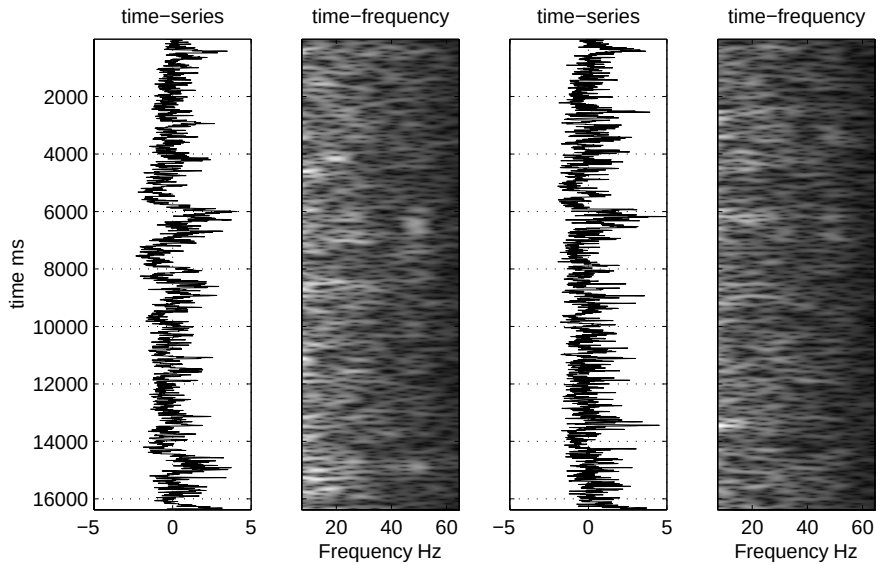
Tovee MJ, Rolls ET, Treves A and Bellis RP (1993) Information encoding and the response of single neurons in the primate temporal visual cortex. *J. Neurophysiology* **70**:640-654

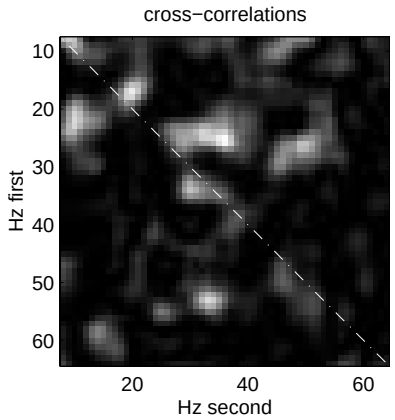
Treue S, Maunsell HR (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**:539-41

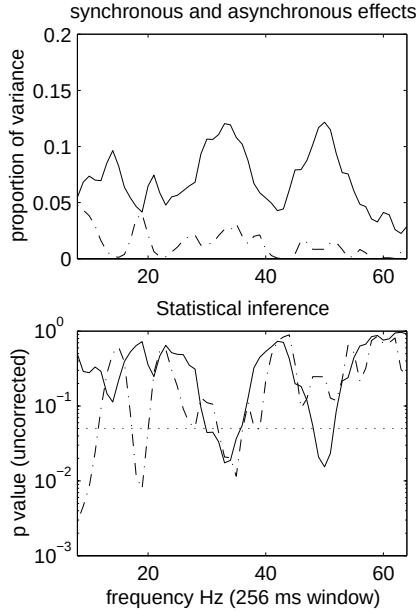
Vaadia E, Haalman I, Abeles M, Bergman H, Prut Y, Slovin H and Aertsen A (1995) Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* **373**:515-518

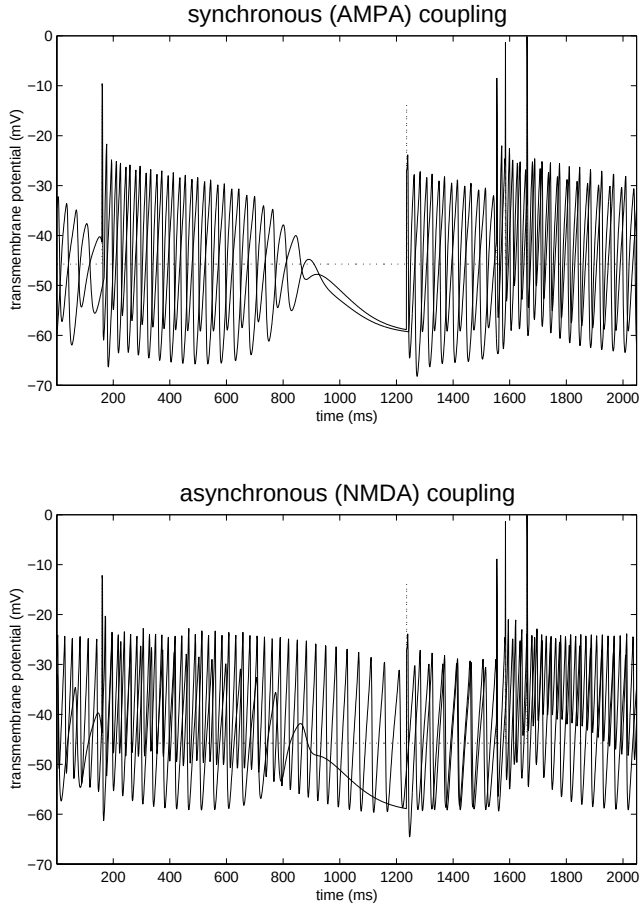
von der Malsburg C (1985) Nervous structures with dynamical links. *Ber Bunsenges. Phys. Chem.* **89**:703-710

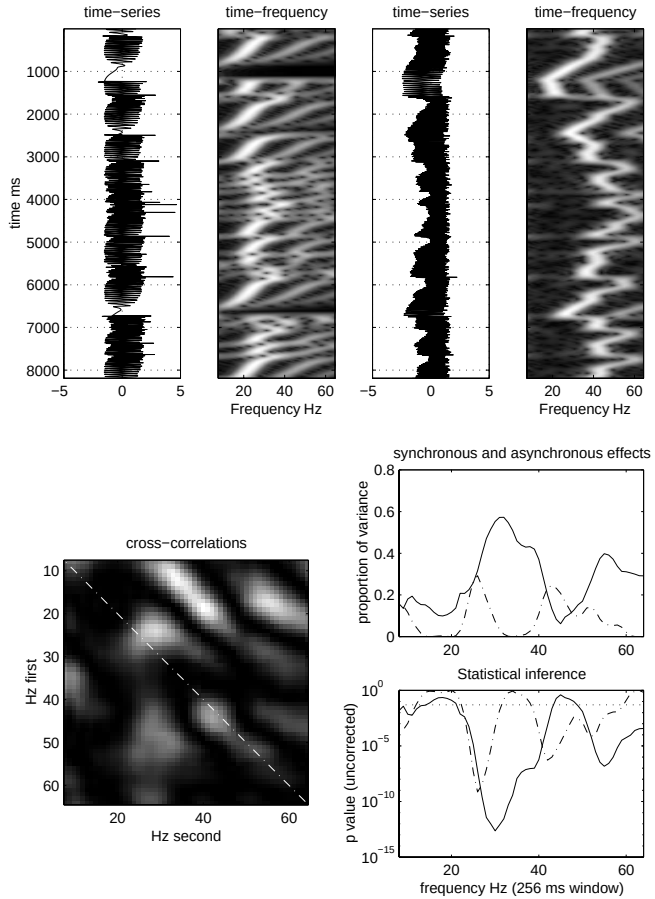
Wray J and Green GGR (1994) Calculation of the Volterra kernels of non-linear dynamic systems using an artificial neuronal network. *Biol. Cybern*. **71**:187-195


time−series
time−frequency
time−series
time−frequency
time ms
2000
4000
6000
8000
10000
12000
14000
16000
−5
0
5
20
40
60
Frequency Hz
Frequency Hz


cross−correlations
Hz first
10
20
30
40
50
60
20
40
60
Hz second


synchronous and asynchronous effects
proportion of variance
0.2
0.15
0.1
0.05
0
20
40
60
Statistical inference
p value (uncorrected)
10^{0}
10^{-1}
10^{-2}
10^{-3}
20
40
60
frequency Hz (256 ms window)


synchronous (AMPA) coupling
transmembrane potential (mV)
0
−10
−20
−30
−40
−50
−60
−70
200
400
600
800
1000
1200
1400
1600
1800
2000
time (ms)
asynchronous (NMDA) coupling
transmembrane potential (mV)
0
−10
−20
−30
−40
−50
−60
−70
200
400
600
800
1000
1200
1400
1600
1800
2000
time (ms)


time−series
time−frequency
time−series
time−frequency
time ms
1000
2000
3000
4000
5000
6000
7000
8000
−5
0
5
20
40
60
Frequency Hz
Frequency Hz
cross−correlations
Hz first
Hz second
synchronous and asynchronous effects
proportion of variance
0.8
0.6
0.4
0.2
0
Statistical inference
p value (uncorrected)
10^{0}
10^{-5}
10^{-10}
10^{-15}
frequency Hz (256 ms window)