# Effective Connectivity

Lee Harrison and 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

In the previous chapter we dealt with functional connectivity and different ways of summarising patterns of correlations among brain systems. In this chapter we turn to effective connectivity and the models of mechanisms that might mediate these correlations.

Brain function depends upon interactions among its components that range from individual cell compartments to neuronal populations. As described in **Chapter 18 (Functional Integration)**, anatomical and physiological studies of connectivity *in vivo* speak to a hierarchy of specialised regions that process increasingly abstract features, from simple edges in V1, through colour and motion in V4 and V5 respectively to face recognition in the fusiform gyrus. The implicitly specialised regions are connected, allowing distributed and reentrant neuronal information processing. The selective responses, or specialisation, of a region is a function of its connectivity. An important theme is this chapter is that these connections can change and show context-sensitivity. We will refer to these changes as plasticity, to describe physiological changes in the influence different brain systems have on each other. The formation of distributed networks, through dynamic interactions, is the basis of functional integration, which is itself, time and context-dependent. Changes in connectivity are important for development, learning, perception and adaptive response to injury.

Given the importance of changes in connectivity, we will consider two classes of experimental factors or input to the brain. The first class evokes responses directly but the second has a more subtle effect, and can induce input-dependent changes in connectivity that modulates responses to the first. We will refer to the second class of inputs as "contextual". For example, augmented neuronal responses associated with attending to a stimulus can be attributed to the changes induced in connectivity by attention. The distinction between inputs that evoke responses and those that modulate effective connectivity is the motivation for developing models that accommodate contextual changes in connection strength.

This chapter is divided into four sections. First, we motivate a Systems Identification approach, where brain responses are parameterised within the framework of a mathematical model. The State-Space representation is used to illustrate coupling within a system and interactions with experimental factors. To demonstrate how connectivity can be modulated experimentally, and how the notion of effective connectivity emerges as a natural metric, a Bilinear State-Space Model (BSSM) is derived to approximate generic nonlinear networks. In Section III we describe the theory behind different approaches to estimating functional integration, looking at Static and Dynamic models. We conclude, in Section IV, with some remarks on strategies and the features of models that have proved useful in modelling connectivity to date.

## A. Notation

Lower case letters will be used for scalars and vectors and upper case for matrices. If $x$ is a normally distributed random variable with mean $\mu$ and variance $\sigma^2$, we write $x \sim$ $x \sim N(\mu, \sigma^2)$. A time series of observations at voxel $i$ is written, $y^i$. An image at time $t$ is written $y_t$ and the value of the $i^{\text{th}}$ voxel at time $t$ is $y_t^i$. The total number of voxels and scans are $N$ and $T$ respectively. We use $exp(X)$, $X^T$, $X^{-1}$ and $X^+$ to denote the matrix exponential, transpose, inverse and pseudoinverse. We also write $x \times y$ to denote the Hadamard product between two vectors (in Matlab this is an array or element-by-element multiplication with $x.*y$). The first-order derivative of time dependent variable, x(t), with respect to time ($\partial x(t)/\partial t$) is denoted by $\dot{x}$.

## II IDENTIFICATION OF DYNAMIC SYSTEMS

System identification (SI) is the use of observed data to estimate the parameters of mathematical models representing a physical system. The mathematical models may be linear or nonlinear, in discrete or continuous time and parameterised in the time or frequency domain. The aim is to construct a mathematical description of a systems response to input. Models may be divided into two main categories: those that invoke hidden states and those that quantify relationships between inputs and outputs without

hidden states, effectively treating the system as a black box (see Juang, 2001) for a comprehensive account). Examples of the former include State-Space Models (SSM) and Hidden Markov Models (HMM), whereas the latter include Generalised Convolution Models (Bendat, 1998) and Autoregressive Models (Chatfield, 1996).

There are two main requirements of a biologically plausible model of functional integration: that it is dynamic and nonlinear. Dynamic, because the brain is a physical system extended in time, meaning that the state of the brain now effects its state in the future. We will see the benefits and problems that stem from relaxing this requirement later. In addition, biological systems depend on nonlinear phenomena for much of their characteristic behaviour (Scott, 1999). Examples include the neuronal dynamics of action potentials (Dayan and Abbott, 2001), population dynamics in co-evolutionary systems (Glass and Kaplan, 2000) and limit cycles in physiological systems (Glass, 2001). The motivation for appealing to nonlinear dynamic models is that their non-additive characteristics enable them to reproduce highly complex behaviour, of the sort we observe in biological systems. However, nonlinear models are often mathematically intractable, calling for approximation techniques.

Linear dynamic models, on the other hand, can be analysed in closed form. Consequently, there exists a large body of theory for handling them. This is due to their adherence to the Principle of Superposition, which means that the systems response to input is additive. There are no interactions between different inputs or between inputs and the intrinsic states of the system such that the response is a weighted linear mixture of inputs. A system that violates this principle would respond in a non-additive manner *i.e.* with more or less than a linear combination of inputs. Such a system is, by definition, nonlinear. However, there is a price for the ease with which linear models can be analysed, as their behavioural repertoire is limited to exponential decay and growth, oscillation or a combination of these. Examples of sub-additive responses are ubiquitous in physiology (*e.g.* saturation). With increasing input many biological systems (*e.g.* biochemical reactions or synaptic input) reach a saturation point where further input does not generate a further response.

A useful compromise is to make linear approximations to a generic nonlinear model. These models have the advantage that they capture essential nonlinear features while

remaining mathematically tractable.  This strategy has engendered bilinear models (Rao, 1992), where nonlinear interaction terms are limited to interactions that can be modelled as the product of two variables (input or intrinsic states). Despite constraints on higher order nonlinearities, bilinear models can easily model plasticity in effective connections.

We will use a bilinear state-space representation to illustrate the concepts of linear and bilinear coupling and how they may be used to model effective connectivity.  The introduction of unknown variables hidden states) may appear to complicate the problem, but this is not the case as long-range order, within observed time series, can be modelled through interactions among the states.

**A Approximating nonlinear functions**

Any sufficiently smooth function, $f(x)$, of a scalar quantity $x$ may be approximated in the neighbourhood of an expansion point, $x_0$, using the Taylor series expansion

$$f(x) \approx f(x_0) + \frac{df(x_0)}{dx}(x - x_0) + \frac{d^2 f(x_0)}{dx^2}\frac{(x - x_0)^2}{2!} + \cdots + \frac{d^n f(x_0)}{dx^n}\frac{(x - x_0)^n}{n!} \qquad 1$$

where the $n^{th}$ order derivatives evaluated at $x_0$. These values are coefficients that scale the contribution of their respective terms. The derivatives are used as they map a local change in $(x - x_0)^n$ onto a change in $f(x)$.  The degree of nonlinearity of $f$ determines the rate of convergence, with weakly nonlinear functions converging rapidly.  The series converges to the *exact* function with inclusion of higher order terms.  A simple example is shown in Figure 1

Figure 1 about here

When these ideas are extended to bivariate nonlinear functions, where $f(x,u)$ depends on two quantities $x$ and $u$, the corresponding Taylor series includes increasingly complex (high-order) terms (products of $x$ and $u$).  The linear and bilinear approximations are given by $f_L$ and $f_{BL}$.

.

$$f_L(x,u) = ax + cu$$
$$a = \frac{\partial f}{\partial x}$$
$$c = \frac{\partial f}{\partial u}$$

2

$$f_{BL}(x,u) = ax + bxu + cu$$
$$b = \frac{\partial^2 f}{\partial x \partial u}$$

For clarity, the expansion point $x_0 = 0$ and the series have been centred so that $f(x_0) = f(u_0) = 0$. The approximation, $f_L$, depends on linear terms in $x$ and $u$ scaled by coefficients $a$ and $c$, calculated from first order derivatives. The first-order terms of $f_{BL}$ are the same as $f_L$, however, the third term is composed of the product of $x$ and $u$, which is scaled by b, the second order derivative with respect to both variables. This term introduces nonlinearity into $f_{BL}$. Note that the bilinear form does not include quadratic functions of $x$ or $u$. The resulting approximation has the appealing property of being both linear in $x$ and $u$, but allowing for a modulation of $x$ by $u$. Changes in $f(x,u)$ are no longer only a linear sum of changes in $x$ and $u$, but include a contribution from a new variable $xu$.

We can now replace the scalar quantities with vectors where $x = [x_1, \ldots, x_n]^T$ and $u = [u_1, \ldots, u_m]^T$. $x$ is an $n \times 1$ vector containing $n$ different state variables and $u$ is an $m \times 1$ vector containing $m$ input variables. The linear and bilinear approximations can be written in matrix form, as

$$f_L(x,u) = Ax + Cu$$

3

$$f_{BL}(x,u) = Ax + \sum_j u_j B^j x + Cu$$

6

The coefficients, are matrices as opposed to scalar terms. They look more complicated, but the same operation is being applied to all the elements of the matrix coefficients. $A$ and $B^j$ are $n \times n$ and $C$ is $n \times m$. For example,

$$A = \frac{\partial f}{\partial x} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix} \qquad 4$$

**B Linear Dynamic Models**

The equations above can be used to model the dynamics of a physical system. See Figure 2 for a simple illustration.

Figure 2 about here

A physical system can be modelled by a number of states and inputs. The states are contained in $x$, called the state vector, and inputs in $u$, the input vector. Generally, $x$ and $u$ can vary with time, denoted by $x(t)$ and $u(t)$. The number of states and inputs in the model are given by $n$ and $m$ respectively. Each state defines a co-ordinate in state-space within which the behaviour of the system is represented as a trajectory. The temporal evolution of the states is modelled by a state equation, which is the first order temporal derivative of the state vector, written as $\dot{x}(t)$ and can therefore be approximated by a Taylor series as above

$$\dot{x} = f_L(x,u) = Ax + Cu \qquad 5$$

A linear Dynamic System (LDS) is shown in Figure 3. The figure consists of two states, $x_1(t)$ and $x_2(t)$, and external inputs, $u_1(t)$ and $u_2(t)$, coupled through a state equation parameterised by matrices $A$ and $C$. As the model contains two states and two inputs, these matrices are both 2×2.

Figure 3 about here

*A* contains parameters that determine interactions among states (labelled inter-state in Figure 3) and the influence a states own activity has on itself (for example, the damping term in a linearly damped harmonic oscillator), while the elements of *C* couple inputs with states. The state equation describes the influences among states and their response to input, thereby providing a complete description of the dynamics of the system, otherwise known as the system's equation of motion. These models are sometimes called Linear Time Invariant (LTI) systems as *A* and *C* do not change with time. Eq(5) can be re-written as

$$z(t) = Jz(t) \qquad\qquad 6$$

Where *z*(t) is a $(n+1) \times 1$ vector and *J* is a $(n+1) \times (n+1)$ matrix. This is a linear equation, which can be solved using standard linear techniques (Boas, 1983), such as the matrix exponential method (see Appendix A.I)
.

**C Bilinear Dynamic Models**

Linear models have dominated scientific models, despite much nonlinearity around us, because they are good first approximations to many phenomena. However, they remain restricted and sometimes unrealistic descriptions. The use of bilinear models represents a substantial break with linearity. The model in Figure 4 has been augmented to illustrate the simple steps needed to formulate a bilinear model. The state equation can be modelled by an equation of the same form as $f_{BL}$, whose essential feature is the bilinear interaction involving the product of an input with a state. This can be written

$$\dot{x} = f_{BL}(x,u)$$
$$= (A + \sum_{j=1}^{m} u_j B^j)x + Cu \qquad\qquad 7$$
$$= \widetilde{A}x + Cu$$

The critical difference is the addition of $B$ that is now modulated by $u(t)$ which, when added to $A$, models input-dependent changes to the intrinsic connectivity of the network. This is illustrated in Figure 4 where the coupling coefficient $a_{21}$ is modulated by the product $b_{21}^2 u_2$. The modified matrix $\widetilde{A}$ operates on the state vector and determines the response of the model. The difference is that $\widetilde{A}$ changes with time because it is a function of time-varying input, which distinguishes it from the LTI model above.

It helps to consider a specific example. If $u_2$ is binary, the model in Figure 4 effectively consists of two LTI models. The models behaviour, i.e. $x(t)$, will be characterised by two linear systems switching from one to the other. The moment the system switches from one linear mode will be determined by changes in $u_2$. For instance, if the two linear models were linearly damped harmonic oscillators, each with markedly different characteristic behaviours, *i.e.* period of oscillation, a switch from one state to another would be accompanied by changes in the oscillation of the states. In short, the dynamics of $x$ represents a simple two-state system.


Figure 4 about here


The benefit of constraining the model to include only bilinear terms is that we have circumvented the issue of intractability of nonlinear models, yet retaining a uniquely nonlinear feature: input-dependent modulation of intrinsic dynamics. Inputs can now be divided into two classes: perturbing and contextual. Perturbing inputs (*e.g.* $u_1$ in Figure 3) influence states directly, without modulating model parameters. The effects of these inputs are distributed according to the intrinsic connections of the model, whereas contextual inputs (*e.g.* $u_2$ in Figure 3) reconfigure the response of the model to perturbations. Time and input-dependent changes in connectivity are a central feature of plasticity and are the motivation for using bilinear state-space models (BSSMs).


**D Metrics of connectivity**
Effective connectivity is defined as the influence a neuron (or neuronal population) has on another (Friston and Price, 2001). It is a dynamic quantity, used to identify degrees of

influence within a physical system, in response to external forces. At the neuronal level this is equivalent to the effect pre-synaptic activity has on post-synaptic responses, otherwise known as synaptic efficacy. Models of effective connectivity are designed to identify a suitable metric of influence among interconnected components (or regions of interest) in the brain. We shall see, throughout the chapter, how measures of effective connectivity identify dynamic structure within data, induced through experimental design and constrained by the operational principles at work within the brain, and how they emerge as a natural metric of plasticity.

Given the two-state BSSM network in Figure 4 each state's equation is given by

$$\dot{x}_1 = a_{11}x_1 + (a_{21} + b_{21}^2 u_2)x_2 + c_{11}u_1$$

$$\dot{x}_2 = a_{22}x_2 + a_{12}x_1 + c_{22}u_2$$

Taking derivatives of $\dot{x}$ with respect to each state

$$\frac{\partial \dot{x}_1}{\partial x_2} = k_{21}(x,u) = a_{21} + b_{21}^2 u_2$$

$$\frac{\partial \dot{x}_2}{\partial x_1} = k_{12}(x,u) = a_{12}$$

This operation discloses the coupling between the two regions (states) *i.e.* the direct influence one region has on another. Generally this may be linear or nonlinear, however, it is reduced to a simple function in a bilinear model, described by $k_{12}$ and $k_{21}$. The coupling from $x_1$ to $x_2$ is linear, represented by a constant term, $a_{12}$. However, the interaction between $u_2$ and $x_2$ induces nonlinearities in the network, rendering $k_{21}$ a function of $u_2$. The degree of influence $x_2$ has on $x_1$ therefore depends on $u_2$. This effect may be quantified by taking derivatives with respect to $u_2$, the contextual input.

$$\frac{\partial^2 \dot{x}_1}{\partial x_2 \partial u_2} = h(x,u) = b_{21}^2$$

As for the first-order derivative, $h(x,u)$, a second order derivative, may be a nonlinear function of the states and input for an arbitrarily nonlinear equation of motion. However, it reduces to a constant term in the bilinear model.

The first and second-order derivatives quantify dynamic characteristics within a network. Therefore they are equivalent to first and second order effective connectivity, or obligatory and modulatory influences (Bûchel and Friston, 2000). This distinction highlights two important issues: the difference between perturbing and contextual input and how to derive a practical measure of connectivity. As illustrated in Figure 4, external input can be categorised by way of its effect on the intrinsic states of a system. Either input modulates a systems intrinsic connectivity or it perturbs the states directly. The former we have coined contextual and the latter perturbing. Both evoke a response. However, contextual inputs enable the model to represent contextual changes by modulation of the intrinsic connectivity. This is a subtle but crucial difference. Practically, models of this nature can be used to infer levels of effective connectivity through estimating parameters such as $a_{12}$ and $b_{21}^2$ from real data.

A caveat is necessary at this point. All models require some form of *a priori* knowledge. Factoring this into a model fairly will always be cause for some debate. This is because what is 'fair' for one model may not be for another. However, for progress to be made a strategy has to be formulated and appraised to assess its value in the context of the data. Models of effective connectivity generally require an anatomical model to specify *which* regions are connected. A simplified but sufficient anatomical model can be based on lesion studies or anatomical data from animal models. A mathematical model, such as that described above *i.e.* an equation of motion representing the brain as a connected physical system, is necessary to model *how* the different regions in the anatomical model interact. These mathematical models can become very complicated and mathematically intractable. Models can be simplified by approximating methods, as we saw with the bilinear model, and making assumptions, such as ignoring temporally distant effects of neuronal events *i.e.* assuming only instantaneous effects of disparate brain regions on each other. This may sound a little abstract at the moment but we will see later how this assumption allows us to make progress by rendering the models

11

tractable. Even though simple models may be criticised for being biological implausible, progress can be incremental. In the next section we will discuss various models for measuring connectivity from PET and fMRI data, highlighting their heuristics, and assumptions.

**E Relevance to neurophysiology**

Figure 5 portrays a model of the visual and attention systems. The model posits photic stimulation as an input, or cause in the environment, which perturbs the brain evoking a response that depends on its current state of connectivity. This connectivity embodies a context, such as attentional set or memory (*i.e.* whether the stimulus is salient). Perceptual learning, or changes in attention induce a reconfiguration of synaptic efficacies and connectivity in terms of ensemble responses (see **Chapter 22**: **Dynamic Causal Modelling**). In the model, these are examples of contextual input that enable the brain to respond differently to the same perturbing stimulus *i.e.* the difference between a novel visual image or the recognition of a stimulus that has recently become salient.

   Having described the basis for modelling the brain as a physical interconnected system and establishing a fundamental distinction between inputs that change states and those that change [parameters] connections, we now turn to some specific examples. We will start with simple models and work back towards the dynamic models described in this section.

<center>Figure 5 about here</center>

III LINEAR MODELS OF EFFECTIVE CONNECTIVITY

The objective of an effective connectivity analysis is to estimate parameters that represent influences among regions that may changes over time, and with respect to experimental tasks. Neuroimaging data is usually processed into a voxel-based time-series representation of an index of neuronal activity (rCBF or BOLD for PET and fMRI respectively). Structure exists in the experimental design, neurophysiological data and critically the theoretical assumptions used to model the observed responses. For

example, the spatial and temporal order within the data provide essential insights into the underlying generative processes and is clearly the rationale of an empirical approach. Models also contain structure, as they represent the theoretical constructs and assumptions needed to identify operational principles responsible for generating the data.

Identifying a complete and biologically plausible mathematical model requires a high level of sophistication. However, some progress can be made by modelling relationships in the data alone (among voxels or regions), without invoking hidden states and ignoring the consequent temporal correlations. This last simplification makes the mathematics much easier but discards temporal information and is biologically unrealistic. We will call these models 'static' as they model interactions among regions as occurring instantaneously and do not encompass the influence previous states have on current responses. Static models are reviewed before turning to more realistic, 'dynamic' models. Both are important in the development of a plausible metric of effective connectivity, the former establishing an historical benchmark to validate the latter. In what follows we describe the development of approaches, demonstrated through different models, designed to represent dynamic interactions as measured through neuroimaging data.

**A Linear Models**

After measuring an index of neuronal activity at each voxel in the brain, over the duration of an experiment, the next step is to assess, on the basis of these data, if there is any reason to believe that different regions of the brain influence each other.

Our first model is linear and assumes that $y_t^i$ is statistically independent of $y_{t-\tau}^i$ for arbitrary $\tau$. This is a valid assumption for PET data as the sampling rate is slow relative to neuronal dynamics. In fact, data are acquired while holding brain states constant using an appropriate task or stimulus. Each measurement is therefore assumed to represent some average brain state. Mathematically, this means the rate of change of the states is assumed to be zero. For fMRI time-series, however, this assumption is generally violated (and certainly for electrophysiological measurements). As the sampling rate of measurement increases, so does the dynamic character of the data. This is the motivation for dynamic models of fMRI responses.

In static linear models the activity of each voxel is modelled as a linear mixture activity in all voxels plus some error. This can be written

$$y_t^i = \beta_{1i} y_t^1 + \beta_{2i} y_t^2 + \ldots + \varepsilon_t^i$$
$$= y_t \beta_i + \varepsilon_t^i$$
$$\beta_i = \begin{bmatrix} \beta_{1i} \\ \vdots \\ \beta_{Ni} \end{bmatrix}$$

$$y_t = y_t \beta + \varepsilon_t$$
$$y_t = [y_t^1, \ldots y_t^N]$$
$$\beta = [\beta_1, \ldots, \beta_N]$$

$$Y = Y\beta + \varepsilon$$
$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_T \end{bmatrix}$$

8

for one region at one time, all regions at one time, and all the data, respectively. The problem with this formulation is that the trivial solution $\beta_{ii} = 1$ completely accounts for the data. We will see later how structural equation modelling deals with this by setting self-connections to zero. However, one also can finesse this problem using singular value decomposition (SVD). Multivariate techniques like SVD are usually thought of as summarising the covariance structure of data or functional connectivity (see the previous chapter). The ensuing eigenimages can be regarded as spatial modes that are functionally and effectively disconnected from each other.

Figure 6 illustrates a SVD of an arbitrary matrix $M$ (of size $n \times m$, where $n > m$) into an equivalent form using the relation $MV = US$, where $V$ is a set of eigenvectors that form a natural co-ordinate system.


Figure 6 about here


14

A plot of the variance components of an SVD, for a typical PET data set (that used in the previous chapter: **Functional Connectivity**), is shown in Figure 6. The spectrum shows a rapid decrease in the eigenvalues after the first few components (Friston *et al.,* 1993). This indicates spatiotemporal order within the data[1]. Such order is a consequence of coherent modes of distributed activity induced by task-related changes that are responsible for generating the data. One could regard the eigenvalues $\lambda = SS^T$ as indices of self-functional connectivity. However, static linear models also afford a perspective on effective connectivity. One can take these eigenimages to represent their influence on a voxel-specific measurement at the $i^{th}$ voxel. If the expression of the modes is given by $Y = US$ then

$$y^i = Y\beta_i = USV_i^T$$
$$\beta_i = V_i^T$$

where $V_i^T$ is the *i*th row of *V*. Compare this with Eq(8). In other words, the eigenimages can be interpreted as the effective connectivity between the voxel in question and the mode corresponding to the eigenimage.

 In summary, this perspective on eigenimage analysis furnishes a measure of connectivity between the voxel i and the rest of the brain in terms of spatiotemporal eigenmodes, as opposed to separate voxels. However, this interpretation is limited to linear, time invariant connection strengths. Next we describe how linear models can be used to approximate nonlinear systems by including interaction, or bilinear, terms.

**B Modelling nonlinearities**

Linear models cannot be used to estimate *changes* in connectivity (induced by modulatory interactions among populations of neurons or contextual inputs). However, introducing interaction terms can attain a dynamic representation. This requires a simple extension to Eq(8); the addition of a new variable, calculated from the product of two

---

[1] A white noise process would have a homogeneous eigenvalue spectrum.

voxels or regions (indexed by $j$ and $k$). We will refer to this new variable as the bilinear term, written as $y_t^j y_t^k$ for scalar and $y^j \times y^k$ (the Hadamard product) for vectors. The idea is that now the model can be used to estimate the effect this new term has on activity in the $i^{th}$ voxel.

$$y_t^i = y_t^j b_1 + y_t^j y_t^k b_2 + e_t^i \qquad\qquad 9$$

where $b_1$ and $b_2$ are model parameters, which scale the effect of their respective terms. The quantity $b_1$ is equivalent to an element in $\beta_{ji}$ (Equation 8), however, $b_2$ is different as it parameterises the effect of a bilinear term $y_t^j y_t^k$. These two coefficients are the obligatory and modulatory effects discussed earlier. Obligatory connections have a direct driving effect, but modulatory connections are subtler and introduce context into the model. This is best illustrated with an example.

The model in Figure 7 (upper panel) is a simple nonlinear model. It consists only of two variables but this is sufficient to illustrate nonlinearity and how we can measure it. It consists of an input $x$, which generates an output $y$. The nonlinearity is due to an interaction between the input and output $i.e.$ the model's response $y$, depends on its current (intrinsic) activity $y$. This simple model consists of a linear and bilinear term parameterised by $b_1$ and $b_2$ respectively.

$$y = xb_1 + xyb_2 \qquad\qquad 10$$

If the model contained the first term only, it would be linear and the relationship between input and output would be represented by a straight line in a plot of $x$ and $y$. The addition of the second term introduces markedly nonlinear behaviour. Plotting $x$ and $y$ for different values of $b_2$ demonstrates this. The input-output behaviour depends on $b_2$ and is reflected in the two different curves in Figure 7 (lower panel). It helps to focus on the model's response to a fixed input. It is easily appreciated that the sensitivity of $y$ to $u$ (the slope) depends on $b_2$. The key point is that data generated from such a process is not

distributed in a linear fashion, *i.e.* the dependence of $y$ on $x$ is not modelled with a straight line. We will describe next a piece-wise local linear approximation to modelling this very simple sort of nonlinearity.

First we need to consider how to estimate the parameters in Eq(10) from measurements of $x$ and $y$. If we assume that the nonlinearities are weak about a local region of the data then the model can be approximated by

$$y_l \approx x(b_1 + \langle y_l \rangle b_2) = xb_l \qquad\qquad 11$$

where $y_l$ is a vector containing values of $y$ within a local range of the data and $\langle y_l \rangle$ is its average. If we assume that $b_1$ and $b_2$ are constant within this region then $b_1 + \langle y \rangle b_2$ is also constant (abbreviated to $b_l$) and can be estimated from values of $x$ and $y_l$. If $y$ is partitioned into $n$ divisions, then this procedure can be repeated for all $n$ domains. Differences in values of $b_l$ for different regions indicate nonlinearity in the data.

Figures 6, 7 and 8 about here

Let us turn to the illustration in Figure 8. These data have been generated from the model in 7, with $b_2 = 0.8$, with added noise. The data is distributed, roughly, into two sub-groups. Modelling the data as a linear process does not capture this feature. The characterisation can be finessed by partitioning the data, on the basis of 'high' and 'low' levels of response (denoted by $y_{high}$ and $y_{low}$). A linear model, given by Eq(11) can then used each partition.

$$y_{high} \approx x(b_1 + \langle y_{high} \rangle b_2) = xb_{high} \qquad\qquad 12$$

$$y_{low} \approx x(b_1 + \langle y_{low} \rangle b_2) = xb_{low}$$

where we have assumed $b_1$ and $b_2$ are constant within the domains of high or low activity. We then use the two values of $\langle y_l \rangle$ and $\mathrm{b}_l$ to estimate values of $b_1$ and $b_2$, where

$$b_1 = \frac{\langle y_{high} \rangle b_{low} - \langle y_{low} \rangle b_{high}}{\langle y_{high} \rangle - \langle y_{low} \rangle}$$

$$b_2 = \frac{b_{low} - b_{high}}{\langle y_{low} \rangle - \langle y_{high} \rangle}$$

13

These coefficients serve as an index of linearity ($b_1$) and nonlinearity ($b_2$), and can be estimated from measured data. Applying these ideas to neuroimaging data, $b_1$ and $b_2$, represent measures of obligatory and modulatory connectivity. This approach of approximating nonlinear interactions was used in (Friston *et al.*, 1995) to demonstrate asymmetrical nonlinear interactions between V1 and V2 from fMRI data during visual stimulation.

The inclusion of a bilinear term enabled us to introduce nonlinear behaviour into our model. The method of piece-wise linear approximation provided a way to measure this nonlinearity. However, the bilinear term was only implicit. Next, we model the bilinear term explicitly in a general linear model. These models have been called psychophysiological (and physio-physiological) interaction (PPI) models.

**C Psychophysiological Interactions**

The partitioning of data required above may seem arbitrary. A way around this is to embed the interaction term into one linear model. (Bûchel *et al.*, 1996) discussed including a series of increasingly high-order interaction terms in a general linear model, each of which is constructed from the products of individual variables. These are introduced as new explanatory variables and provide a means of modelling the difference in regression slopes without partitioning or splicing the data by hand. In this way, standard linear regression techniques, implemented in SPM, can be used to estimate the magnitude and significance of these bilinear effects directly.

Changing the symbols representing the $j^{th}$ and $k^{th}$ voxels in the previous equation to $u_1$ and $u_2$, denoting input to a system whose response is $y^i$, the prototype linear model is

$$
\begin{aligned}
y^i &= X_L \beta_L + e^i \\
&= [u_1, u_2, X_0] \beta_L + e^i
\end{aligned}
$$

14

The equation has been simplified by dividing the explanatory variables into those of interest and $X_0$, which contains all the other covariates. The model's parameters quantify the influence each explanatory variable has on $y^i$ and can be estimated using standard linear techniques. Extending the model to include a bilinear variable, denoted by $u_1 \times u_2$ (the Hadamard product), gives

$$
\begin{aligned}
y^i &= X_{BL} \beta_{BL} + e^i \\
&= [u_1 \times u_2, u_1, u_2, X_0] \begin{bmatrix} \beta_I \\ \beta_L \end{bmatrix} + e^i
\end{aligned}
$$

15

The model is divided into bilinear term and linear terms. The corresponding parameters, $\beta_I$ and $\beta_L$ can be estimated as before. Both the main-effect and interaction terms are included because the main effect of each covariate has to be modelled to properly assess the additional explanatory power afforded by the bilinear or PPI term. Standard hypothesis testing of the bilinear term (testing $H_0 : \beta_I = 0$) can be used to estimate the significance of its effect. In this analysis, the data does not have to be divided, as input-dependent changes in regression slope are modelled by the bilinear term.

Figure 9 illustrates two examples of bilinear effects in real data. The study was a fMRI experiment investigating the modulatory effects of attention on visual responses to radial motion (see the figure legend and Büchel and Friston, 1997 for experimental details). The aim of both models was to quantify a top-down modulatory effect of attention on V1 to V5 connectivity. The left-hand model combines psychological data (attentional set) with physiological data (V1 and V5 activity) to model the interaction, whereas the right-

hand model uses a physiological measure (PPC activity) a surrogate for the psychological effect. These analyses correspond to psychophysiological and physio-physiological interactions respectively. Both demonstrate a significant modulatory effect of attention (Büchel and Friston, 1997; Friston et al., 1997). The lower diagram in Figure 9 is a regression analysis of the same data, divided according to attentional set, to demonstrate the difference in regression slopes.

Figure 9 about here

In this example, attention was modelled as a 'contextual' variable, while visual stimulation perturbed the system. The latter evoked a response within the context of the former. Using the example in Figure 9, visual stimuli evoke different responses depending on attentional set, modelled as a change in connectivity. Attention appears to reconfigure connection strengths among prefrontal and primary cortical areas (Mesulam, 1998). The bilinear effect may take any appropriate form in PPI models, including, for example, psychological, physiological or pharmacological indices. These models emphasise the use of factorial experimental designs (Friston *et al*., 1997) and allow us to consider experimental inputs in a different light, distinguishing contextual input (one factor) from direct perturbation (another factor). PPI models have provided important evidence for the interactions among distributed brain systems and enabled inferences about task-dependent plasticity using a relatively simple procedure. The next model we consider was developed explicitly with path analysis in mind but adopts a different approach to the estimation of model parameters. This approach rests on the specification of priors or constraints on the connectivity.

**D Structural Equation Modelling**

Structural Equation Modelling (SEM), or path analysis, is a multivariate tool that is used to test hypotheses regarding the influences among interacting variables. Its roots go back to the 1920s when Path Analysis was developed to quantify unidirectional causal flow in genetic data and developed further by social scientists in the 1960s (Maruyama, 1998). It received criticism for the limitations inherent in the least squares method of estimating

model parameters, which motivated a general linear modelling approach from the 1970s onwards. It is now available in commercial software packages including LISREL, EQS and AMOS. See (Maruyama, 1998) for an introduction to the basic ideas. Researchers in functional imaging started to use it in the early 1990s (McIntosh and Gonzalez-Lima, 1991;McIntosh and Gonzalez-Lima, 1992a; McIntosh and Gonzalez-Lima, 1992b;McIntosh and Gonzalez-Lima, 1994). It was applied first to animal auto-radiographic data and later to human PET data, where, among other experiments, it was used to identify task-dependent differential activation of the dorsal and ventral visual pathways (McIntosh et al., 1994). Many investigators have used SEM since then. An example of its use to identify attentional modulation of effective connectivity between prefrontal and premotor cortices can found in (Rowe *et al*., 2002).

A SEM relates to the general linear model above in that it has the same form. There are, however, a number of modifications, some of which are illustrated in Figure 10. The coupling matrix $\beta$, has been 'pruned' to include only paths of interest. Critically, self-connections are precluded. The data matrix, *Y*, contains responses from regions of interest and possibly experimental or bilinear terms. The model is

$$ y_t = y_t \beta + z_t \qquad\qquad 16 $$

The regional time-series *Y are* known but $\beta$ contains free parameters to be estimated. To simplify the model the residuals $z$ are assumed to be independent. They are interpreted as driving each region stochastically from one measurement to another and, to reflect this, are sometimes called innovations.

Figure 10 about here

The free parameters are estimated using the covariance structure of the data, instead of minimising the sum of squared errors as described above. The rationale is that the former reflects the global behaviour of the data, *i.e.* capturing relationships among variables, in contrast to the latter, that reflects the goodness of fit from the point of view of each region. Practically, a cost function is constructed from the actual and implied covariance,

which is used as an objective function to estimate parameters. The implied covariance, $\langle y_t^T y_t \rangle$, is easily computed by rearranging Eq(16) and assuming some value for the covariance of the innovations, $\langle z_t^T z_t \rangle$

$$
\begin{aligned}
y_t(1-\beta) &= z_t \\
y_t &= z_t(1-\beta)^{-1} \\
\langle y_t^T y_t \rangle &= (1-\beta)^{-1T} \langle z_t^T z_t \rangle (1-\beta)^{-1}
\end{aligned}
\tag{17}
$$

Details are provided in Appendix A.2. A gradient descent such as a Newton-Raphson scheme may be used to estimate parameters, where starting values can be estimated using OLS (McIntosh and Gonzalez-Lima, 1994).

Inferences about path coefficients rest on the notion of nested, or stacked, models. A nested model consists of a free model within which any number of constrained models are 'nested'. In a free model, all parameters are free to take values that optimise the objective function, whereas a constrained model has one, or a number of parameters omitted, constrained to be zero or equal across models (*i.e.* attention and non-attention). By comparing the goodness of fit of each model against the others, $\chi^2$ statistics can be derived (Bollen, 1989). Hypotheses testing proceeds using this statistic. For example, given a constrained model, which is defined by the *omission* of a pathway, hypothesis testing may be construed as evidence for or against the pathway by 'nesting' it in the free model. If the difference in goodness of fit is highly unlikely to have occurred by chance, the connection can be declared significant. Examples of models used by Büchel *et al.* using the attentional data set are shown in Figure 11. Nonlinear SEM models are constructed by adding a bilinear term as an extra node. A significant connection from a bilinear term represents the modulation of influence in exactly the same way as in a PPI. Büchel *et al* used SEM (Büchel and Friston, 1997) on the visual attention data set, validating the method by confirming conclusions reached using other regression models.

Figure 11 about here

SEM is a regression analysis, which means that it shares the same deficiencies as the linear model approach described above *i.e.* temporal information is discounted[2]. However, it has enjoyed relative success and become established over the past decade, due in part, to its commercial availability as well as its intuitive appeal. However, it usually requires a number of rather *ad hoc* procedures, such as partitioning the data to create nested models, or pruning the coupling matrix to render the solution tractable. These problems are confounded with an inability to capture nonlinear features and temporal dependencies. By moving to more sophisticated models we acknowledge the effect of the history of input and embed *a priori* knowledge into models at a more plausible and mechanistic level. These issues will be addressed in the following section.

## IV DYNAMIC MODELS

The static models described above discount temporal information. Consequently, permuted data sets produce the same path coefficients as the original permutation. Models that use the order in which data are produced are more natural candidates for the brain and include those that attempt to model its equations of motion, *e.g.* state-space models and generalised convolution models. In this section we will review Kalman Filtering, autoregression and generalised convolution models.

### A Kalman Filter

The Kalman filter is used extensively in engineering to model dynamic data (Juang, 2001). It is based on a state-space model that invokes an extra set of [hidden] variables to generate data. These models are a powerful as long-range order, within observed data, is modelled through interactions among hidden states, instead of mapping input directly onto output (see below). It is an 'online' procedure consisting of two steps: prediction and correction. The hidden states are estimated (prediction step) using the current information, which is updated (correction step) on receipt of each new measurement.

---

[2] There exist versions of SEM that do model dynamic information, see (Cudeck, 2002) for details of Dynamic Factor Analysis

These two steps are repeated recursively as new information arrives. A simple example demonstrates intuitively how the filter works. This example is taken from the source in reference (Ghahramani, 2002).

Consider a series of data points, which we receive one at a time. Say we wanted to calculate a running average with each new data point. Given that the $t^{th}$ variable is $x_t$ and the estimate of the mean (which we will call the 'state') after $t$ values of $x$ is $\hat{x}_t$

$$\hat{x}_t = \frac{1}{t} \sum_t x_t$$

for $t$-1 samples
$$\hat{x}_{t-1} = \frac{1}{t-1} \sum_{t-1} x_{t-1}$$

$\hat{x}_t$ is related to $\hat{x}_{t-1}$ by
$$\hat{x}_t = \frac{t-1}{t} \hat{x}_{t-1} + \frac{1}{t} x_t$$

This can be rearranged, giving $\hat{x}_t = \hat{x}_{t-1} + K(x_t - \hat{x}_{t-1})$. where $K$ is the Kalman Gain = $1/t$. This illustrates the general form of a Kalman Filter: a prediction and a weighted correction. This general mathematical form may be expressed in words as

$$\begin{pmatrix} Current\ estimate \\ \hat{x}_t \end{pmatrix} = \begin{pmatrix} Previous\ estimate \\ \hat{x}_{t-1} \end{pmatrix} + \begin{pmatrix} Weighted\ (K)\ prediction\ error \\ (x_t - \hat{x}_{t-1}) \end{pmatrix}$$

The filter balances two types of information from a prediction (based on previous data) and an observation, weighted by their respective precisions. This is performed optimally as the weighting using Bayes rule. If the measured data is not reliable $K$ goes to zero, weighting the prediction error less and relying more on the preceding prediction to afford a current one. Conversely, if the dependence on sequential state values is unreliable, then $K$ is large. This emphasises information provided by the data when constructing an estimate of the current state.

Quantities that require calculation in the forward recursion are Kalman Gain, mean and co-variance of prediction and correction matrices. A backward recursive algorithm called the Kalman Smoother calculates the mean and co-variance of the states at time t, given all the data, which is a *post hoc* procedure to improve the estimates. Maximum likelihood estimators may be used to calculate these quantities. A demonstration of the filter is shown in Figure 12, applied to the visual attention data.

Static models can only model 'snap-shots' of path coefficients (although a 'snap-shot' of a bilinear term introduces time-dependence through, for example, changes in PFC activity or attentional set). In contrast, by modelling the path coefficient as a hidden state, to be estimated from observed data, the filter exposes fluctuations in the coupling, between V1 and V5, which rises and falls with attention (even though attentional status never entered the model). To understand how the filter was applied we start with our familiar linear observation model:

$$
\begin{aligned}
y_t &= x_t \beta_t + \varepsilon_t \\
\varepsilon_t &\sim N(0,R)
\end{aligned}
\qquad 18
$$

To simplify notation *y* and *x* are univariate and are both known *e.g.* BOLD activity from V1 and V5. $\beta$ is modelled as a state variable, observed vicariously through the BOLD responses, that is allowed to change with time (influenced by its own internal states and input) according to the update equation

$$
\begin{aligned}
\beta_t &= \beta_{t-1} + \eta_t \\
\eta_t &\sim N(0,Q)
\end{aligned}
\qquad 19
$$

Where the subscript indexes the scan number, as $\beta_t$ can vary from scan to scan. Given this extra state variable, the Kalman Filter can be used to estimate its dynamics. Figure 12 illustrates the model and plots the time dependent path coefficient, $\beta_t$, for V1 to V5 connectivity. Note that if $\eta_t = 0$ then $\beta_t = \beta_{t-1}$. This is the static estimate from an ordinary regression analysis.

The results of this analysis, also known as variable parameter regression, agree with previous analyses, in that there is task dependent variation in inter-regional connectivity. This variation has the same form as described by a BSSM of contextual input, but the variation is treated as random and unknown. See Eq(19). However, in general these changes in connectivity are induced experimentally by known and deterministic causes.

In **Chapter 21** (**Dynamic Causal Modelling**) we return to state space models and reanalyse the attentional data set in a way that allows designed manipulations of attention to affect the models hidden states. In dynamic casual modelling the states are dynamic variables (*e.g.* neuronal activity) and the effective connectivity corresponds to fixed parameters than can interact with time-varying states and inputs[3]. The remainder of this chapter focuses on approaches that do not refer to hidden states, such as autoregression and Generalised Convolution Models.

Figure 12 about here

**B Multivariate Autoregressive Models**

We are familiar with the notion that a time series of sequential measurements may contain temporal information*: i.e.* the order in which the data are generated is important. Temporal correlations therefore provide insight into the physical mechanisms generating them. A simple and intuitive model of temporal order is an autoregressive (AR) model, where the value of a variable at time $t$ depends on preceding values, up to time lag $t$-$p$ (where $p < t$).

Parameters of AR models comprise regression coefficients, at successive time lags, that estimate the characteristic sequential dependencies of the system in a simple and effective manner, using measured data only. This can be extended to include several variables with dependencies among variables at different lags. These dependencies may be interpreted as the influence of one variable on another and can be recruited as measures of effective connectivity. Models involving many variables are called Multivariate

---

[3] This should be contrasted with the above application in which the connectivity itself was presumed to be a time-varying state.

Autoregressive (MAR) models and have been used to measure dependencies among regional activities as measured with fMRI (Harrison *et al*., 2003).

MAR models do not invoke hidden states. Instead, correlations among measurements at different time lags are used to quantify the relationships. This incorporates history into the model in a similar way to the Volterra approach described below. MAR models are linear but can be extended to include bilinear interaction terms. To understand MAR we will build up a model from a univariate AR model and eventually see that they conform to GLMs with time-lagged explanatory variables.

Consider one data at voxel *i* at time *t* modelled as a linear combination of time lags from *t*-1 to *t*-*p*.

$$
y_t^i = \left[ y_{t-1}^i, \dots, y_{t-p}^i \right] \begin{bmatrix} w_1 \\ \vdots \\ w_p \end{bmatrix} + e_t
\qquad\qquad 20
$$

*w* is a *p*×1 column vector containing the model parameters (AR coefficients) and $e_t$ is Gaussian noise. The model of $y_t$ is a linear mixture of *preceding* values. This has the same form as the model in Eq(8) where the explanatory variables are now values over different time lags, instead of voxels. We can extend the model to *d* regions contained in the row vector $y_t = \left[ y_t^1, \dots, y_t^d \right]$

$$
y_t = \sum_{j=1}^{p} \left[ y_{t-j}^1, \dots, y_{t-j}^d \right] \begin{bmatrix} w_j^{11} & \cdots & w_j^{1d} \\ \vdots & \ddots & \\ w_j^{d1} & & w_j^{dd} \end{bmatrix} + \left[ e_1, \dots, e_d \right]
\qquad 21
$$

$$
= \sum_{j=1}^{p} y_{t-j} W_j + \varepsilon
$$

which has *d*×*d* model parameters at each time lag, describing interactions among all pairwise combinations of variables. This is simple a GLM whose parameters can be

estimated in the usual way to give *W*, which is a *p*×(*d*×*d)* array of AR coefficient matrices.  As in previous sections the model can be augmented with bilinear interaction terms as in SEM.  See Figure 13 for a schematic of the model.  There are no inputs to the model, except for the error terms, which play the role of innovations (*c.f.* SEM).  This means that experimentally designed effects have no explicit role (unless they enter through bilinear terms).  However, the model attempts to identify relations between variables over time, which distinguishes it from static linear models of effective connectivity.


Figure 13 about here


The magnitude of *p*, or order of the model, becomes an issue when trying to avoid over fitting the data.  This is a common problem because a higher-order model will explain more of the data, in a least squares sense, without necessarily capturing the dynamics of the system any better than a more parsimonious, optimal model.  A procedure for choosing an optimal value of *p* is therefore necessary.  This can be achieved using a Bayesian approach (Penny and Roberts, 2002).  A Bayesian framework also allows for inferences about connection strengths to be made based on posterior probabilities (see **Chapter 17: Classical and Bayesian Inference**).

   MAR was used to model the visual attention data.  The results are shown in Figure 14. Two models were estimated using three regions in each.  The motive was to validate the method against established procedures that demonstrated a modulatory influence of PFC on V5 to PPC connectivity and PPC on V1 to V5 connectivity.  The posterior densities of the weight matrix *W* are shown using their conditional means and variances. The probability that an individual parameter is different from zero can be inferred from these posterior distributions.  Parameters whose conditional density encompasses zero are less likely to have any influence.  Conversely, the more distal a density mass is from zero, the greater our certainty that the model supports an effect.  Non-zero parameters that characterise second order connectivity are circled in the figure.

Figure 14 about here

MAR models have not been used as extensively as other models of effective connectivity. However they are an established technique for quantifying temporal dependencies within time series (Chatfield, 1996). They are simple and intuitive models requiring no *a priori* knowledge of connectivity, as in SEM. However, this could also be construed as a shortcoming, in that simple MAR models cannot harness *a priori* knowledge.

**C Generalised Convolution Models**

Up until now we have considered models based on the general linear model and simple state space models. The former may be criticised for not embracing temporal information within data, which the Kalman Filter (an example of the latter) resolved by invoking hidden states. An alternative approach is to exclude hidden states and formulate a function that maps the history of input directly onto output. This can be achieved by characterising the response (output) of a physical system over time to an idealised input (an impulse), called an Impulse Response Function (IRF) or Transfer Function (TF). In the time domain this function comprises a kernel that quantifies the idealised response. This is convenient as it bypasses any characterisation of possible internal states generating the data. However, it renders the system a 'black box', within which we have no model. This is both the methods strength and weakness.

Once the IRF has been characterised from experimental data it can be used to model responses to arbitrary inputs. For linear systems, adherent to the Principle of Superposition, this reduces to convolving the input with the IRF. The modelled response depends on the input, without any reference to the interactions that may have produced it. An example, familiar to neuroimaging, is the Hemodynamic Response Function (HRF) used to model the hemodynamic response of the brain to experimental tasks. However, we are interested in nonlinear models, which are obtained by generalising the notion of convolution models to include high-order interactions among inputs, an approach originally developed by Volterra in 1930 (Rieke et al., 1997).

The generalised nonlinear state and observation equations are, respectively

$$\dot{x}(t) = f(x(t), u(t))$$
$$y(t) = g(x(t), u(t))$$

22

These can be reformulated to relate output, $y(t) = h(u(t - \tau))$ to input $u(t)$, without reference to the states $x(t)$. Where $h$ is a nonlinear function, which can be expanded into a series of functionals (functions of functions, see below)

$$y(t) = h_0 + \int_{-\infty}^{\infty} h_1(\tau_1) u(t - \tau_1) \partial \tau_1 + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_2(\tau_1, \tau_2) u(t - \tau_1) u(t - \tau_2) \partial \tau_1 \partial \tau_2 + \cdots$$

$$+ \int_{-\infty}^{\infty} h_n(\tau_1, \cdots, \tau_n) u(t - \tau_1) \cdots u(t - \tau_n) \partial \tau_1 \cdots \partial \tau_n$$

23

where $n$ is the order of the series and may take any positive integer to infinity. This is known as a Volterra series. Under certain conditions, $h$ converges as $n$ increases (Fliess *et al.*, 1983) and can provide a complete description of a system given enough terms. To understand this we need to consider the Taylor Series expansion as a means of constructing an approximation to a general nonlinear function (see Figure 1). Any sufficiently smooth nonlinear function can be approximated, within the neighbourhood of an expansion point $x_0$, by scaling increasingly high-order terms of coefficients computed from derivatives of the function about $x_0$ (see Equation 1). The Volterra series is a Taylor series expansion, where high order terms are constructed from variables modelling interactions and scaled by time-varying coefficients. The Volterra series is a power-series expansion where the coefficients of Eq(1) are now functions, known as kernels. The kernels are functions of time and as the series involves functions of functions they are known as functionals.

An increase in accuracy of the approximation is achieved by considering higher order terms, as demonstrated by deriving linear and bilinear SSMs. The same is true of the linear and bilinear convolution models

$$y(t) \approx h_0 + \int h_1(\tau_1) u(t - \tau_1) \partial \tau_1$$

$$y(t) \approx h_0 + \int h_1(\tau_1) u(t - \tau_1) \partial \tau_1 + \int \int h_2(\tau_1, \tau_2) u(t - \tau_1) u(t - \tau_2) \partial \tau_1 \partial \tau_2$$

The HRF derives from a linear convolution model. The systems IRF and the occurrences of experimental trials are therefore represented by $h_1$ and $u(t)$, respectively. The linear model is distinguished by its compliance with the Principle of Superposition: given two input impulses, the response is simply the sum of the two responses. By including the second-order kernel, non-additive responses can be modelled. Practically, this means that the timing of inputs is important in that different pairs of inputs may produce different responses. The Volterra approach is a generalisation of the convolution method, which convolves increasingly high-order interactions with multidimensional kernels to approximate the non-additive component of a systems response.

   Kernels scale the effect each input, in the past, has on the current value of $y(t)$. As such, Volterra series have been described as 'power series with memory'. Sequential terms embody increasingly complex interactions among inputs up to arbitrary order. The series converges with increasing terms, which, for weakly nonlinear systems, is assumed to occur after the second order term. Nonlinear behaviour is modelled using these interactions, scaled throughout their history. A diagram of a bilinear convolution model is shown in Figure 15.


Figure 15 about here


The first and second-order kernels quantify the linear and bilinear responses, consequentially they are equivalent to first and second-order effective connectivity respectively (Friston, 2000). The kernels are also related mathematically to the bilinear State-space representation (see **Chapter 23: Mathematical Appendix**). For every state-space representation there is an equivalent set of kernels and an equivalent generalised convolution representation.

Having established that Volterra kernels are a metric of effective connectivity, we need to estimate them from experimental data. By reformulating the model using an appropriate basis set, kernels can be reconstructed from estimated coefficients. The HRF is modelled well by Gamma functions and this is the reason for choosing them to approximate Volterra kernels. The Volterra kernels, for a general dynamic system (of any arbitrary form or complexity), are difficult to compute unless the underlying generative process leading to the data is well characterised, as is the HRF.

A bilinear convolution model can be reformulated by convolving the basis set $b_i$ with the inputs.

$$x_i(t) = \int b_i(\tau_1) u(t - \tau_1) \partial \tau_1$$

These are then used as explanatory variables in a GLM

$$y(t) = g_0 + \sum_{i=1}^{n} g_i x_i(t) + \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} x_i(t) x_j(t) \qquad\qquad 24$$

where $y(t)$ and $x_i(t)$ are known and $g_0, g_i$ and $g_{ij}$ are to be estimated. The kernels are given by

$$h_0 = g_0$$
$$h_1(\tau_1) = \sum_{i=1}^{n} g_i b_i(\tau_1) \qquad\qquad 25$$
$$h_2(\tau_1, \tau_2) = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} b_i(\tau_1) b_j(\tau_2)$$

The unknowns in Eq(24) can be estimated using a GLM and used to reconstruct the original kernels from Eq(25). This method was applied to the attentional data set used previously. The model consisted of inputs from three regions: putamen, V1/2 complex and PPC, to V5, as shown in Figure 16. BOLD recordings from these regions were used

as an index of neuronal activity, representing input to V5.  The lower panel illustrates responses to simulated inputs, using the empirically determined kernels.  It shows the response of V5 to an impulse from V2 and provides a direct comparison of V5 responses to the same input from V2, but with and without PFC activity.  The influence of PFC is clear.  This is due to its modulatory influence on V2 to V5 connectivity and is an example of second order effective connectivity.

Figure 16 about here

The Volterra method has many useful qualities.  It approximates nonlinear behaviour without arbitrarily partitioning the data using temporal information within the familiar framework of a generalised convolution model.  Kernels may be estimated using a GLM and inferences made under parametric assumptions.  Furthermore, kernels contain the dynamic information we require to measure effective connectivity.  However, kernels characterise an ideal response, generalised to accommodate nonlinear behaviour, which is effectively a summary of the system as a whole.  It should also be noted that Volterra series are only local approximations around an expansion point.  Although they may be extremely good approximations for some systems they may not be for others.  For instance, Volterra series cannot capture the behaviour of periodic or chaotic dynamics.  A major weakness of the method is that we have no notion of the internal mechanisms that generated the data, and this is one motivation for returning to state-space models (see **Chapter 22: Dynamic Causal Modelling**).

**V CONCLUSION**

This chapter has described different methods of modelling inter-regional coupling using information in neuroimaging data.  The development and application of these methods is motivated by the central importance of changes in effective connectivity in development, cognition and pathology.  We have portrayed the models in a historical fashion; from the first linear observation models to bilinear terms of regression and convolution models.  In

the next two chapters we will revisit Volterra-based and bilinear state space models. The emphasis has been on the bilinearity within imaging data and how different models attempt to extract this information. Bilinear models are a practical extension to linear models, capturing plasticity induced by environmental and neurophysiological changes, while retaining mathematical tractability.

**APPENDICES**

## A.1: Matrix exponential method for integrating state equations

Variables $x$, $u$ and $z$ are time dependent (*i.e.* are functions of time), however, for simplicity, the time dependence has been omitted. Consider any the differential equation

$$
\begin{aligned}
\dot{x} = f(x,u) &\approx Ax + uBx + Cu + D \\
A &= \frac{\partial f}{\partial x} \\
B &= \frac{\partial^2 f}{\partial x \partial u} \\
C &= \frac{\partial f}{\partial u} \\
D &= f(0,0)
\end{aligned}
\qquad \text{A.1}
$$

A time dependent general solution for $z(t)$ is

$$
\begin{aligned}
z(t + \Delta t) &= \exp(J\Delta t)z(t) \\
z &= \begin{bmatrix} 1 \\ x \end{bmatrix} \\
J &= M + Nu \\
M &= \begin{bmatrix} 0 & 0 \\ D & A \end{bmatrix} \\
N &= \begin{bmatrix} 0 & 0 \\ C & B \end{bmatrix}
\end{aligned}
\qquad \text{A.2}
$$

*for constant input.* $z(t)$ can be calculated iteratively over epochs of time during which u is stationary. The stability of the solution depends on the eigenvalues of $J$, with negative values resulting in stable solutions (Boas, 1983).


## A.2: SEM objective function

The observed covariance calculated from the data is

$$S = \frac{1}{N\text{-}1} Y^T Y \qquad\qquad \text{A.3}$$

where N is the number of observations. The covariance implied by the SEM

$$\Sigma = (1 - \beta)^{-1T} \left\langle z_t^T z \right\rangle (1 - \beta)^{-1} \qquad\qquad \text{A.4}$$

where $\left\langle z_t^T z \right\rangle$ is diagonal, as the innovations are assumed to be independent. An objective function comparing $S$ and $\Sigma$ is the Maximum Likelihood function shown below. Note that a Weighted Least Squares function may be used for non-Gaussian data (Büchel and Friston, 1997) to reduce discrepancy between the implied and estimated covariance matrices

$$F_{ML} = \log|\Sigma| - trace(S\Sigma^{-1}) - \log|S| - p \qquad\qquad \text{A.5}$$

**Figure Legends**

**Figure 1**

Approximations to $f(x) = \ln(x)$ about $x_0 = 1$ using Taylor series expansions, where $f_1(x) = x - 1$ and $f_2(x) = (x-1) - 1/2(x-1)^2$ are the first and second-order approximations. The improvement about $x_0$ for higher order approximations is apparent.

**Figure 2**

The function *f(x)* models a simple one-state (*x*) linear dynamical system *i.e.* $\dot{x} = f(x)$. The state starts at the value $x_0$ where it *decreases* (*i.e.* the value of *f(x)* < 0) at the rate $\dot{x}_0$. After a period of time the state has decreased to $x_1$ with a rate of $\dot{x}_1$. The state continues to change until *f(x)* = 0, when $\dot{x}_n = 0$. The overall behaviour of the model is that the state decreases exponentially with time. A familiar example is radioactive decay, where the state is the number of radioactive atoms. Their rate of decay is not uniform, but varies linearly with the number of atoms.

**Figure 3**

Linear Dynamic System (LDS) with two states, $x_1$ and $x_2$, and inputs, $u_1$ and $u_2$, determined by the time invariant matrices *A* and *C* in the state equation. The state equation contains terms for intrinsic connectivity and how states are connected to external inputs. These are parameterised by the elements in matrices *A* and *C* respectively. *A* contains elements that model influences among states (inter-state) and on themselves (for example, the damping term in the model of a linearly damped harmonic oscillator).

**Figure 4**

A bilinear dynamic model similar to Figure 3. However, input $u_2$ can interact with coupling coefficients $a_{21}$ rendering matrix $\widetilde{A}(t)$ time-dependent. This induces input-dependent modulation of the coupling parameters, with the consequence of different

responses to other inputs *i.e.* $u_1$. All connections that are not shown correspond to zero elements in the matrices.

**Figure 5**

A model of functional integration between the visual and attention systems. Sensory input has an effect directly on the primary visual cortex, while contextual inputs, such as motion or attention to the stimulus, modulate pathways between nodes in the extended network. In this way, contextual input (*e.g.* induced by instructional set) may activate (or deactivate) pathways, which in turn determine the response of the system to stimulus-bound inputs.

**Figure 6**

Upper panel: Singular Value Decomposition (SVD). The technique converts an arbitrary matrix, such as *M*, into an equivalent form, consisting of a mixture of eigenimages (eigenvectors) and scaled by a variance component. Lower panel: Eigenspectrum of a typical PET data set is shown to demonstrate that the majority of variance is captured by the first few eigenvectors.

**Figure 7**

Piece-wise local linear approximation of a simple nonlinear relation between *x* (input) and *y* (response). **(a)** A simple nonlinear model is shown. Nonlinearity in the response is generated from a bilinear term $xy$, which models a non-additive interaction between input and intrinsic activity. The model is noise free for simplicity. The interaction term is scaled by $b_2$, effectively quantifying the model's sensitivity to input at different levels of intrinsic activity. **(b)** Plots of input and output data at different values of $b_2$ disclose the model's sensitivity to $b_2$. At a fixed input, $x = u$, the response varies dependent on its value. The key point is that data generated from such processes are not distributed in a linear fashion.

**Figure 8**

Data has been generated from the model in Figure 7 ($b_2 = 0.8$) and noise added. This illustrates how an index of nonlinearity can be constructed by partitioning the data. Data in $y$ is divided into two ranges, where all values of $y$ greater than 1 were grouped into $y_{high}$ and all values lower than 1 into $y_{low}$ (their averages are denoted by $\langle y_{high} \rangle$ and $\langle y_{low} \rangle$). The linear relationship between $x$ and $y$ within each of these ranges can be estimated using the approximation at the bottom of the figure. The value of $b_l$ reflects local sensitivity to global nonlinearities. An index of linear and nonlinear features of the data can be approximated through estimates of $b_1$ and $b_2$, which are calculated from values of $\langle y_l \rangle$ (known data) and $b_l$ (calculated from data).

**Figure 9**

**(a)** Extended statistical model including interaction terms used to model Psychophysiological and Physio-physiological Interactions (PPI). Subjects were asked to make a judgement regarding changes in velocity of a radially moving stimulus or to just observe the stimulus. The velocity of the actual stimulus remained constant, so that only the attentional set was manipulated. Comparisons of connectivity among the primary visual cortex (V1/2 complex and V5) and Posterior Parietal Cortex (PPC) during the different cognitive states were assessed. Bilinear terms are denoted by PPI. The left model examines the modulatory influence of attentional set (*U*) on V1 and V5 coupling, while the right assesses PPC activity-dependent modulation of the same connection. **(b)** This panel demonstrates the change in sensitivity of V5 to V1 input, depending on attentional set, using the method illustrated in Figure 8 *i.e.* partitioning the data and regressing V5 on V1. The lower graph shows the difference in regression slopes and their variance (2 standard deviations).

**Figure 10**

A SEM is used to estimate path coefficients for a specific network of connections, by 'pruning' the coupling matrix. The figure illustrates that a particular connectivity is

specified, which is usually based on a prior anatomical model. $x_t$ may contain physiological, psychological data or interaction terms (to estimate to the influence of 'contextual' input on first-order coupling). The innovations $z_t$ are assumed to be independent, and can be interpreted as driving inputs to each node.

**Figure 11**

Inference about connection strengths proceeds using nested models. Parameters from Free and Constrained models are compared with a $\chi^2$ statistic. Two examples are given, first comparing coupling coefficients during attention and non-attention and testing if they are the same. The second tests for the significance of a connection strength between an interaction term and PPC activity (see Büchel and Friston, 1997).

**Figure 12**

State-space model of the path coefficient between V1 and V5. The connection strength is modelled as a hidden variable that changes with time according to the update equation (upper panel). Changes in $\beta_t$ are estimated using the Kalman Filter and Smoother, which reveal fluctuations that match changes in the attentional set (lower panel).

**Figure 13**

Temporal and inter-variable relationships may be modelled as a Multivariate Autoregressive process. The figure shows time-lagged data where the arrows imply statistical dependence. The equation representing the model, including all time points, is given beneath the figure. $W$ contains estimates of temporal dependence, which may be used as a metric of coupling. $Y$ may contain physiological, psychological data or interaction terms.

**Figure 14**

Results of two MAR models applied to the visual attention data set. Each panel contains posterior density estimates of $W$ over time lags (x-axis) for each connection. The mean and 2 standard deviations for each posterior density are shown. Diagonal elements quantify autocorrelations and off diagonals the cross-correlations. The regions used in

each model are: V1/2, V5 and $PPI_{V1 \times PPC}$ and V5, PPC and $PPI_{V5 \times PFC}$. The models support coupling between the interaction terms and V5 and PPC respectively.

**Figure 15**

A bilinear convolution model of a simple network consisting of two inputs and an output. Linear contributions from each input are estimated by the first order kernel $h_1$. However, non-additive responses, due to nonlinear interactions within the system, are modelled by the second order kernel, $h_2$.

**Figure 16**

Top: Brain regions and connections comprising the model. Bottom: Characterisation of the effects of V2 inputs on V5 and their modulation by posterior parietal cortex (PPC). The broken lines represent estimates of V5 responses when PPC activity is zero, according to a second order Volterra model of effective connectivity with inputs to V5 from V2, PPC and the pulvinar (PUL). The solid curves represent the same response when PPC activity is one standard deviation of its between-condition variation. It is evident that V2 has an activating effect on V5 and that PPC increases the responsiveness of V5 to these inputs. The insert shows all the voxels in V5 that evidenced a modulatory effect ($p < 0.05$ uncorrected). These voxels were identified by thresholding a statistical parametric map of the F statistic testing for the contribution of second order kernels involving V2 and PPC (treating all other terms as nuisance variables). The data were obtained from with fMRI under identical stimulus conditions (visual motion subtended by radially moving dots) whilst manipulating the attentional component of the task (detection of velocity changes).

**References**

Bendat, J.S. 1998. *Nonlinear Systems Techniques and Applications*. John Wiley & Sons, New York.

Boas,M.L. 1983. *Mathematical models in the physical sciences*. John Wiley & Sons.

Bollen,K.A. 1989. *Structural equations with latent variables 3*. John Wiley & Sons, New York.

Büchel, C. and Friston,K. 2000. Assessing interactions among neuronal systems using functional neuroimaging. *Neural Netw.* **13:** 871-882.

Büchel, C. and Friston,K.J. 1997. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cereb.Cortex* **7:** 768-778.

Büchel, C., Wise,R.J., Mummery,C.J., Poline,J.B., and Friston,K.J. 1996. Nonlinear regression in parametric activation studies. *NeuroImage.* **4:** 60-66.

Büchel, C. and Friston,K.J. 1997. Characterising functional integration. In *Human Brain Function*, (R. S. J. Frackowiak, ed.) pp. 108-127. Academic Press, San Diego.

Chatfield,C. 1996. *The Analysis of Time Series: An Introduction*. CRC Press, Florida.

Cudeck,R. 2002. *Structural equation Modelling: Present and Future*.

Dayan,P. and Abbott,L. 2001. *Theoretical Neuroscience*. MIT.

Fliess,M., Lamnabhi,M., and Lamnabhi-Lagarrigue,F. 1983. An algebraic approach to nonlinear functional expansions. *IEEE Trans Circ Syst* **30:** 554-70.

Friston,K.J. 2000. The labile brain. I. Neuronal transients and nonlinear coupling. *Philos.Trans.R.Soc Lond B Biol.Sci.* **355:** 215-236.

Friston,K.J., Buechel,C., Fink,G.R., Morris,J., Rolls,E., and Dolan,R.J. 1997. Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage.* **6:** 218-229.

Friston,K.J., Frith,C.D., and Frackowiak,R.S.J. 1993. Time-Dependent Changes in Effective Connectivity Measured with PET. *Hum.Brain Mapp.* **1:** 69-79.

Friston,K.J. and Price,C.J. 2001. Dynamic representations and generative models of brain function. *Brain Res.Bull.* **54:** 275-285.

Friston,K.J., Ungerleider,L.G., Jezzard,P., and Turner,R. 1995. Characterizing Modulatory Interactions Between Areas V1 and V2 in Human Cortex: A New Treatment of Functional MRI Data. *Hum.Brain Mapp.* **2:** 211-224.

Ghahramani,Z. Unsupervised Learning course (lecture 4). www.gatsby.ucl.ac.uk/~zoubin/index.html. 2002.

Glass,L. 2001. Synchronization and rhythmic processes in physiologyn 3. *Nature* **410:** 277-284.

Glass,L. and Kaplan,D. 2000. *Understanding Nonlinear Dynamics*.

Harrison,L.M., Penny,W., and Friston,K.J. 2003. Multivariate autoregressive modelling of fMRI time series. *NeuroImage* **submitted**.

Juang,J.-N. 2001. *Identification and Control of Mechanical Systems*. Cambridge University Press, Cambridge, England.

Maruyama,G.M. 1998. *Basics of Structural Equation Modelling*. Sage Publications, Inc..

McIntosh,A.R. and Gonzalez-Lima,F. 1991. Structural modelling of functional neural pathways mapped with 2- deoxyglucose: effects of acoustic startle habituation on the auditory system 7. *Brain Res.* **547:** 295-302.

McIntosh,A.R. and Gonzalez-Lima,F. 1992b. Structural modelling of functional visual pathways mapped with 2- deoxyglucose: effects of patterned light and footshock 5. *Brain Res.* **578:** 75-86.

McIntosh,A.R. and Gonzalez-Lima,F. 1992a. The application of structural equation modelling to metabolic mapping of functional neural systems. In *Advances in metabolic mapping techniques for brain imaging of behavioural and learning functions*, pp. 219-255. Kluwer Academic Publishers, Dordrecht, Germany.

McIntosh,A.R. and Gonzalez-Lima,F. 1994. Structural equation modelling and its application to network analysis in functional brain imaging. *Hum.Brain Mapp.* **2:** 2-22.

McIntosh,A.R., Grady,C.L., Ungerleider,L.G., Haxby,J.V., Rapoport,S.I., and Horwitz,B. 1994. Network analysis of cortical visual pathways mapped with PET. *J.Neurosci.* **14:** 655-666.

Mesulam,M.M. 1998. From sensation to cognition. *Brain* **121:** 1013-1052.

Penny,W. and Roberts,S.J. 2002. Bayesian multivariate autoregressive models with structured priors. *IEE Proc.-Vis Image Signal Process* **149:** 33-41.

Rao,T.S. 1992. Identification of Bilinear Time Series Models 5. *Statistica Sinica* **2:** 465-478.

Rieke,F., Warland,D., de Ruyter van Steveninck,R., and Bialek,W. 1997. Foundations. In *Spikes: Exploring the Neural Code*, pp. 38-48. MIT.

Rowe,J., Friston,K., Frackowiak,R., and Passingham,R. 2002. Attention to action: specific modulation of corticocortical interactions in humans. *NeuroImage.* **17:** 988-998.

Scott,A. 1999. *Nonlinear Science: Emergence & Dynamics of Coherent Structures*. Oxford University Press, Oxford, England

$f_1(x)$

$f_2(x)$

$\ln(x)$

$x_0$

$x$

$f(x) = \dot{x}$

$\dot{x}_n$
$\dot{x}_3$
$\dot{x}_2$

$\dot{x}_1$

$\dot{x}_0$

$x_n$ $x_3$ $x_2$ $x_1$ $x_0$

$x$



$x$

$x_0$
$x_1$
$x_2$
$x_3$
$x_n$

$time$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ 0 & c_{22} \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Inter-state & self

$$\dot{x} = Ax + Cu$$

Input $u_1$

Input $u_2$

$b_{21}^2$

$c_{11}$

$c_{22}$

$x_1$

$\dot{x}_1 = a_{11}x_1 + (a_{21} + b_{21}^2 u_2(t))x_2 + c_{11}u_1$

$a_{21}$

$x_2$

$\dot{x}_2 = a_{22}x_2 + a_{12}x_1 + c_{22}u_2$

$a_{11}$

$a_{22}$

$a_{12}$
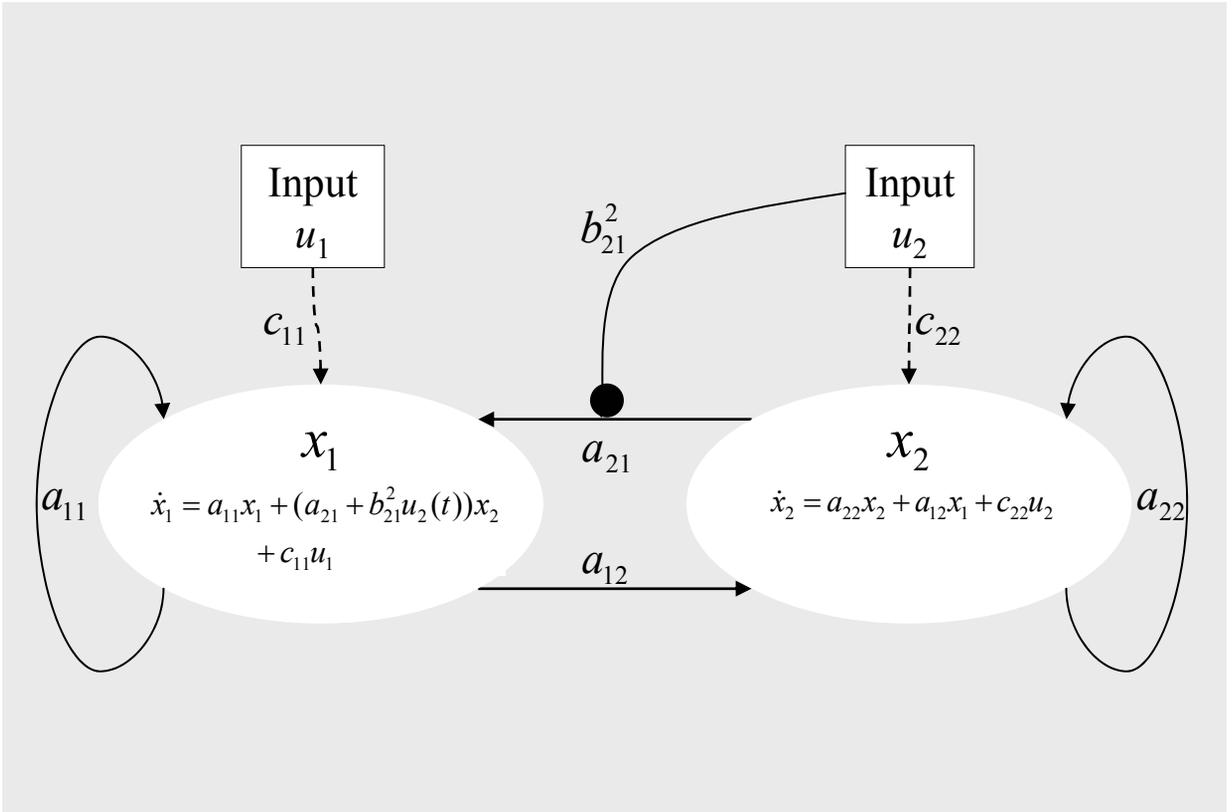
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \left( \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ b_{21}^2 & 0 \end{bmatrix} u_2 \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ 0 & c_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\dot{x} = \widetilde{A}x + Cu$$

$$\widetilde{A} = A + \sum_{j=1}^{m} u_j B^j$$

Induced connectivity

4
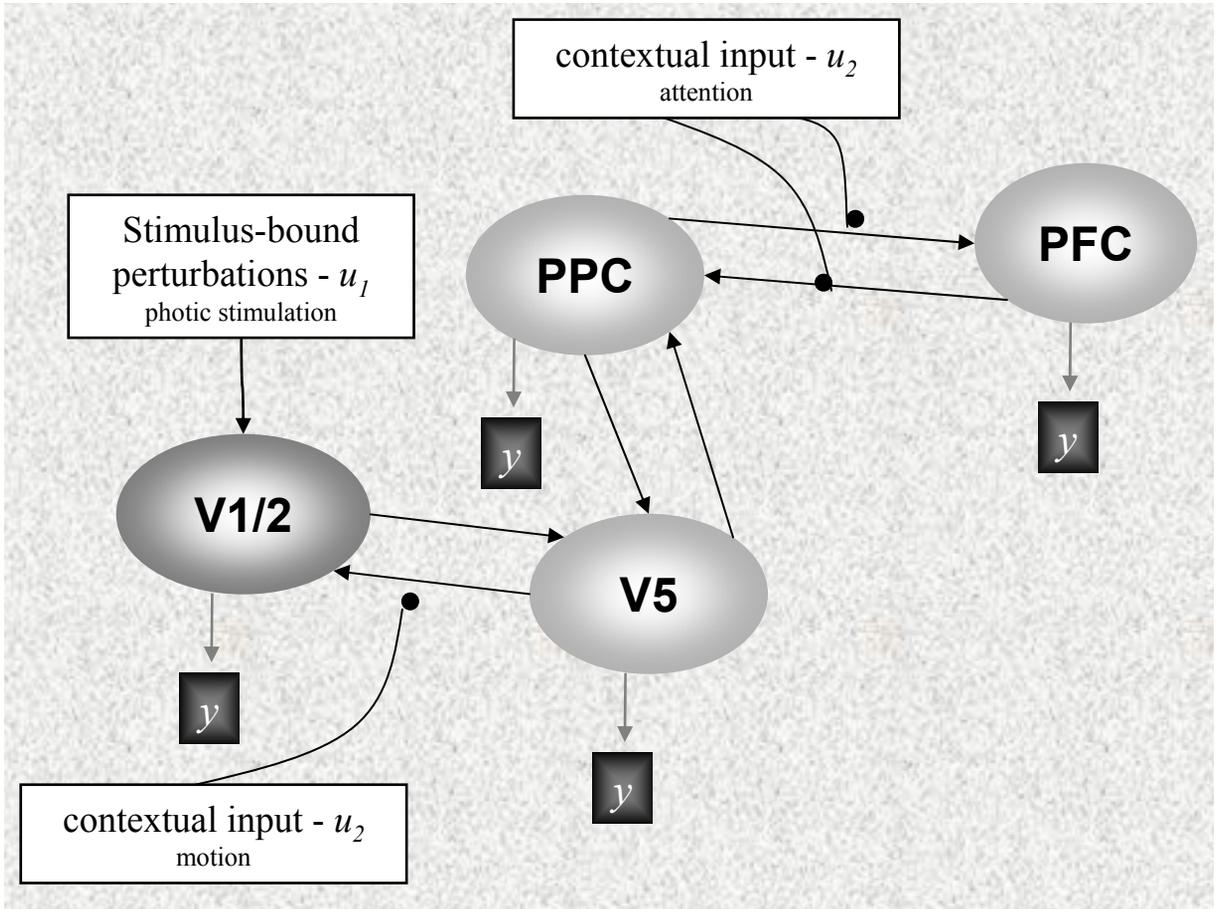
Functional integration within the visual and attention systems

$$M \; = \; s_1 U_1 \; V_1^T \; + \; U_2 \; V_2^T \; +... \; U_m \; V_m^T$$

$$\widetilde{M} \; = \; \sum_{i=1}^{r} \; s_i U_i V_i^T \; = \; \sum_{i=1}^{r} \; s_i M_i$$



64.4 percent of variance
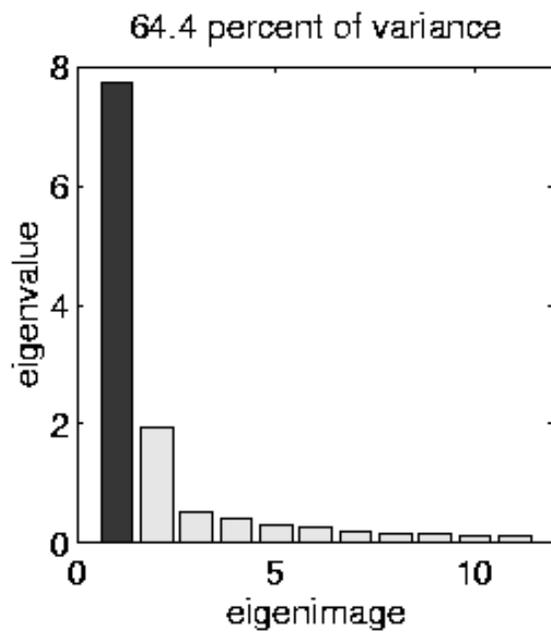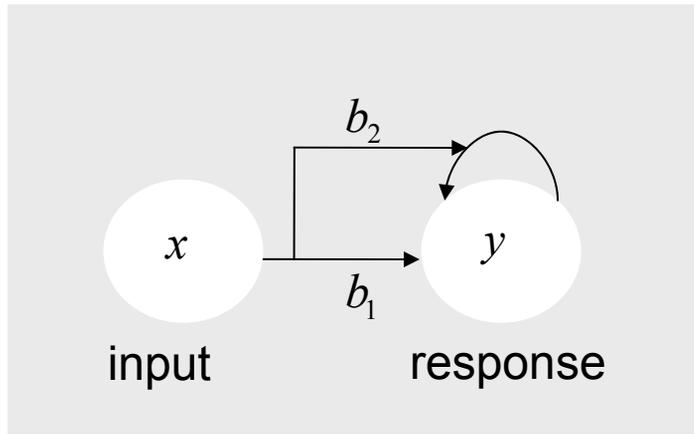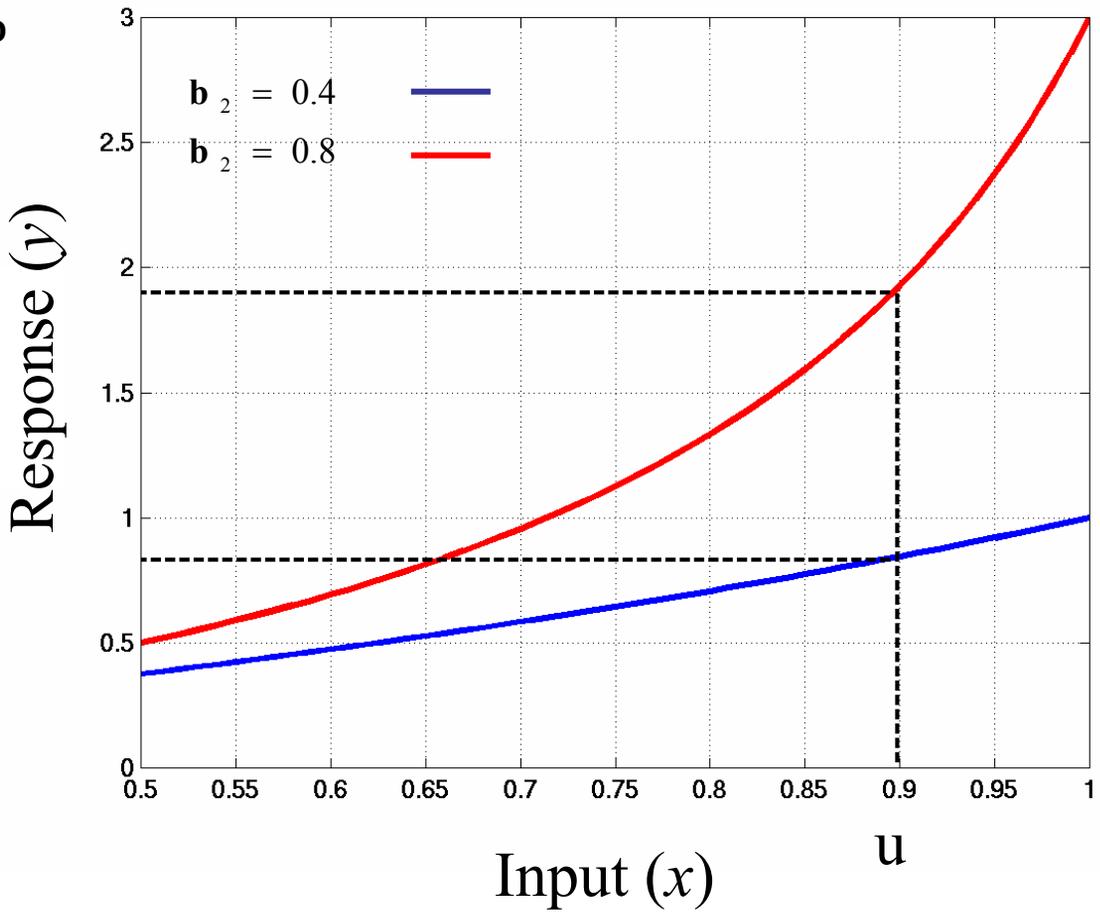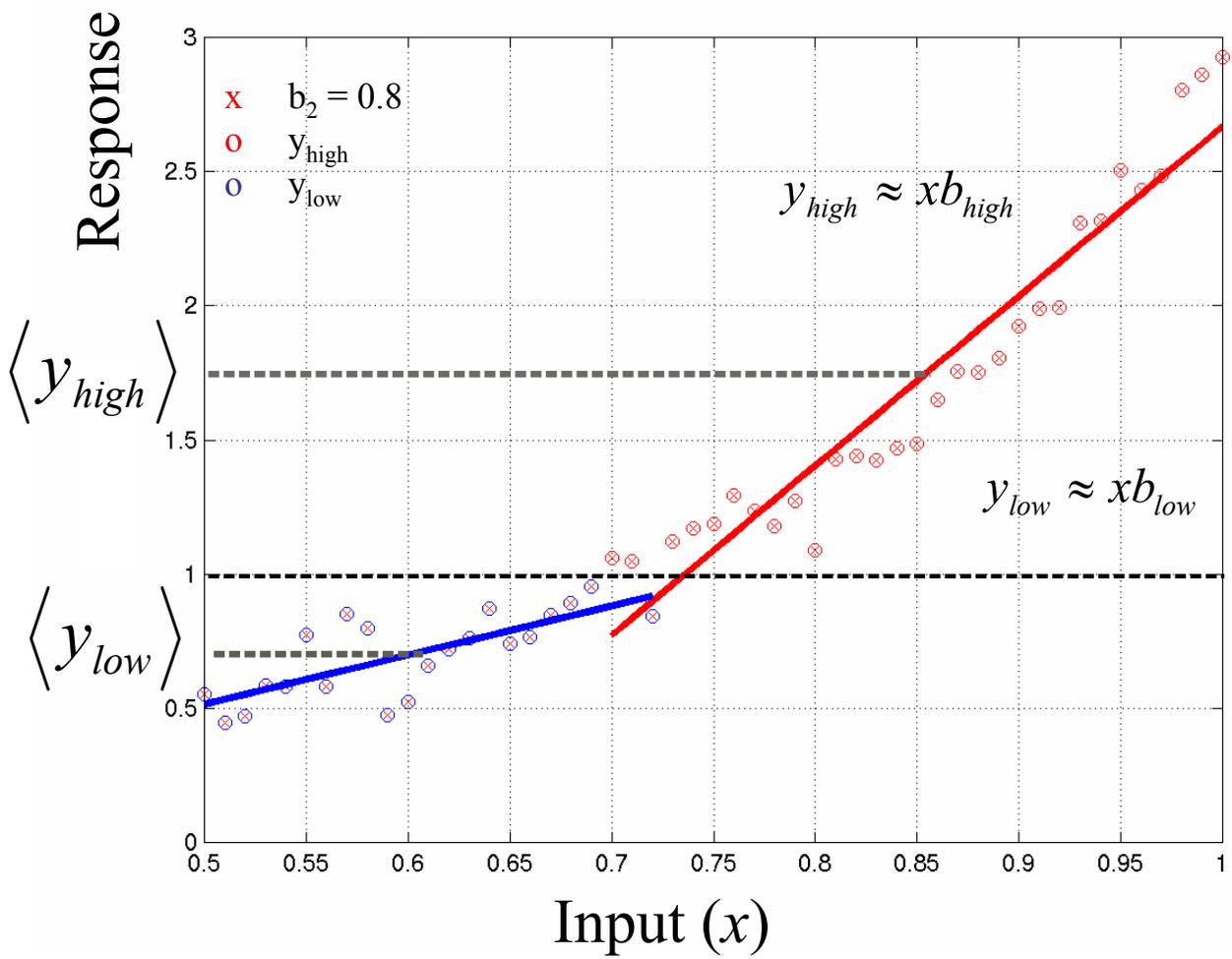
$$y_l \approx x(b_1 + b_2 \langle y \rangle)$$
$$= xb_l$$

**a**

Psychophysiological

$$V5 = (V1 \times U)\beta_{PPI} + [V1 \quad U]\beta + e$$
$$U = \text{attentional set}$$

Physio-physiological

$$V5 = (V1 \times PPC)\beta_{PPI} + [V1 \quad PPC]\beta + e$$
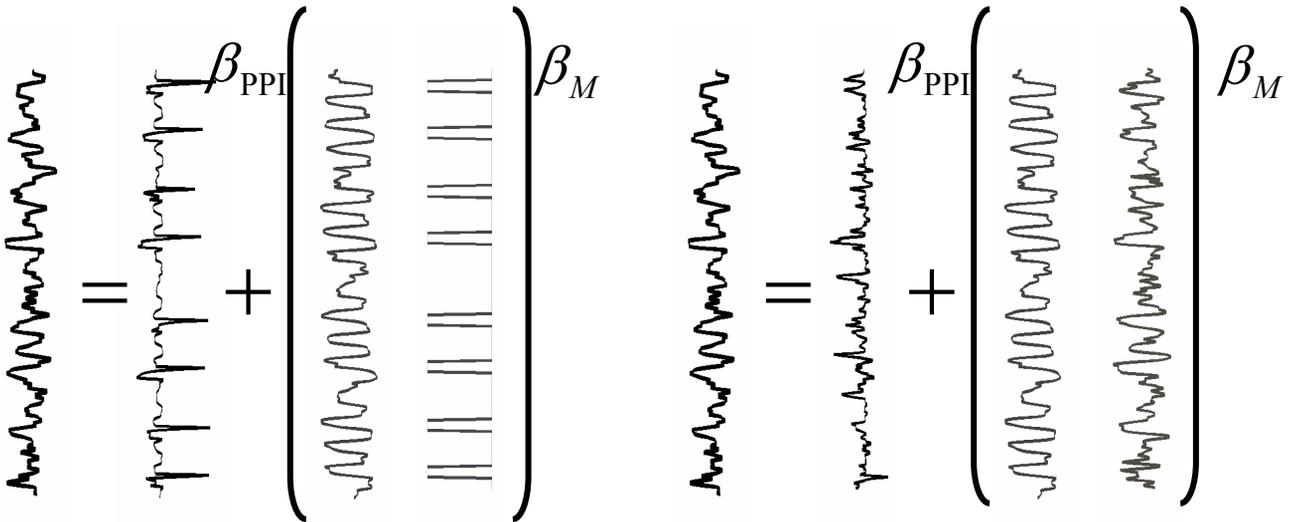


**b**

V5 vs V1 activity during
attention and no-attention
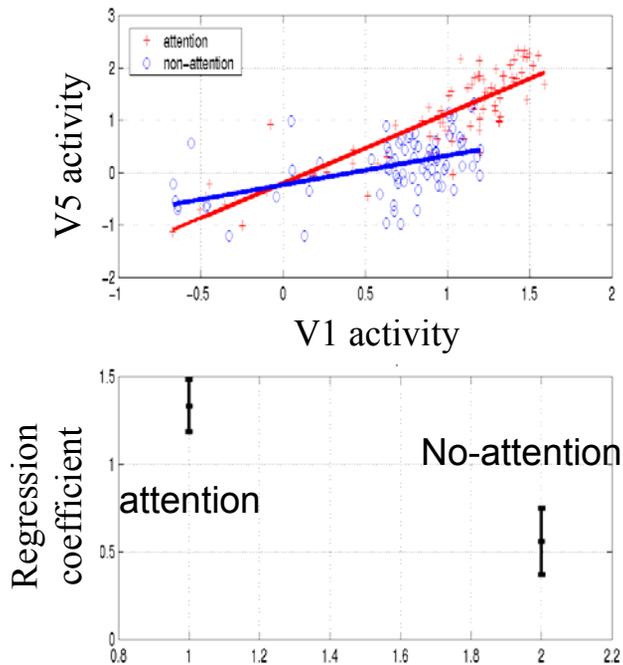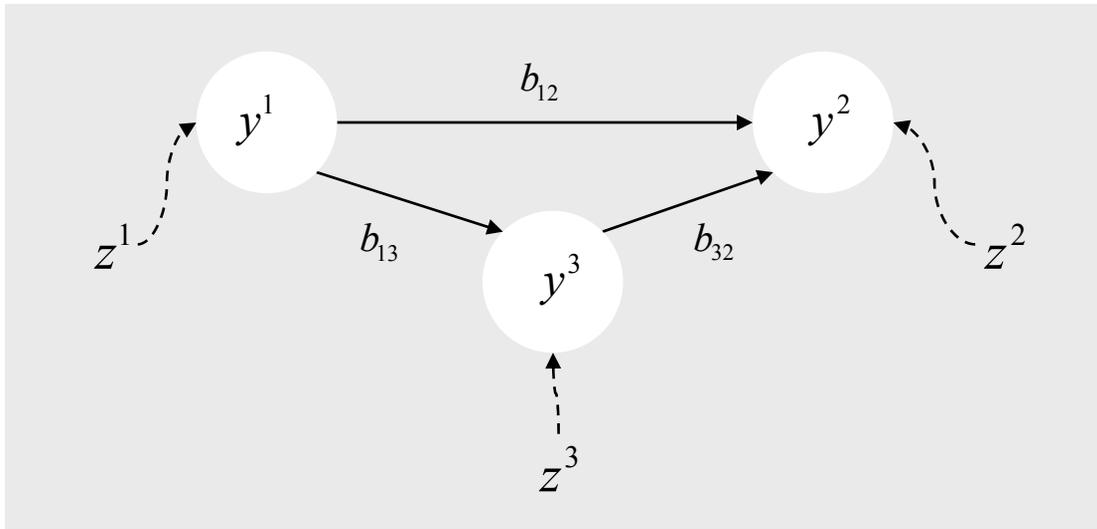
$$[y_t^1 \quad y_t^2 \quad y_t^3] = [y_t^1 \quad y_t^2 \quad y_t^3] \begin{bmatrix} 0 & b_{12} & b_{13} \\ 0 & 0 & 0 \\ 0 & b_{32} & 0 \end{bmatrix} + [z_t^1 \quad z_t^2 \quad z_t^3]$$

$$y_t = y_t \beta + z_t$$
$$z_t \sim N(0, \Sigma)$$
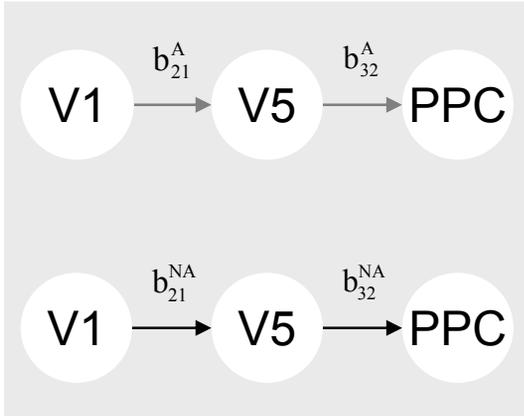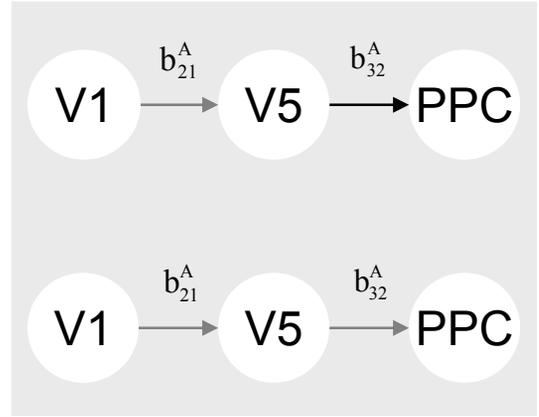
Free model[s] vs Constrained model[s]

$$H_0 : b_{21}^A = b_{21}^{NA}, \quad b_{32}^A = b_{32}^{NA}$$

$$H_0 : b_{35} = 0$$

states

$$\beta_{t-2} \rightarrow \beta_{t-1} \rightarrow \beta_t \rightarrow \beta_{t+1}$$

$$y_{t-2} \qquad y_{t-1} \qquad y_t \qquad y_{t+1}$$
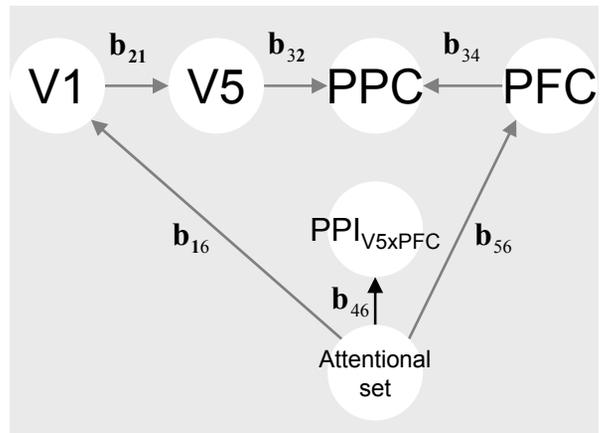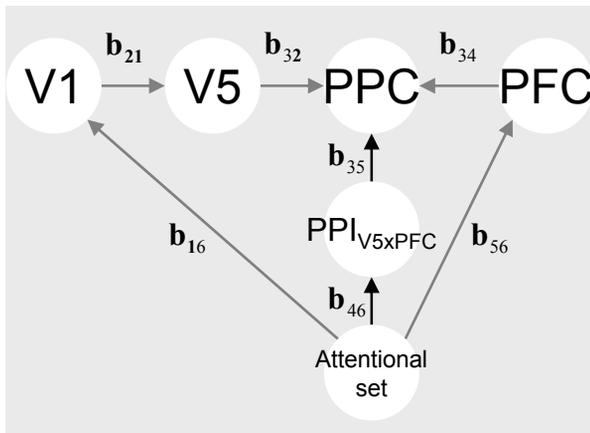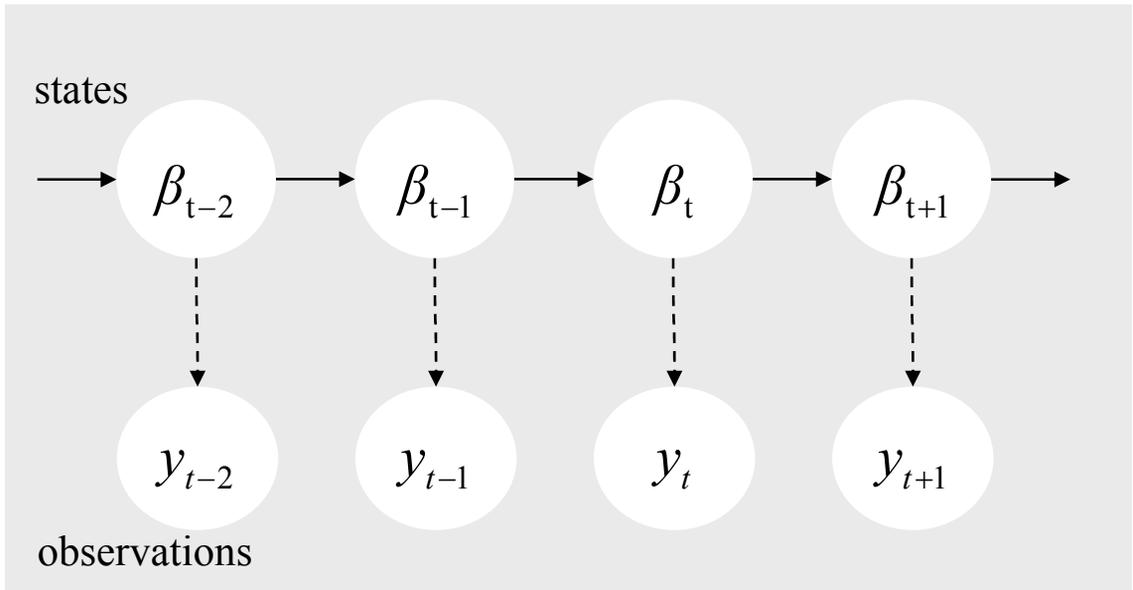
observations

Update (of state) equation

$$\beta_t = \beta_{t-1} + \eta_t, \qquad \eta_t \sim N(0, Q)$$

Observation equation

$$y_t = x_t \beta_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, R)$$



Time varying regression coefficient $\beta_t$

0.8

0.5

Time (scans)
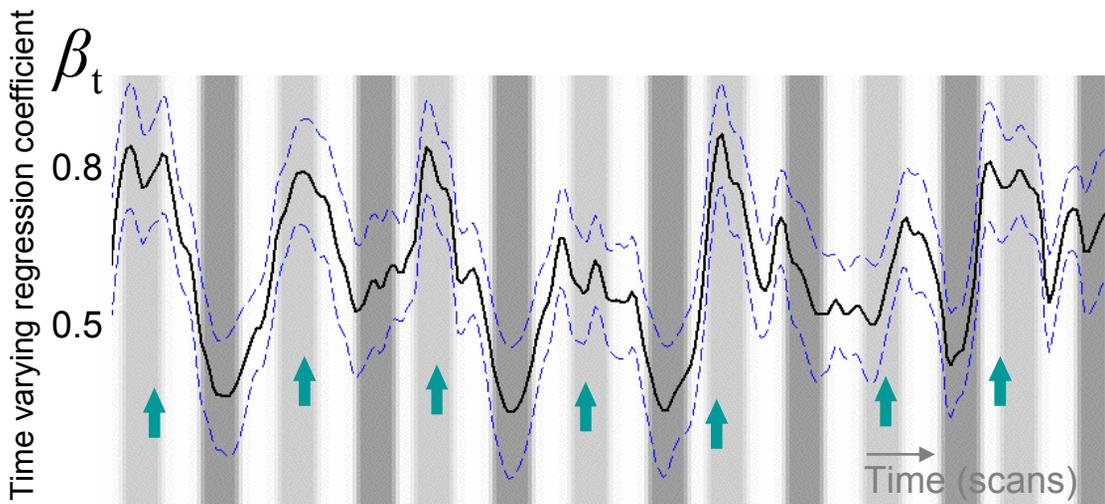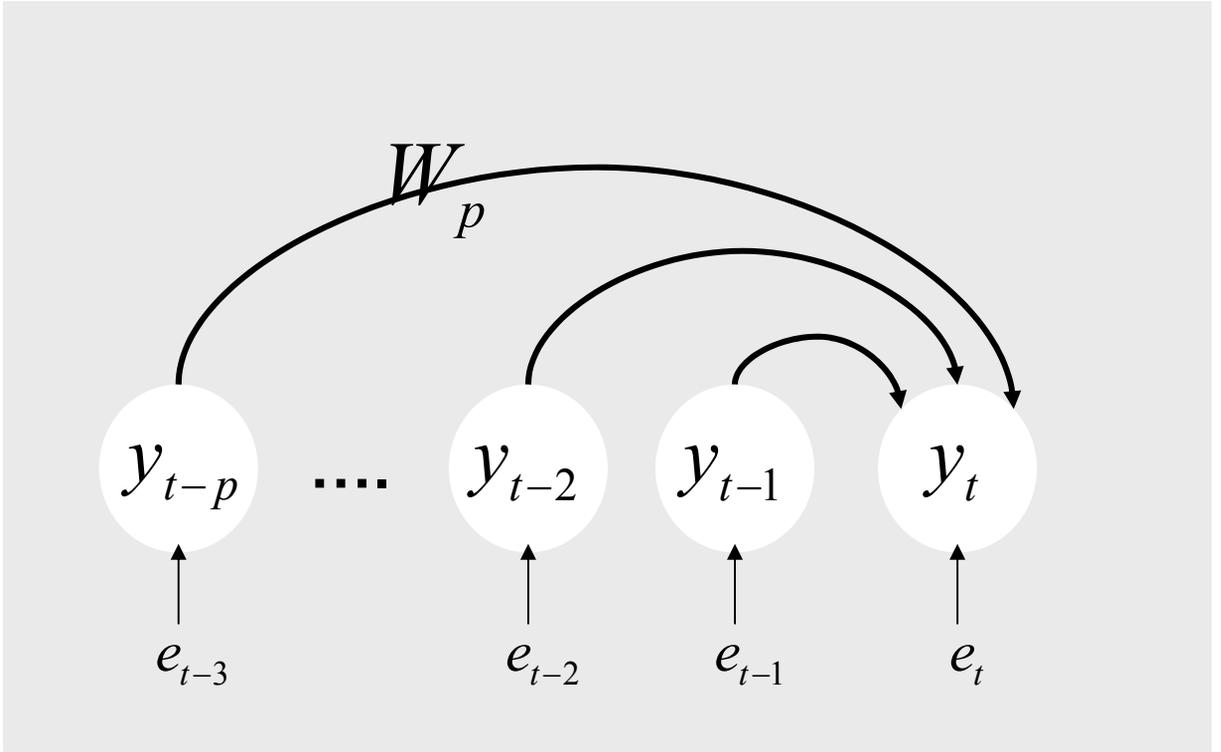
$$y_t = \sum_{j=1}^{p} y_{t-j} W_j + e$$

$$y(t) \approx h_0 + \int h_1(\tau_1)u(t-\tau_1)\partial\tau_1 + \int\int h_2(\tau_1,\tau_2)u(t-\tau_1)u(t-\tau_2)\partial\tau_1\partial\tau_2$$