Mathematical Appendix

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

1 Introduction

A. Overview

This chapter presents a theoretical review of models that are used for effective connectivity. In this discussion we focus on the nature and form of the models themselves and less on estimation or inference issues. The aim is to relate the various models commonly employed and to make their underlying assumptions and requirements more transparent.

As we have seen in the preceding chapters there are a number of models for estimating effective connectivity using neuroimaging time series (PET, fMRI, EEG and MEG). By definition, effective connectivity depends on a model, through which it is defined operationally (Friston *et al* 1995). This chapter reviews the principal models that could be adopted and how they relate to each other. We consider dynamic causal models (DCM), Generalised Convolution Models (GCM), [bi-]coherence, structural equation models (SEM) and multivariate autoregression models (MAR). In brief, we will show that they are all special cases of each other and try to emphasise their points of contact. However, some fundamental distinctions arise that guide the selection of the appropriate models in different situations.

Figure 1 about here

1 Single or multiple regions?

The first distinction rests upon whether the model is used to explain the coupling between the inputs and the responses of one cell, assembly or region, or whether the model encompasses interactions among the states of multiple regions. In terms of models this distinction is between input-output models (e.g. multiple-input, single-output models (MISO) and multiple-input multiple-output models (MIMO)) and explicit input-*state*-output models. Usually the input-output approach is concerned with the nonlinear transformation of inputs, by a region, to produce its outputs. The implicit states correspond to hidden states of a single region and the effective connectivity concerns the *vertical* link between inputs and outputs (see Figure 1a). In contradistinction, the input-state-output approach is generally concerned with characterising the *horizontal* coupling among variables that represent the states of different regions. These states are observed vicariously though the outputs (see Figure 1b). Examples of input-output models include the Volterra formulation of effective connectivity, and related coherence analyses in the spectral domain. An example of a model that tries to estimate horizontal coupling among hidden states is DCM. A critical aspect of vertical, input-output models of effective connectivity is that they can proceed without reference to the hidden states. Conversely, the horizontal interactions require indirect access to the states or some strong assumptions about how they produce outputs. In short, analyses of effective connectivity can be construed as trying to characterise the input-output behaviour of a single region or the coupling among the states of several regions using an explicit input-state-output model. Below we start by reviewing input-output models and then turn to input-state-output models.

2 Deterministic or stochastic inputs?

The second key distinction is between models, where the input is known and fixed (*e.g.* DCM) and those in which it is not (MAR and SEM). Only the former class of models affords direct measures of effective connectivity. The remaining models are useful for establishing the presence of coupling under certain assumptions about the input (usually that it is white noise that drives the system). This distinction depends on whether the inputs enter as known and deterministic quantities (*e.g.* experimentally designed causes of evoked responses) or whether we know (or can assume) something about the density function of the inputs (*i.e.* its statistics up to second or higher orders). Most models of the stochastic variety assume the inputs are Gaussian, i.i.d. and stationary. Some stochastic models (*e.g.* coherence) use local stationarity assumptions to estimate high order moments from observable but noisy inputs. For example, polyspectral analysis represents an intermediate case in which the inputs are observed but only their statistics are used. However, the key distinction is not whether one has access to the inputs but whether those inputs have to be treated as stochastic or not. Stationarity assumptions in stochastic models are critical because they preclude full analyses of evoked neuronal responses or transients that, by their nature, are non-stationary. Despite this, there are situations where the input is not observable or under experimental control. These situations preclude the estimation of the parameters of DCMs. Approaches like MAR and SEM can be used to proceed if the inputs can be regarded as stationary. The distinction between deterministic and stochastic inputs is critical in the sense that it would be inappropriate to adopt one class of model in a context that calls for the other.

3 Connections or statistical dependencies?

The final distinction is in terms of what is being estimated or inferred. Recall that functional connectivity is defined by the presence of statistical dependencies among remote neurophysiological measurements. Conversely, effective connectivity is a parameter of a model that specifies the casual influences among brain systems. It is useful to distinguish *inferences* about statistical dependencies and *estimation* of effective connectivity in terms of the distinction between functional and effective connectivity. Examples of approaches that try to establish statically dependencies include coherence analyses and MAR. This is because there these techniques do not presume any model of how hidden states interact to produce responses. They are interested only in establishing [usually linear] dependencies among outputs over different frequencies or time lags. Although MAR may employ some model to assess dependencies, this is a model of dependencies among outputs. There is no assertion that outputs *cause* outputs. Conversely SEM and DCM try to estimate the model parameters and constitute analyses of effective connectivity proper. Generalised convolution approaches fall into this class because they rest on the *estimation* of kernels that are an equivalent representation of some input-state-output model parameters.

B Effective connectivity

Effective connectivity is the influence that one neuronal system exerts over another at a synaptic or ensemble level. This should be contrasted with functional connectivity, which implies a statistical dependence between two neuronal systems that could be mediated in any number of ways. Operationally, effective connectivity can be expressed as the response induced in an ensemble, unit or region by input from others, in terms of partial derivatives of the target activity x_i, with respect to the source activities. First E^i_j and second E^i_{jk} order connectivities are then

$$E^i_j(\sigma_1) = \frac{\partial x_i(t)}{\partial x_j(t-\sigma_1)}, \qquad E^i_{jk}(\sigma_1,\sigma_2) = \frac{\partial^2 x_i(t)}{\partial x_j(t-\sigma_1)\partial x_k(t-\sigma_2)}, \qquad \ldots \tag{1}$$

First-order connectivity 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. Second-order connectivity reflects the *modulatory* influence of the input at $t-\sigma_1$ on the response evoked at $t-\sigma_2$. And so on for higher orders. Note that, in this general formulation, effective connectivity is a function of current input and inputs over the recent past[1].

[1] In contrast, functional connectivity is model-free and simply reflects the mutual information $I(x_i, x_j)$. In this paper we are concerned only with models of effective connectivity

Furthermore, implicit in Eq(1) is the fact that effective connectivity is casual, unless σ_1 is allowed to be negative. It is useful to introduce the dynamic equivalent, in which the response of the target is measured in terms of *changes in* activity

$$\dot{E}^i_j = \frac{\partial \dot{x}_i}{\partial x_j}, \qquad \dot{E}^i_{jk} = \frac{\partial^2 \dot{x}_i}{\partial x_j \partial x_k} \cdots \qquad 2$$

where $\dot{x}_i = \partial x_i / \partial t$. In this dynamic form all influences are casual and instantaneous. Before considering specific models of effective connectivity we will review briefly their basis (see **Chapter 20: Effective Connectivity**).

C Dynamical systems

The most general and plausible model of neuronal systems is a nonlinear dynamical model that corresponds to an analytic multiple-input multiple-output (MIMO) system. The state and output equations of a analytic dynamical system are

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), \theta) \\ y(t) &= \lambda(x(t)) + \varepsilon \end{aligned} \qquad 3$$

Typically the inputs $u(t)$ correspond to designed experimental effects (*e.g.* stimulus functions in fMRI), or represent stochastic drives or system perturbations. Stochastic observation error $\varepsilon \sim N(0, \Sigma)$ enters linearly in this model. For simplicity, the expressions below deal single-input, single-output (SISO) systems, and will be generalised later. The measured response *y* is some nonlinear function of the states of the system *x*. These state variables are usually unobserved or hidden (*e.g.* the configurational status of all ion channels, the depolarisation of every dendritic compartment, *etc.*). The parameters of the state equation embody effective connectivity, either in terms of mediating the coupling between inputs and outputs (MISO models of a single region) or through the coupling among state variables (MIMO models of multiple regions). The objective is to estimate and make inferences (usually Bayesian) about these parameters, given the outputs and possibly the inputs. Sometimes this requires one to specify the form of the state equation. A ubiquitous and useful form is the bilinear approximation to (3); expanding around x_0

$$
\begin{aligned}
\dot{x}(t) &\approx Ax + uBx + Cu \\
y &= Lx
\end{aligned}
\tag{4}
$$

$$
A = \frac{\partial f}{\partial x}, \quad B = \frac{\partial^2 f}{\partial x \partial u}, \quad C = \frac{\partial f}{\partial u}, \quad L = \frac{\partial \lambda}{\partial x}
$$

For simplicity, we have assumed $x_0 = 0$ and $f(0) = \lambda(0) = 0$. This bilinear model is sometimes expressed in a more compact form by augmenting the states with a constant

$$
\begin{aligned}
\dot{X} &= (M + uN)X \\
y &= HX
\end{aligned}
\tag{5}
$$

$$
X = \begin{bmatrix} 1 \\ x \end{bmatrix}, \quad M = \begin{bmatrix} 0 & 0 \\ f(0) & A \end{bmatrix}, \quad N = \begin{bmatrix} 0 & 0 \\ C & B \end{bmatrix}, \quad H = \begin{bmatrix} \lambda(0) & L \end{bmatrix}
$$

(see Friston 2002). Here the model's parameters comprise the matrices $\theta \in \{A, B, C, L\}$. We will use the bilinear parameterisation when dealing with MIMO models and their derivatives below. We will first deal with MISO models, with and without deterministic inputs.

II. INPUT-OUTPUT MODELS FOR SINGLE REGIONS

A Models for deterministic inputs - The Volterra formulation

In this section we review the Volterra formulation of dynamical systems. This formulation is important because it allows the input-output behaviour of a system to be characterised in terms of kernels that can be estimated without knowing the states of the system.

The Fliess fundamental formula (Fliess *et al* 1983) describes the causal relationship between the outputs and the history of the inputs in (3). This relationship conforms to a Volterra series, which expresses the output $y(t)$ as a generalised convolution of the input $u(t)$, critically without reference to the state variables $x(t)$. This series is simply a functional Taylor expansion of the outputs with respect to the inputs (Bendat 1990). The reason it is a *functional* expansion is that the inputs are a function of time.

$$y(t) = h(u,\theta) + \varepsilon$$
$$h(u,\theta) = \sum_i \int_0^t \dots \int_0^t \kappa_i(\sigma_1,\dots,\sigma_i)u(t-\sigma_1),\dots,u(t-\sigma_i)d\sigma_1,\dots,d\sigma_i \qquad 6$$
$$\kappa_i(\sigma_1,\dots,\sigma_i) = \frac{\partial^i y(t)}{\partial u(t-\sigma_1),\dots,\partial u(t-\sigma_i)}$$

were $\kappa_i(\sigma_1,\dots\sigma_i)$ is the *i*th order kernel. In Eq (6) the integrals are restricted to the past or history of the inputs. This renders Eq (6) causal. In some situations an acausal formulation may be appropriate (*e.g.* in which the kernels have non-zero values for future inputs - see Friston and Büchel 2000). One important thing about (6) is that it is linear in the unknowns, enabling unbiased estimates of the kernels using least squares. In other words, (6) can be treated as a general linear observation model enabling all the usual estimation and inference procedures (see **Chapter 20: Effective Connectivity** for an example). Volterra series are generally thought of as a high-order or generalised nonlinear convolution of the inputs to provide an output. To ensure estimability of the kernels they can be expanded in terms of some appropriate basis functions $q_j^i(\sigma_1,\dots,\sigma_i)$ to give the general linear model

$$y(t) = \sum_{ij} \beta_j^i h_j^i(u) + \varepsilon$$
$$h_j^i(u) = \int_0^t \dots \int_0^t q_j^i(\sigma_1,\dots,\sigma_i)u(t-\sigma_1),\dots,u(t-\sigma_i)d\sigma_1,\dots,d\sigma_i \qquad 7$$
$$\kappa_i(\sigma_1,\dots,\sigma_i) = \sum_j \beta_j^i q_j^i(\sigma_1,\dots,\sigma_i)$$

The Volterra formulation is useful as a way of characterising the influence of inputs on the responses of a region. The kernels can be regarded as a re-parameterisation of the bilinear form in Eq(4) that encodes the impulse response to input. The kernels for the states are

$$
\begin{aligned}
\kappa_0 &= X(0) \\
\kappa_1(\sigma_1) &= e^{\sigma_1 M} N e^{-\sigma_1 M} X(0) \\
\kappa_2(\sigma_1,\sigma_2) &= e^{\sigma_2 M} N e^{(\sigma_1-\sigma_2)M} N e^{-\sigma_1 M} X(0) \\
\kappa_2(\sigma_1,\sigma_2,\sigma_3) &= \ldots
\end{aligned} \qquad 8
$$

The kernels associated with the output follow from the chain rule

$$
\begin{aligned}
h_0 &= H\kappa_0 \\
h_1(\sigma_1) &= H\kappa_1(\sigma_1) \\
h_2(\sigma_1,\sigma_2) &= \ldots
\end{aligned} \qquad 9
$$

(see Friston 2002 for details). If the system is fully nonlinear, then the kernels can be considered local approximations. If the system is bilinear they are globally exact. It is important to remember that the estimation of the kernels does not assume any form for the state equation and completely eschews the states. This is the power and weakness of Volterra-based analyses.

The Volterra formulation can be used directly in the assessment of effective connectivity if we assume the measured response of one region (*j*) constitutes the input to another (*i*) *i.e.* $u_j(x) = y_j(t)$. In this case the Volterra kernels have a special interpretation; they are synonymous with effective connectivity. From (6) the first order kernels are

$$
\kappa_1(\sigma_1) = \frac{\partial y_i(t)}{\partial y_j(t-\sigma_1)} = E_j^i(\sigma_1) \qquad 10
$$

Extensions of Eq(6) to multiple inputs (MISO) models are trivial and allow high-order interactions among inputs to a single region to be characterised. This approach was used in Friston and Büchel (2000) to examine parietal modulation of V2 inputs to V5, by estimating and making inferences about the appropriate second order kernel. The advantage of the Volterra approach is that nonlinearities can be modelled and estimated in the context of highly nonlinear transformations within a region and yet the estimation and inference proceed in a standard linear least squares setting. However, one has to assume that the inputs conform to measured responses elsewhere in the brain. This may be tenable for EEG but the hemodynamic responses measured by fMRI make this a more questionable approach. Furthermore, there is no causal model of the interactions among areas that would otherwise offer useful constraints on the estimation. The direct application of Volterra estimation, in this fashion, simply examines each node, one at a time, assuming the activities of other nodes are veridical measurements of the inputs to the node in question. In summary, although the Volterra kernels are useful characterisations of the input-output behaviour of single regions, they are not constrained by any model of interactions among regions. Before turning to DCMs, that embody these interactions, we will deal with the SISO situation in which the input is treated as stochastic.

B Models for stochastic inputs – Coherence and Polyspectral analysis

In this section we deal with systems in which the input is stochastic. The aim is to estimate the kernels (or their spectral equivalents) given only statistics about the joint distribution of the inputs and outputs. When the inputs are unknown one generally makes assumption about their distributional properties and assumes [local] stationariness. Alternatively the inputs may be measurable but too noisy to serve as inputs in Eq(7). In this case they can be used to estimate the input and output densities in terms of higher order cumulants or polyspectral density. The *n*th order cumulate of the input is

$$c_u\{\sigma_1,\ldots,\sigma_{n-1}\} = \langle u(t)u(t-\sigma_1),\ldots,u(t-\sigma_{n-1})\rangle \qquad 11$$

where we have assumed here and throughout that $E\{u(t)\}=0$. It can be seen that cumulants are a generalisation of auto-covariance functions. The second-order cumulant is simply the auto-covariance function of lag and summarises the stationary second-order behaviour of the input. Cumulants allow one to formulate (6) in terms of the second order statistics of input and outputs. For example,

$$c_{yu}(\sigma_a) = \langle y(t)u(t-\sigma_a)\rangle$$
$$= \sum_i \int_0^t \ldots \int_0^t \kappa_i(\sigma_1,\ldots,\sigma_i)\langle u(t-\sigma_a)u(t-\sigma_1)\ldots u(t-\sigma_i)\rangle d\sigma_1 \ldots d\sigma_i \qquad 12$$
$$= \sum_i \int_0^t \ldots \int_0^t \kappa_i(\sigma_1,\ldots,\sigma_i) c_u(\sigma_a-\sigma_1,\ldots,\sigma-\sigma_i) d\sigma_1 \ldots d\sigma_i$$

Eq(12) says that the cross-covariance between the output and the input can be decomposed into components that are formed by convolving the *i*th order kernel with the input's (*i*+1)th cumulant. The important thing about this is that all cumulants, greater than second order, of Gaussian processes are zero. This means that if we can assume the input is Gaussian then

$$c_{yu}(\sigma_a) = \int_0^t \kappa_i(\sigma_1) c_u(\sigma_a - \sigma_1) d\sigma_1 \qquad 13$$

In other words, the cross-covariance between the input and output is simply the auto-covariance function of the inputs convolved with the first-order kernel. Although it is possible to formulate the covariance between inputs and outputs in terms of cumulants, the more conventional formulation is in frequency space using polyspectra. The *n*th polyspectrum is the Fourier transform of the corresponding cumulant

$$g_u(\omega_1,\ldots,\omega_{n-1}) = \left(\tfrac{1}{2\pi}\right)^{n-1} \int \ldots \int c_u\{\sigma_1,\ldots,\sigma_{n-1}\} e^{-j(\omega\sigma_1,\ldots,\omega\sigma_{n-1})} d\sigma_1,\ldots,d\sigma_{n-1} \qquad 14$$

Again, polyspectra are simply a generalisation of spectral densities. For example, the second polyspectrum is spectral density and the third polyspectrum is bispectral density. It can be seen that these relationships are generalisations of the Wiener-Khinchine theorem, relating the auto-covariance function and spectral density through the Fourier transform. Introducing the spectral density representation

$$u(t) = \int s_u(\omega) e^{-j\omega} d\omega \qquad 15$$

we can now rewrite the Volterra expansion, Eq(6) as

$$h(u,\theta) = \sum_i \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e^{j(\omega_1 +,\dots,+\omega_i)t} \Gamma_1(\omega_1,\dots,\omega_i) s_u(\omega_1),\dots,s_u(\omega_i) d\omega_1,\dots,d\omega_i \qquad 16$$

where the functions

$$\Gamma_1(\omega_1) = \int_0^\infty e^{-j\omega_1\sigma_1} \kappa_1(\sigma_1) d\sigma_1$$

$$\Gamma_2(\omega_1,\omega_2) = \int_0^\infty \int_0^\infty e^{-j(\omega_1\sigma_1+\omega_2\sigma_2)} \kappa_2(\sigma_1,\sigma_2) d\sigma_1 d\sigma_2$$

$$\dots$$

are the Fourier transforms of the kernels. These functions are called *generalised transfer functions* and mediate the expression of frequencies in the output given those in the input. Critically, the influence of higher order kernels, or equivalently generalised transfer functions means that a given frequency in the input can induce a *different* frequency in the output. A simple example of this would be squaring a sine wave input to produce an output of twice the frequency. In the Volterra approach the kernels were identified in the time domain using the inputs and outputs directly. In this section system identification means estimating their Fourier transforms (*i.e.* the transfer functions) using second and higher order statistics of the inputs and outputs. Generalised transfer functions are usually estimated through estimates of polyspectra. For example, the spectral form for (13), and its high-order counterparts are

$$\begin{aligned} g_{uy}(-\omega_1) &= \Gamma_1(\omega_1) g_u(\omega_1) \\ g_{uuy}(-\omega_1,-\omega_2) &= 2\Gamma_2(\omega_1,\omega_2) g_u(\omega_1) g_u(\omega_2) \\ &\vdots \\ g_{u\dots y}(-\omega_1,\dots,-\omega_n) &= n!\Gamma_n(\omega_1,\dots,\omega_n) g_u(\omega_1)\dots g_u(\omega_n) \end{aligned} \qquad 17$$

Given estimates of the requisite [cross]-polyspectra these equalities can be used to provide estimates of the transfer functions (see Figure 2). These equalities hold when the Volterra expansion contains just the nth order term and are a generalisation of the classical results for the transfer function of a linear system [first equality in Eq(17)]. The importance of these results, in terms of effective connectivity, is the implicit meaning conferred on *coherence* and *bi-coherence* analyses. Coherence is simply the second-order cross spectrum $g_{uy}(\omega)$ between the input and output and is related to first-order effects (*i.e.* the first-order kernel or transfer function) though Eq(17). Coherence is therefore a surrogate marker for first-order or linear connectivity. Bi-coherence or the cross-bispectrum $g_{uuy}(\omega_1,\omega_2)$ is the third-order cross-polyspectrum and implies a non-zero second-order kernel or transfer function. Bispectral analysis was used (in a simplified form) to demonstrate nonlinear coupling between parietal and frontal regions using MEG in **Chapter 21** (**Volterra kernels and effective connectivity**). In this example cross-bispectra were estimated, in a simple fashion, using time-frequency analyses.

C Summary

In summary, Volterra kernels (generalised transfer functions) characterise the input-output behaviour of a system. The *n*th order kernel is equivalent to *n*th order effective connectivity when the inputs and outputs conform to processes that mediate interactions among neuronal systems. If the inputs and outputs are known, or can be measured precisely, the estimation of the kernels is straightforward. In situations where stochastic inputs and outputs are less precisely observed, kernels can be estimated indirectly through their generalised transfer functions using cross-polyspectra. The robustness of kernel estimation, conferred by expansion in terms of temporal basis functions, is recapitulated in the frequency domain by smoothness constraints during estimation of the polyspectra. The spectral approach is limited because it assumes (i) the system contains only the kernel of the order estimated and (ii) stationariness. The intuition behind the first limitation relates to the distinction between parameter estimation and variance partitioning in standard regression analyses. Although it is perfectly possible to estimate the parameters of a regression model given a set of non-orthogonal explanatory variables it is not possible to uniquely partition variance in the output caused by these explanatory variables.

III INPUT-STATE-OUTPUT MODELS FOR MULTIPLE REGIONS

In this section we address models for multiple interconnected regions where one can measure the responses of these regions to input that may or may not be known. Although it is possible to extend the techniques of the previous sections to cover MIMO systems, the ensuing inferences about the influence of input to one region, on the response of another are not sufficiently specified to constitute an analysis of effective connectivity. This is because these influences may be mediated in many ways and are not parameterised in terms of the effective connectivity among the regions themselves. In short, one is not interested in the vertical relationship between multiple inputs and multiple outputs, but in the horizontal interactions among the state variables of each region (Figure 1). A parameterisation that encodes this inter-regional coupling is therefore required. All the models discussed below assume some form or model for the interactions among the state variables and attempt to estimate the parameters of this model, sometimes without actually observing the states themselves.

A Models for known inputs – Dynamic Causal Modelling.

The most direct and generic approach is to estimate directly the parameters of Eq(3) and use them to compute effective connectivity as described in Eq (1) and Eq(2). Although there are many forms one could adopt for Eq(3) we will focus on the bilinear approximation, which is possibly the most parsimonious but useful nonlinear approximation available. Furthermore, as shown below, the bilinear approximation re-parameterises the state equations of the model directly in terms of effective connectivity. Dynamic causal modelling does not necessarily entail the use of a bilinear model. Indeed DCMs can be specified to any degree of biological complexity and realism supported by the data. However, bilinear approximations represent the simplest form to which all DCMs can be reduced. This reduction allows analytic derivation of kernels and other computations, like integrating the state equation, to proceed in an efficient fashion.

Each region may comprise several state variables whose casual interdependencies are summarised by the bilinear form in Eq(4). Here the key connectivity parameters of the state equation are matrices M and N. For a given set of inputs or experimental context the bilinear approximation to any set of state equations is

$$
\begin{aligned}
\dot{X}(t) &= JX(t) \\
X(t+\sigma) &= e^{J\sigma}X(t) \\
J &= M + \sum_i N_i u_i
\end{aligned}
\qquad 18
$$

Notice that there are now as many N matrices as there are [multiple] inputs. The bilinear form reduces the model to first-order connections that can be modulated by the inputs. In MIMO models the effective connectivity is among the states such that first-order effective connectivities are simply

$$
\begin{aligned}
\dot{E} &= \frac{\partial \dot{X}}{\partial X} = J \\
E &= \frac{\partial X(t)}{\partial X(t-\sigma)} = e^{J\sigma}
\end{aligned}
\qquad 19
$$

(this includes connections with the constant term in Eq(5)). Note that these are context-sensitive in the sense that the Jacobian J is a function of experimental context or inputs $u(t) = [u_1(t), \ldots, u_m(t)]$. A useful way to think about the bilinear parameter matrices is to regard them as the intrinsic or latent dynamic connectivity, in the absence of input, and changes induced by each input (see the previous chapter for a fuller description)

$$
\begin{aligned}
\dot{E}(0) &= N = \begin{bmatrix} 0 & 0 \\ f(0) & A \end{bmatrix} \\
\frac{\partial \dot{E}}{\partial u_i} &= M_i = \begin{bmatrix} 0 & 0 \\ C_i & B_i \end{bmatrix}
\end{aligned}
\qquad 20
$$

The latent dynamic connectivity among the states is A. Often one is more interested in the B_i as embodying changes in this connectivity induced by different cognitive set, time or drugs. Note that C_i is treated as the input-dependent component of the connection from the constant term or drive. Clearly it would be possible to introduce other high order terms to model interactions among the states but we will restrict ourselves to bilinear models for simplicity.

The fundamental advantage of DCM over alternative strategies is that the casual structure is made explicit by parameterising the state equation. The estimation of effective connectivity and ensuing inferences are usually through posterior mode analysis based on normality assumptions about the errors and some suitable priors on the parameters. The parameters of the bilinear form are $\theta = \{A, B, C, L\}$. If the priors are also specified under Gaussian assumptions, in terms of their expectation η_θ and covariance C_θ, Gauss-Newton EM scheme can be adopted to find the posterior mode $\eta_{\theta|y}$ (see the previous chapter for details).

In essence, dynamic causal modelling comprises (i) specification of the state and output equations of an ensemble of region-specific state variables. A bilinear approximation to the state equation reduces the model to first-order coupling and bilinear terms that represent the modulation of that coupling by inputs. (ii) Posterior density analysis of the model parameters then allows one to estimate and make inferences about inter-regional connections and the effect of experimental manipulations on those connections.

As mentioned above, the state equations do not have to conform to the bilinear form. The bilinear form can be computed automatically given any state equation. This is important because the priors may be specified more naturally in terms of the original biophysical parameters of the DCM, as opposed to the bilinear form. The choice of the state variables clearly has to accommodate their role in mediating the effect of inputs on responses and the interactions among areas. In the simplest case the states variables could be reduced to mean neuronal activity per region, plus any biophysical state variables needed to determine the output (*e.g.* the states of hemodynamic models for fMRI). Implicit in choosing such state variables is the assumption that they model all the dynamics to the level of detail required. Mean field models and neural mass models are useful here in motivating the number of state variables and the associated state equations. Constraints on the parameters of the model are implemented through their priors. These restrict the parameter estimates to plausible ranges. An important constraint is that the system is dissipative and does not diverge exponentially in the absence of input. In other words, the priors ensure that the largest eigenvalue of J is less than zero.

1 Summary

In summary, DCM is the most general and direct approach to identifying the effective connectivity among the states of MIMO systems. The identification of DCMs usually proceeds using Bayesian schemes to estimate the posterior mode or most likely parameters of the model given the data. Posterior mode analysis requires only the state equations and priors to be specified. The state equations can be arbitrarily complicated and nonlinear. However, a Bilinear approximation to the causal influences among state variables serves to reduce the complexity of the model and parameterises the model directly in terms of first order connectivity and its changes with input (the bilinear terms). In the next section we deal the situations in which the input is unknown. This precludes DCM because the likelihood of the responses cannot be computed unless we know what caused them.

B Models for stochastic inputs – SEM and regression models

When the inputs are treated as unknown, and the statistics of the outputs are only considered to second order, one is effectively restricted to linear or first-order models of effective connectivity. Although it is possible to deal with discrete-time bilinear models, with white noise inputs, they have the same covariance structure as ARMA (autoregressive moving average) models of the same order (Priestly 1988 p66). This means that, in order to distinguish between linear and nonlinear models, one would need to study moments higher than second order (*c.f.* the third order cumulants in bi-coherence analyses). Consequently, we will focus on linear models of effective connectivity under white stationary inputs. These inputs are the innovations introduced in the last chapter. There are two important classes of model here: These are structural equation models and ARMA models. Both are finite parameter linear models that are distinguished by their dependency on dynamics. In SEM the interactions are assumed to be instantaneous whereas in ARMA the dynamic aspect is retained explicitly in the model.

SEM can be derived from DCMs by assuming the inputs vary slowly in relation to neuronal and hemodynamics. This is appropriate for PET experiments and possibly some epoch-related fMRI designs but not for event-related designs in ERP or fMRI. Note that this assumption pertains to the inputs or experimental design, not to the time constants of the outputs. In principle, it would be possible to apply DCM to a PET study.

Consider a linear DCM where we can observe the states precisely and there was only one state variable per region

$$\begin{aligned}\dot{x} &= f(x,u)\\ &= Ax + u = (A^0 - 1)x + u\end{aligned} \qquad 21$$

$$y = \lambda(x) = x$$

Here we have discounted observation error but allow stochastic inputs $u \sim N(0,Q)$. To make the connection to SEMs more explicit, we have expanded the connectivity matrix into off-diagonal connections and a leading diagonal matrix, modelling unit decay $A = A^0 - 1$. For simplicity, we have absorbed *C* into the covariance structure of the inputs *Q*. As the inputs are changing slowly relative to the dynamics, the change in states will be zero at the point of observation and we obtain the regression model used by SEM.

$$\begin{aligned}\dot{x} &= 0 \Rightarrow\\ (1 - A^0)x &= u\\ x &= (1 - A^0)^{-1}u\end{aligned} \qquad 22$$

This should be compared with Eq(17) in **Chapter 20** (**Effective Connectivity**). The more conventional motivation for Eq(22) is to start with an instantaneous regression equation $x = A^0 x + u$ that is formally identical to the second line above. Although this regression model obscures the connection with dynamic formulations it is important to consider because it is the basis of commonly employed methods for estimating effective connectivity in neuroimaging to data. These are simple regression models and SEM.

1 Simple Regression models

$x = A^0 x + u$ can be treated as a general linear model by focussing on one region at a time, for example the first, to give

$$x_1 = [x_2, \ldots, x_n] \begin{bmatrix} A_{12} \\ \vdots \\ A_{1n} \end{bmatrix} + u_1 \qquad 23$$

c.f. Eq(8) in **Chapter 20** (**Effective Connectivity**) The elements of A can then be solved in a least squares sense by minimising the norm of the unknown stochastic inputs u for that region (*i.e.* minimising the unexplained variance of the target region given the states of the remainder). This approach was proposed in Friston *et al* (1995) and has the advantage of providing precise estimates of connectivity with high degrees of freedom. However, these least square estimators assume, rather implausibly, that the inputs are orthogonal to the states and, more importantly, do not ensure the inputs to different regions conform to the known covariance Q. Furthermore, there is no particular reason that the input variance should be minimised just because it is unknown. Structural equation modelling overcomes these limitations at the cost of degrees of freedom for efficient estimation

2 Structural equation modelling

In SEM estimates of A^0 minimise the difference (KL divergence) between the observed covariance among the [observable] states and that implied by the model and assumptions about the inputs.

$$\begin{aligned} \langle xx^T \rangle &= \langle (1 - A^0)^{-1} uu^T (1 - A^0)^{-1T} \rangle \\ &= (1 - A^0)^{-1} Q (1 - A^0)^{-1T} \end{aligned} \qquad 24$$

This is critical because the connectivity estimates implicitly minimise the discrepancy between the observed and implied covariances among the states induced by stochastic inputs. This is in contradistinction to the instantaneous regression approach (above) or ARMA analyses (below) in which the estimates simply minimise unexplained variance on a region by region basis.

C Quasi-bilinear models – PPIs and moderator variables

There is a useful extension to the regression model implicit in Eq(22) that includes bilinear terms formed from known inputs that are distinct from stochastic inputs inducing [co]variance in the states. Let these known inputs be denoted by *v*. These usually represent some manipulated experimental context such as cognitive set (*e.g.* attention) or time. These deterministic inputs are also known as moderator variables in SEM. The underlying quasi-bilinear DCM, for one such input, is

$$\dot{x} = (A^0 - 1)x + Bvx + u \tag{25}$$

Again, assuming the system has settled at the point of observation

$$\begin{aligned} \dot{x} &= 0 \\ (1 - A^0 - Bv)x &= u \\ x &= A^0 x + Bvx + u \end{aligned} \tag{26}$$

This regression equation can be used to form least squares estimates as in Eq(23) in which case the additional bilinear regressors *vx* are known as *psychophysiological interaction* (PPI) terms (for obvious reasons). The corresponding SEM or path analysis usually proceeds by creating extra 'virtual' regions whose dynamics correspond to the bilinear terms. This is motivated by rewriting the last expression in Eq(26) as

$$\begin{bmatrix} x \\ vx \end{bmatrix} = \begin{bmatrix} A^0 & B \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ vx \end{bmatrix} + \begin{bmatrix} u \\ 0 \end{bmatrix} \tag{27}$$

It is important to note that psychophysiological interactions and moderator variables in SEM are exactly the same thing and both speak to the importance of bilinear terms in casual models. Their relative success in the neuroimaging literature is probably due to the fact that they model changes in effective connectivity that are generally much more interesting than the connection strengths *per se*. Examples are changes induced by attentional modulation, changes during procedural learning and changes mediated pharmacologically. In other words, bilinear components afford ways of characterising *plasticity* and as such play a key role in methods for functional integration. It is for this reason we focussed on bilinear approximations as a minimal DCM in the previous section.

D Summary

In summary, SEM is a simple and pragmatic approach to effective connectivity when (i) dynamical aspects can be discounted, (ii) a linear model is sufficient and (iii) the state variables can be measured precisely and (iv) the input is unknown but stochastic and stationary. These assumptions are imposed by ignorance about the inputs. Some of these represent rather severe restrictions that limit the utility of SEM in relation to DCM or state-space models considered next.. The most profound criticism of simple regression and SEM in imaging neuroscience is that they are models for interacting brain systems in the context of unknown input. The whole point of designed experiments is that the inputs are known and under experimental control. This renders the utility of SEM for designed experiments somewhat questionable.

IV MULTIVARIATE ARMA MODELS

ARMA models can be generally represented as *state-space* (or *Markovian*) models that provide a compact description of any finite parameter linear model. From this state-space representation MAR models can be derived and estimated using a variety of well-established techniques. We will focus on how the state-space representation of linear models of effective connectivity can be derived from the dynamic formulation and the assumptions required in this derivation.

As in the previous section let us assume a linear DCM in which inputs comprise stationary white noise $u \sim N(0,Q)$ that are offered to each region in equal strength (*i.e.* $C = \mathbf{1}$). This renders Eq(3) a linear stochastic differential equation (SDE)

$$\begin{aligned} \dot{x} &= Ax + u \\ y &= Lx \end{aligned} \qquad (28)$$

The value of x at some future lag comprises a deterministic and a stochastic component η that obtains by regarding the effects of the input as a cumulation of local linear perturbations

$$x(t+\tau) = e^{\tau A}x(t) + \eta$$
$$\eta = \int_0^\tau e^{\sigma A}u(t+\sigma)d\sigma \qquad 29$$

Using the assumption that the input is serially uncorrelated

$$\left\langle u(t+\sigma_1)u(t+\sigma_2)^T \right\rangle = \begin{cases} Q, & \sigma_1 = \sigma_2 \\ 0, & \sigma_1 \neq \sigma_2 \end{cases}$$

the covariance of the stochastic part is

$$\begin{aligned} W &= \left\langle \eta\eta^T \right\rangle \\ &= \left\langle \int_0^\tau e^{\sigma_1 A}u(t+\sigma_1)d\sigma_1 \int_0^\tau u(t+\sigma_2)^T e^{\sigma_2 A^T} d\sigma_2 \right\rangle \\ &= \int_0^\tau e^{\sigma A}\left\langle u(t+\sigma)u(t+\sigma)^T \right\rangle e^{\sigma A^T} d\sigma \\ &= \int_0^\tau e^{\sigma A} Q e^{\sigma A^T} d\sigma \end{aligned} \qquad 30$$

It can be seen that when the lag is small $e^{\sigma A} \to 1$ and $W \approx Q$.

Equation (29) is simply a MAR(1) model that could be subject to the usual analysis procedures.

$$x_{t+1} = e^{\tau A}x_t + \eta_t \qquad 31$$

By incorporating the output transformation and observation error we can augment this AR(1) model to a full state-space model with system matrix $F = e^{\tau A}$, input matrix $G = \sqrt{W}$ and observation matrix *L*.

$$\begin{aligned} x_t &= Fx_{t-1} + Gz_t \\ y_t &= Lx_t + \varepsilon_t \end{aligned} \qquad 32$$

where z is some white innovation that models dynamically transformed stochastic input u. This formulation would be appropriate if the state variables were not directly accessible and observation noise ε_t was large in relation to system noise z_t.

A first-order AR(1) model is sufficient to completely model effective connectivity if we could observe all the states with reasonable precision. In situations where only some of the states are observed it is possible to compensate for lack of knowledge about the missing states by increasing the order of the model.

$$\begin{aligned} x_t &= F_1 x_{t-1} + \ldots + F_p x_{t-p} + G z_t \\ y_t &= L x_t + \varepsilon_t \end{aligned} \qquad 33$$

Similar devices are using in the reconstruction of attractor using temporal embedding at various lags. Note that increasing the order does not render the model nonlinear, it simply accommodates the possibility that each region's dynamics may be governed by more than one state variable. However, increasing model order looses any direct connection with formal models of effective connectivity because it is not possible to transform an AR(p) model into a unique DCM. Having said that AR(p) models may be very useful in establishing the presence of coupling even if the exact form of the coupling is not specified (*c.f.* Volterra characterisations).

In summary, discrete-time linear models of effective connectivity can be reduced to multivariate AR(1) (or, more generally ARMA(1,1)) models, whose coefficients can be estimated given only the states (or outputs) by assuming the inputs are white Gaussian and enter with the same strength at each node. They therefore operate under the same assumptions as SEM but are true time-series models. The problem is that MAR coefficients in F can only be interpreted as effective connections when (i) the dynamics are linear and (ii) all the states can be observed though the observation matrix. In this case

$$E = \frac{\partial x_t}{\partial x_{t-1}} = F = e^{\tau A} \qquad 34$$

Compare with Eq(19). However, high-order MAR(p) do represent a useful way of establishing statistical dependencies among the responses, irrespective of how they are caused.

Figure 2 about here

V CONCLUSION

We have reviewed a series of models, all of which can be formulated as special cases of DCMs. Two fundamental distinctions organise these models. The first is whether they pertain to the coupling of inputs and outputs by the nonlinear transformations enacted among hidden states of a single region or whether one is modelling the lateral interactions among the state variables of several systems, each with its own inputs and outputs. The second distinction (see Figure 2) is that between models that require the inputs to be fixed and deterministic as in designed experiments and those where the input is not under experimental control but can be assumed to be well behaved (usually *i.i.d.* Gaussian). Given only information about density of the inputs, or the joint density of the inputs and outputs imposes limitations on the model of effective connectivity adopted. Unless one embraces moments greater than second order only linear models can be estimated.

Many methods for non-linear system identification and casual modelling have been developed in situations where the systems' input was not under experimental control and, in the case of SEM, not necessarily for time-series-data. Volterra kernels and DCMs may be especially useful in neuroimaging because we deal explicitly with time-series data generated by designed experiments.

References

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

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

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

Friston KJ and Büchel C (2000) Attentional modulation of V5 in human *Pro Natl Acad. Sci* USA **97**:7591-7596

Friston KJ (2002) Bayesian estimation of dynamical systems: An application to fMRI. *NeuroImage* **16**:465-483

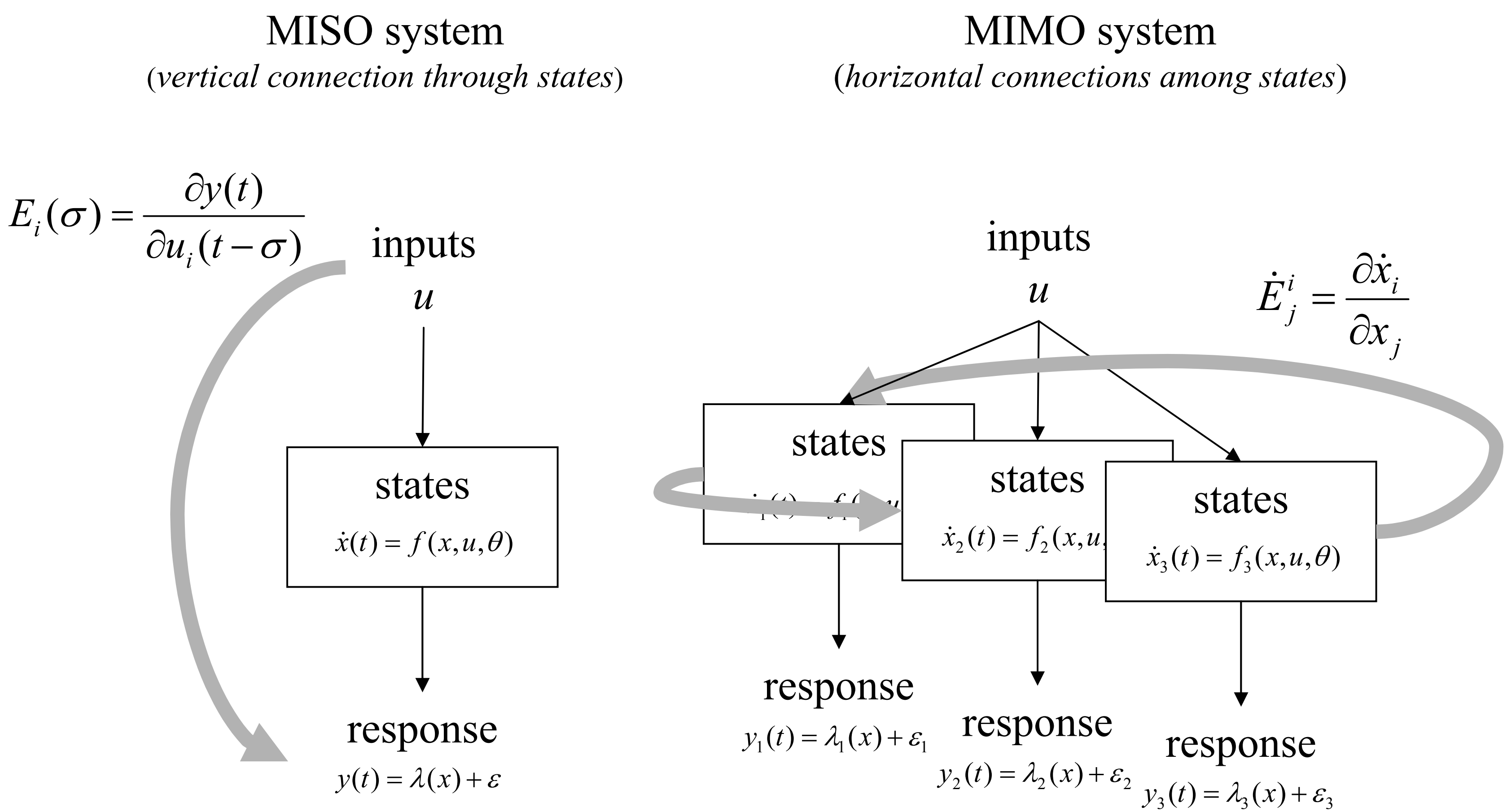
Legends for Figures

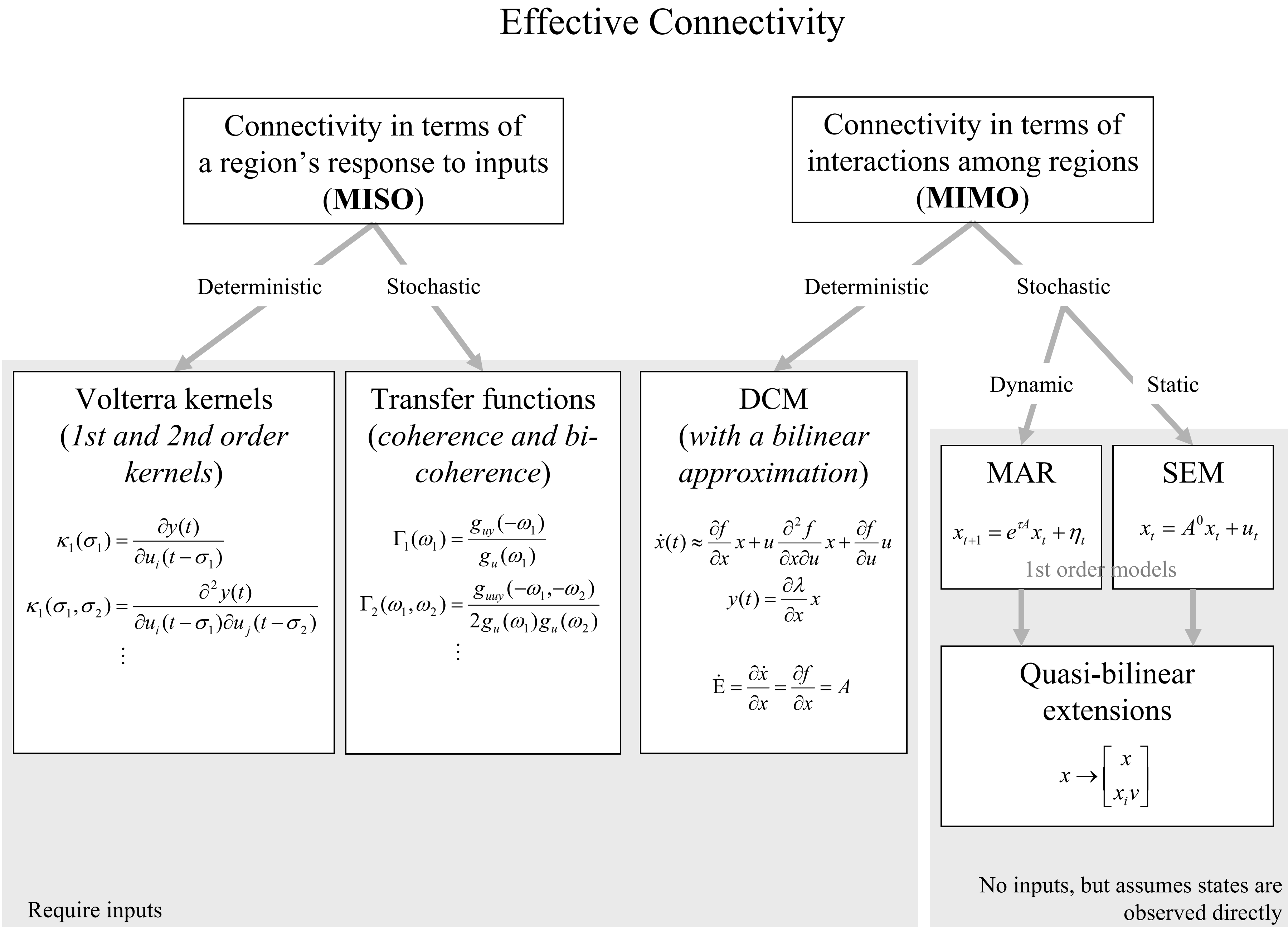
Figure 1

Schematic depicting the difference between analyses of effective connectivity that address the input-output behaviour of a single region and those that refer explicitly to interaction among the states of multiple regions.

Figure 2

Overview of the models considered in this chapter. They have been organised to reflect their dependence on whether the inputs are known or not and whether the model is a time-series model or not.


MISO system
(*vertical connection through states*)
MIMO system
(*horizontal connections among states*)
$E_i(\sigma) = \frac{\partial y(t)}{\partial u_i(t-\sigma)}$
inputs
u
states
$\dot{x}(t) = f(x,u,\theta)$
response
$y(t) = \lambda(x) + \varepsilon$
inputs
u
$\dot{E}_j^i = \frac{\partial \dot{x}_i}{\partial x_j}$
states
states
states
$\dot{x}_3(t) = f_3(x,u,\theta)$
response
$y_1(t) = \lambda_1(x) + \varepsilon_1$
response
$y_2(t) = \lambda_2(x) + \varepsilon_2$
response
$y_3(t) = \lambda_3(x) + \varepsilon_3$


Effective Connectivity
Connectivity in terms of a region's response to inputs (MISO)
Connectivity in terms of interactions among regions (MIMO)
Deterministic
Stochastic
Deterministic
Stochastic
Dynamic
Static
Volterra kernels (1st and 2nd order kernels)
$\kappa_1(\sigma_1) = \frac{\partial y(t)}{\partial u_i(t-\sigma_1)}$
$\kappa_1(\sigma_1,\sigma_2) = \frac{\partial^2 y(t)}{\partial u_i(t-\sigma_1)\partial u_j(t-\sigma_2)}$
Transfer functions (coherence and bi-coherence)
$\Gamma_1(\omega_1) = \frac{g_{uy}(-\omega_1)}{g_u(\omega_1)}$
$\Gamma_2(\omega_1,\omega_2) = \frac{g_{uuy}(-\omega_1,-\omega_2)}{2g_u(\omega_1)g_u(\omega_2)}$
DCM (with a bilinear approximation)
$\dot{x}(t) \approx \frac{\partial f}{\partial x}x + u\frac{\partial^2 f}{\partial x \partial u}x + \frac{\partial f}{\partial u}u$
$y(t) = \frac{\partial \lambda}{\partial x}x$
$\dot{\mathrm{E}} = \frac{\partial \dot{x}}{\partial x} = \frac{\partial f}{\partial x} = A$
MAR
$x_{t+1} = e^{\tau A}x_t + \eta_t$
SEM
$x_t = A^0 x_t + u_t$
1st order models
Quasi-bilinear extensions
$x \rightarrow \begin{bmatrix} x \\ x_i v \end{bmatrix}$
Require inputs
No inputs, but assumes states are observed directly