Middle East Technical University
Biomedical Engineering Department, Bioelectric

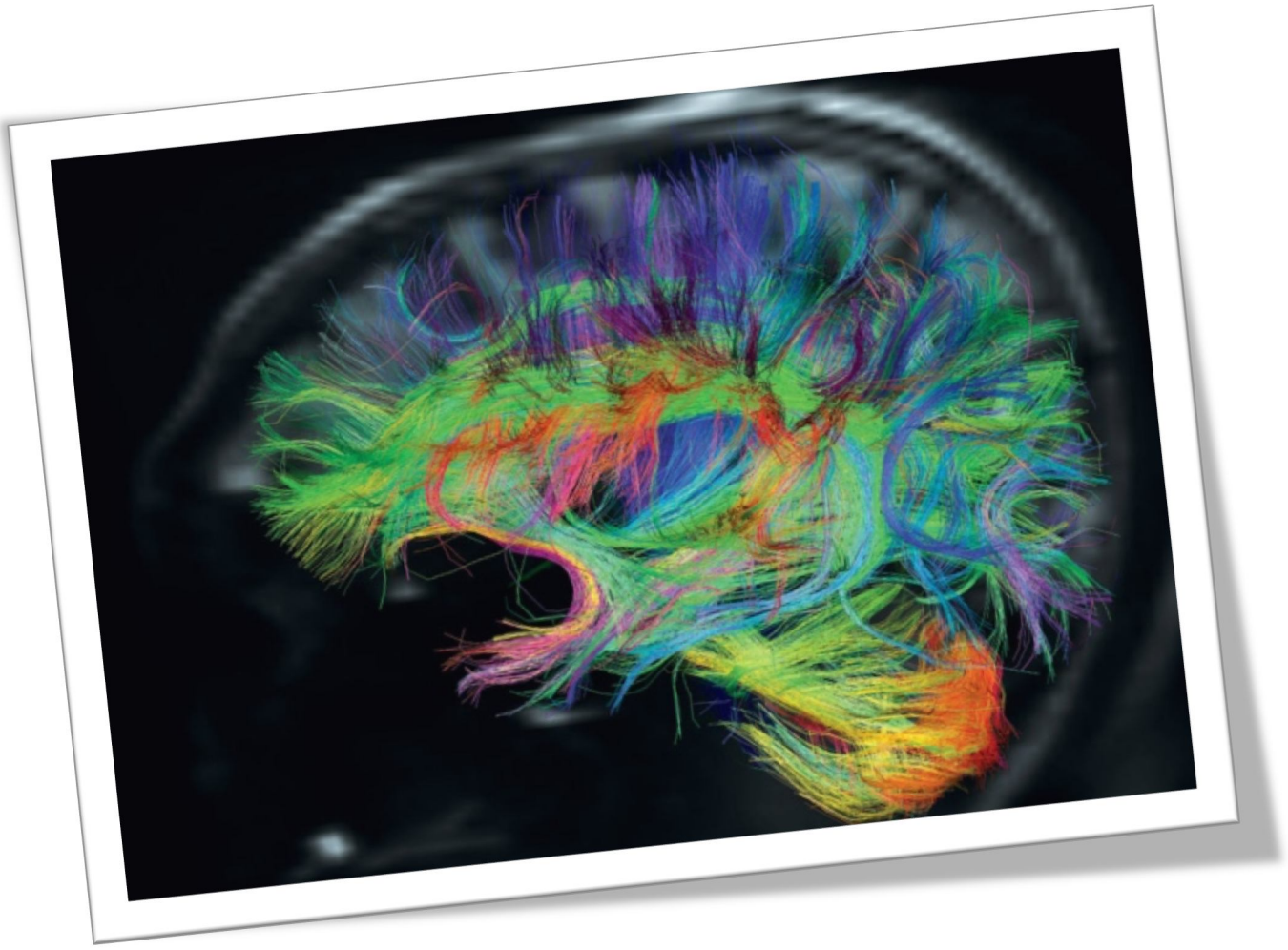
Dynamic Casual Modeling

Report

vol. II

Researcher: Arman Afrasiyabi

Adviser: Assoc. Prof. Ilkay Ulusoy



Abstraction

In this report (volume II), I will try to explain the **parameter estimation** method of the DCM. Then, the inference about the DCM parameters will be covered. We will see the difference between GLM and DCM, and you will also see multifactorial design. Finally, the eight different modes of the DCM will be covered. However, I will try to explain the **Model Selection** of DCM in the third DCM report.

How DCM estimate parameters?

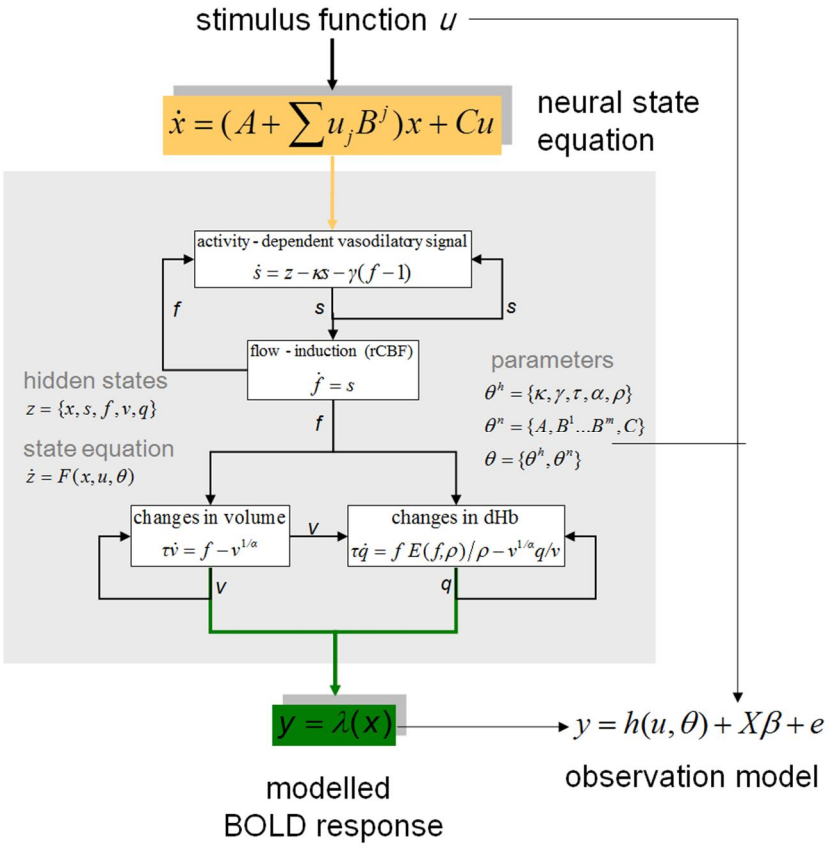
To answer this question, DCM uses a Bayesian approach θ for given some measurements. According to this approach, we must apply two steps in order to do that estimation:

I) We must specify the likelihood function which requires the assumption about the noise.

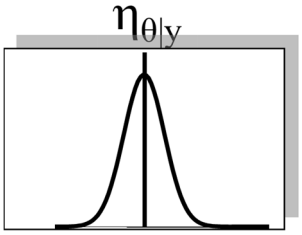
II) We must specify the prior on the parameters by writing down our belief (prior) about the likely arrange of parameter which we are interested.

There are many strategies to implement two conditions, but we use some empirical priors for **hemodynamic priors**. That means we use some of the principle nature priors. A good example of this would be the consideration of blood volume as positive parameter. On the other hands, the other parameters can be **shrinkage priors** which are the priors that are set to be zero and have relative tight variation.

The figure below shows an overview of parameter estimation. We discuss this scheme in the previous report which starts from stimulus function, u. Then it proceeds through a cascade of hemodynamic differential equations. Finally, it gives us a nonlinear equation for the BOLD response. We can add this model with the observation model that adds to observed BOLD response (h(u, θ)) a confound variable ($X\beta$) and an error term (e).

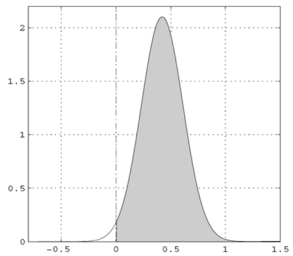


Inference about DCM parameters: Bayesian single-subject analysis



The Bayesian inversion can be done by estimation of parameters, and it is the result of means of variational EM under Laplacian approximation. The output of this approximation would be a posterior distribution for each parameter of a Gaussian shape.

We can ask questions about our parameters given such distribution. For example, how likely is it that a particular effect that is encoded by parameter is bigger than zero? This is very simple question because you can simply compute the cumulative probability under the distribution in relation with some threshold.



$$p = \phi_N \left(\frac{c^T \eta_{\theta|y} - \gamma}{\sqrt{c^T C_{\theta|y} c}} \right)$$

Where, γ is chosen as zero ("does the effect exist?")

The discussed method is the approach when you deal with a single subject. However, you may interest with dealing groups of subjects instead of single one. In this case, you would have some choice for **stochastic analysis**. DCM uses two of these choices: **fixed effects** and **random effects**.

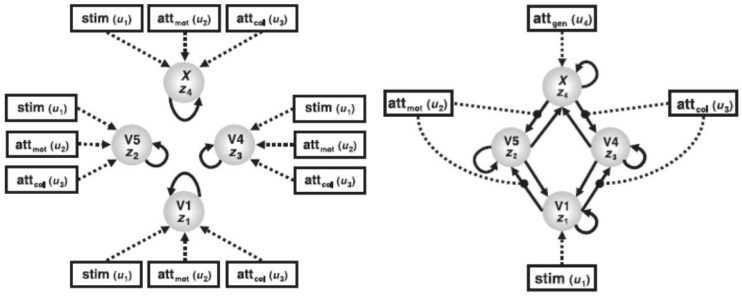
Fixed effects is the simplest method, and it uses the technique called Bayesian parameter averaging. Simply speaking, it uses the posterior of one subject as the prior for the next. The philosophy is that "*Today's posterior is tomorrow prior*". This method can implement over many subjects, and it is cumulative. That means the order is not important.

$$\begin{aligned} p(\theta \mid y_1 \dots y_N) &\propto p(y_1 \dots y_N \mid \theta)\, p(\theta) \\ &\propto p(\theta) \prod_{i=1}^{N} p(y_i \mid \theta) \\ &\propto p(\theta \mid y_1) \prod_{i=2}^{N} p(y_i \mid \theta) \\ &\propto p(\theta \mid y_1, y_2) \prod_{i=3}^{N} p(y_i \mid \theta) \\ &\propto p(\theta \mid y_1 \dots y_{N-1})\, p(y_N \mid \theta) \end{aligned}$$

In **random effects**, you can adapt a very simple frequent perspective. You can take the parameter(s) of interest. Then you can harvest them to each subject. Next, you may plug them to into the second level test which is a T-test. We will see these two choices from practical point of view.

GLM vs. DCM

The question that came to mind at this point is "What type of design is good for DCM?" A good design for DCM is correlated with the **General Linear Models (GLM)**. It means that a good GLM design is also good for DCM. GLM is a mass- univariate which is used in the statistical analysis of fMRI. Both DCM and GLM are trying to do same things; they try to provide an ***explanation*** for local BOLD responses. However, unlike GLM (figure below left), the explanation of DCM (figure below right) is not voxel by voxel. Instead, we want to have regional explanation in DCM.

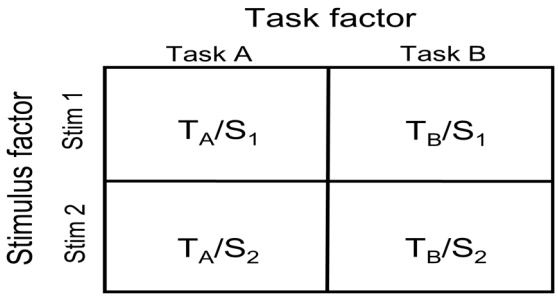


In GLM, we try to explain the activity of the voxels as linear combination of our predicted variables or inputs. On the other hand, in DCM, these inputs do not act on nodes alike. They only enter to certain nodes which are connected to each other in some way. Some of the inputs will impact as much the input on the connections.

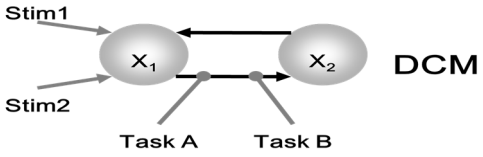
The very important note that we should remember is that when one deal with a deterministic DCM and you cannot be able to detect the activation using GLM, then there is no reason to try a model that absence of that effect with the DCM. Therefore, GLM allows you to specify to decide whether or not that there is something to explain. Once you have done that, then you can use the DCM. In DCM, you cannot test that whether that effect is on or not because you have already done that. Instead, you can try to test a particular multiple ideas of how that effect could have a reason.

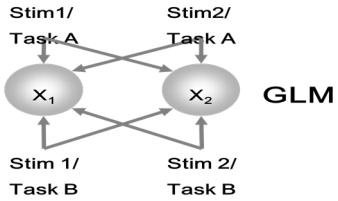
Multifactorial design: explaining interactions with DCM

You can set any design which is suits in terms of GLM. There are some designs that are particularly useful because they allow for specifying questions in a very strong forward way. The best design is factorial one. To understand that why the factorial design is so useful, let us to describe an example. Imagine that there is a 2*2 design of figure below. There are two tasks and two stimuli in this design. For example, task A is “attention on”, B is “attention off”, and stimuli 1 is motion, stimuli 2 is color.

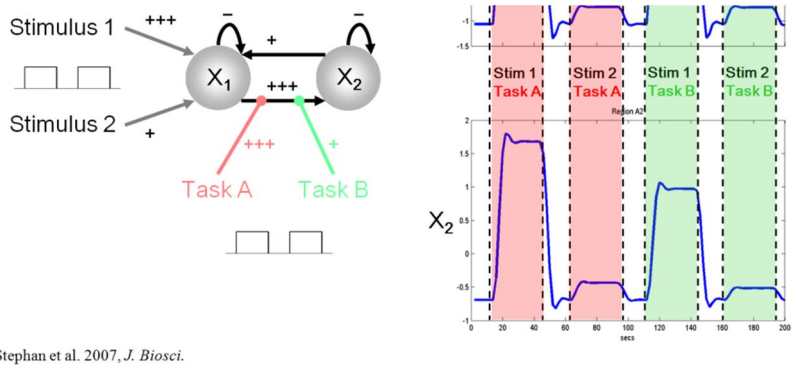


Now, imagine the two voxel in the brain, and you want to apply the GLM. Each of the voxels in that brain would be modeled as linear combination of all **four regresses of design matrix**. Additionally, suppose that you found (by the use of SPM) main effect of stimulus in the first region X1 and interaction stimulus by task in the second region X2. The question, here, is that how would you model this and try to explain these findings using DCM?



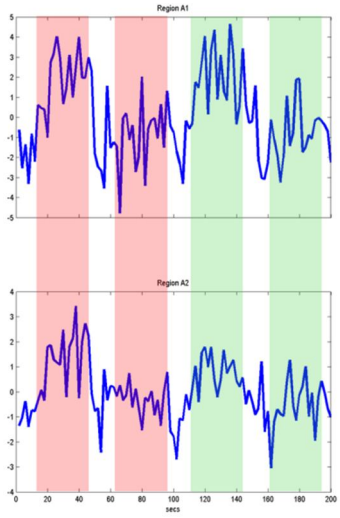


On way of answering this question is to model a differential sensivity of first region to the stimulus as a driving input. Then, you can allow that differential effect to convey to the second region “X2”. The degree in which it conveys depends on which of the two tasks you currently engage in. It is important to note that to visualize more in depth we consider our example in the absence of noise to clarify.



Stephan et al. 2007, *J. Biosci.*

As figure above shows, we have one strong stimuli (+++) input and one weak (+) one. The degree which it is convey depends on the task which has strong modulatory effect for task “A” and weak one for task “B”. In this case, the explanation of that interaction would be that there is a task dependent transfer of differential stimulus responsivity from the first area to the second area. In the real world, it would be like figure below.

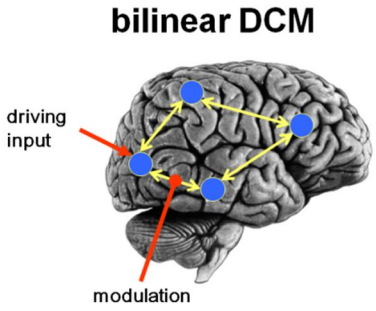


However, DCM is not a fixed model, but it is the framework. It is framework for specifying models. What the software gives us is default models which in principle you can change the ideas such as prior knowledge. Over the lifetime of SPM, changes occurred; therefore, telling the version of the SPM is very important. The detail about the evolutionary change of DCM is given the following web address: www.fil.ion.ucl.ac.uk/spm/software/spm8/SPM8_Release_Notes_r4010.pdf . In DCM10, SPM8, you specify a functional structure by yourself. In other words there are eight ways grouped in three choices of choosing the structure of your model before specifying the connections in your model.

I) Bilinear vs. nonlinear model
II) Single state DCM two state
III) Deterministic vs. stochastic DCM

Bilinear vs. Nonlinear Model

As figure below shows, the **bilinear DCM** is what we have seen before. It allows you to describe the strength of connection changes as a function of some none variables (modulation). As we discussed, this can be described in terms of Taylor Series up to the bilinear terms. The result of this would be the equation that is called state equation.



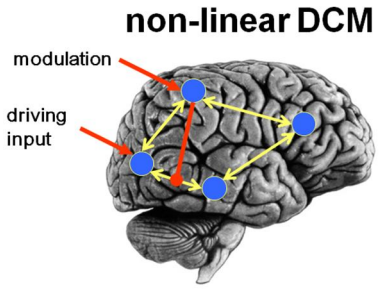
Two-dimensional Taylor series (around x_0=0, u_0=0):

$$\frac{dx}{dt} = f(x,u) \approx f(x_0,0) + \frac{\partial f}{\partial x}x + \frac{\partial f}{\partial u}u + \frac{\partial^2 f}{\partial x \partial u}ux + \ldots$$

Bilinear state equation:

$$\frac{dx}{dt} = \left(A + \sum_{i=1}^{m} u_i B^{(i)}\right)x + Cu$$

However, you may want to describe the way of coupling which may change as a function of activity in some region. In other words, how a system gets the coupling between other regions may be interested. In this case, the second tem of Taylor Series must be added and that will give us a **nonlinear differential equation**. In other words, you must add the matrixes that include a multiplicative change in coupling strength.

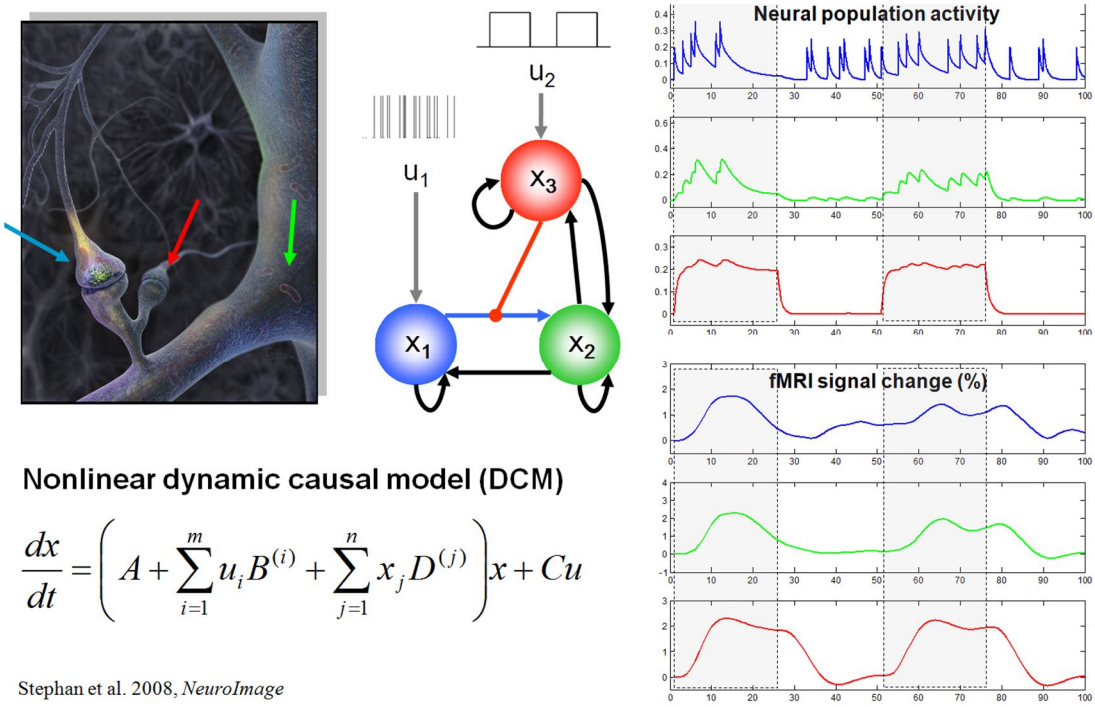


$$\frac{dx}{dt} = f(x,u) \approx f(x_0,0) + \frac{\partial f}{\partial x}x + \frac{\partial f}{\partial u}u + \frac{\partial^2 f}{\partial x \partial u}ux + \frac{\partial^2 f}{\partial x^2}\frac{x^2}{2} + \ldots$$

Nonlinear state equation:

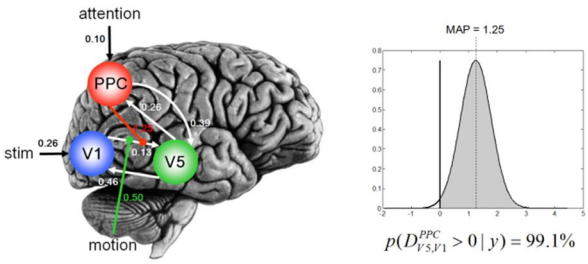
$$\frac{dx}{dt} = \left(A + \sum_{i=1}^{m} u_i B^{(i)} + \sum_{j=1}^{n} x_j D^{(j)}\right)x + Cu$$

To visualize this case, please imagine the figure below (the neurons). The green population of cells that can have input from blue population of cells and the synaptic input is regulated by the activity coming from red input. If you generated the stimuli like the blue one, then the green populations will response to it. But, this response will occur only in the presence of high red input. When the red population is off (shown as the second period of the figure below), the green population will not response to the blue even though the blue provide rather strong input. The reason of this would be that the red population enables the strength in the multiplication fashion.



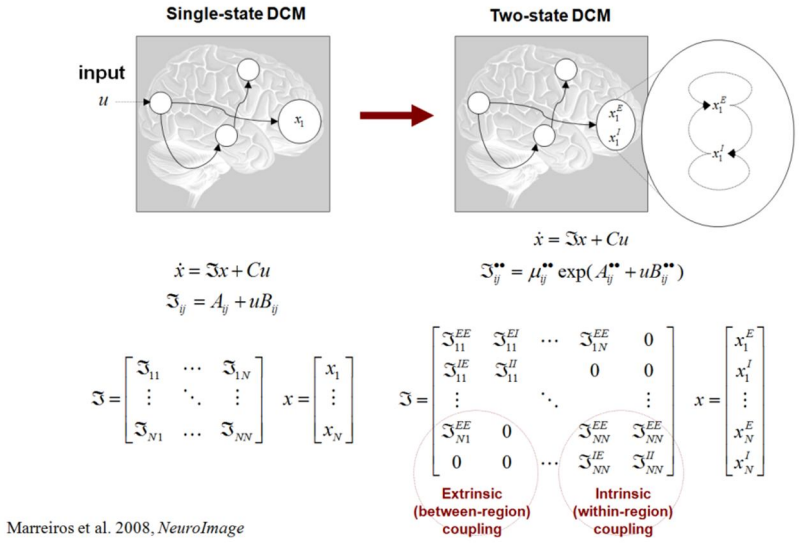
The result is shown in the figure above is in the form of the BOLD signal. You also can see non-linear response in fMRI signal change rate. You can, for example, see that the green BOLD signal is washed out. DCM tries to bring back this washed out BOLD signal.

Figure in the opposit shows a very simple example using attention and motion dataset. In this example, the attention to motion raises in V5 such that attention introduce a base line shift in perioral cortex. For example, this could be a projection of brain stem, reasoning activity in anatomic fashion in parietal cortex. This activity changes the labeling V1 to V5 and causes that this connection to become stronger. You can now trigger the posterior distribution of that gating parameter and find that it has high probability under that model given that model.



Single state DCM vs. Two states

The other two parameters (**Single state DCM vs. two states** and **Deterministic vs. stochastic DCM**) can be used through the GUI. The first variant is to that have two states per region instead of one state. These two states are map on to putative excitatory and inhibitory population in your region.



In other words, you now model each region that show how excitatory and inhibitory neurons connect to each other. When you couple regions to each other, you only allow for connections that connect to other regions that are in the cortex which are the way that long rem projection made up of. However, it is not the case for subcortical region. It is clearly dependent on what is you trying to model. If you are not sure, you have to use the model selection.

Deterministic vs. stochastic DCM

The final development of DCM has this option which includes a **stochastic (w)** component at neuronal level. "w" can account for neuronal fluctuations. For example, when the code is at the resting state data, this model can be applied for two different purposes. First, it absorbs of unknown influence in the task data and accounts for things that you have not modeled. Second, it freezes you for necessity to provide driving inputs in the resting state case. However, we need more parameter estimation for this model, more time and more complexity.

$$\frac{dx}{dt} = f\left(x,u,\theta\right) + \omega$$