Tractography-based priors for dynamic causal models

Klaas Enno Stephan [a,b,*,1], Marc Tittgemeyer [c,1], Thomas R. Knösche [d], Rosalyn J. Moran [b], Karl J. Friston [b]

[a] Laboratory for Social and Neural Systems Research, Institute for Empirical Research in Economics, University of Zurich, Switzerland
[b] Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, WC1N 3BG, London, UK
[c] Max Planck Institute for Neurological Research, Cologne, Germany
[d] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

ARTICLE INFO

ABSTRACT

Functional integration in the brain rests on anatomical connectivity (the presence of axonal connections) and effective connectivity (the causal influences mediated by these connections). The deployment of anatomical connections provides important constraints on effective connectivity, but does not fully determine it, because synaptic connections can be expressed functionally in a dynamic and context-dependent fashion. Although it is generally assumed that anatomical connectivity data is important to guide the construction of neurobiologically realistic models of effective connectivity; the degree to which these models actually profit from anatomical constraints has not yet been formally investigated. Here, we use diffusion weighted imaging and probabilistic tractography to specify anatomically informed priors for dynamic causal models (DCMs) of fMRI data. We constructed 64 alternative DCMs, which embodied different mappings between the probability of an anatomical connection and the prior variance of the corresponding of effective connectivity, and fitted them to empirical fMRI data from 12 healthy subjects. Using Bayesian model selection, we show that the best model is one in which anatomical probability increases the prior variance of effective connectivity parameters in a nonlinear and monotonic (sigmoidal) fashion. This means that the higher the likelihood that a given connection exists anatomically, the larger one should set the prior variance of the corresponding coupling parameter; hence making it easier for the parameter to deviate from zero and represent a strong effective connection. To our knowledge, this study provides the first formal evidence that probabilistic knowledge of anatomical connectivity can improve models of functional integration.

© 2009 Elsevier Inc. All rights reserved.

Introduction

One of the key themes in biology is the characterisation of structure–function relationships. For example, the range of functional interactions a protein can engage in depends on its three-dimensional structure. In the brain, a similar relationship exists between anatomical and effective connectivity. The former denotes the presence of axonal connections among neurons or neuronal populations and the latter refers to the causal influences that are mediated by these connections (Aertsen and Preißl, 1999; Friston, 1994). It is generally accepted that anatomical connectivity provides important constraints on effective connectivity. For example, it has been shown that the functional repertoire of a cortical area (the "functional fingerprint") is closely related to the pattern of its anatomical connections (the "connectional fingerprint") (Passingham et al., 2002). Similarly, network analyses of anatomical connectivity (Hilgetag et al., 2000) and functional interactions in the Macaque cortex (Stephan et al., 2000) have revealed

similar clustering in the cortical network and indicated that, both from an anatomical and functional perspective, the Macaque cortex possesses small world properties. Finally, a number of analytical and simulation studies have recently started to explore how different types of structural network topologies are linked to different types of neuronal dynamics (Bullmore and Sporns, 2009; Honey et al., 2007; Sporns et al., 2002; Strogatz, 2001).

It is important to note that anatomical connectivity constrains but does not determine effective connectivity. There are several reasons for this. First, the function of a synapse depends on its recent history. For example, in the absence of any structural changes of the synapse *per se*, marked facilitation or depression of synaptic transmission can occur at a timescale of milliseconds (Zucker and Regehr, 2002). A second, and probably more important, reason why there is no one-to-one mapping between anatomical and effective connectivity is that the structural presence of a synaptic connection does not determine whether it will be engaged during a particular process or not. Various mechanisms exist by which synaptic connections can be enabled or disabled in a dynamic fashion at the timescale of milliseconds. These mechanisms include gating and gain control mechanisms, which render synaptic transmission dependent on the current membrane potential and the history of other synaptic inputs nearby (see

---

* Corresponding author. Laboratory for Social and Neural Systems Research, Institute for Empirical Research in Economics, University of Zurich, Switzerland. Fax: +41 44 6344907.
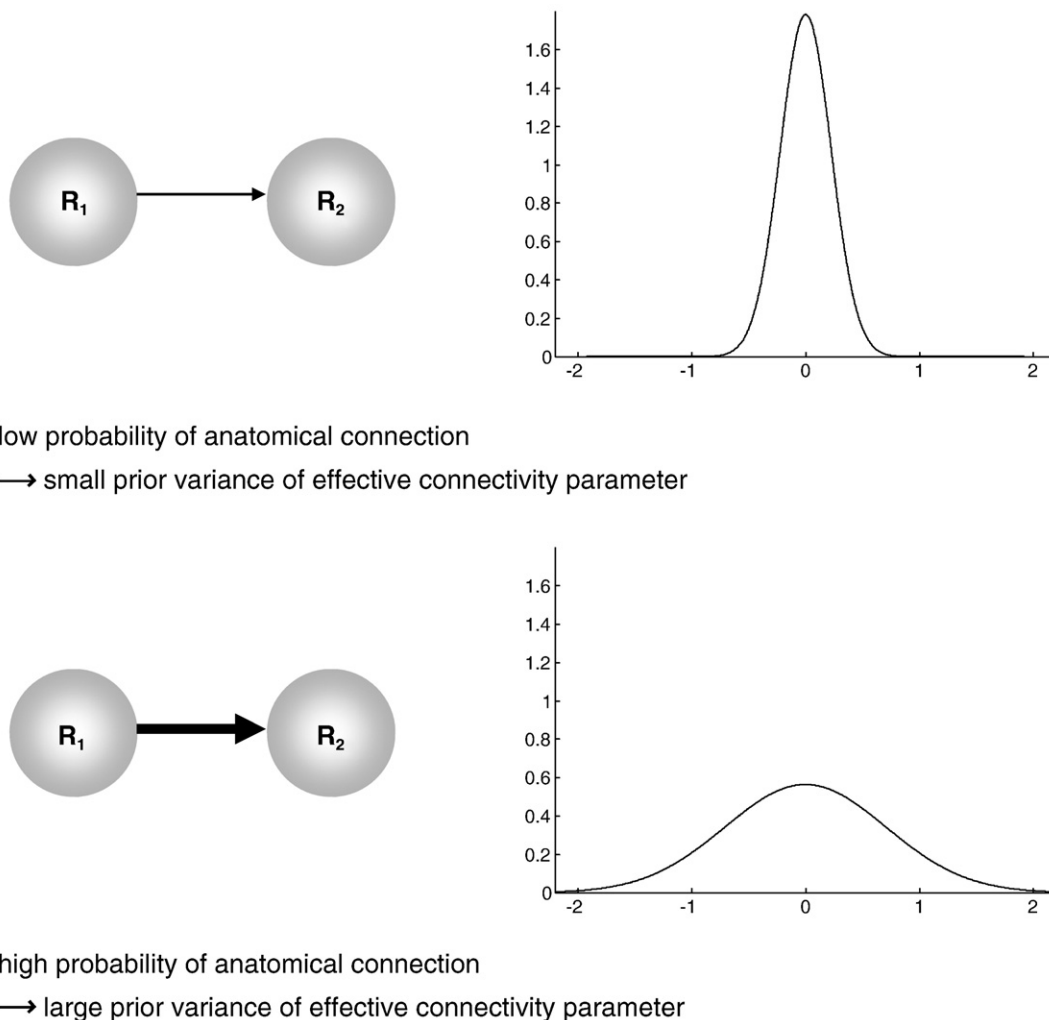E-mail address: k.stephan@iew.uzh.ch (K.E. Stephan).
[1] These authors contributed equally to this work.

Discussion). These transient and nonlinear effects are important for explaining the dynamics of functional interactions among neuronal populations (Friston et al., 1995; McIntosh, 2000; Salinas and Sejnowski, 2001; Stephan et al., 2008a).

Major efforts have been made over the last decade to make anatomical connectivity data available for constructing better models of brain function, ranging from construction of large scale databases of anatomical tract tracing data (Bota et al., 2005; Burns and Young, 2000; Scannell et al., 1999; Stephan et al., 2001) to analyses of human brain connectivity by non-invasive diffusion weighted imaging (Alexander, 2005; Behrens et al., 2003; Jones et al., 1999; Kaden et al., 2008; Le Bihan, 2003; Mori and Zhang, 2006; Parker et al., 2002; Tuch et al., 2005). So far, however, no study has formally investigated to what degree models of effective connectivity profit from detailed quantitative knowledge of anatomical connectivity. Here, we present such an investigation by combining probabilistic tractography and dynamic causal models (DCMs) (Friston et al., 2003) of fMRI data. Our approach rests on a simple idea: we use an estimate of the probability that a given anatomical connection exists, as provided by tractography, to constrain the likely range of the effective connection strength. More formally, the anatomical likelihood of a given connection is used to inform the prior variance of the corresponding coupling parameter in the DCM. Such anatomically informed priors can have different forms. The most intuitive notion is that the higher the likelihood that a

given connection exists anatomically, the larger one should set the prior variance of the corresponding effective connectivity, making it easier for the parameter to deviate from its prior mean of zero and therefore represent a strong (negative or positive) connection. This is shown schematically by Fig. 1.

In the present study, we constructed 64 alternative DCMs, each of which embodied a different mathematical mapping between the anatomical probability of a given connection and the prior variance of the corresponding coupling parameter. Some models embodied the (intuitive) notion that the prior variance should increase with anatomical likelihood; these models only differed with regard to the mathematical relationship between anatomical connection likelihood and prior variances in DCM. Other models represented the counter-intuitive notion that the prior variance of coupling parameters in DCM should decrease with increasing anatomical likelihood. As a reference, we also used several naive models in which the prior variances of the coupling parameters in DCM were independent of the anatomical likelihood of the corresponding connections. Using Bayesian model selection (Penny et al., 2004; Stephan et al., 2007b), we tested which of these models best explained experimentally measured fMRI data in a DCM of four visual areas (Stephan et al., 2007a), using data from a group of twelve healthy volunteers. Our comparisons showed that the best model (i.e., the model with the highest evidence) is one in which anatomical probability increases the prior variance of DCM coupling

low probability of anatomical connection
⟶ small prior variance of effective connectivity parameter

high probability of anatomical connection
⟶ large prior variance of effective connectivity parameter

Fig. 1. This figure provides a schematic summary of the intuitive notion that a higher probability that a connection between two regions R_1 and R_2 exists anatomically should be associated with a larger prior variance of the corresponding effective connectivity parameter in DCM, hence making it easier for the parameter to deviate from its prior mean of zero and therefore represent a strong (negative or positive) effective connection.

parameters in a nonlinear and monotonic (sigmoidal) fashion. In contrast, models that ignore anatomical connectivity have substantially less evidence. To our knowledge, this study provides the first empirical evidence that probabilistic knowledge of anatomical connectivity can improve models of functional integration in the brain.

This paper comprises three sections. In the first we review DCM, with a special focus on the construction of anatomically informed priors on the coupling parameters. In the second section, we describe the data and how probabilistic tractography measures were derived. In the final section, we present the results of Bayesian model comparisons used to establish if and how anatomical constraints are evident in functional data.

Theory

Dynamic causal modelling (DCM)

DCM for fMRI uses a bilinear model of neural dynamics in a system of n distributed brain regions, where neural population activity in each region is represented by a single state variable. DCM models the change of this neural state vector x using the following differential equation:

$$\frac{dx}{dt} = \left(A + \sum_{j=1}^{m} u_j B^{(j)} \right) x + Cu. \tag{1}$$

Here, the A matrix represents the endogenous (context-independent or fixed) strength of connections between the regions, and the matrices $B^{(j)}$ represent the modulation of these connections (e.g. due to learning, attention, etc.) induced by the jth input u_j. Finally, the C matrix represents the influence of direct (exogenous) inputs to the system (e.g. sensory stimuli). Note that all parameters are rate constants and are thus in units of s^{-1}.

To explain regional BOLD responses, DCM for fMRI combines this model of neural dynamics with a biophysically motivated hemodynamic model; for details see Friston et al. (2000) and Stephan et al. (2007b). Together, the neural and hemodynamic state equations

Fig. 2. This figure shows different Gaussian priors for coupling parameters in the DCM shown by Fig. 3. The dotted black line represents the prior variance, $v = 0.0405$, that was originally suggested as a default value for a four-area DCM (Friston et al., 2003). The solid grey lines represent connection-specific prior variances that result from transforming the anatomical connection probabilities φ_{ij} (Fig. 3) using Eq. (6) and the hyperparameters of the optimal model m_{45} (with hyperparameters $\alpha = 4$, $\beta = 12$; see Fig. 4).

furnish a deterministic forward model with hidden biophysical states. For any given combination of parameters θ and inputs u, the measured BOLD response y is modelled as the predicted BOLD signal $h(u,\theta)$ plus a linear mixture of confounds $X\beta$ (e.g., signal drift) and Gaussian observation error e:

$$y = h(u,\theta) + X\beta + e. \tag{2}$$

DCM uses a fully Bayesian approach to parameter estimation, with empirical priors for the hemodynamic parameters and conservative shrinkage priors for the coupling parameters; see Friston (2002) and Friston et al. (2003) for details. Briefly, the posterior moments are updated iteratively using variational Bayes, under a fixed-form Laplace (i.e., Gaussian) approximation, $q(\theta)$, to the conditional density $p(\theta|y)$. This uses gradient ascent on a free-energy bound on the log-marginal likelihood, $\ln p(y|m)$, for a particular model, m. This optimises the maximum a posteriori (MAP) estimates of the parameters in the E-step of an EM algorithm, whereas the M-step optimises hyperparameters λ that control the covariance components of the observation error e.

Of particular interest are the Gaussian shrinkage priors $p(\theta|m)$ that constrain the estimates of the coupling matrices (i.e. the A and B matrices in Eq. (1)). They are called shrinkage priors because they have a prior mean of zero and a small prior variance (i.e. high prior precision) and thus shrink the posterior estimates of coupling parameters in DCM towards zero. The choice of shrinkage priors in DCM was motivated by two reasons: (i) they enforce conservative parameter estimation, and (ii) ensure the system is dissipative.[2] These considerations led to a quantitative heuristic for the prior variance of coupling parameters, which depends on the number of areas in the system (Friston et al., 2003). For example, for a four-area model as in the present study, the prior variance of coupling parameters is set to 0.0405. Fig. 2 shows this default prior (bold dotted line) and contrasts it with anatomically informed priors (specifically, those which were found to be optimal by our model selection procedure described below).

Critically, the higher the prior variance of coupling parameters, the easier it is for posterior estimates to deviate, in either direction, from the prior mean of zero. In DCM, the shrinkage priors on the coupling parameters are identical for all connections in the system, regardless of whether they are anatomically likely or not. In the next section, we motivate a scheme by which anatomical measures from probabilistic tractography can be used to define the prior variance for individual connection strengths and thus constrain dynamic causal models anatomically.
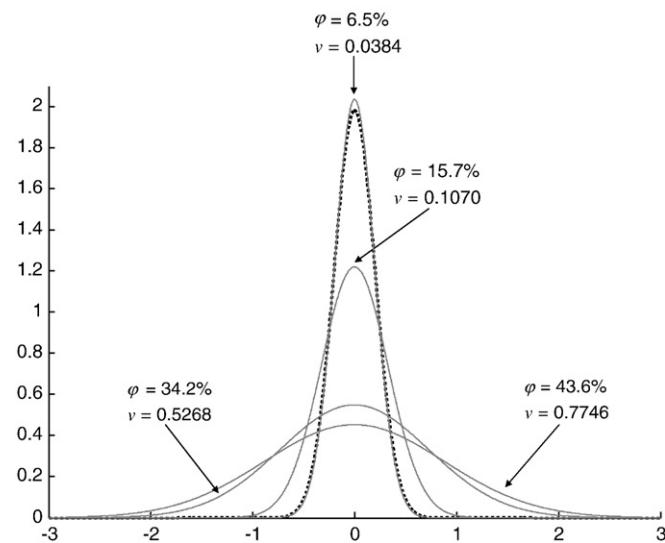
Anatomically informed priors

Anatomical measures

Imagine that we are given some (probabilistic) measure φ_{ij}^{*} of anatomical connectivity between regions i and j in the DCM. Generally, these will describe a probability density; for example, the probability of a streamline from a seed region to reach a target region. Because these measures are densities their absolute values depend on their units of measurement (e.g., per voxel). We can resolve this dependency by working with the relative probabilities on n connections

$$\varphi_{ij} = \frac{\varphi_{ij}^{*}}{\sum_{ij} \varphi_{ij}^{*}}. \tag{3}$$

This ensures that $\sum \varphi_{ij} = 1$ and furnishes a measure of anatomical strength for any one connection, relative to all others. In Fig. 3B, φ_{ij} and

---

[2] In DCM as implemented in SPM5/SPM8, stability of the system is guaranteed by using lognormal priors on the negative self-connections; see the discussion on stability in Stephan et al. (2008).

φ_{ij}^* are shown alongside the connections of the DCM considered in this study. We now have to consider how this relative measure might enter a probabilistic model of effective connectivity like a DCM.

Anatomically informed priors

In DCM, priors on the effective connectivity θ_{ij} between regions i and j are Gaussian shrinkage priors (see above)

$$p(\theta_{ij}) \propto \exp\left(\tfrac{1}{2}\Pi_{ij}\theta_{ij}^2\right) \qquad (4)$$

where Π_{ij} is the prior precision of the connection. When the precision is large, the connections are constrained to be small; i.e., they shrink to their prior expectation of zero. This form allows us to specify anatomically informed priors $p(\theta_{ij}) \rightarrow p(\theta_{ij}|\varphi_{ij})$, which change monotonically as a function of the anatomical connection strength, by making the precisions a function of φ_{ij}:

$$\Pi_{ij} = \Pi_0 + \exp(\alpha - \beta\varphi_{ij}). \qquad (5)$$

If $\beta > 0$, as the relative strength φ_{ij} increases, the precision will decrease toward a lower bound Π_0. In other words, if the anatomical measure suggests that a given connection is more probable than other connections, then the prior precision will be small, relaxing the shrinkage on the posterior estimates which, consequently, can take large values in either positive or negative direction. The relative effects of anatomical constraints will be less and less pronounced as β approaches zero. At $\beta = 0$, the anatomical measures play no role, and

we are dealing with "anatomically uninformed" priors. Here, the degree of shrinkage is the same for all connections and is determined entirely by $\exp(\alpha) \geq 0$. Finally, for $\beta < 0$, the prior precision increases with the anatomical likelihood of a given connection, resulting in counterintuitive priors that shrink the posterior estimates more strongly for probable connections than for improbable ones. In summary, one can regard the hyperparameters α and β as controlling generic and anatomical constraints, respectively. Because these hyperparameters are unknown, they have to be optimised with respect to the model evidence, as described below.
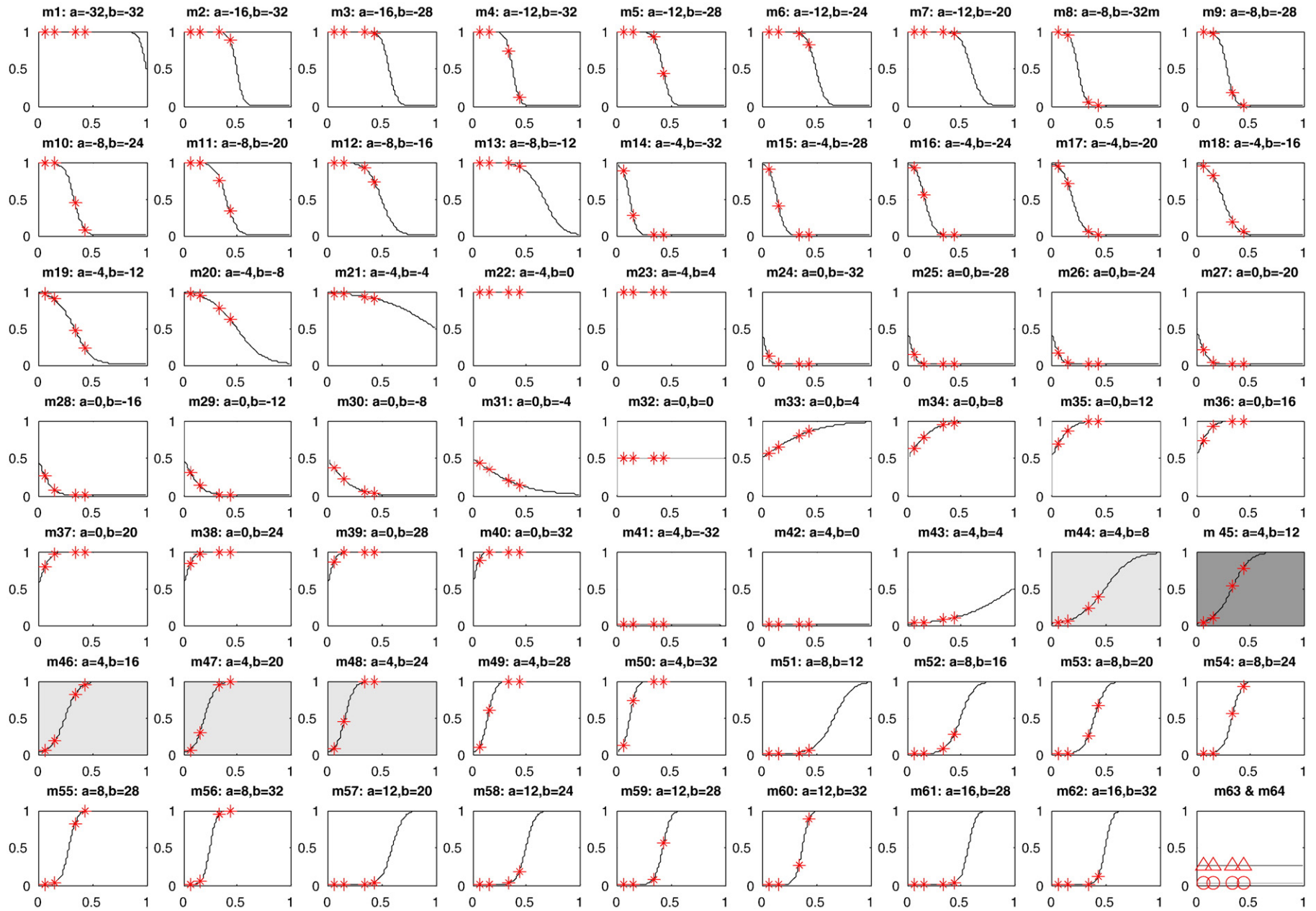
Prior variances

Usually, priors are specified not in terms of precisions but in terms of variances $\sum_{ij} = \Pi_{ij}^{-1}$. In the case of anatomically informed priors we have

$$\Sigma_{ij} = \frac{1}{\Pi_0 + \exp(\alpha - \beta\varphi_{ij})} = \frac{\Sigma_0}{1 + \Sigma_0 \exp(\alpha - \beta\varphi_{ij})}. \qquad (6)$$

This is simply a logistic sigmoid function of φ_{ij}. Here, the sigmoid is bounded by the upper limit on variance $\sum_0$ (e.g., imposed by dynamical stability priors), where $\sum_0$ is the inverse of Π_0 in Eq. (5). The point of inflection (i.e., shift) is determined by α, while β controls the slope (i.e., gain) of the sigmoid. Most importantly, the prior variance; (i) increases monotonically with φ_{ij} if $\beta > 0$, (ii) is unaffected by anatomical information if $\beta = 0$, and (iii) decreases monotonically with φ_{ij} if $\beta < 0$. This perspective on anatomically informed priors $p(\theta_{ij}|\varphi_{ij})$ highlights its flexibility and form (see also Fig. 4).
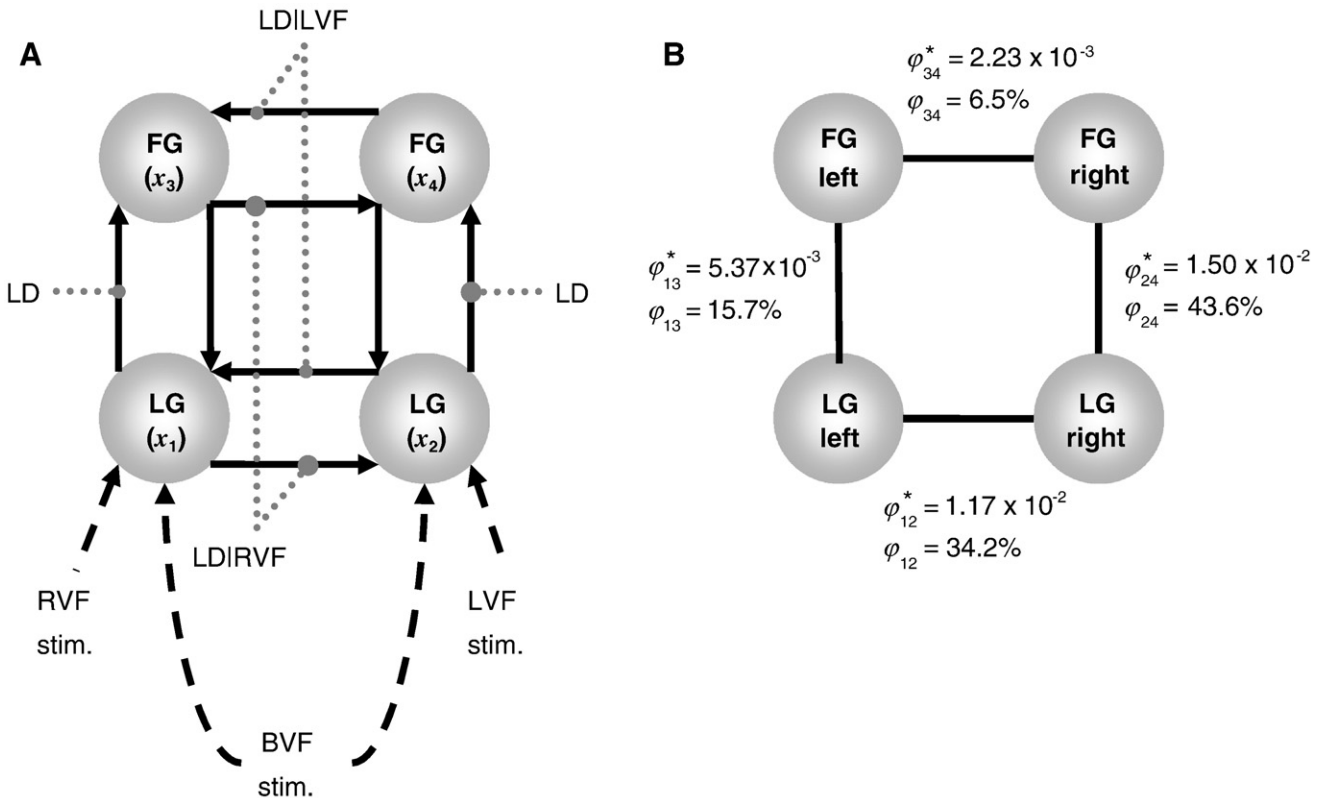
Fig. 3. (A) Summary of the DCM used in this evaluation study; for details see (Stephan et al., 2007a). This four-area model included the reciprocally connected lingual gyrus (LG) and fusiform gyrus (FG) in both hemispheres (black solid lines). Non-foveal visual stimuli (words) were presented in either the right (RVF) or left (LVF) visual field with a randomized stimulus onset asynchrony between 1.5 and 2.5 s during 24 s blocks; these were modelled as individual events driving contralateral LG activity (black dashed lines). During the instruction periods, bilateral visual field (BVF) input was provided for 6 s; this was modelled as a box-car input to LG, in both hemispheres. Connections were modulated by task and stimulus properties (grey dotted lines). Intra-hemispheric LG→FG connections were allowed to vary during a letter decision (LD) task, regardless of visual field. In contrast, inter-hemispheric connections were modulated by task conditional on the visual field (LD|LVF and LD|RVF). (B) Mean anatomical connection probabilities at the group level. For each connection, its probability φ_{ij}^* was computed using each participating region as a seed region once and then computing the average. This was done for each subject and the results were then averaged across all 42 subjects. For all connections the standard error across subjects was considerably smaller than the mean connection strength: $SE(\varphi_{12}) = 3.18 \times 10^{-3}$, $SE(\varphi_{13}) = 1.08 \times 10^{-3}$, $SE(\varphi_{24}) = 2.90 \times 10^{-3}$ and $SE(\varphi_{12}) = 7.68 \times 10^{-4}$. φ_{ij} represent normalised connection probabilities (c.f. Eq. (3)).

Fig. 4. This figure summarises all 64 models tested by showing their mapping functions (specified by hyperparameters α and β; see Eq. (6)) that converted anatomical connection probabilities into prior variances of effective connectivity parameters in the DCM. Red stars represent the prior variances that result from transforming the anatomical connection probabilities φ_{ij} shown in Fig. 3. The plot in the right lower corner shows the prior variances for two models, m_{63} (red triangles) and m_{64} (red circles). The top 5 models, as established by Bayesian model selection (see Fig. 5 for details), are shown with a grey background. Note that these were the only models with a non-negligible posterior probability (larger than 10^{-4}; compare Fig. 5C). The best model, m_{45}, is highlighted by a dark grey background; its prior variances for the effective connectivity parameters are shown in Fig. 2.

Are anatomically informed priors useful for DCM?

The critical issue is, of course, whether DCMs with anatomically informed priors are better than DCMs which discount anatomical connectivity. This can be addressed by comparing the model evidence of DCMs with different hyperparameters $\{\alpha, \beta, \sum_0\}$, using Bayesian model selection (Penny et al., 2004; Stephan et al., 2007b). In particular, the value of β which optimises the model evidence indicates whether and how anatomically informed priors improve the model.

In this study, we compared different anatomically informed and uninformed priors using an established four-region DCM of visual responses measured during a paradigm requiring letter decisions and spatial decisions about identical word stimuli (Stephan et al., 2007a); here, we focus on the modulation of connectivity induced by letter decisions. The fMRI data came from a group of twelve healthy volunteers (all male and right-handed, mean age 24.9 years, SD 3.4).

Using this DCM, we fixed the upper bound hyperparameter $\sum_0$ to unity and optimised the remaining hyperparameters with respect to the model evidence by searching over model space; where each model had different values of α and β. Specifically, both these hyperparameters were varied in steps of 4 over the range $[-32, 32]$, resulting in 289 different models. For each of these models, we computed the associated prior covariance matrix according to Eq. (6), based on anatomical connectivity estimates which are described in the next section and summarised by Fig. 3B. In a subsequent pruning step, we

compared these matrices and removed those showing only small differences (i.e., where the maximal difference between corresponding elements in the covariance matrices was less than 10^{-2}). This resulted in 62 sets of hyperparameters and associated models, which will be referred to as $m_1 \cdots m_{62}$. These DCMs were distinguished only by different anatomically informed ($\beta \neq 0$) or anatomically uninformed ($\beta = 0$) priors. As additional controls, we considered two further models with anatomically uninformed priors: one model, m_{63}, used the default shrinkage priors in DCM (i.e. a prior variance of 0.0405 for each connection; see red circles in Fig. 4). The second, m_{64} was constructed by assuming that each connection was equally likely (i.e., $\varphi_{ij} = 25\%$), resulting in the same prior covariance for each connection (i.e. 0.2689; see red triangles in Fig. 4). All 64 models were fitted to the fMRI data from each of the 12 subjects and subsequently evaluated using Bayesian model selection (BMS), as described next.

Bayesian model selection (BMS)

A decision about which of several competing models is optimal cannot be based only on the relative fit to the data (i.e., accuracy) but also needs to take into account differences in model complexity (i.e., the number of free parameters and the functional form of the generative model, Pitt and Myung, 2002). Penalizing for model complexity is important because while accuracy increases monotonically with complexity, at some point the model will start fitting
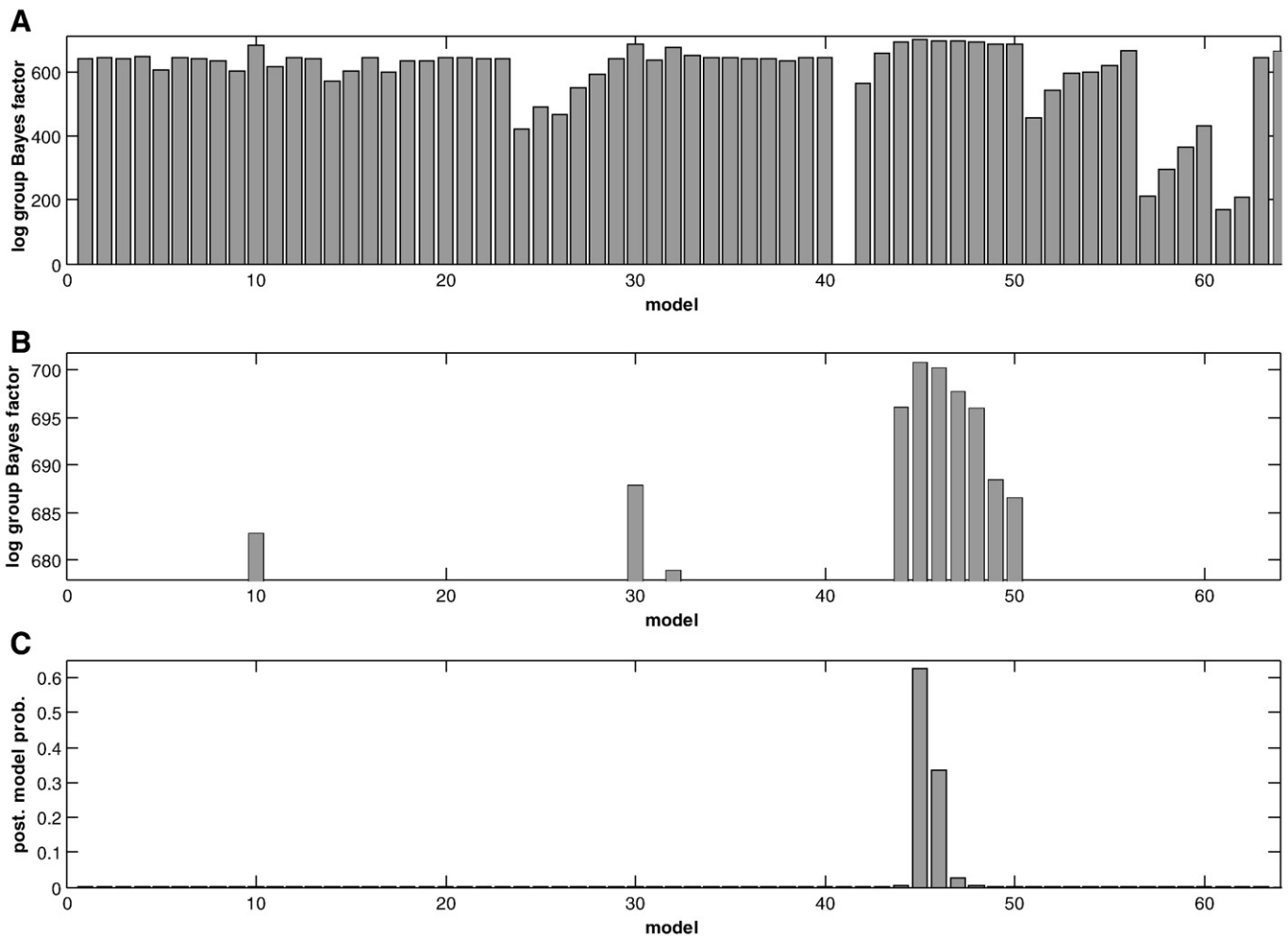
Fig. 5. The log-evidence (pooled over subjects) for all 64 models, in relation to the worst model (m_{41}), are shown in subplot A. Because the scaling makes it difficult to recognize the best model at a glance, the graph is redrawn in subplot B and thresholded such that only the top 10 models are plotted. It can now be seen easily that the best model is model m_{45}. Furthermore, comparing this result with Fig. 4, it becomes apparent that 7 out of the 10 best models possess anatomically informed priors that accord with the intuitive notion that prior variance should increase with the probability or strength of an anatomical connection. The posterior probabilities of all models are shown by subplot C.

noise that is specific to the particular data (i.e., "over-fitting"). Therefore, models that are too complex show poor generalisability. Given that all models are equally likely a priori, the question "which is the optimal model?" can be reformulated as "which model represents the best balance between fit and complexity?" This is the model that maximizes the marginal likelihood or model evidence:

$$p(y|m) = \int p(y|\theta, m)p(\theta|m)d\theta. \tag{7}$$

Here, the integration subsumes the number and conditional dependencies among model parameters θ (Friston et al., 2007; Penny et al., 2004). In this study, we approximate the log-evidence with the same free-energy bound, F, that is used to optimise the parameters above (Friston et al., 2007; Mac Kay, 2003; Neal and Hinton, 1998). As detailed elsewhere (Stephan et al., 2009), the advantage of F over other approximations (like the Akaike or Bayesian Information Criteria), is that it accounts properly for conditional dependencies among the parameter estimates. The derivation of F and its mathematical interpretation with regard to model accuracy and complexity have been described in detail in previous publications (Friston et al., 2007; Penny et al., 2004; Stephan et al., 2007b).

To quantify the relative evidence for two models m_i and m_j at the group level, we report differences in their log-evidence summed across subjects (c.f. Fig. 5); this is equivalent to using log group Bayes factors (Stephan et al., 2007b)

$$\ln \prod_n BF_{ij}^{(n)} = \ln \prod_n \frac{p(y_n|m_i)}{p(y_n|m_j)} \approx \sum_n F_i^{(n)} - F_j^{(n)} \tag{8}$$

where n is an index over subjects. A difference in log-evidence (free-energy) of three or more is generally considered to be strong evidence for one model over another (Kass and Raftery, 1995). It is worth mentioning that although random-effects procedures for group-level BMS are available (Stephan et al., 2009), the fixed-effects BMS implicit in Eq. (8) is more appropriate here. This is because we are characterising a basic relationship between anatomical and effective connectivity that is fixed over subjects.

When each model is equally likely a priori (i.e. flat priors on models) the posterior probability of each model is proportional to the model evidence: $p(m_i|y) \propto p(y|m_i)$. This means one can normalise the evidence for each model by dividing by the sum of evidences across models to give its posterior probability. The posterior probabilities for all models tested are shown in Fig. 5C.

Methods

Probabilistic tractography

Subjects

42 healthy volunteers (21 male) took part in the diffusion data acquisition, which was carried out at the Max Planck Institute for Human Cognitive and Brain Sciences at Leipzig, Germany. Subjects were on average 26.5 years old (range: 22–34; standard deviation: 2.8 years) and no subject had a history of neurological, psychiatric, or other major medical disorder. The study was approved by the local ethics committee of the University of Leipzig (Leipzig, Germany), and participants gave written informed consent. Data were handled anonymously.

Data acquisition and pre-processing

Diffusion weighted data and high-resolution three-dimensional T1 and T2 weighted images were acquired on a Siemens 3 T Trio scanner with an 8-channel array head coil and maximum gradient strength of 40 mT/m. The diffusion weighted data were acquired using twice-refocused spin-echo echo planar imaging (Reese et al., 2003) (TR = 12 s, TE = 100 ms, 72 axial slices, resolution $1.72 \times 1.72 \times$

1.7 mm). We used a GRAPPA technique (with a reduction factor of 2.0) for parallel imaging. Diffusion weighting was isotropically distributed along 60 directions (Jones et al., 1999) with a b-value of 1000 s/mm². The high angular resolution of the diffusion weighting directions improves the robustness of probability density estimation by increasing the signal-to-noise ratio and reducing directional bias. Additionally, seven data sets with no diffusion weighting (b0) were acquired initially and interleaved after each block of 10 diffusion weighted images as anatomical reference for motion correction. To further increase signal-to-noise, we acquired three consecutive scans, which were subsequently averaged together. The entire data acquisition protocol lasted approximately 45 min. Motion correction for the diffusion weighted images was applied to all images using 7-parameter global re-scale registration (Jenkinson et al., 2002) as implemented in the FSL software (FMRIB Software Library, University of Oxford, http://www.fmrib.ox.ac.uk/fsl). All baseline b0 images were aligned to a reference b0 image and the resulting linear transformation matrices were then applied to the diffusion weighted images following each baseline b0 image. The gradient direction for each volume was corrected using the rotation parameters. Then, the three scan repetitions were averaged to improve the signal-to-noise ratio.

Tractography

We applied the tractography approach as described by Kaden et al. (2007, 2008). This approach, which is based on the local fibre orientation density, computed by spherical deconvolution of the diffusion weighted signal, yields an estimate of the spatial probability distribution of connectivity from given seed regions. This approach is particularly useful for studies interested in anatomical connectivity between larger brain regions, which may not be adequately represented by single voxels (Koch et al., 2002) or single points (e.g. the centre of voxels; Behrens et al., 2003; Parker and Alexander, 2003). Instead, the concept of Kaden et al. extends the idea of connectivity to arbitrarily defined areas or volumes and defines anatomical connectivity as the proportion of fibre pathways originating in a specific source region that intersect a target region (Kaden et al., 2007). If the area or volume of the source region approaches a point, this measure reduces to the existence formulation proposed by Behrens et al. (Behrens et al., 2003), which only takes values on the discrete subset {0, 1}. It should be noted that none of the available tractography approaches makes it possible to determine the directionality of synaptic transmission along a given fibre tract; this is a general limitation of any connectivity metric based on diffusion weighted MRI.

Seed regions

Tractography was performed for each subject individually in his/her native (non-normalised) space. The resulting connectivity maps were then warped into a standard space (using the MNI 1 mm isotropic brain as a reference) for cross-subject averaging and comparison. To create seed masks for each subject, MNI coordinates were normalised to each subject's native space, using the inverse of the normalisation parameters. All resulting images were visually inspected to ensure that normalisation was successful and that each image was acceptable for analysis (e.g., in the correct orientation and not distorted).

To ensure that the computed tractograms were dominated by long-range connections, seed points were placed at the grey matter/white matter interface (white matter: fractional anisotropy; FA > 0.1). As the regional coordinates of fMRI time series used for the DCMs were located at the cortical surface, these coordinates were projected to the grey/white matter interface following the shortest (geodesic) path. Subsequently a seed region was defined by all points on the white matter surface within a radial distance of 3 voxels from the projected coordinates.

Connectivity measure

Considering all fibres originating in a given source region *S*, its structural connectivity with a given target region *T* can be defined in terms of the proportion of those fibres that intersect *T* while running within the brain white matter, yielding a number in the interval between zero (no fibres intercept *T*) and one (all fibres starting in *S* reach *T*) (Kaden et al., 2007). This quantity gives no information about the absolute number of connections between two regions, but reflects the degree of connectedness or relative connection density. It can be considered as a measure of the likelihood of a connection in the sense that it can be interpreted as the frequency at which one would reach *T* by randomly seeding a fibre starting within *S*. In our framework, the notions of anatomical connection strength and anatomical connection likelihood are therefore interchangeable.

One should note, however, that this connectivity metric may differ depending on whether *S* or *T* is chosen as the source region for fibre tracking (noting that this does not reflect the directionality of synaptic transmission along the pathways). For any given connection, one can remove this dependency on the seed region by computing the connectivity metric using each region as a seed region once and then averaging the result. Following this procedure, anatomical connectivity was estimated in each subject for all of the connections in the DCM (Stephan et al., 2007a) above and shown in Fig. 3. The subject-specific anatomical connection probabilities were then averaged across subjects; the resulting group values are reported alongside the corresponding connections in Fig. 3B. Note that the anatomical connection probabilities φ_{ij}^* resulting from our procedure are fairly small, as each region is most likely connected to many more brain areas than those included in the DCM. Fig. 3B also shows the normalised connection probabilities φ_{ij} for each connection (c.f. Eq. (3)).

Our numerical implementation for computing the above connectivity measures was identical to that described in Kaden et al. (2007). Streamline tractography was run from 100 randomly sampled starting points per voxel in the seed region; i.e. about 2400 times on average. A distribution of the connectivity value was then obtained by repeating this procedure 1000 times with different realizations of the local model, sampled using the Metropolis–Hastings algorithm. This means about 2.4 million fibres were computed for each connectivity value.

Results

Fig. 4 shows the relationship between the anatomical connection probability and prior variance of coupling parameters in DCM for all 64 models tested (note that the anatomically uninformed models m_{63} and m_{64} are shown in the same subplot). The log-evidences for all 64 models, in relation to the worst model (m_{41}) are shown in Figs. 5A and B. Fig. 5C shows the corresponding posterior probabilities for all models. It can be seen that the best model is model m_{45}. In this model, the DCM priors are anatomically informed, such that the prior variance of a coupling parameter in DCM increases monotonically with the anatomical likelihood of the connection. This result is pleasing because it confirms the intuitive notion that (i) probabilistic knowledge of anatomical connectivity should inform coupling strengths in dynamic models, and, more specifically, (ii) the more likely the anatomical connection, the easier it should be for the corresponding coupling to be expressed functionally and take large values.

Comparing Fig. 5B with Fig. 4, it is apparent that the next 6 best models, m_{44} and m_{46}–m_{50}, use anatomically informed priors with a similar form to the optimal model, i.e. priors that relax shrinkage when the strength of an anatomical connection increases. It is also worth noting that only the five best models, m_{44}–m_{48}, show a non-negligible posterior probability (i.e. larger than 10^{-4}; see Fig. 5C). These models are highlighted with a grey background in Fig. 4.

Two of the top 10 models, m_{10} and m_{30}, possessed counterintuitive priors (where anatomical likelihood decreased prior variance). When comparing model m_{45} to model m_{30}, the best counterintuitive model, the difference in log-evidence was 12.89 in favour of m_{45}. This means that the group fMRI data are $\exp(12.89) \approx 4 \times 10^5$ times more likely under model m_{45} than under model m_{30}.

Several of the models tested had anatomically uninformed priors, with the same prior variance for all coupling parameters, regardless of the anatomical likelihood of the respective connection. These models either (i) resulted from our systematic search over model space (for $\beta = 0$ or for large values of α that shifted the sigmoid such that all probabilities came to lie on the asymptote), (ii) were included as an additional control case (m_{63}, assuming that each connection was equally likely, i.e. $\varphi_{ij} = 25\%$), or (iii) used the default prior variances as originally defined for DCM (m_{64}). Notably, the corresponding uninformed prior covariances varied considerably across models, ranging from very high (e.g. 1 in m_1) to very low (e.g. 0.018 in m_{42}). As shown by Fig. 5, m_{45} performed considerably better than all models with uninformed priors, regardless of whether these used coupling parameters with high, intermediate or low prior variances. Specifically, when comparing model m_{45} to model m_{32} (the best of all models with anatomically uninformed priors), the relative log-evidence was 21.83 in favour of m_{45}. This difference means that the observed group fMRI data are $\exp(21.83) \approx 3 \times 10^9$ times more likely under model m_{45} than under model m_{32}.

Discussion

In this study, we used probabilistic tractography based on diffusion weighted imaging data to obtain a measure of the anatomical likelihood of connections among visual areas. We then instilled these measures into an established dynamic causal model of interacting visual areas by making the prior variance of the DCM coupling parameters a function of the anatomical likelihood. A series of competing dynamic causal models was constructed and compared, using Bayesian model selection, to identify the most likely mapping between anatomical estimates of connectivity and prior variances in the DCM. In particular, we compared DCMs with anatomically informed and anatomically uninformed priors. Our results showed that the best DCM used anatomically informed priors where the prior variances of coupling parameters increased as a monotonic nonlinear function of anatomical likelihood.

This study is novel in two ways. To our knowledge, this is the first formal demonstration that knowing anatomical connectivity improves inference about effective connectivity. Although we cannot generalise to other data or paradigms, these results provide a sufficiency proof that anatomical information is useful in the context of modelling functional integration. Secondly, this study shows how anatomical and functional data can be integrated and analysed within the same inference framework for dynamic systems. In this context, we have demonstrated a generic model selection procedure that allows one to quantify the evidence for anatomical constraints in other settings. In the following, we discuss these two issues in detail.

Decades of anatomical and physiological work have shown that anatomical connectivity provides critical constraints on effective connectivity. In particular, the construction of large scale databases of neuroanatomical data over the last decade have made it possible to establish global relations between anatomical connectivity and brain function (Bota et al., 2005; Burns and Young, 2000; Hilgetag et al., 1996, 2000; Honey et al., 2007; Kötter and Stephan, 2003; Kötter et al., 2001; Passingham et al., 2002; Scannell et al., 1999; Sporns et al., 2000; Stephan et al., 2001; Young, 1992). However, it is equally clear that effective connectivity is only constrained, and not fully determined, by anatomical connections. There are numerous reasons for this, which are largely related to short-term synaptic plasticity and neuromodulation. For example, synapses can alter their transmission properties depending on the recent history of presynaptic and postsynaptic events. These phenomena include synaptic facilitation

and depression and a variety of NMDA receptor-dependent mechanisms that change postsynaptic responsiveness at very short time-scales, ranging from milliseconds to minutes (Bagal et al., 2005; Citri and Malenka, 2008; Montgomery and Madison, 2004; Passafaro et al., 2001; Zucker and Regehr, 2002). Short-term synaptic plasticity also has structural correlates at the level of synaptic proteins (which may be altered by very fast processes like phosphorylation), the number of membrane-bound NMDA and AMPA receptors, and molecular changes in intracellular signal transduction cascades. However, these structural changes live at a much smaller spatial scale than axonal connections *per se*. In other words, anatomical connectivity is usually described at a cellular scale, whereas neuronal communication and synaptic efficacy lie at the subcellular and molecular scale. An example of changes in synaptic efficacy is the phenomenon that a synaptic connection can dramatically alter its strength due to nonlinear dendritic integration of multiple synaptic inputs. For example, in the presence of non-inactivating dendritic sodium conductances (Schwindt and Crill, 1995) or dendritic calcium conductances activated by back-propagating action potentials (Larkum et al., 2004) the postsynaptic response at a given dendritic synapse can depend strongly on the temporal and spatial distribution of other synaptic inputs. The fact that synaptic efficacy and effective connectivity are highly context-dependent has been described numerous times in the context of gain control (Salinas and Sejnowski, 2001), theories of neural context (McIntosh, 2000), functional integration (Friston, 2002b) and nonlinear gating of connections among neuronal populations (Stephan et al., 2008a).

For these reasons and others, it has been clear for a long time that even a perfect knowledge of the anatomical layout of brain connections would not enable us to predict the brain's functional integration or context-dependent dynamics (see also Ghosh et al., 2008). It seems equally clear, however, that although anatomical connectivity does not predict effective connectivity, it might provide an important constraint (Stephan et al., 2008b). Driven by this notion, a variety of studies have recently started to use empirical data sets of anatomical connectivity data as a basis for large scale simulations of brain function (Honey et al., 2007; Jirsa, 2004; Sporns et al., 2000; Tagamets and Horwitz, 1998) or have investigated what types of network structures give rise to certain types of neuronal dynamics (Strogatz, 2001). Also, since their first application (McIntosh and Gonzalez-Lima, 1991), models for inferring effective connectivity from neuroimaging data have exploited existing knowledge of anatomical connectivity to specify model structure. This assumed that model quality should be improved by incorporating anatomical knowledge. While this notion is plausible, it has not been tested formally. The present study has provided evidence in favour of this assumption.

A variety of attempts have been made over the last few years to use estimates of anatomical connectivity as derived from diffusion weighted imaging, to constrain or refine analyses of fMRI data. For example, some studies have shown that specific brain areas can be delineated on the basis of their anatomical connectivity pattern (Anwander et al., 2007; Tomassini et al., 2007), and that this delineation is well matched by fMRI activity in tasks that activate the area in question (Behrens et al., 2006; Friederici et al., 2006; Johansen-Berg et al., 2004, 2005; Klein et al., 2007). Some work has also addressed joint analyses of anatomical and functional connectivity. For example, Koch et al. (2002) correlated the results from functional connectivity analyses of fMRI data and tractography based on diffusion weighted imaging data. They found that the relation between anatomical and functional connectivity did not follow a simple rule but varied considerably across regions. More recently, it has been suggested that anatomical and functional connectivity, as inferred from fMRI data, could be analysed within a single mathematical framework (Jbabdi et al., 2007). Our study extends these efforts by linking anatomical connectivity estimates to inferences about effective connectivity. The proposed use of anatomically informed priors in dynamic causal models represents a simple approach that may prove useful in other contexts.

Clearly, there are many ways that anatomically informed priors in DCMs could be formulated mathematically. Nevertheless, even in this initial study, we investigated a relatively large range of different models, using a generic formulation that was able to model a variety of different relationships between the anatomical likelihood of a connection and the prior variance of the associated DCM coupling parameter. Notably, this set of 64 models included DCMs in which the prior variance increased with anatomical likelihood (as one would expect intuitively), DCMs in which the prior variance decreased with anatomical likelihood (i.e. counterintuitive models), and models in which the anatomical probability did not inform the priors (i.e., all connections were treated equally). All models were then compared using Bayesian model selection, a technique that takes into account not only the fit, but also the relative complexity of competing models. Differences in model complexity arise not only by inclusion of additional parameters, but also through differences in their prior variance. Simply speaking, increasing the prior variance of parameters endows a model with more effective degrees of freedom to fit the data and therefore renders it more complex (formally speaking, model complexity increases with the Kullback-Leibler divergence between the posterior density and the prior density; for details, see Stephan et al., 2009).

Figs. 4 and 5 summarise the results of our model comparison procedure. It is striking that the seven best models, m_{44}–m_{50}, use anatomically informed priors of an intuitive sort, i.e. priors that relax shrinkage when the strength of an anatomical connection increases. One may wonder why these particular anatomically informed priors fared better than formally similar priors. One explanation may be that optimal models tended to preclude very low prior variances, whereas in other models with priors of a similar form the less likely connections (i.e. left LG ↔ left FG and left FG ↔ right FG) had very low prior variances (compare Fig. 4). Low prior variances have opposing effects on model evidence (c.f. the appendix to Stephan et al., 2009): on one hand, as mentioned above, low prior variance increases model evidence by reducing the effective degrees of freedom and thus model complexity. On the other hand, because any difference between the posterior and prior means also contributes to model complexity, low prior variance means that any deviation of posterior estimates from zero will incur a higher cost (higher model complexity) than in models with high prior variance. These opposing effects on model complexity prohibit trivial strategies for choosing priors, such as generally minimising or maximising prior variance for all connections, and illustrate the necessity of establishing procedures for defining connection-specific prior variances, as suggested in this paper. In the present study, our BMS results imply that explaining the observed data required that connections with comparatively small anatomical likelihood nevertheless possessed non-negligible effective strengths.

After the seven best models, the next-best three models exhibited either anatomically informed but counterintuitive priors (in which anatomical likelihood decreased the prior variance) or anatomically uninformed priors, with the same prior variance for all coupling parameters, regardless of the anatomical likelihood of the respective connection (Fig. 5). Why these particular models were better than some models with anatomically informed and intuitive priors is not clear. However, on the whole, the relative evidence for these models was very low. They were considerably worse than the best model (with group Bayes factors of 10^5–10^9 in favour of the latter). Additionally, without exception, their posterior probabilities were very close to zero (i.e. below 10^{-4}; see Fig. 5C), while the posterior probability of the best model was 63%.

The present study has a number of limitations. First, the functional and structural data were not obtained from the same subjects. The diffusion weighted MRI data were acquired from a group of 42 healthy

volunteers, and the tractography values used in this study represent averages across this population. Anatomical connection probabilities were estimated for regions defined by fMRI activations in a second population of 12 subjects (Stephan et al., 2007a); these functional data were used to fit the DCMs. It is possible that our model selection approach might have given different hyperparameter estimates if we had used structural and functional data from the same subjects. We do not think a potential difference is likely to be substantial, because we averaged the tractography results across a relatively large group of 42 volunteers, thus diminishing any impact of inter-individual variation. We will revisit this issue in future studies that obtain structural and functional data from the same subjects. Concerning the present study, we would like to stress that its purpose was not to obtain exact hyperparameter estimates. We wanted to see whether tractography-informed priors improved the model evidence; and if so, what form the mapping from anatomical probability to prior constraint might take.

A second limitation, as in all modelling studies, is that only a limited number of models could be tested. We addressed this issue by defining a relatively large model space (defined by two parameters in the sigmoidal mapping) and systematically searching this model space. A third limitation, shared by all studies that rely on diffusion weighted imaging, is that our estimates of anatomical connectivity are non-directional. That is, having established (probabilistically) that there is a connection between region S and region T, we do not know whether the direction of synaptic transmission is from S to T, from T to S, or in both ways. In our study, this constraint is only relevant in that the prior variances for a pair of reciprocal connections are the same; in contrast, our DCMs allow for separate estimates of effective connectivity in the two directions.

Third, a general problem of many probabilistic tractography procedures, as with the one used in this study, is that the measure of connectivity between two regions is influenced by their spatial distance. This is due to the accumulation of errors over space when deriving connectivity measures from sequential tractography. In the network examined in this study, the distances traversed by intra-hemispheric connections are shorter than for inter-hemispheric connections. It is thus possible that the anatomical connectivity measures for the latter are underestimated relative to the former. However, for highly coherent fibre bundles, like the corpus callosum, this effect may not be critical: although no formal analysis exists, this was demonstrated anecdotally in previous work using the same tractography method (c.f. Figs. 7 and 8 in Kaden et al., 2007). This issue is a potential confound that could be addressed in future studies by including a distance hyperparameter in the prior covariance model and evaluating its contribution using Bayesian model comparison as above.

Finally, it should be noted that our proposed method is only useful when dealing with a network of more than two regions. This is because of the non-directional nature of diffusion data (see above) and because we are dealing with relative probabilities; where $\varphi_{ij} = 100\%$ for a two-area network (c.f. Eq. (3)). This is not a severe limitation because most applications of DCM concern networks with more than two areas.

Our approach could be extended to overcome some of the limitations mentioned above. In this paper, we optimised hyperparameters controlling the way tractography measures provide prior constraints on effective connectivity. This optimisation was with respect to a free-energy bound on the ensuing log model evidence, using a systematic search of hyperparameter space. In principle, this optimisation could be part of a hierarchically extended model that included the hyperparameters as unknown quantities. This would entail formulating priors on these hyperparameters (i.e., a prior on a prior or hyperprior) and extending the model inversion scheme to accommodate covariance functions that are nonlinear in the hyperparameters. Then, the explicit search over hyperparameters used in

this paper could be replaced by optimisation of a single model. An additional advantage of this procedure would be that we would have a posterior density on the hyperparameters themselves, allowing us to quantify how certain one was about the contribution of anatomical constraints. As noted by one of our reviewers, another advantage of such hierarchically extended models is that they would allow for different priors on the two directions of a reciprocal connection. The latter is not possible within the present framework because the non-directional nature of tractography data enforces the same anatomical connection probability for both directions, giving identical prior variances. However, if we used the anatomical connection probability to specify a hyperprior, the prior variances could change adaptively, in a direction-specific fashion. This would account, for example, for possible asymmetries in forward and backward connections.

In summary, using a combination of probabilistic tractography, dynamic causal modelling and Bayesian model selection, we have demonstrated that probabilistic estimates of anatomical connection strengths can be used to improve models of effective connectivity. We expect that this type of approach will prove useful in future studies of functional integration that have access to both fMRI and diffusion weighted imaging data.

Acknowledgments

References

Aertsen, A., Preißl, H., 1999. Dynamics of activity and connectivity in physiological neuronal networks. In: HG, S. (Ed.), Nonlinear Dynamics and Neuronal Networks. VCH Publishers, New York, pp. 281–302.

Alexander, D.C., 2005. Multiple-fiber reconstruction algorithms for diffusion MRI. Ann. N. Y. Acad. Sci. 1064, 113–133.

Anwander, A., Tittgemeyer, M., von Cramon, D.Y., Friederici, A.D., Knosche, T.R., 2007. Connectivity-based parcellation of Broca's area. Cereb. Cortex 17, 816–825.

Bagal, A.A., Kao, J.P., Tang, C.M., Thompson, S.M., 2005. Long-term potentiation of exogenous glutamate responses at single dendritic spines. Proc. Natl. Acad. Sci. U S A 102, 14434–14439.

Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088.

Behrens, T.E., Jenkinson, M., Robson, M.D., Smith, S.M., Johansen-Berg, H., 2006. A consistent relationship between local white matter architecture and functional specialisation in medial frontal cortex. Neuroimage 30, 220–227.

Bota, M., Dong, H.W., Swanson, L.W., 2005. Brain architecture management system. Neuroinformatics 3, 15–48.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.

Burns, G.A., Young, M.P., 2000. Analysis of the connectional organization of neural systems associated with the hippocampus in rats. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 355, 55–70.

Citri, A., Malenka, R.C., 2008. Synaptic plasticity: multiple forms, functions, and mechanisms. Proc. Neuropsychopharmacology 33, 18–41.

Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I., Anwander, A., 2006. The brain differentiates human and non-human grammars: functional localization and structural connectivity. Proc. Natl. Acad. Sci. U S A 103, 2458–2463.

Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: a synthesis. Hum. Brain Mapp. 2, 56–78.

Friston, K.J., 2002a. Bayesian estimation of dynamical systems: an application to fMRI. Neuroimage 16, 513–530.

Friston, K.J., 2002b. Beyond phrenology: what can neuroimaging tell us abut distributed circuitry? Ann. Rev. Neurosci. 25, 221–250.

Friston, K.J., Mechelli, A., Turner, R., Price, C.J., 2000. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. NeuroImage 12, 466–477.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.

Friston, K.J., Ungerleider, L.G., Jezzard, P., 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.

Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation. Neuroimage 34, 220–234.

Ghosh, A., Rho, Y., McIntosh, A.R., Kotter, R., Jirsa, V.K., 2008. Noise during rest enables the exploration of the brain's dynamic repertoire. PLoS. Comput. Biol. 4, e1000196.

Hilgetag, C.C., O'Neill, M.A., Young, M.P., 1996. Indeterminate organization of the visual system. Science 271, 776–777.

Hilgetag, C.C., Burns, G.A., O'Neill, M.A., Scannell, J.W., Young, M.P., 2000. Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 355, 91–110.

Honey, C.J., Kotter, R., Breakspear, M., Sporns, O., 2007. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. Proc. Natl. Acad. Sci. U. S. A. 104, 10240–10245.

Jbabdi, S., Woolrich, M.W., Andersson, J.L., Behrens, T.E., 2007. A Bayesian framework for global tractography. Neuroimage 37, 116–129.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17, 825–841.

Jirsa, V.K., 2004. Connectivity and dynamics of neural information processing. Neuroinformatics 2, 183–204.

Johansen-Berg, H., Behrens, T.E., Robson, M.D., Drobnjak, I., Rushworth, M.F., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M., 2004. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. Proc. Natl. Acad. Sci. U S A 101, 13335–13340.

Johansen-Berg, H., Behrens, T.E., Sillery, E., Ciccarelli, O., Thompson, A.J., Smith, S.M., Matthews, P.M., 2005. Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. Cereb. Cortex 15, 31–39.

Jones, D.K., Horsfield, M.A., Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn. Reson. Med. 42, 515–525.

Kaden, E., Anwander, A., Knösche, T., 2008. Variational inference of the fiber orientation density using diffusion MR imaging. Neuroimage 42, 1366–1380.

Kaden, E., Knösche, T.R., Anwander, A., 2007. Parametric spherical deconvolution: inferring anatomical connectivity using diffusion MR imaging. Neuroimage 37, 474–488.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.

Klein, J.C., Behrens, T.E., Robson, M.D., Mackay, C.E., Higham, D.J., Johansen-Berg, H., 2007. Connectivity-based parcellation of human cortex using diffusion MRI: Establishing reproducibility, validity and observer independence in BA 44/45 and SMA/pre-SMA. Neuroimage 34, 204–211.

Koch, M.A., Norris, D., Hund-Georgiadis, M., 2002. An investigation of functional and anatomical connectivity using magnetic resonance imaging. Neuroimage 16, 241–250.

Kötter, R., Stephan, K.E., 2003. Network participation indices: characterizing component roles for information processing in neural networks. Neural. Netw. 16, 1261–1275.

Kötter, R., Stephan, K.E., Palomero-Gallagher, N., Geyer, S., Schleicher, A., Zilles, K., 2001. Multimodal characterisation of cortical areas by multivariate analyses of receptor binding and connectivity data. Anat. Embryol. (Berl) 204, 333–350.

Larkum, M.E., Senn, W., Luscher, H.R., 2004. Top-down dendritic input increases the gain of layer 5 pyramidal neurons. Cereb. Cortex 14, 1059–1070.

Le Bihan, D., 2003. Looking into the functional architecture of the brain with diffusion MRI. Nat. Rev. Neurosci. 4, 469–480.

Mac Kay, D.J.C., 2003. Information Theory, Inference, and Learning Algorithms. Cambridge University Press, Cambridge.

Montgomery, J.M., Madison, D.V., 2004. Discrete synaptic states define a major mechanism of synapse plasticity. Trends Neurosci. 27, 744–750.

McIntosh, A.R., 2000. Towards a network theory of cognition. Neural Netw. 13, 861–870.

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

Mori, S., Zhang, J., 2006. Principles of diffusion tensor imaging and its applications to basic neuroscience research. Neuron 51, 527–539.

Neal, R.M., Hinton, G.E., 1998. A view of the EM algorithm that justifies incremental sparse and other variants. In: Jordan, M.I. (Ed.), Learning in Graphical Models. Kluwer Academic Publishers, Dordrecht.

Parker, G.J.M., Alexander, D.C., 2003. Probabilistic Monte Carlo based mapping of cerebral connections utilising whole-brain crossing fibre information. Lect. Notes Comput. Sci. 2737, 684–695.

Parker, G.J., Stephan, K.E., Barker, G.J., Rowe, J.B., MacManus, D.G., Wheeler-Kingshott, C.A., Ciccarelli, O., Passingham, R.E., Spinks, R.L., Lemon, R.N., Turner, R., 2002. Initial demonstration of in vivo tracing of axonal projections in the macaque brain and comparison with the human brain using diffusion tensor imaging and fast marching tractography. Neuroimage 15, 797–809.

Passafaro, M., Piech, V., Sheng, M., 2001. Subunit-specific temporal and spatial patterns of AMPA receptor exocytosis in hippocampal neurons. Nat. Neurosci. 4, 917–926.

Passingham, R.E., Stephan, K.E., Kotter, R., 2002. The anatomical basis of functional localization in the cortex. Nat. Rev. Neurosci. 3, 606–616.

Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Comparing dynamic causal models. Neuroimage 22, 1157–1172.

Pitt, M.A., Myung, I.J., 2002. When a good fit can be bad. Trends Cogn. Sci. 6, 421–425.

Reese, T.G., Heid, O., Weisskoff, R.M., Wedeen, V.J., 2003. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magn. Reson. Med. 49, 177–182.

Salinas, E., Sejnowski, T.J., 2001. Gain modulation in the central nervous system: where behavior, neurophysiology, and computation meet. Neuroscientist 7, 430–440.

Scannell, J.W., Burns, G.A., Hilgetag, C.C., O'Neil, M.A., Young, M.P., 1999. The connectional organization of the cortico-thalamic system of the cat. Cereb. Cortex 9, 277–299.

Schwindt, P.C., Crill, W.E., 1995. Amplification of synaptic current by persistent sodium conductance in apical dendrite of neocortical neurons. J. Neurophysiol. 74, 2220–2224.

Sporns, O., Tononi, G., Edelman, G.M., 2000. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. Cereb. Cortex 10, 127–141.

Sporns, O., Tononi, G., Edelman, G.M., 2002. Theoretical neuroanatomy and the connectivity of the cerebral cortex. Behav. Brain Res. 135, 69–74.

Stephan, K.E., Hilgetag, C.C., Burns, G.A., O'Neill, M.A., Young, M.P., Kotter, R., 2000. Computational analysis of functional connectivity between areas of primate cerebral cortex. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 355, 111–126.

Stephan, K.E., Kamper, L., Bozkurt, A., Burns, G.A., Young, M.P., Kötter, R., 2001. Advanced database methodology for the Collation of Connectivity data on the Macaque brain (CoCoMac). Philos. Trans. R. Soc. Lond. B. Biol. Sci. 356, 1159–1186.

Stephan, K.E., Marshall, J.C., Penny, W.D., Friston, K.J., Fink, G.R., 2007a. Interhemispheric integration of visual processing during task-driven lateralization. J. Neurosci. 27, 3512–3522.

Stephan, K.E., Weiskopf, N., Drysdale, P.M., Robinson, P.A., Friston, K.J., 2007b. Comparing hemodynamic models with DCM. Neuroimage 38, 387–401.

Stephan, K.E., Kasper, L., Harrison, L.M., Daunizeau, J., den Ouden, H.E., Breakspear, M., Friston, K.J., 2008a. Nonlinear dynamic causal models for fMRI. Neuroimage 42, 649–662.

Stephan, K.E., Riera, J.J., Deco, G., Horwitz, B., 2008b. The Brain Connectivity Workshops: moving the frontiers of computational systems neuroscience. Neuroimage 42, 1–9.

Stephan, K.E., Penny, W.D., Daunizeau, J., Moran, R.J., Friston, K.J., 2009. Bayesian model selection for group studies. Neuroimage 46, 1004–1017.

Strogatz, S.H., 2001. Exploring complex networks. Nature 410, 268–276.

Tagamets, M.A., Horwitz, B., 1998. Integrating electrophysiological and anatomical experimental data to create a large-scale model that simulates a delayed match-to-sample human brain imaging study. Cereb. Cortex 8, 310–320.

Tomassini, V., Jbabdi, S., Klein, J.C., Behrens, T.E., Pozzilli, C., Matthews, P.M., Rushworth, M.F., Johansen-Berg, H., 2007. Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. J. Neurosci. 27, 10259–10269.

Tuch, D.S., Wisco, J.J., Khachaturian, M.H., Ekstrom, L.B., Kotter, R., Vanduffel, W., 2005. Q-ball imaging of macaque white matter architecture. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 360, 869–879.

Young, M.P., 1992. Objective analysis of the topological organization of the primate cortical visual system. Nature 358, 152–155.

Zucker, R.S., Regehr, W.G., 2002. Short-term synaptic plasticity. Annu. Rev. Physiol. 64, 355–405.