Dynamic causal modelling of lateral interactions in the visual cortex

D.A. Pinotsis *, D.S. Schwarzkopf, V. Litvak, G. Rees, G. Barnes, K.J. Friston

The Wellcome Trust Centre for Neuroimaging, University College London, Queen Square, London WC1N 3BG, UK

A R T I C L E   I N F O

Article history:
Accepted 25 October 2012
Available online 2 November 2012

Keywords:
Neural field theory
Dynamic causal modelling
MEG
Connectivity
Visual cortex
Electrophysiology

A B S T R A C T

This paper presents a dynamic causal model based upon neural field models of the Amari type. We consider the application of these models to non-invasive data, with a special focus on the mapping from source activity on the cortical surface to a single channel. We introduce a neural field model based upon the canonical microcircuit (CMC), in which neuronal populations are assigned to different cortical layers. We show that DCM can disambiguate between alternative (neural mass and field) models of cortical activity. However, unlike neural mass models, DCM with neural fields can address questions about neuronal microcircuitry and lateral interactions. This is because they are equipped with interlaminar connections and horizontal intra-laminar connections that are patchy in nature. These horizontal or lateral connections can be regarded as connecting macrocolumns with similar feature selectivity. Crucially, the spatial parameters governing horizontal connectivity determine the separation (width) of cortical macrocolumns. Thus we can estimate the width of macro columns, using non-invasive electromagnetic signals. We illustrate this estimation using dynamic causal models of steady-state or ongoing spectral activity measured using magnetoencephalography (MEG) in human visual cortex. Specifically, we revisit the hypothesis that the size of a macrocolumn is a key determinant of neuronal dynamics, particularly the peak gamma frequency. We are able to show a correlation, over subjects, between columnar size and peak gamma frequency — that fits comfortably with established correlations between peak gamma frequency and the size of visual cortex defined retinotopically. We also considered cortical excitability and assessed its relative influence on observed gamma activity. This example highlights the potential utility of dynamic causal modelling and neural fields in providing quantitative characterisations of spatially extended dynamics on the cortical surface — that are parameterised in terms of horizontal connections, implicit in the cortical micro-architecture and its synaptic parameters.

© 2012 Elsevier Inc. All rights reserved.

Introduction

This work combines neural field models — that model the activity of layers of cells in cortical patches — with a Bayesian framework for optimising model parameters — known as Dynamic Causal Modelling (DCM). This combination allows one to address questions about lateral cortical interactions in terms of optimal models and model parameters. DCM has been applied extensively to fMRI and electrophysiological data (David et al., 2006; Friston et al., 2003; Penny et al., 2004) and has been used recently to model spatiotemporal dynamics on the cortical surface (Pinotsis et al., 2012). Combining generative models with Bayesian optimisation techniques enables one to characterise the functional architectures that generate empirical data. In this paper, we show that DCM can disambiguate between alternative spatiotemporal models of cortical activity — using Bayesian model comparison — and furnish quantitative explanations of observed responses, in terms of the biophysical properties of lateral cortical connections. This work focuses on two classes of biophysical models describing mesoscale brain activity: neural mass and field models. Neural field models describe how hidden neuronal states (such as the average depolarisation of a neural population or layer) evolve over both space and time. In contrast, neural mass models only characterise dynamics over time, under the assumption that all the neurons of a particular population are located at (approximately) the same point.

In previous work (Pinotsis et al., 2012), we considered the relationship between neural mass and field models and showed that field models can be reduced to neural masses by applying shrinkage priors to spike propagation times — such that lateral interactions were effectively instantaneous. We introduced a DCM that provides an explicit model of spatially extended cortical activity that allows one to make inferences about key parameters controlling the topographic distribution of cortical activity using LFP data, like the extent of lateral cortical connections and the conduction velocity of spike propagation. We were able to show that including spatial parameters enables one to explain effects that other models — such as neural mass models — attribute to variations in temporal parameters, like synaptic rate constants. Dynamic causal models based on neural fields enable one to characterise the propagation of activity on the cortex and provide a formal understanding of the mechanisms generating spatiotemporal responses. In our previous work above, we considered...
model optimisation under the assumption that one could measure local signals (local field potentials) that are sensitive to all spatial frequencies. In this paper, we consider neural field models of non-invasive (EEG or MEG) data. This development allows one to make inferences about the nature of lateral interactions in cortical sources, without spatially resolved measurements.

Spectral responses of neural fields

The characterisation of electrophysiological signals depends upon models of how they are generated in source space and how the resulting (hidden) neuronal states are detected by sensors. At the source level we consider a model based upon a canonical microcircuit that allows one to separate the sources of forward and backward connections in cortical hierarchies. In terms of the mapping from source to sensor space — we use a conventional lead field formulation that is expanded in terms of spatial basis functions. As in previous work, we focus on the modelling of power spectra. There is a long history of modelling steady-state (or ongoing) activity spectra, associated with neural fields, usually in models of the whole cortex, e.g., (Jirsa, 2009), Robinson (2006) has developed a neurophysiologically grounded model of corticothalamic activity, which reproduces many properties of empirical EEG signals; such as the spectral peaks seen in various sleep states and seizure activity. Technically, the spectra summarising the response of cortical sources can be defined in terms of transfer functions, mapping endogenous neuronal fluctuations to observed responses (Freeman, 1972; Nunez, 1995; Robinson et al., 2001; Robinson et al., 2003). In Pinotsis and Friston (2011), we derived the transfer function — and an expression for the spectral responses — for a source described by a classical neural field equation, while in Pinotsis et al. (2012) we extended our approach to a cortical source that comprises multiple layers. This allowed us to model the spectral activity of cortical fields as measured on the cortical surface. Here, we pursue a similar approach and generate the corresponding spectral responses, as measured by non-invasive sensors. These predictions are generated in an efficient manner that exploits the nature of the mapping from sources to sensors in EEG and MEG.

The resulting scheme can be regarded as inverting or fitting population models of the Amari type, using real data and Bayesian model inversion. Previous work in a similar vein includes the use of Kalman filters to develop estimation schemes for both neural mass (Riera et al., 2007; Valdes et al., 1999) and neural field models of a single population (Galka et al., 2008; Schiff and Sauer, 2008). In a related approach, Daunizeau et al. (2009)) replaced the standard dipole source — used in neural mass models — with the principal Fourier mode of a neural field, for the particular case of exponentially decaying synaptic density over the cortical surface. Finally (Markoukian et al. (2010)) used a combination of linear and nonlinear optimisation methods to invert a two-layered neural field model of voltage-sensitive dye data, describing inhibitory and excitatory populations (without conduction delays). The neural field model considered here has four layers and is based on canonical cortical microcircuitry that accounts for several aspects of local cortical computations in theoretical neurobiology. This model provides an extension of the well known Jansen and Rit neural mass model and incorporates conduction delays associated with the propagation of neuronal spikes.

Lateral interactions and neural fields

Modelling lateral interactions with neural field models has a long history. Pioneering work was introduced in papers by Amari, Wilson and Cowan, Grossberg and colleagues (Amari and Arbib, 1977; Amari and Takeuchi, 1978; Grossberg and Levine, 1975; Wilson and Cowan, 1973). These developments can be traced back to the work of physicists in the 19th century — such as Helmholtz and Mach — on visual perception. The first neural field models considered spontaneous pattern formation, by analysing the steady-state behaviour of underlying field equations: for example, Wilson and Cowan developed a treatment of Turing instabilities in the context of neural fields (Wilson and Cowan, 1972). In a similar vein, Grossberg initiated a line of work on shunting interactions — via nonlinearly coupled inputs — and considered limits as the solutions approached steady state (Grossberg, 1973). About the same time, Amari proved that systems of neural fields typically approach steady-state, in which some parts remain active, thus providing a metaphor for short-term memory.

From an anatomical viewpoint, the functional specialisation of visual (and auditory) cortex is reflected in its patchy or modular organisation — in which local cortical structures share common response properties. This organisation may be mediated by a patchy distribution of horizontal intrinsic connections that can extend up to 8 mm, linking neurons with similar receptive fields: see, e.g., Angelucci and Bressloff (2006), Burkhalter and Bernardo, (1989), and Wallace and Bajwa (1991). The existence of patchy connections in different cortical areas (and species) has been established with tracer studies in man, macaque and cat: Burkhalter and Bernardo (1989), Stettler et al. (2002) and Wallace and Bajwa (1991), respectively. It has been shown that such connections can have profound implications for neural field dynamics: see Baker and Cowan (2009). The precise form of such connections may be explained by self-organisation under functional and structural constraints; for example, minimising the length of myelinated axons to offset the cost of transmitting action potentials (Cherniak, 1994; Wen and Chklovskii, 2008). Generic constraints of this sort have been used to motivate general principles of connectivity; namely, that evolution attempts to optimise a trade-off between metabolic cost and topological complexity (Bassett et al., 2010). In short, visual and auditory cortices can be characterised by a patchy organisation that is conserved over the cortex and which allows for both convergence and divergence of cortical connections. Synaptic densities can then be approximated by isotropic distributions with an exponential decay over the cortical surface. In this work, we use a combination of patchy and isotropic distributions, using connectivity kernels with non-central peaks to model sparse intrinsic connections in cortical circuits that mediate both local and non-local interactions. In other words, we consider models in which neurons receive signals both from their immediate neighbours and remote populations that share the same functional selectivity (Pinotsis and Friston, 2011). We focus on the particular problem of identifying the parameters of lateral connections within a bounded cortical patch or manifold, as measured at a distance, with an array of non-invasive (MEG or EEG) sensors.

This paper comprises three sections. The first reviews a canonical neural mass model based on anatomical data and theoretical constraints from the theory of predictive coding. The second section describes a neural field model based upon this canonical micrometricity. Our focus here is on equipping the resulting neuronal model with an electromagnetic forward model to predict responses in non-invasive sensors. In the third section, we use Bayesian model comparison to adjudicate between various formulations of the ensuing neural field model (DCM) and establish its construct validity by optimising parameters pertaining to GABAergic concentrations, that have been shown to correlate with the peak gamma frequency of steady-state activity (Muthukumaraswamy et al., 2009).

A canonical model of cortical activity

This section introduces the neuronal field model used in subsequent sections on dynamic causal modelling. The particular neuronal model of source activity used here is based upon a refinement of conventional (convolution–based) neuronal models that explicitly model the neuronal sources of forward and backward connections in cortical hierarchies — these are the superficial and deep pyramidal cell
populations respectively. We develop the corresponding neural field model from established neural mass models.

The canonical microcircuit

Recent work suggests that superficial layers of visual cortex oscillate at gamma frequencies, while deep layers primarily oscillate at lower frequencies (Buffalo et al., 2011). Since forward connections originate predominately from superficial layers and backward connections primarily originate in deep layers, these spectral asymmetries suggest that forward connections use faster (gamma) temporal frequencies, while backward connections may employ lower (alpha or beta) frequencies—a suggestion that has experimental support (Roopun et al., 2010). These asymmetries mandate something quite remarkable: namely, a synthesis and segregation of forward and backward efferents from afferent input. This segregation can only arise from local non-linear neuronal computations that are structured and precisely interconnected. Non-linear neuronal transformations are necessary to account for the implicit coupling between different frequencies and arise from both synaptic mechanisms (e.g. non-linear dendritic integration) and population dynamics (e.g. sigmoid activation functions). The canonical microcircuit is a detailed proposal for such a laminar-specific intracortical architecture that describes how information flows through the cortical column. This model is based on findings in the primary visual cortex (Douglas and Martin, 1991) but recent work (Lefort et al., 2009; Weiler et al., 2008) indicates that similar microwirucircuits exist in other regions, such as somatosensory and motor cortex.

Douglas and Martin recorded intracellular potentials from cells in area 17 of the cat while they stimulated cortical afferents and noticed a strong compartmentalisation of the superficial and deep cell properties—reflected in slow superficial responses and fast input layer responses. The authors created a conductance-based model that reproduced the evolution of excitation and inhibition through the cortical circuit with great precision. This model contained three groups of cells: superficial and deep pyramidal cells, and a common pool of inhibitory cells. All three pools of neurons receive thalamic drive—although the thalamic drive to deep layer cells was weaker than the others. All neuronal populations expressed significant self-connectivity, and interconnected with the other populations. This model reproduced the features observed in their electrophysiological recordings—including the latency difference between superficial and deep layer neurons—and has served to establish several basic properties that are now believed to be replicated in other cortical areas: first, although superficial and deep compartments are strongly interconnected, their response properties are also strongly segregated. Second, cortex is not under tonic inhibition, rather, both excitation and inhibition are generated by afferent thalamic input and both shape ongoing cortical responses. Third, the canonical microcircuit can amplify thalamic inputs to generate self-sustaining activity, while also maintaining a delicate balance between excitation and inhibition—so as to prevent runaway excitation.

Previous computational modelling studies indicate that this circuitry allows the cortex to optimally organise and integrate bottom-up, lateral, and top-down information (Raizada and Grossberg, 2003). Douglas and Martin suggest that the rich anatomical connectivity of superficial layer (2/3) pyramidal cells allows them to collect information from all relevant top-down, lateral, and bottom-up inputs, and—through processing in the dendritic tree—select the most likely interpretation of its subcortical inputs. For a relevant discussion and more details on the canonical microcircuit and its potential role in predictive coding we refer the reader to Bastos et al. (under review).

Haeusler and Maass used Hodgkin and Huxley neurons to build a realistic microcircuit model and showed that a cortical column—whose connectivity conforms to the canonical microcircuit—can perform various computations efficiently, in relation to a column with random connectivity (Haeusler and Maass, 2007). By collapsing two pairs of cell types in the Haeusler and Maass model—while preserving the topology of the connectivity—one obtains the canonical microcircuit depicted in Fig. 1: this circuit comprises four populations: excitatory spiny stellate input cells (1), inhibitory interneurons (2), deep excitatory output pyramidal cells (3) and superficial excitatory pyramidal cells (4). In what follows, we describe a mathematical model of how the neuronal states of this these populations evolve over time. This model provides the basis of our dynamic causal model: We first define the vector-valued function

$$V = (v_1, v_2, v_3, v_4)^T$$

where \(v_a(t)\) denotes the expected depolarisation of the \(a\)-th population \((a = 1, ..., 4)\) at time \(t\). The dynamics of depolarisation are then described by the following second-order differential equation:

$$\dot{V} + 2BV = -B^2V + ABFV + GU$$

where \(U(t) \in \mathbb{R}^4\) is a vector of external inputs (exogenous neuronal fluctuations) and \(A\) and \(B\) are \(4 \times 4\) matrices of synaptic parameters controlling the maximum postsynaptic responses for excitatory \((m_e)\) and inhibitory \((m_i)\) populations and the rate-constants of postsynaptic filtering \((\kappa_1, ..., \kappa_4; \text{c.f., decay})\):

$$A = \text{diag}(m_e, m_e, m_e, m_e)$$

$$B = \text{diag}(\kappa_1, \kappa_2, \kappa_3, \kappa_4)$$

$$F : \mathbb{R}^4 \rightarrow \mathbb{R}^4$$ is a nonlinear mapping from postsynaptic depolarisation to presynaptic firing rates and \(G : \mathbb{R}^4 \rightarrow \mathbb{R}^4\) maps the inputs to the motion of hidden neuronal states; namely

$$G = (\kappa_1 m_e, 0, 0, 0)^T$$

In summary, Eq. (2) expresses the rate of change of expected depolarisation in each population as a sum of three terms; the first is a simple decay, the second is due to presynaptic input from other parts of the cortex and the final part is due to external inputs \(U(t)\). Writing out Eq. (2) in component form, we have

$$\dot{v}_1 + 2\kappa_1 \dot{v}_1 + \kappa_1^2 v_1 = \kappa_1 m_e (-d_{14} \cdot \sigma(v_{4}) + d_{11} \cdot \sigma(v_{1}) - d_{12} \cdot \sigma(v_{2}) + U)$$

$$\dot{v}_2 + 2\kappa_2 \dot{v}_2 + \kappa_2^2 v_2 = \kappa_2 m_e (d_{23} \cdot \sigma(v_{3}) + d_{22} \cdot \sigma(v_{2}) + d_{23} \cdot \sigma(v_{3}))$$

$$\dot{v}_3 + 2\kappa_3 \dot{v}_3 + \kappa_3^2 v_3 = \kappa_3 m_e (-d_{32} \cdot \sigma(v_{2}) + d_{33} \cdot \sigma(v_{3}))$$

$$\dot{v}_4 + 2\kappa_4 \dot{v}_4 + \kappa_4^2 v_4 = \kappa_4 m_e (d_{41} \cdot \sigma(v_{1}) + d_{44} \cdot \sigma(v_{4}))$$

where the sigmoid firing rate function is

$$\sigma(v_a) = \frac{1}{1 + \exp(\eta - v_a)}$$

Here, \(r\) and \(\eta\) are parameters that determine the shape of this sigmoid activation function. In particular, \(r\) is synaptic gain and \(\eta\) is the postsynaptic potential at which the half of the maximum firing rate is elicited. In Eq. (5), \(d_{ab} \cdot \sigma(v_b)\) is (endogenous) presynaptic input to the \(a\)-th population from the \(b\)-th and corresponds to the mapping \(F\)-\(V\). This is a sigmoid function \(\sigma(v_a)\) of postsynaptic depolarisation in the \(b\)-th population, multiplied by the intrinsic connection strength \(d_{ab}\) between the two populations (Jansen and Rit, 1995). See Fig. 1 for a schematic of this model.

A canonical microcircuit field model and its transfer function

In the case of neural field models, we consider spatially extended sources occupying bounded manifolds (patches) in different layers that lie beneath the cortical surface. In this setting, each population
where $\upsilon$ is the inverse speed with which spikes propagate along distance $k_{ab}$ (a measure of synaptic connectivity) and $c_{mn}$ are as a measure of the range and dispersion of lateral connections, summarised in terms of the parameters $h_a$ and $c_{mn}$ respectively, namely

$$K^{(i)} = \begin{bmatrix}
    k^{(i)}_{11} & k^{(i)}_{12} & 0 & k^{(i)}_{14} \\
    k^{(i)}_{21} & k^{(i)}_{22} & k^{(i)}_{23} & 0 \\
    0 & k^{(i)}_{32} & k^{(i)}_{33} & 0 \\
    k^{(i)}_{41} & 0 & 0 & k^{(i)}_{44}
\end{bmatrix}$$

$$K^{(e)} = \begin{bmatrix}
    k^{(e)}_{11} & 0 & 0 & 0 \\
    0 & k^{(e)}_{22} & 0 & 0 \\
    0 & 0 & k^{(e)}_{33} & 0 \\
    0 & 0 & 0 & k^{(e)}_{44}
\end{bmatrix}$$

$K^{(i)}$ is an exponentially decaying kernel commonly used in the literature to account for excitatory and inhibitory interactions (see e.g. Pinotsis et al., 2012), while the extrinsic part of the kernel $K^{(e)}$ was introduced in Grindrod and Pinotsis, (2011) and Pinotsis and Friston (2011) to model patchy lateral connections. This kernel is characterized by non-central peaks allowing for differences in (and estimation of) the range and dispersion of lateral connections, summarised in terms of the parameters $h_a$ and $c_{mn}$ respectively, namely

$$K^{(i)}_{ab} = \frac{1}{d_{ab}} e^{-c_{ab}|x|}$$

$$K^{(e)}_{ab} = \frac{1}{c_{ab}} \left( e^{-c_{ab}|x-h_a|} + e^{-c_{ab}|x+h_a|} \right)$$

Here, the parameters $d_{ab}$ and $c_{ab}$ encode the strength (analogous to the number of synaptic connections) and extent of intrinsic connections between cortical layers. The intrinsic connections can be regarded as interlaminar connections within a macrocolumn, while the extrinsic (between macrocolumn) connections correspond to horizontal connections and connect layers of the same type at a distance $h_a$. Later, we will use the extent parameter $c_{ab}$ as a measure.
of the size of cortical macrocolumns: see Fig. 2 for an illustration of the spatial parameters and Fig. 3 for a distinction between interlaminar and horizontal intra-laminar connections.

On comparing Eqs. (2) and (7) we notice that presynaptic input to the α-th population from the β-th is now expressed in terms of a spatiotemporal convolution as opposed to the simple nonlinear (sigmoid) mapping of neural mass models. For connectivity kernels modelling homogeneous (local) interactions, this convolution gives rise to wave equations describing propagation of the presynaptic input on the cortical manifold. In essence, Eq. (7) with connectivity defined by Eq. (8) defines a class of neural field models that can accommodate the intrinsic connectivity rules prescribed by the canonical microcircuit.

As in our earlier work (Pinotsis and Friston, 2011; Pinotsis et al., 2012), we assume that the neural field defined by Eq. (7) is perturbed around a spatially homogeneous steady-state $V_0$ (attained in the absence of external or exogenous perturbations)

$$V_0 = B^{-1}A \cdot F(V_0) \int K(x)dx$$

Using linear systems analysis, we now define the transfer function of the field model described above by the following relation

$$T(k, \omega) = \frac{P(k, \omega)}{U(k, \omega)}$$

where $U(k, \omega)$ is the two-dimensional Fourier transform of external input:

$$U(k, \omega) = FT(U(x, t)) = \int \int U(x, t)e^{-i k x - i \omega t} dx$$

and $P(k, \omega)$ is the Fourier transform of the perturbations around the steady-state solution. Given the transfer function, we can characterize the spectral response of the system to any external input, in terms of the underlying connectivity kernel, propagation velocities and post-synaptic response function. By analogy to a single population, substituting $V(x,t) = V_0 + P(x,t)$ into Eq. (7) and expanding $F \cdot V$ around $V_0$, we obtain a second-order expression for the perturbations $P(x,t)$

$$\dot{\tilde{P}} + 2B \tilde{P} = -B^2 \tilde{P} + AB \cdot D \cdot \gamma \cdot P + GU$$

$$\gamma = \partial F / \partial V$$

Here, $\gamma$ is the gain of the nonlinear mapping between depolarisation and firing rate:

$$\gamma_{ab} = \frac{\partial \psi(v_a, 0)}{\partial v_b} = \left\{ \begin{array}{ll}
\frac{re^n}{(1 + e^n)^2} & a = b \\
n & a \neq b
\end{array} \right.$$

Eqs. (10) and (12) provide the transfer function of our canonical microcircuit neural field model. Taking the Fourier transform of Eq. (12) and substituting into Eq. (10) gives:

$$T(k, \omega) = \left( -\omega^2 k^2 - 2i \omega B + B^2 - J(k, \omega) \right)^{-1} G$$

$$J(k, \omega) = ABD(k, \omega) \gamma$$

where $J(k,\omega)$ is a $4 \times 4$ matrix incorporating the synaptic parameters, connectivity parameters and gain matrix and $D(k,\omega)$ is the Fourier transform of the spatiotemporal connectivity. In summary, Eq. (14) provides a mathematical model or transfer function mapping from exogenous inputs or fluctuations acting upon each neuronal layer and the resulting spatiotemporal response in source space. This transfer function is specified completely by synaptic and connectivity parameters implicit in the neural field model. We next consider the mapping from source space to sensor space that completes the forward or generative model.
A neural field DCM

A generative model for power spectra

In what follows, we describe the generative or forward mapping from external inputs (exogenous fluctuations) to observed spectral responses. This allows one to compare the predictions of our model with real data and requires a mapping of neuronal states (the depolarisation fields above) to sensors. This mapping is called the lead field. The lead field allows one to infer hidden parameters characterising the deployment of sources on the cortical surface, even when there is no explicit spatial information in the data, see also Robinson et al. (2004). The lead field samples particular spatio-temporal frequencies, depending on the sensitivity profile of the sensors used. For example, if the lead field has a narrow spatial support (e.g., when using LFP electrodes), its spatial Fourier transform will be broad and it will be sensitive to a wide range of spatial frequencies. Conversely, when the lead field sees a large part of the cortical surface (e.g., non-invasive EEG sensors), the spatial Fourier transform will be narrow and only fluctuations in low spatial frequencies will contribute to the observed cross-spectra.

In the illustrative examples below we used an adaptive spatial filter or beamformer (Van Veen et al., 1997) to obtain estimates of ongoing neuronal activity in primary visual cortex (Schwarzko et al., 2012). This provides an estimate of electrical cortical activity based on a weighted combination of sensors — sometimes referred to as a virtual electrode. In other words, we will consider a single (virtual) sensor — so that the lead field maps from hidden neuronal states to a single data channel. We first describe the lead field and its parameterisation in terms of a Fourier basis set and coefficients that scale the contribution of neuronal states in each layer. We then describe how the predicted spectral response at the sensor depends upon the (Fourier transforms) of the lead field and the transfer function \( T(k, \omega) \) above that maps external inputs to neuronal states. Given the particular form of the lead field, we can simplify the expressions for the resulting transfer function from external inputs to spectral responses. This transfer function provides the basis of our probabilistic model. In detail: The lead field is parameterised by \( \varphi \) as a continuous gain function \( L(x, \varphi) \) over the cortical patch that is applied to a mixture of neuronal (depolarisation) states at each point on the patch. This mixture is determined by four coefficients \( \bar{Q} = [\bar{q}_1 \bar{q}_2 \bar{q}_3 \bar{q}_4] \), while the gain function is parameterised in terms of the coefficients \( L(k, \varphi) \) of a spatial Fourier basis set:

\[
L(x, \varphi) = \sum_k L(k, \varphi) e^{ikx} \tag{15}
\]

The predicted response at the virtual electrode — for a given set of all the neural and lead field parameters \( \theta \) — is obtained by integrating over the cortical patch

\[
y(t, \theta) = \int L(x, \varphi) (\bar{Q} \cdot V(x, t)) \, dx \tag{16}
\]

which leads to a spectral response of the form

\[
Y(\omega, \theta) = \sum_k L(k, \varphi) (\bar{Q} \cdot T(k, \omega) U(k, \omega)) \tag{17}
\]

The predicted spectral response measured by the sensor is therefore

\[
g(\omega, \theta) = Y(\omega, \theta) Y^*(\omega, \theta) \tag{18}
\]

where \( g_\theta(k, \omega) = |U(k, \omega)|^2 \) is the auto-power spectrum of external input. We will consider a gain function with a simple Gaussian form, which we parameterise in terms of its dispersion \( \varphi \) such that

\[
L(x, \varphi) = e^{-\varphi^2/2 \gamma^2}, \quad \text{noting that the amplitude is fixed to avoid redundancy with the parameters } \bar{Q} = [\bar{q}_1 \bar{q}_2 \bar{q}_3 \bar{q}_4]. \]

This leads to Fourier coefficients of the form \( L(k, \varphi) = e^{-2 \pi^2 \varphi^2 k^2} \) and Eq. (19) becomes

\[
g(\omega, \theta) = \sum_k \left| e^{-2 \pi^2 \varphi^2 k^2} \right|^2 \bar{Q}^T T(k, \omega) g_\theta(k, \omega) T(k, \omega)^* \bar{Q}^T \tag{19}
\]

The second equality follows by substituting the transfer function \( T(k, \omega) \) in Eq. (14) into Eq. (19), to express the prediction as a mixture of contributions from each population weighted by \( \bar{q}_i \):

\[
W_\theta(k, \omega) = \left| e^{-2 \pi^2 \varphi^2 k^2} \kappa_1 m_2 S_j(k, \omega) R^{-1}(k, \omega) \right|^2 g_\theta(k, \omega) \tag{20}
\]

The term \( S_j(k, \omega) R^{-1}(k, \omega) \) in (20) expresses the relative contribution of each population to the predictions at source level and depends upon the particular form of the connections among these populations. It can be seen from Eq. (8), that this ratio depends upon the (Fourier transforms of) intrinsic and extrinsic connectivity (see also, Pinotsis and Friston, 2011; Pinotsis et al., 2012):

\[
D\|_\theta(k, \omega) = \frac{\bar{q}_i (\bar{c}_a - \bar{\eta}_a)}{\bar{c}_a^2 - \bar{c}_a^2 \bar{\eta}_a^2 - 2 \bar{\eta}_a \bar{c}_a} + k^2 \tag{21}
\]

In particular, \( R(k, \omega) \) and \( S_\theta(k, \omega) \) are given by

\[
R(k, \omega) = -V_\text{in}(k, \omega) \left( -V_\text{in}(k, \omega) + Q_1(k, \omega) Q_1(k, \omega) \right) + Q_2(k, \omega) \left( -V_\text{in}(k, \omega) + Q_3(k, \omega) \left( -V_\text{in}(k, \omega) + Q_4(k, \omega) \right) \right) + Q_5(k, \omega) \left( -V_\text{in}(k, \omega) + Q_6(k, \omega) \left( -V_\text{in}(k, \omega) + Q_7(k, \omega) \right) \right) + Q_8(k, \omega) \left( -V_\text{in}(k, \omega) + Q_9(k, \omega) \right) \tag{22}
\]

where the functions \( Q_\theta(k, \omega) \) and \( V_\text{in}(k, \omega) \) depend on the Fourier transforms \( D\|_\theta(k, \omega) \) and \( D\|_\theta(k, \omega) \) as follows:

\[
Q_\theta(k, \omega) = -\kappa_2^2 + \gamma \left( D\|_\theta(k, \omega) + D\|_\theta(k, \omega) \right) \kappa_2 m_2 + 2 i \omega \kappa_2 + \omega^2 \tag{23}
\]

These expressions may look complicated but can be obtained in a fairly straightforward way from Eq. (19). In summary, the predicted spectral response at the sensor is given by:

\[
g(\omega, \theta) = \sum_k \left| e^{-2 \pi^2 \varphi^2 k^2} \kappa_1 m_2 \sum_q Q_q S_q(k, \omega) R^{-1}(k, \omega) g_\theta(k, \omega) \right|^2 \tag{24}
\]

Eq. (24) reflects the fact that the predicted spectral responses of the system are coupled to its spatial as well as its temporal properties; these properties are encoded in the transfer functions \( S_\theta(k, \omega) \) and \( R(k, \omega) \) through the underlying connectivity functions \( D\|_\theta(k, \omega) \). In turn, these are specified by the synaptic parameters associated with the canonical microcircuit \( \theta = (m_1, m_2, n_1, n_2, \gamma, \kappa_1, \kappa_2) \) and the spatial parameters \( \theta = [\bar{q}_1 \bar{q}_2 \bar{q}_3 \bar{q}_4] \) that encode intrinsic and extrinsic connections among different layers and neighbouring columns or points on the cortical manifold.
To complete our specification of a generative model, we assume that the observed cross-spectra \( g_{ij} \) are a mixture of predicted spectra, channel and Gaussian observation noise

\[
g_{ij}(\omega) = g(\omega, \theta) + g_n(\omega, \theta) + \epsilon
\]

\[
g(x_{ij}) = \alpha_n + \frac{\beta_n}{\alpha_n}
\]

\[
g_n(\omega, \theta) = \alpha_n + \frac{\beta_n}{\alpha_n}, \quad \text{(25)}
\]

The spectra of the neuronal fluctuations or input \( g_n(x_{ij}) \) are assumed to be spatially white; namely, they do not depend on spatial frequency. However both input and noise spectra are modelled as a mixture of white and coloured fluctuations over time. In completing the model, we have introduced extra free parameters \( \theta = [\alpha_n, \beta_n, \omega_n, \alpha_n, \beta_n] \) controlling the spectra of the inputs and channel noise. Eq. (25) provides the basis for our generative model and entails free parameters controlling the spectra of the inputs and channel noise as well as the amplitude of observation error. Gaussian assumptions about the observation error mean that we have a probabilistic mapping from all unknown (free) parameters to observed (spectral) data features. Inversion of this model means estimating, probabilistically, the free parameters given data.

Model inversion and fitting

Having prescribed the generative model of our DCM, we can now turn to its optimisation via Bayesian techniques. Almost universally, the fitting or inversion of Dynamic Causal Models optimises variational free-energy. Variational free-energy serves as a bound approximation to the log-evidence \( \text{ln} p(M) \) for a model \( M \). This optimisation is carried out with respect to a variational density \( q(\theta) \) over the unknown model parameters. By construction, the free-energy bound ensures that when the variational density maximises free-energy, it approximates the true posterior density over parameters, \( q(\theta) \approx p(\theta | M) \).

At the same time, the free-energy bound ensures that when the variational density maximises free-energy, it approximates the true posterior density over parameters, \( q(\theta) \approx p(\theta | M) \).

The underlyng generative model generally admits a unique solution during model inversion; this follows from the use of biophysically plausible priors over the biophysical parameters. Table 1 describes the priors over synaptic parameters (that are also used in the classical Jansen and Rit model) as well as parameters pertaining to the spatial structure of cortical sources. These priors are based on the modelling literature. Others come from the experimental literature; e.g. the prior for the conduction velocity is assumed to be 1.5 m/s (Kandel et al., 2000; Steriade et al., 1997). In general, priors are chosen to restrict parameter estimates in a physiologically meaningful range. However, it should be noted that the precise values of the priors are not important; the inversion scheme has the latitude to accommodate deviations from these values to optimise model evidence.

### Prior expectations of model parameters (The spatial parameters assume the cortical patch has a diameter of 25 mm).

| Parameter | Physiological interpretation | Prior mean |
|-----------|-----------------------------|------------|
| \( m_r, m_i \) | Maximum postsynaptic depolarisation | 8, 32 (mV) \* |
| \( k_{1r}, k_{2r}, k_{3r}, k_{4r} \) | Postsynaptic rate constants | 1/2, 1/2, 1/16, 1/28 (ms^{-1}) \* |
| \( a_{21}, a_{32}, a_{21} \) | Amplitude of intrinsic connectivity kernels | 3200 |
| \( d_{12}, a_{12}, a_{12}, a_{21} \) | Connectivity kernels | 800,800,1600,1600 |
| \( c_{ab} \) | Spatial decay of connectivity kernels | 8000,8000,4800 |
| \( h_0 \) | Separation between columns | 4.5 (mm)^{10} |
| \( r, \eta \) | Parameters of the postsynaptic firing rate function | 0.54, 0^{4} |
| \( v \) | Inverse conduction speed | 0.6 s/m^{b} |
| \( \sigma \) | Dispersion of the lead field | \( \sqrt{2/20} \) |
| \( q_1, q_2, q_3, q_4 \) | Relative layer contributions | 0, 1, 0, 80 |

\* Wendling et al., 2000.

### Results — an empirical illustration and validation

In previous work (Schwarzkopf et al., 2012), we recorded visually induced gamma oscillations in a group of healthy human subjects using magnetoencephalography (MEG) and used functional magnetic resonance imaging (fMRI) to measure the surface area of central primary visual cortex (V1) with standard retinotopic mapping techniques (Dale et al., 1999; Sereno et al., 1995). We showed that the retinotopically defined surface areas of central V1 and V2 are correlated with the peak frequency of visually induced oscillations in the MEG gamma band. This led to the proposal that individual differences in macroscopic gamma frequency may be attributed to inter-individual variability in the microscopic architecture of visual cortex (Schwarzkoepf et al., 2012). In this section, we use the neural field model of the previous section to address this proposal.

### Empirical data and validation

Here, we used MEG data and beamforming to summarise the spectral expression of endogenous activity in the visual cortex of 16 subjects, see Fig. 4 and Schwarzkoepf et al. (2012). Each participant was stimulated in both left and right visual fields — so we characterised separate sources in right and left hemispheres, resulting in 32 data sets. Subjects were seated in a MEG system and viewed the stimulus on a projection screen in front of them. During the entire recording run, they were required to maintain fixation on a small red dot in the centre of the screen. The

---

**Table 1**

| Parameter | Physiological interpretation | Prior mean |
|-----------|-----------------------------|------------|
| \( m_r, m_i \) | Maximum postsynaptic depolarisation | 8, 32 (mV) \* |
| \( k_{1r}, k_{2r}, k_{3r}, k_{4r} \) | Postsynaptic rate constants | 1/2, 1/2, 1/16, 1/28 (ms^{-1}) \* |
| \( a_{21}, a_{32}, a_{21} \) | Amplitude of intrinsic connectivity kernels | 3200 |
| \( d_{12}, a_{12}, a_{12}, a_{21} \) | Connectivity kernels | 800,800,1600,1600 |
| \( c_{ab} \) | Spatial decay of connectivity kernels | 8000,8000,4800 |
| \( h_0 \) | Separation between columns | 4.5 (mm)^{10} |
| \( r, \eta \) | Parameters of the postsynaptic firing rate function | 0.54, 0^{4} |
| \( v \) | Inverse conduction speed | 0.6 s/m^{b} |
| \( \sigma \) | Dispersion of the lead field | \( \sqrt{2/20} \) |

---

**Fig. 4.** Observed MEG spectra from a selection of subjects with different V1 surface area. Schwarzkopf et al., 2012 found a strong positive correlation between gamma peak frequency (marked with arrows) and V1 surface area. We used a biophysical model to investigate whether individual differences in macroscopic gamma frequency may reflect inter-individual variability in the architecture of visual cortex.
stimulus was a static, high-contrast, square-wave, vertical grating, which measured 4° × 4° of visual angle and had a spatial frequency of 3 cycles per degree. The grating was presented on a mean luminance uniform grey background in the lower visual field, such that the corner closest to the centre of gaze was horizontally and vertically displaced from the fixation spot by 0.5°. There were 180 trials; on half the grating was shown in the lower right quadrant; on the other half, it was shown in the lower left. We recorded MEG data using a whole-head CTF axial gradiometer system with 275 channels, sampled at 600 Hz. Three electrical coils were placed at fiducial locations and used to monitor subject head movement. Data were analysed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm). Recordings were divided into epochs from 1.5 s before stimulus onset until 1.5 s after stimulus onset (i.e., the earliest time for stimulus offset that preceded the participant’s behavioural response). Employing a beamforming approach, we estimated induced MEG responses to a visual stimulus in a virtual electrode placed in the medial occipital cortex, in a location consistent with primary visual cortex. We used an LCMV beamformer algorithm (Van Veen et al., 1997) implemented in SPM8 to quantify source power in the time window between 0.5 and 1.5 s after stimulus onset relative to baseline power over one second preceding stimulus onset. Source orientation at each voxel was determined using the method of Sekihara et al. (2004). We located peak gamma activity in the medial occipital cortex, and at this peak location we used the beamformer weights to extract the time series of the virtual electrode. Power spectra for frequencies between 30 and 80 Hz during the stimulation and the baseline were calculated using a multitaper spectral estimate (Thomson, 1982) using seven discrete prolate spheroidal sequences as data tapers. To summarise the peak gamma frequency — and the amplitude and bandwidth of the gamma response — we fitted a Gaussian function to the percentage power change in each frequency bin. A full description of the paradigm, recording and processing can be found in Schwarzkopf et al. (2012).

We modelled these spectral data — using DCM — to address the following questions: (i) Are neural masses or fields more appropriate for explaining these data? (ii) Can differences in the width of cortical columns explain the intersubject differences in gamma frequency? (iii) How does cortical excitability relate to the peak gamma frequency? And (iv) what are the important determinants of spectral gamma activity? The first question illustrates the use of Bayesian model comparison to compare DCMs based on neural fields and neural masses. We will see below that — for the MEG data and generative models considered here — the evidence for neural field models is relatively weak, which itself is interesting in relation to previous results. However, our primary aim was to show how DCM can answer anatomical questions about brain topography and architecture; by appealing to neural field models that embody plausible assumptions about source deployment and microcircuitry. We illustrate this point using subject-specific estimates of intrinsic connections from DCM and establish their predictive validity in terms of their ability to predict inter-subject variations in the size of V1 and gamma activity. We consider that the visual cortex is tiled with (overlapping) macrocolumns and assume that the major influences on each macrocolumn come from its nearest neighbours. We also assume rotational symmetry so that the spatial organisation of macro columns can be modelled by the field model described in the previous section (see Figs. 2 and 3). These connections are characterised by parameters describing local synaptic arbores (like the strength \( a_{\text{in}} \) and extent \( 1/c_{\text{in}} \) of intralaminar connections) and the range of horizontal connections (\( h_{\text{h}} \)). Our objective was to provide a mechanistic link between observed cross-spectral densities and local microstructure: we used the spatial decay rate of horizontal interlaminar connections \( c_{\text{ab}} \) as an estimate of column width.

**Neural masses or fields?**

Clearly, the choice of an appropriate model depends upon the question of interest; in particular, neural fields are appropriate for addressing questions about the deployment of sources on the cortical surface and induced spatial dynamics. However, neural field models might still be more appropriate from a Bayesian perspective, even if the spatial parameters of a neuronal model are not the focus of study: In the context of our Bayesian scheme, each model is scored using a free energy bound on model-evidence, where better models have a higher free energy (assuming free energy is a good approximation to model evidence). The model with the highest evidence implies the model has an optimal balance between accuracy and complexity — in the sense that the model provides an accurate explanation for the data in the simplest way. In our earlier work, we showed that Bayesian model selection can distinguish between neural mass and field variants of the same microcircuitry; where — for LFP data from the rat auditory cortex — the neural field variant had more evidence. This could be explained by the fact that neural field models provide an augmented repertoire of predictions for dynamics resulting from the propagation of activity on the cortical surface. In short, our earlier model comparisons showed that the neural field model yields a better fit to the LFP data — while accounting for the complexity cost due to its extra free parameters.

Here, we performed a similar analysis, by computing the relative log-evidence (using the free energy approximation) for the canonical microcircuit model, comparing field and mass variants at the group level. Synaptic (\( \kappa_1, \kappa_2, a_{\text{ab}} \)) spatial (\( t, c_{\text{ab}} \)) and sigmoid (\( r, \eta \)) parameters were optimised, while the remaining parameters in Table 1 including the horizontal distance \( h_{\text{h}} \) were fixed to physiologically plausible values. As with our previous model comparison, the neural mass model was formulated as a special (and limiting) case of the neural field model by shrinking the propagation delays (conduction times) times to zero. See Fig. 5 for an example of a model fit to a single subject’s response, using the neural field and mass variants. Contrary to our earlier result, we found inconsistent evidence in favour of the neural mass model (the log-evidence for the neural mass model was on average 2.37 greater than that of the field model); see Fig. 6 (top panel). In relation to our previous model comparisons, this result highlights the fact that the best model depends upon the data analysed. It also underlines the importance of combining a neuronal model with a spatial forward model: although both auditory and visual cortices are thought to conform to the local homogeneity constraints implicit in neural field models, the loss of spatial frequency resolution — with non-invasive data — might render neural field models unnecessary, in relation to neural mass models. In brief, model evidence depends crucially upon both the chosen models and the particular dataset used for their comparison. Our failure to establish a greater evidence for neural field models, in the present model comparison, is intuitively sensible because non-invasive MEG data have much lower spatial resolution than the LFP data we used in the previous model comparison. This observation speaks to the potential importance of using spatially resolved data to take full advantage of neural field models: data with high signal to noise ratio and wide brain coverage — such as those afforded by ECoG sensors or multi-array grids — can, in principle, disclose a full spectrum of spatiotemporal dynamics at different scales. This may be important for an informed (efficient) estimate of spatial parameters in neural field models. We will illustrate this point further in future work.

To assess conditional dependencies among parameter estimates, we computed the mean of posterior correlation matrices across all subjects, see Fig. 6 (right) and considered a correlation between a particular pair of parameters (one spatial and one neuronal — bottom left): the key thing to notice here are the relatively weak correlations between posterior estimates of columnar width and \( a_{\text{ab}} \) (the strength of connections from deep pyramidal cells to inhibitory interneurons). We will focus on this connection later when assessing the importance of spatial and synaptic parameters in predicting peak gamma frequency.
Finally, we characterised the variations in spectral profile produced by changes in these two parameters about their posterior estimates. Our aim here was not to provide a systematic characterisation of the neural field model; for this we refer the reader to earlier work (Nunez, 1995; Robinson et al., 2001, 2003, 2004; Valdes et al., 1999). We just wanted to provide a demonstration of the complicated but smooth contribution to spectral responses made by various parameters. It is this contribution that enables us to obtain informed estimates of underlying microcircuitry. Note that increasing both the spatial and neuronal parameters increases peak gamma frequency. In other words, increasing the spatial extent of local horizontal connections and increasing the excitatory drive to inhibitory interneurons increases peak gamma frequency. However, these parameters show (weak) negative conditional correlations. This means it is an open question as to which parameter is best able to account for changes in peak gamma frequency over subjects.

Can individual differences in peak gamma frequency be attributed to wider columns?

Our conventional analyses (Schwarzkopf et al., 2012), showed a strong positive correlation between gamma peak frequency and V1 surface area. This implies that individual differences in V1 area might reflect differences in cortical architecture. A larger cortex could either comprise a similar number of wider columns or simply contain a greater number of columns of constant width, which would be consistent with the explanation offered by the model of Prothero, (1997). Anatomical evidence suggests that the microarchitecture of V1 in humans resembles a scaled version of macaque V1, implying that column width increases with V1 size across species (Adams et al., 2007). However, across individuals of the same species both the width and number of columns can be highly variable (Horton and Hocking, 1996). Crucially, we can resolve this...
issue directly given the parameterisation of the intrinsic connectivity within the neural field model. In our model, the width of a macrocolumn, corresponds to the dispersion of horizontal synaptic connections $1/c_{ab}$.

We found a correlation between columnar width and peak gamma frequency that reached trend significance (see top panel of Fig. 8). Given our previous empirical finding (that peak gamma frequency and V1 surface area are positively correlated) we considered a one-tailed test and found Pearson $r = 0.271, p = 0.06$, (d.f. 30). This correlation implies that variability in the observed spectral responses in the gamma range can be attributed to individual differences in columnar width. This finding is more than simply noting that increasing the size of a column increases peak gamma frequency under the neural field model (see Fig. 7). This is because all the other free parameters of the model were optimised in a subject-specific fashion. In short, the hypothesis that intersubject variations in columnar width are associated with peak gamma frequencies was confirmed thereby providing a microscopic explanation for the correlation between peak gamma frequency and macroscopic (retinotopic) measurements of visual cortex anatomy — an explanation based purely on non-invasive MEG data.

We also found a significant positive correlation between columnar width and V1 size as measured with retinotopic mapping (see bottom panel of Fig. 8, where Pearson $r = 0.36, p = 0.02$, 30 d.f., — one-tailed test). This suggests that subjects with a larger V1 have wider as opposed to more macrocolumns. The results in Fig. 8 are consistent with our previous empirical finding of a significant correlation between observed gamma peak and size of the visual cortex (Schwarzkopf et al., 2012); however, the mediation of this correlation can now be associated with the organisation underlying cortical circuitry. This highlights the use of a generative (mechanistic model) to characterise cortical microstructure that would otherwise be hard or impossible to disclose.

**Is gamma activity related to GABAergic connections?**

We next considered the relation of individual differences in GABA concentration and peak gamma frequency (Gaetz et al., 2011; Muthukumaraswamy et al., 2009,2010) in terms of parameters describing synaptic transmission: Muthukumaraswamy and colleagues showed that individual differences in gamma oscillation frequency are positively correlated with resting GABA concentration in visual cortex — as measured with magnetic resonance spectroscopy. Furthermore, they showed that fMRI responses are inversely correlated with resting GABA and that gamma oscillation frequency is strongly inversely correlated with the magnitude of the BOLD response. These results were taken to suggest that the excitation/inhibition balance in visual cortex is reflected in peak gamma frequencies at rest. We address this hypothesis by looking for correlations involving the connections to inhibitory interneurons. We found that posterior estimates of the excitatory connections between deep pyramidal cells and inhibitory interneurons correlated negatively with gamma peak using both the neural field model (Pearson $r = -0.37, p = 0.02$, 30 d.f, two-tailed test) and the neural mass model (Pearson $r = -0.36, p = 0.04$, 30 d.f, two-tailed test), see Fig. 9. This confirms the hypothesis that intersubject variation in inhibitory drive in visual cortex is associated with characteristic changes in the peak frequency of gamma oscillations — and provides a mechanistic link from a synaptic level description to spectral behaviour that can be measured noninvasively. In summary, significant correlations were found between $\alpha$ and peak gamma frequency and columnar width and V1 size. Also, trend significance was observed for gamma peak to width correlations.

**What are the important determinants of gamma peak frequency?**

Previous studies suggested two possible causes for the inter-subject variability in gamma peak frequency. Muthukumaraswamy et al. (2009) suggested that peak gamma frequency is determined by the level of inhibition in V1 as described by resting GABA concentration measured with MR spectroscopy. In our previous study (Schwarzkopf et al., 2012), we found a correlation between V1 size and peak gamma frequency and suggested that the size of V1 and associated differences in microanatomy could be true determinants of peak gamma frequency. This suggests that both GABA concentration and V1 size can influence gamma frequency; however they these factors may or may not be causally linked.

Biophysical parameters estimated using DCM provide an opportunity to investigate alternative explanations of phenotypic differences, like gamma peak frequency: our contribution analysis above (Fig. 7) shows that the excitatory drive to inhibitory neurons ($a_{23}$) and macrocolumn width could mediate differences in peak gamma frequency. We therefore looked at the correlations over subjects between peak gamma frequency ($f$), V1 surface area and the posterior estimates of these parameters. These correlations are summarised in Table 2:

Interestingly, the partial correlation between $a_{23}$ and gamma peak remained significant when controlling for V1 size and width $1/c_{ab}$ ($r = -0.332, p = 0.037$). This suggests that the correlation between gamma peak and V1 inhibition cannot be accounted for completely by the spatial parameters (at the microscopic or macroscopic level). To elucidate the relationship between key model parameters and the phenotypes of V1 size and peak gamma frequency, we used Structural Equation Modelling (SEM) as implemented in SPSS Amos software (IBM). Each candidate model corresponded to a particular hypothesis about the causal relationships between peak gamma frequency and other variables. The Akaike Information Criterion (AIC) was used to compare models, while accounting for differences in model complexity.

Fig. 10 shows models tested and the associated AIC values. To simplify the model space, we assumed that V1 size is determined genetically (or epigenetically) and is not influenced by peak gamma...
frequency or microscopic parameters. Furthermore, we assumed that the peak gamma frequency is determined by the microscopic or macroscopic anatomy. Although one could argue that these assumptions are too simplistic — given the circular causality implied by experienced dependent plasticity — they are sufficient for our illustrative purposes. The key question here is whether peak gamma frequency is determined by (macroscopic on microscopic) spatial parameters, connectivity parameters or both.

It can be seen immediately from Fig. 10 that models without the \( a_{23} \rightarrow f \) link (the top row) all have lower evidence (higher AIC) than...
models in which peak gamma frequency depends upon the drive to inhibitory cells. This is consistent with the significant partial correlations we found between $a_{23}$ and $f$ above. The best model (Model 8) captured the correlation structure between the variables very well, as can be seen from Table 3 showing the predicted and observed correlations (in parentheses).

In this model, the correlation between V1 size and peak gamma frequency is mediated by a direct path (reflecting the influences of V1 size on peak gamma frequency that can be attributed to variables not considered in these models) and an indirect path through columnar width and excitatory connections. In other words, peak gamma frequency is mediated proximally by excitatory drive to inhibitory (GABAergic) interneurons and the strength of this drive is determined, in part, by the size of macrocolumns. In turn, the size of the macro columns is constrained by the macroscopic (retinotopic) size of V1. This structural equation modelling suggests a causal link between wider macrocolumns and an increase in inhibitory drive — a hypothesis which we will come back to in the discussion.

### Discussion and conclusions

By exploiting a combination of neural field modelling and Bayesian inference, we have shown that dynamic causal modelling can answer the following sorts of questions: which is the best biophysical model for explaining electrophysiological data? What are the important determinants of gamma peak frequency — in terms of synaptic parameters and horizontal interactions? Can MEG beamformed data help us access local cortical microstructure? And how can we distinguish between competing hypotheses about structure-function relationships? We have focused on two classes of biophysical models of brain...
activity; so-called neural field and mass models and have considered their application to modelling empirical MEG data. Bayesian model comparison — using a variational free energy approximation to model evidence — suggests neural field models provide a better explanation of empirical data if, and only if, there is sufficient spatial frequency information in the data. In other words, we found greater evidence for neural field models in previous analyses of LFP data but failed to find more evidence for neural field models, relative to neural mass models, in the current study of MEG (virtual electrode) data. The key distinction between these different modalities is that the LFP data is sensitive to a wide range of spatial frequencies and the temporal fluctuations that these frequencies contain. In contrast, the lead fields inherent in non-invasive electromagnetic recordings are necessarily broader and suppress temporal dynamics that are expressed in high spatial frequencies.

However, the use of neural field models may be necessary when testing hypotheses that are framed in terms of the spatial parameters of neural fields — like lateral or horizontal connections: we illustrated this using a neural field DCM to explain inter-subject variations in spectral activity in terms of synaptic and connectivity parameters. We showed that the posterior estimates of cortical anatomy — in particular columnar width — are associated with the spectral peak in the gamma range. Indeed, the estimates of column width made from a single time-series estimate in the foveal portion of V1 correlates directly with size of V1 as estimate from retinotopic mapping. This constitutes a compelling illustration of how non-invasive data can provide quantitative estimates of the spatial properties of neural sources and explain systematic variations in the dynamics those sources generate. We also found correlations, over subjects, between the peak gamma frequency and cortical inhibition as parameterised by the excitatory drive to inhibitory cell populations. This correlation was motivated by previous studies looking at the three way relationship between resting GABA concentrations in visual cortex, characteristic gamma activity and excitability — as measured with evoked responses using fMRI (Muthukumaraswamy et al., 2009).

Implications for visual perception

While visually induced gamma oscillations have received increasing attention and appear to play a role in visual perception (Herrmann et al., 2010), it remains unclear what determines the spectral properties of an individual’s gamma response, and how it relates to underlying visual cortex microcircuitry. Characterising individual differences in cortical micro-architecture may have important implications for research on visual processing. A few studies have claimed to directly visualise columnar architecture in human visual cortex (Goodyear and Menon, 2001; Goodyear et al., 2002; Vacoule et al., 2008). These studies rely on ultra-high field fMRI and are technically challenging and therefore not available for widespread use. Furthermore, while the spatial resolution of these experiments may be sufficient to identify cortical columns it may be unable to reliably resolve individual differences in the column width. This limits such investigations to invasive optical imaging of intrinsic signals or voltage-sensitive dyes in animal models; yet typically animal studies do not focus on individual differences across large groups. A reliable biomarker of columnar architecture — readily measurable with non-invasive techniques — may therefore be very useful for investigating individual differences in cortical micro-architecture in human subjects.

Previous research shows that macroscopic measures of V1 size predict visual perceptual ability such as Vernier discrimination (Duncan and Boynton, 2003). We recently showed that the size of V1 is negatively correlated with the strength of visual illusions, where the conscious perception of a visual stimulus is modified by the spatial context in which it is presented (Schwarzkopf et al., 2011a, 2011b). We interpreted these results as evidence that local contextual interactions in V1 are weaker in individuals with a large V1 area — because they have to be transmitted laterally over greater distances. Furthermore, the size of population receptive fields (pRF) — which reflect a combination of the size and positional scatter of the receptive fields of single neurons — as well as the extent of contextual interactions beyond the classical receptive field — are smaller in subjects with a large V1 (Harvey and Dumoulin, 2011). This implies a greater tendency to favour local processing in a large V1. Consistent with this, orientation discrimination is also superior in individuals with a large V1 (Schwarzkopf et al., 2011a). Our results imply that better orientation discrimination could be explained by both wider cortical columns and stronger lateral inhibition. This hypothesis is also supported by the empirical work of (Edden et al., 2009) who found that GABA concentration correlates with both gamma peak frequency and orientation discrimination ability; we will focus on this issue in future work.

In summary, individuals with larger retinotopic visual cortices might have weaker contextual interactions and more extended spatial processing. Both these findings are consistent with a greater dispersion of horizontal connections, inferred on the basis of our DCM results. Interestingly, gamma oscillation frequency is also correlated with orientation discrimination ability, which suggests that orientation discrimination, V1 size, pRF size, GABA concentration and gamma oscillations are all inter-related. We hoped to have shown how Dynamic Causal Modelling with neural fields may provide a quantitative and mechanistic approach to such correlations among cortical structure and function.

Acknowledgments

This work was funded by the Wellcome Trust. We thank W. Penny and A. Bastos for useful discussions.

References

Duncan, D.L., Sincich, L.C., Horton, J.C., 2007. Complete pattern of ocular dominance columns in human primary visual cortex. J. Neurosci. 27, 10391–10403.
Amari, S., Arbib, M.A., 1977. Competition and cooperation in neural nets. Syst. Neurosci. 119–165.
Amari, S., Takeuchi, A., 1978. Mathematical theory on formation of category detecting nerve cells. Biol. Cybern. 29, 127–136.
Angelucci, A., Bressloff, P.C., 2006. Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1. Visual Perception, Part 1, Fundamentals of Vision: Low and Mid-Level Processes in Perception, 154, pp. 93–120.
Baker, T.I., Cowan, J.D., 2009. Spontaneous pattern formation and pinning in the primary visual cortex. J. Physiol. Paris 103, 52–68.
Bassett, D.S., Greenfield, D.L., Meyer-Lindenberg, A., Weinberger, D.R., Moore, S.W., Bullmore, E.T., 2010. Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. PloS Comput. Biol. 6, e1000748.
Bastos, AM, Usrey, WM, Adams, RA, Mangun, GR, Fries, P and Friston, KJ, (under review) Canonical microcircuits for predictive coding.
Buffalo, EA, Fries, P, Landman, R, Buschman, TJ, Desimone, R., 2011. Laminar differences in gamma and alpha coherence in the ventral stream. Proc. Natl. Acad. Sci. 108, 11262.
Burkhalter, A., Bernado, K.L., 1989. Organization of corticocortical connections in human visual cortex. Proc. Natl. Acad. Sci. U. S. A. 86, 1071.
Cherniak, C., 1994. Component placement optimization in the brain. J. Neurosci. 14, 2418–2427.
Dale, AM, Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.
Daunizeau, J., Kiebel, S.J., Friston, K.J., 2009. Dynamic causal modelling of distributed electromagnetic responses. Neuroimage 47, 590–601.
David, D., Kiebel, S.J., Harrison, LM, Mattout, J., Kiehne, JM, Friston, K.J., 2006. Dynamic causal modelling of evoked responses in EEG and MEG. Neuroimage 30, 1255–1272.
Douglas, R.J., Martin, K.A., 1991. A functional microcircuit for cat visual cortex. J Physiol. 440, 735.
Duncan, R.O., Boynton, G.M., 2003. Cortical magnification within human primary visual cortex correlates with acuity thresholds. Neuron 38, 659–671.
Edden, R.A.E., Muthukumaraswamy, S.D., Freeman, T.C.A., Singh, K.D., 2009. Orientation discrimination performance is predicted by GABA concentration and Gamma oscillation frequency in human visual cortex. J. Neurosci. 29 (50), 15721–15726.
Freeman, W.J., 1972. Linear analysis of dynamics of neural masses. Annu. Rev. Biophys. Bioeng. 1, 225–256.
Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.
Friston, K., Mattout, J., Trujillo-Barreto, N., Ashburner, J., Penny, W., 2007. Variational free energy and the Laplace approximation. NeuroImage 34, 220–234.

Gaerttner, P., Roberts, T.P.L., Singh, K.D., Muthukumaraswamy, S.D., 2011. Functional and structural correlates of the aging brain: relating visual cortex (V1) gamma band responses to age-related structural change. Hum. Brain Mapp.

Galka, A., Ozaki, T., Muhle, H., Stephani, U., Sinatchkin, M., 2008. A data-driven model of the generation of human EEG based on a spatially distributed stochastic wave equation. Cogn. Neurodyn. 2, 101–113.

Goodyear, B.G., Menon, R.S., 2001. Visual evoked potential generation. Clin. Neurophysiol. 112, 2133–2147.

Goodyear, B.G., Menon, R.S., 2002. High resolution fMRI of ocular dominance columns with linear and non-linear models. J. Theor. Biol. 215, 377–389.

Haeusler, S., Maass, W., 2007. A statistical analysis of information-processing properties of lamina-specific cortical microcircuits. J. Cereb. Circ. 17, 149.

Harriman, K., 2011. The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture. J. Neurosci. 31, 13604–13612.

Herrmann, C.S., Fründ, I., Lenz, D., 2010. Human gamma-band activity: a review on cognitive and behavioral correlates and network models. Neurosci. Biobehav. Rev. 34, 981–992.

Horton, J.C., Hocking, D.R., 1996. Intrinsic variability of ocular dominance columns in normal macaque monkeys. J. Neurosci. 16, 7228–7239.

Kandel, E.R., Schwartz, J.H., Jessell, T.M., 2000. Principles of Neural Science, 4th ed. McGraw-Hill, New York.

Lefort, S., Tomm, C., Floyd Sarria, J.C., Petersen, C.C., 2009. The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. Neuron 61, 301–316.

Markoumikau, V., Iglo, C., Grinvald, A., Jancke, D., Spons, D., 2010. A dynamic neural field model of mesoscopic cortical activity captured with voltage-sensitive dye imaging. PLoS Comput. Biol. 6.

Muthukumaraswamy, S.D., Eddin, R.A., Jones, D.K., Sweetenham, J.B., Singh, K.D., 2009. Resting GABA concentration predicts peak gamma frequency and MRI amplitude in response to visual stimulation in humans. Proc. Natl. Acad. Sci. 106, 8356–8360.

Muthukumaraswamy, S.D., Singh, K.D., Sweetenham, J.B., Jones, D.K., 2010. Visual gamma oscillations and evoked responses: variability, repeatability and structural MRI correlates. NeuroImage 49, 3349–3357.

Nunez, P.L., 1995. Neocortical Dynamics and Human EEG Rhythms. Oxford University Press, USA.

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

Pinotsis, D.A., Friston, K.J., 2011. Neural fields, spectral responses and lateral connections. NeuroImage 55, 39–48.

Pinotsis, D.A., Moran, R.J., Friston, K.J., 2012. Dynamic causal modeling with neural fields. NeuroImage 59 (2), 1261–1274.

Prothero, J., 1997. Cortical scaling in mammals: a repeating units model. J. Hirnforsch. 38, 195.

Raizada, R.D., Grossberg, S., 2003. Towards a theory of the laminar architecture of cerebral cortex: computational clues from the visual system. Cereb. Cortex 13, 100.

Riera, J., Jimenez, J.C., Wae, X., Kawashima, R., Ozaki, T., 2007. Nonlinear local field potentials. Hum. Brain Mapp. 28, 335–354.

Robinson, P.A., 2006. Patchy propagators, brain dynamics, and the generation of spatially structured gamma oscillations. Phys. Rev. E 73, 041904.

Robinson, P.A., Loxley, P.N., O'Connor, S.C., Rennie, C.J., 2001. Modal analysis of cortical dynamics, electroencephalographic spectra, and evoked potentials. Phys. Rev. E 6304, 041900.

Robinson, P.A., Rennie, C.J., Rowe, D., O'Connor, S.C., Wright, J.J., Gordon, E., Whitehouse, R.W., 2003. Neurophysiological modeling of brain dynamics. Neurophysiolog. Biobehav. Rev. 28, 574–579.

Robinson, P.A., Rennie, C.J., Rowe, D.L., O'Connor, S.C., 2004. Estimation of multiscale neurophysiologic parameters by electroencephalographic means. Hum. Brain Mapp. 23, 53–72.

Roopun, A.K., LeBleu, F.E., Rammell, J., Cunningham, M.O., Traub, R.D., Whittington, M.A., 2010. Cholinergic neuromodulation controls directed temporal communication in neocortex in vitro. Front. Neural. Circuits 4.

Schiff, S., Sauer, T., 2008. Kalman filter control of a model of spatiotemporal cortical dynamics. BMC Neurosci. 9, 01.

Schwarzkoopf, D.S., Song, C., Rees, G., 2011a. Linking perceptual experience with the functional architecture of the visual cortex. J. Vis. 11, 844.

Schwarzkoopf, D.S., Song, C., Rees, G., 2011b. The surface area of human V1 predicts the subjective experience of object size. Nat. Neurosci. 14, 28–30.

Sekihara, K., Nagarajan, S.S., Poeppel, D., Marantz, A., 2004. Asymptotic SNR of scalar and vector minimum-variance beamformers for neuromagnetic source reconstruction. IEEE Trans. Biomed. Eng. 51, 1726–1734.

Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893.

Steriade, M., Jones, E.G., McCormick, D.A., 1997. Thalamus. Elsevier, Amsterdam.

Stettler, D.D., Das, A., Bennett, J., Gilbert, C.D., 2002. Lateral connectivity and contextual interactions in macaque primary visual cortex. Neuron 36, 739–750.

Thomson, D.J., 1982. Spectrum estimation and harmonic analysis. Proc. IEEE 70, 1055–1096.

Valdes, P.A., Jimenez, J.C., Riera, J., Biscay, R., Ozaki, T., 1999. Nonlinear EEG analysis based on a neural mass model. Biol. Cybern. 81, 415–424.

Van Veen, B.D., Van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880.

Wallace, M.N., Bajwa, S., 1991. Pathy intrinsic connections of the ferret primary auditory cortex. Neuron 7, 378.

Wen, Q., Chklovskii, D.B., 2008. A cost–benefit analysis of neuronal morphology. J. Neurophysiol. 99, 2320–2328.

Wendling, F., Bellanger, J.J., Bartolomei, F., Chauvel, P., 2000. Relevance of nonlinear lumped-parameter models in the analysis of depth-EEG epileptic signals. Biol. Cybern. 83 (4), 367–378.

Wilson, H.R., Cowan, J.D., 1972. Excitatory and inhibitory interactions in localized populations of model neurons. Biophys. J. 12, 1–24.

Wilson, H.R., Cowan, J.D., 1973. Mathematical theory of functional dynamics of cortical and thalamic nervous tissue. Kybernetik 13, 55–80.

Yacoub, E., Harel, N., Ugurbil, K., 2000. High-field fMRI unveils orientation columns in cingulate cortex: computational clues from the visual system. Cereb. Cortex 13, 100.

Yacoub, E., Harel, N., Ugurbil, K., 2000. High-field fMRI unveils orientation columns in humans. Proc. Natl. Acad. Sci. U. S. A. 105, 10607–10612.