A Stimulus-Locked Vector Autoregressive Model for Slow Event-Related fMRI Designs

Wesley K. Thompson and
Department of Psychiatry, University of California, San Diego, La Jolla, CA 92093, USA, email: wes.stat@gmail.com, telephone: 858.534.3684, fax: 858.534.7653

Greg Siegle
Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15213, USA

Summary

Neuroscientists have become increasingly interested in exploring dynamic relationships among brain regions. Such a relationship, when directed from one region toward another, is denoted by "effective connectivity." An fMRI experimental paradigm which is well-suited for examination of effective connectivity is the slow event-related design. This design presents stimuli at sufficient temporal spacing for determining within-trial trajectories of BOLD activation, allowing for the analysis of stimulus-locked temporal covariation of brain responses in multiple regions. This may be especially important for emotional stimuli processing, which can evolve over the course of several seconds, if not longer. However, while several methods have been devised for determining fMRI effective connectivity, few are adapted to event-related designs, which include non-stationary BOLD responses and multiple levels of nesting. We propose a model tailored for exploring effective connectivity of multiple brain regions in event-related fMRI designs - a semi-parametric adaptation of vector autoregressive (VAR) models, termed "stimulus-locked VAR" (SloVAR). Connectivity coefficients vary as a function of time relative to stimulus onset, are regularized via basis expansions, and vary randomly across subjects. SloVAR obtains flexible, data-driven estimates of effective connectivity and hence is useful for building connectivity models when prior information on dynamic regional relationships is sparse. Indices derived from the coefficient estimates can also be used to relate effective connectivity estimates to behavioral or clinical measures. We demonstrate the SloVAR model on a sample of clinically depressed and normal controls, showing that early but not late cortico-amygdala connectivity appears crucial to emotional control and early but not late cortico-cortico connectivity predicts depression severity in the depressed group, relationships that would have been missed in a more traditional VAR analysis.

Introduction

Interest among neuroscientists has increasingly centered on the determination of dynamic relationships among specialized brain regions in processessing information. The determination of dynamic relationships among brain regions has been termed "connectivity" analysis (Friston, 1994; Horwitz, 2003). In particular, an "effective connectivity" analysis attempts to determine directed influences from one brain region to another. However, researchers often have little prior knowledge about the nature of effective connectivity among regions, including

the strength or direction (excitatory or inhibitory) of such relationships. Moreover, even less may be known about how these effective connectivity relationships may vary across time after the application of an experimental stimulus. We propose an exploratory model for effective connectivity analysis which is specifically tailored to elucidate directed dynamic relationships among multiple brain regions in slow event-related fMRI experimental designs, while making minimal assumptions about the nature of these relationships. We were particularly motivated by observations of unpredicted lagged and non-linear relationships between regions (Wagner et al 2001; Siegle et al 2007) as well as frequently observed responses that do not match a canonical hemodynamic model (e.g., Siegle et al 2002; 2007) to develop a more exploratory approach. This exploratory approach for effective connectivity is in contradistinction to structural equation models (SEM; McIntosh and Gonzales-Lima, 1995) and dynamic causal models (DCM; Friston et al 2003), which generally require *a priori* the specification of effective connectivity relationships. Our exploratory model's estimates of effective connectivity can be used to assist in building confirmatory parametric models (including SEMs or DCMs). Additionally, because we explicitly allow for variation in effective connectivity across groups and subjects, indices derived from the connectivity coefficients can be used to predict individual variation in other variables, such as behavioral or clinical measures. We provide an example of this in the analysis of the application dataset below.

Vector autoregressive (VAR) models have proven popular for effective connectivity analysis over the last several years (Harrison et al 2003; Goebel et al 2003; Ho et al 2005; Bhattacharya et al 2006; Rykhlevskaia et al 2006). One reason for this popularity is that in VAR modeling, the direction and valence of effective connectivity relationships do not need to be pre-specified. Specifically, the standard VAR model relates the time-dependent structure of fMRI activation of multiple regions through a multivariate regression of prior levels of activation on current levels. Directed connectivity coefficients between regions are derived from the off-diagonal elements of the autoregressive matrices. The VAR approach has been utilized for the construction of graphical models for effective connectivity through the use of Granger causality analysis (Goebel et al 2003; Eichler, 2005; Roebroeck et al 2005; Valdes-Sosa et al 2005). Granger causality (Granger, 1969) is based on the common-sense idea that causes precede effects. With this in mind, activation in one region is said to "Granger cause" activation in another if prior levels of activation in the first region improve prediction of current values in the second; for stationary Gaussian VAR processes this is equivalent to testing whether the off-diagonal elements of the autoregressive matrices corresponding to these two regions are nonzero (Eichler, 2005). However, as noted by a referee, one must be careful not to over-interpret Granger causal relationships as truly causal; among other difficulties, VAR modeling is open to confounding from unmeasured relationships, as is any statistical modeling procedure (including SEM and DCM). Nonetheless, VAR models have proven a useful methodology for the effective connectivity analysis of neuroimaging data.

To date, however, VAR models have not been adapted to exploit specific characteristics of event-related fMRI experimental designs, which include nonstationary BOLD response curves and potentially dynamic effective connectivity relationships among brain regions. Adaptation of VAR models to slow event-related designs is particularly desirable because these designs are well-suited for exploring temporal dynamics of activity in regions involved in complex cognitive and emotional tasks, which may take several seconds to process. In the slow event-related experimental paradigm, brief stimuli are separated by an inter-trial interval usually on the order of 10 to 20 seconds. This duration is generally sufficient to allow a hemodynamic response associated with the stimulus to occur and fully decay (Glover, 1999), allowing for the examination of post-stimulus temporal dynamics among brain regions. Experimental stimuli are usually presented in a randomized order and mean stimulus-locked hemodynamic response curves can be computed by averaging over repetitions of trials of the same type. Locking activation curves to stimulus onset enables exploration of BOLD response trajectories

and dynamic interrelationships among regions which can be contrasted across stimulus types, individual subjects, and/or experimental subgroups.

An example of this approach is presented in Figure (1); data in this figure are group mean stimulus-locked response trajectories for three regions of interest (ROIs) for 24 mentally-healthy controls and 32 acutely depressed subjects. As described in more detail below, these subjects were presented with emotionally-negative words (200 ms) with an inter-stimulus interval of 10.8 seconds. The mean fMRI response depends both on group and, as is to be expected, time since stimulus. These data thus exhibit first-order nonstationarity.

While curve averaging within subjects or trial types provides a simple method for studying stimulus-locked first-order nonstationarities in BOLD responses, it is less readily apparent how to handle potential nonstationarities in effective connectivity relationships. Such nonstationarities arise if effective connectivity varies in a systematic fashion across the time course of a given trial type. For example, two regions may have a directed relationship near the beginning of a trial but not toward the end or an initially excitatory relationship may become less so or even inhibitory. Several studies have shown that connectivity is time and context dependent (Aertson and Preissl, 1991; Friston, 1994; Sato et al 2006; Bhattacharya et al 2007). Thus, current VAR methods which assume stationarity and otherwise do not explicitly incorporate the structure of event-related designs do a poor job at modeling important aspects of regional dynamics following application of a stimulus.

Empirical evidence also indicates that both activation and effective connectivity varies strongly from person to person (Mechelli et al 2002) and from group to group (e.g., Anand et al 2005). Indeed, testing for individual or group differences in connectivity among regions is often of considerable interest in and of itself (Mayberg et al 1999; Siegle et al 2007). Hence, another desirable property for a effective connectivity analysis is the ability to explicitly account for between-subject and between-group variability in effective connectivity relationships.

We therefore propose an adaptation of VAR models tailored for the exploration of effective connectivity in multi-subject slow event-related designs. This model, termed the "stimulus-locked VAR" (SloVAR) model, allows for fixed-effect differences in effective connectivity across stimulus types or experimental subgroups and random effect differences in connectivity across subjects within groups. Stochastic variation in BOLD response trajectories across scans and trials is modeled via multivariate Gaussian inputs, while effective connectivity is modeled via a generalized VAR framework. This generalized framework allows for a nonstationary VAR covariance structure on the regional BOLD responses by allowing the autoregression parameters to vary as a function of lag and time since last stimulus onset. This highly parametrized effective connectivity model is regularized via a low-rank basis expansion using a bivariate tensor product of cubic B-splines. We implement this using a mixed model formulation (Wand, 2003) with roughness penalty parameters treated as components of variance and estimated from the data. The mixed model is implemented within a Bayesian framework by specifying priors on the parameters, including roughness penalty parameters and using a Markov Chain Monte Carlo (MCMC) Gibbs sampling algorithm.

The output of SloVAR modeling is thus a set of surfaces representing the directed relationships of each structure in the model, at each lag, to each other structure. We envision two primary uses for this information. First, visual inspection of observed relationships facilitates exploration of non-trivial time-dependent relationships among structures. These relationships could otherwise be missed or not accounted for in analyses. Following an inspection, parameters capturing relevant aspects of variation can be extracted for testing in parametric models (e.g., the strength of association of early activity in one structure with later activity in

another structure). This useage follows common practice in using nonparametric models as aides in building parametric ones (Ruppert et al 2003). Formal inference from such empirically-determined models can be sought in the subsequent analysis of data obtained from additional subjects or by data splitting (Picard and Berk, 1990). Second, posterior estimates of connectivity parameters can be related to known non-fMRI cognitive or clinical information to better understand the relationships of fMRI-connectivity to these variables. If quantitative relationships are observed, features of the connectivity surfaces that generated the relationships can be explored post-hoc, again, necessitating extraction of relevant parameters to allow hypothesis testing.

Following a demonstration of the utility of the SloVAR model with simulated data, each of these applications is illustrated in the Results section below, using a sample psychiatric neuroimaging dataset in which we had previously examined functional connectivity using more conventional measures. As shown in the Results, use of SloVAR increases both intuitive understanding of the data as well as its clinical applicability. These findings suggest that early (within the first two seconds) but not late within-trial cortical to amygdala connectivity appears crucial to emotional control. This result matches theoretical explanations for mutually inhibitory influences of the prefrontal cortex on the amygdala and vice-versa but for the first time adds temporal specificity to the theoretical picture. Moreover, temporal specificity is important for relating connectivity to depression severity within the depressed group: early but not late cortico-cortico connectivity predicts depression severity on a clinical measure. This predictive relationship would have been missed in more traditional VAR analyses which assume stationarity and hence do not account for the stimulus-locked temporal variation in effective connectivity observed in these data across the course of a trial.

A MATLAB toolbox implementing the SloVAR model is available from the first author upon request.

Materials and Methods

The SloVAR Model

Suppose there are N subjects, with subject i completing M_i trials. Each trial consists of a short stimulus followed by an intertrial period consisting of T fMRI scans. Stimulus-locked fMRI activation profiles for each trial are obtained from P pre-specified brain regions. Note, we assume that one activation time course is obtained per region per trial by, e.g., computing a per-trial average time course of all voxels contained within the region. For trial j nested within subject i, denote the P-dimensional BOLD response at the tth time (scan) following stimulus presentation by $f_{ij}(t)$. The SloVAR model decomposes the P-dimensional BOLD response trajectories $f_{ij}(\cdot)$ into two parts, giving effective connectivity model

$$f_{ij}(t)=g_{ij}(t)+\omega_{ij}(t), \tag{1}$$

Where

$$g_{ij}(t)=\sum_{s=1}^{t-1}\Gamma_i(s,t)\,f_{ij}(t-s) \quad 1 \leq t \leq T. \tag{2}$$

The $g_{ij}(t)$ represent the portion of the BOLD response attributable to the internal dynamics of the system of P regions up to time t, with the $\omega_{ij}(t) \sim \text{MVN}_P(\mathbf{0}, \Sigma_\omega)$ allowing for stochastic

external inputs from other regions and otherwise noisy system dynamics. The $P \times P$ connectivity coeffecent matrices Γ_i vary across subjects to account for subject-specific differences in connectivity. The Γ_i also vary as a function of lag s and time t, so that

$$\Gamma_i(s,t) = \begin{bmatrix} \gamma_{i11}(s,t) & \dots & \gamma_{i1P}(s,t) \\ \dots & \dots & \dots \\ \gamma_{iP1}(s,t) & \dots & \gamma_{iPP}(s,t) \end{bmatrix}. \tag{3}$$

The connectivity coeffecents $\gamma_{ip_1p_2}(s, t)$, quantify the effect of region p_2 on region p_1 at time t and lag s, with time locked to stimulus (trial) onset. Since the connectivity coeffecients are allowed to vary as a function of time, the effect of region p_2 on region p_1 at lag s may differ across the course of a single trial. This may be the case if, for example, region p_2 sends inputs to region p_1 near the beginning of a trial but not near the end, or if the input is inhibitory at one timepoint and excitatory at another. Since $1 \leq s < t \leq T$, the connectivity coefficients $\gamma_{ip_1p_2}(\cdot, \cdot)$ in (3) are real-valued functions of two variables with triangluar support. The resulting NP^2 connectivity coefficient surfaces are estimated and regularized using linear combinations of bivariate basis functions with roughness penalties as described in the next subsection.

Regularization of Autoregression Surfaces

Equations (1)-(3) allow for the nonstationary evolution of underlying BOLD trajectory dynamics as a function of time since stimulus onset. However, without regularization this level of flexibility results in an over-parametrized model: with P regions and T times per trial this would result in $T(T-1)P^2/2$ parameters to estimate for each connectivity matrix Γ_i. One way of handling this is by treating the the connectivity surfaces $\gamma_{ip_1p_2}(\cdot, \cdot)$ as smooth functions of time and lag; this assumption was justified in the example dataset by plotting and inspecting raw lagged relationships. To implement this approach the connectivity coefficients $\gamma_{ip_1p_2}(s, t)$ are approximated using a tensor product of univariate bases, with $\psi_1(s) = (\psi_{11}(s), \dots, \psi_{1H_1}(s))'$ a basis for the first argument (lag) and $\psi_2(t) = (\psi_{21}(t), \dots, \psi_{2H_2}(t))'$ a basis for the second argument (scan, or time). To be specific, we choose $\psi_1(s)$ to consist of H_1 cubic B-splines with support on the interval $[1, T-1]$ and $\psi_2(t)$ to consist of H_2 cubic B-splines with support on the interval $[2, T]$. Considerations for choosing H_1 and H_2 are outlined in the Discussion section; briefly, the strategy is to include enough basis functions for the for flexible, data-determined fits while keeping the number small enough to be computationally feasible (Crainiceanu et al, 2005). Let $\psi(s, t)$ denote the H-dimensional tensor product of these two bases evaluated at lag s and time t, where $1 \leq s < t \leq T$ and $H = H_1H_2$. The connectivity coefficients are approximated by a linear combination of the components of $\psi(s, t)$

$$\gamma_{ip_1p_2}(s,t) = \alpha_{ip_1p_2}^T \psi(s,t). \tag{4}$$

This results in a model with $HP^2/2$ parameters to be estimated per connectivity matrix Γ_i. The reduced-rank smoothing spline strategy we implement is to include more than enough basis functions to obtain a good fit to the data and then penalize the roughness of the estimates. The roughness penalty is allowed to vary across time and lag arguments and implemented by the appropriate prior specification in a Bayesian mixed model.

Specifically, define the integrated squared second derivative penalty matrices as

$$J_{\psi 1} = \int_1^{T-1} \psi_1^{(2)}(s) \psi_1^{(2)}(s)' \, ds$$
$$J_{\psi 2} = \int_2^{T} \psi_2^{(2)}(t) \psi_2^{(2)}(t)' \, dt$$

Following Wood (2006), we can approximate the integrated squared second derivative penalty for the total surface curvature by defining $\tilde{S}_{\psi 1} = (J_{\psi 1} \otimes I_{H_2})$ and $\tilde{S}_{\psi 2} = (I_{H_1} \otimes J_{\psi 2})$, where $\otimes$ denotes the Kronecker product and I_{H_j} denotes the identity matrix of dimension H_j, $j = 1, 2$, and letting $\tilde{S} = \tilde{S}_{\psi 1} + \tilde{S}_{\psi 2}$ (an $H \times H$ matrix). The total surface curvature penalty matrix $\tilde{S}$ is not strictly positive definite because it has 4 eigenvalues equal to zero, corresponding to the linear subspace spanned by the cubic B-spline tensor product basis.

Thus, to utilize $\tilde{S}$ in the prior specification for the tensor product basis coefficients α_{ip1p2} we proceed by computing its spectral decomposition $\tilde{S} = VDV'$, where the columns of V are orthonormal eigenvectors and D is a diagonal matrix of eigenvalues whose first 4 diagonal elements are zero. Partition $V = [V_F \ V_R]$, where V_F is composed of the first 4 columns and V_R the last $H - 4$ columns of V. The tensor product basis is linearly transformed to $\psi^{\star}(s,t) = [\ V_F' \psi(s,t) \ \ V_R' \psi(s,t)] = [\psi_F(s,t) \ \ \psi_R(s,t)]$. This linear transformation of the tensor product basis partitions it to perfectly smooth (linear) basis functions $\psi_F(\cdot)$ and "wiggly" basis functions $\psi_R(\cdot)$. The coefficients for $\psi_F(\cdot)$ are unpenalized, since they are already smooth, whereas the coefficients for $\psi_R(\cdot)$ are penalized.

After this linear transformation of the basis functions, model (4) for the connectivity surfaces $\gamma_{ip1p2}(\cdot, \cdot)$ becomes

$$\gamma_{ip_1 p_2}(s,t) = \eta_{ip_1 p_2}^T \psi_F(s,t) + \delta_{q_i p_1 p_2}^T \psi_R(s,t). \tag{5}$$

With priors as specified in the next subsection, this becomes a Gaussian random effects model with subject-specific random effects η_{ip1p2} for the completely smooth basis functions and penalized fixed effects δ_{qip1p2} for the "wiggly" basis functions. These fixed effects vary across levels of a categorical variable $q_i \in \{1, \ldots, Q\}$, so that the mean shape of the connectivity coefficient surfaces are allowed to depend on contextual factors such as subgroup; a similar formulation could account for other categorical variables, including, e.g., within-subject differences in experimental stimulus type.

Prior Specification

The parameters of the connectivity model are denoted by $\Phi = \{\eta_{ip1p2}, \bar{\eta}_{qp1p2}, \delta_{qp1p2}, \Sigma_\omega, \Sigma_\eta, \sigma_{\delta_1}^2, \sigma_{\delta_2}^2\}$, where indices $1 \leq i \leq N$, $1 \leq p_1, p_2 \leq P$, and $1 \leq q \leq Q$ range over subjects, regions, and subgroups, respectively. Prior specification for these parameters follows common practice for Bayesian mixed models (Gelman et al 2003). In particular, priors are conjugate where possible. These priors have been shown to work well as long as hyperpriors for random effects variance components are chosen carefully (Kass and Natarajan, 2006; Zhou et al 2006).

Specifically, the $P \times P$ variance-covariance matrix Σ_ω for $\omega_{ij}(t)$ in Equation (1) is inverse Wishart (IW) with scale matrix S_ω and degrees of freedom d_ω. Hyperparameters S_ω and d_ω can be chosen to give a diffuse prior distribution by, e.g., setting $S_\omega = I_P$ and $d_\omega = P$. The hierarchical prior distributions of the 4-dimensional random effects η_{ip1p2} and the $(H - 4)$-dimensional fixed effects δ_{qp1p2} are given by

$$\eta_{ip_1\,p_2} \sim \mathrm{MVN}_4(\overline{\eta}_{q_i\,p_1\,p_2}, \Sigma_\eta), \quad \overline{\eta}_{qp_1\,p_2} \sim \mathrm{MVN}_4(0, cI),$$

$$\delta_{qp_1\,p_2} \sim \mathrm{MVN}_{H-4}\left(0, \left(\frac{1}{\sigma_{\delta_1}^2}S_{\varphi_1} + \frac{1}{\sigma_{\delta_2}^2}S_{\varphi_2}\right)^{-1}\right),$$

with c a fixed constant (say $c = 1000$) chosen to give a diffuse but proper prior. The 4×4 variance-covariance matrix Σ_η is given a conjugate IW(S_η, d_η) distribution, with S_η and d_η chosen to give a diffuse prior, if desired, using for example the results of Kass and Natarajan (2006). Finally, $\sigma_{\delta_1}^2$ and $\sigma_{\delta_2}^2$ are smoothing parameters for the $\varphi_R(\cdot, \cdot)$ basis functions. These two parameters are given inverse gamma (IG) distributions with hyperparameters τ_1 and τ_2 fixed; these are set to small values (say $\tau_1 = \tau_2 = .01$) to give diffuse yet proper priors.

The MCMC Gibbs sampling scheme and the derivation of the conditional posterior distributions using these prior distributions on the parameters are given in the Appendix.

Results

Simulation Study

Here we evaluate the performance and utility of the proposed model in estimating the connectivity coefficients through a simulation study. In this study we randomly generated 100 datasets according to model (1)-(3) with 25 subjects, 20 trials per subject (each consisting of 7 scans) and 2 ROIs. We generated data so that the region 1 demonstrates an effective connectivity relationship directed to the region 2, whereas there is no connectivity from region 2 toward region 1 (the connectivity surface is zero at all points). The self-connectivity surfaces and the connectivity surface for region 1 to region 2 are nonstationary: connectivity surfaces vary strongly as a function of time since stimulus onset. The true between-region connectivity coefficients, along with the simulation results, can be viewed in Table (1).

The SloVAR model was fitted for each of the 100 simulated datasets. Table (1) gives the mean off-diagonal connectivity surface estimates and actual coverage of the corresponding 95% posterior credible intervals. The proportion of times the posterior credible intervals include the true coefficient values is slightly lower than the nominal level, with median 94.8% for all points. Departures from the nominal level at individual points are due to biases induced from the smoothing algorithm. This may be due to choice of basis or to using a particular number and placement of knots. This subject is discussed in more detail in the conclusions. Nevertheless, these biases are far less severe than biases from fitting a stationary VAR model, as we show below.

To compare the SloVAR estimates with estimates obtained from usual practice, we also fit VAR models assuming stationary, time-invariant autoregression coefficients. The number of lags was set to 6. Nominal 95% posterior credible intervals for VAR coefficients were determined as above. The mean estimates of connectivity coefficients and the coverage of the corresponding confidence intervals is also given in Table (1). Unsurprisingly, the stationary VAR model provides highly biased parameter estimates and very low coverage.

To assess the impact of model misspecification (i.e., falsely assuming stationarity), we performed bivariate Granger causality tests at the 0.05 level (Granger, 1969) using the results of the standard VAR model fits to the simulated data. These tests correctly concluded for all 100 simulations that there was a significant Granger causal relationship from region 1 to region 2. However, these tests also incorrectly concluded that for all 100 simulations there was a significant Granger causal relationship from region 2 to region 1. These substantively false conclusions result from the highly-biased estimates seen in Table (1) due to model

misspecification. In general, the SloVAR model will provide a more acurate picture of effective connectivity than does the standard VAR model when the connectivity is strongly depedent on stimulus-locked time. We suggest possible ways for implementing more formal Granger causal tests based on the SloVAR model in the Discussion.

Application

Ruminative processing in depression—We present data from a psychiatric neuroscience experiment designed to test for functional differences in brain activation between unipolar depressed subjects and never-depressed controls (Siegle et al 2007). The study was constructed to understand brain correlates of sustained elaborative or "ruminative" processing of emotional information in depression, which is linked to depressive severity and persistence (Nolen-Hoeksema et al 1993). In particular, we examined relationships between a candidate mechanism for this sustained processing, activity in the left amygdala, a brain region associated with recognizing the emotionality of information and generating emotional reactions (LeDoux, 1996), and a two cortical regions: the left dorso-lateral prefrontal cortex (DLPFC), a region associated with executive control and initiating emotion-regulation, and the rostral portion Brodmann's Area 24 (BA24), a region associated with processing self-relevant information and emotion regulation, particularly inhibition of the amygdala. Abnormalities in activity of each of these areas have been observed in depression (for a review, see Mayberg, 2002).

Relationships of functioning among these areas is less well studied but has been hypothesized extensively. For example, if executive control is necessary for emotion regulation (Metcalfe and Mischel, 1999), and specifically, if participants use prefrontal cortex to initiate a process of emotion regulation that results in inhibition of limbic regions such as the amygdala (Mayberg et al 1999), sustained emotional reactivity might result indirectly from decreased function in brain regions subserving executive control such as the DLPFC. There are no direct connections from the DLPFC to the amygdala. Rather, inhibitory connections exist from ventromedial regions such as BA24 to the amygdala (Ghashghaei and Barbas, 2002). The DLPFC has been more formally associated with cognitive aspects of executive control but has been hypothesized to be involved in effortful initiation of regulatory processes that are eventually implemented by the more medial regions (e.g., Davidson et al 2000). Inhibitory connections from the amygdala back to multiple regions of the prefrontal cortex (Amaral et al 1992) could also allow excessive amygdala activity to contribute to decreased prefrontal activity (Moore and Grace, 2000), even without endogenously disrupted DLPFC activity. Ventromedial regions, in particular, have been hypothesized as key to integrating limbic and cortical functioning (Siegle et al 2006).

Experimental Design and Data Acquisition—To assess relationships between cortical and amygdala activity, 32 unmedicated depressed and 24 healthy participants completed tasks designed to provoke limbic reactivity to emotional stimuli in depression. In 60 slow event-related trials, participants viewed a fixation cue (1 s) followed by a positive, negative, or neutral word (200 ms), followed by a mask (row of Xs; 10.8 s). Participants were required to push a button for whether presented words were not, somewhat, or very relevant to them or their lives. For concision, the analyses presented in this paper include only the 20 trials using negative words. Thirty-four 3.2mm slices were acquired parallel to the AC-PC line using a reverse spiral pulse sequence (3T GE scanner, T2*-weighted images depicting BOLD contrast; TR=1500ms, TE=25ms, FOV=24cm, flip=60), yielding 12 whole-brain images per 18 second trial. More details on experimental design and subject characteristics can be found in Siegle et al (2007).

Data were pre-processed for analyses in several standard steps, as described in Siegle et al (2007). Following motion correction using the 6 parameter AIR algorithm (Woods et al 1993), linear trends within runs were removed to eliminate effects of scanner drift. Outliers

were winsorized. In contrast to Siegle et al (2007), data were not temporally smoothed before cross-registration. Rather, the detrended fMRI data were cross-registered to a reference brain using the 12 parameter AIR algorithm and spatially smoothed.

The left amygdala was identified anatomically in the functional data (using the mask shown in Siegle et al (2007), Figure 1) because it is small and boundaries with functionally distinct regions are hard to identify on functional times. In contrast, DLPFC and cingulate regions were identified empirically because they encompass large regions of potential functional heterogeneity and because relevant subregions are reliably differentiated on exploratory analyses of tasks involving cognitive control and emotional information processing. Specifically, the DLPFC region derived from a voxelwise analysis of group differences in the time-course of response to putting digits in order on a sorting-task was employed (from a group × scan random-effects analysis of variance (ANOVA)), as shown in Siegle et al (2007), Figures 3 and 5). A VMPFC region which differentiated depressed and healthy individuals in the time course of response to negative words, in the context of a group *x* valence *x* scan ANOVA (shown in Siegle et al 2007, Figure 2) was used for this analysis.

The averaged fMRI signal from each ROI was temporally smoothed similar to Pasquale et al (2008) to obtain regional BOLD response profiles. Within-trial regional activation profiles were normalized by subtracting the regional BOLD activation on the first time within each trial from all subsequent trial times and dividing by the within-subject median regional activation across the whole time series. Because the resulting normalized BOLD activation trial profiles necessarily begin at zero, the first time in each trial was subsequently discarded from the analysis, resulting in seven times per trial which did not start with zero. Figure (1) shows the regions of interest and the employed mean time-series for each region for each group.

Statistical Analysis—The connectivity coefficients were allowed to vary across groups (control or depressed) as in Equation (5) with $Q = 2$. The tensor basis functions for the connectivity surfaces consisted of cubic B-splines with one internal knot for the time and one internal knot for the lag dimension, giving 5 basis functions for each dimension. In the notation of Section (3.2), $H_1 = H_2 = 5$, and hence the tensor product of the univariate bases has a total of $H = 25$ basis functions. This was deemed more than sufficient to provide a good fit for a surface consisting of 21 points; sensitivity analyses (varying the number and placement of knots), not shown here, showed little difference in fits. The MCMC algorithm for the connectivity model (as described in the Appendix) was applied for 10,000 iterations with a burn in of 5,000 iterations. Convergence was monitored by initializing the chains at multiple random starting values and observing that the posterior distributions of parameters had converged to the same space (Gelman et al 2003).

Posterior estimates $\hat{\gamma}_{qp1P2}(s, t)$ for group-level connectivity surfaces were obtained by computing the mean over MCMC iterations

$$\widehat{\gamma}_{qp_1 p_2}(s,t) = \frac{1}{L}\sum_{l=1}^{L}\psi'_F(s,t)\overline{\eta}^{[l]}_{qp_1 p_2} + \psi'_R(s,t)\overline{\eta}^{[l]}_{qp_1 p_2},$$

(6)

where l indexes the MCMC iteration and the total number of posterior draws $L = 5,000$. Here $q = 1$ indicates control group coefficients and $q = 2$ indicates depressed group coefficients. The posterior connectivity surfaces determined by (6) were evaluated at $1 \leq s < t \leq 7$. Other indices derived from the connectivity coefficients were obtained similarly.

The matrices of posterior estimates of connectivity surfaces are plotted in Figures (2) and (3) for the control and depressed groups, respectively. (the surfaces in these figures are reparameterized as functions of time (t) and time minus lag ($t − s > 0$) for better visual presentation).

In general, the strongest connectivities are of a region with itself. These "self-connectivity" surfaces (depicted on the diagonal panels of Figures (2) and (3)) show a strong positive value at lag 1 across all time points, averaging 1.40 for all three regions in both groups followed by a dampened oscillating pattern of negative/positive values with means at lag 2=-.53, lag 3=-.42, lag 4=.26, lag 5=.46, and lag 6=-.51. This pattern is consonant with an initial amplification of a signal that subsequently levels off at later time points, which in a general sense describes the hemodynamic response to neural activity. The 95% pointwise posterior credible intervals indicate that these self-connectivity coefficients are significantly different from zero at the .01 level at all times and lags.

Of more interest are the between-region connectivity surfaces, depicted on the off-diagonal panels of Figures (2) and (3). These exhibit a greater degree of variation across regions and groups in terms of stength, shape, and direction (excitatory or inhibitory).

In the control group, there were no observed significant connectivity estimates from DLPFC to amygdala or from amygdala to either DLPFC or BA24. The DLPFC and BA24 have a mutually excitatory relationship in the control group. The mean lag 1 connectivity over scans 2-7 from BA24 into DLPFC is .23 with 95% posterior credible interval (PCI) given by [.12, .34], and the mean lag one connecitivity from DLPFC to BA24 is .09 (PCI= [.01, .17]). The BA24 to amygdala connectivity relationship exhibits a different pattern. Here, it is early activation, primarily time 1, which inhibits amygdala activation, with BA24 activity at time 1 damping amygdala activation at scans 3-7 (mean= −.23, PCI= [−.23, −.08]).

The depressed group qualitatively exhibits similar patterns of connectivity across the regions, with some exceptions. For example, the BA24 to amygdala relationship in depressed participants appears similar to that observed in controls, with BA24 activity at time 1 damping amygdala activation at time 3-7 (mean= −.26, PCI= [−.40, −.14]), and BA24 remains excitatory for DLPFC in the depressed group for all time points, lag 1 (mean= .16, PCI= [.05, .26]). Yet, unlike controls, the directed excitatory influence of DLPFC on BA24 is not significant in the depressed group.

In addition to exploring group differences in mean effective connectivity, it is of considerable interest to determine whether individual variation in effective connectivity estimates are predictive of clinical measures of depression. We therefore explored the relationship in the depressed group only between differences in subject-level connectivity and scores on the Beck Depression Inventory (BDI), a clinical measure of depressive severity. The BDI is a 21-question self-report inventory of depressive symptoms: higher scores indicate more severe depressive symptoms. Pre-treatment BDI scores were obtained on all 32 depressed subjects. We fit regression models with BDI scores as response variables and the posterior estimates of connectivity parameters η_{i, p_1, p_2} as predictors.

A stepwise model selection procedure was performed using AIC as the criterion for adding and dropping predictors. This procedure resulted in a model with only one predictor, $\eta_{i, \text{ba24, dlpfc}}^{[1]}$, or the first component of the parameter $\eta_{i, \text{ba24, dlpfc}}$, which has a correlation of .52 ($p = .004$ with BDI score. The parameter $\eta_{i, \text{ba24, dlpfc}}^{[1]}$ determines subject differences in effective connectivity from DLPFC to BA24 at early scans.

To illustrate the effect of this parameter on DLPFC to BA24 connectivity, we computed the mean connectivity surface for the 8 depressed subjects with values of $\eta_{i,\,ba24,\,dlpfc}^{[1]}$ in the top quartile and subtracted from it the mean connectivity surface for the 8 subjects in the bottom quartile. This difference in DLPFC to BA24 connectivity is depicted in Figure (4). The biggest impact is at early scans, especially scan 2 lag 1.

Figure (5) presents a scatterplot of BDI vs. $\gamma_{i\,ba24\,dlpfc}(1, 2)$. From this plot, it can be seen that subjects with high lag 1 DLFC to BA24 inhibition at time 2 have higher BDI scores, whereas subjects with high lag 1 DLPFC to BA24 excitation at time 2 have lower BDI scores. Thus higher levels of depression are associated with early DLPFC inhibition of BA24 and lower levels of depression are associated with early DLPFC excitation of BA24. The correlation of $\gamma_{i\,ba24\,dlpfc}(1, 2)$ with BDI score is .48 ($p = .008$).

Discussion

To date, the most common methods applied to effective connectivity analyses have been SEM (Bullmore et al 2000; McIntosh and Gonzalez-Lima, 1994) and DCM (Friston et al 2003; Penny et al 2004). SEM was originally developed for econometric analysis (see Bollen, 1989), whereas DCM was specifically tailored for studies of fMRI effective connectivity (for a comparison of SEM and DCM, see Penny et al (2004)). While both methods have been successfully applied to effective connectivity analysis in various contexts, they also have some disadvantgages in the context of multi-subject event-related experimental designs. Neither SEM nor DCM typically models the temporal correlation of observed fMRI data directly and neither DCM nor SEM explicitly models individual subject differences in effective connectivity directly. Most importantly, both SEM and DCM are essentially confirmatory in nature and are difficult to apply in situations where a well-developed theory of regional interaction is absent (Bullmore et al 2000). For exploring effective connectivity, the method of VAR modeling has become more prevalent as it is agnostic as to the directionality of dynamic relationships. We have proposed SloVAR as an adaptation of usual VAR models to exploit important characteristics of slow event-related fMRI designs in exploring for effective connectivity relationships. These adaptations allow for nonstationarites in dynamic relationships locked to time since stimulus onset and for flexible data-driven estimates of these stimulus-locked connectivity coefficents. Allowing for flexible data-driven fits and for individual differences in connectivity enables researchers to obtain subject-level indices which can isolate specific aspects of effective connectivity relationships among regions, e.g., early versus late connectivity. The incremental utility of the SloVAR technique above and beyond standard methods for examining functional relationships between brain regions was demonstrated by re-examining a published fMRI dataset in which we had previously examined functional connectivity. Previous analyses had suggested that there were group differences in connectivity but their character was not probed.

SloVAR modeling suggested that both DLPFC and BA24 have inhibitory influences on the amygdala, as has been hypothesized (e.g., Phillips et al 2003; 2008). Moreover, it is likely that the inhibitory function of the DLPFC on the amygdala is mediated by excitation of BA24 early on in a trial. Early activation in BA24 inhibits the amygdala for the entire remainder of the trial, suggesting that emotion regulation is affectuated within the first two seconds after stimulus presentation. In effect, if amygdala activity is to be inhibited, it must occur before it can rise to a high enough level that it can hijack emotion regulatory systems. This formulation is consistent as well with theoretical explanations for mutually inhibitory influences of the prefrontal cortex on the amygdala and vice-versa but for the first time, adds intuitively plausible temporal specificity to the theoretical picture.

Note, in the simulations we specifically considered the situation where effects present in the early part of one event-related waveform were systematically related effects in the later part of another waveform. These simulations demonstrated that traditional VAR methods would 1) give highly biased estimates in this situation, and 2) not detect the actual nature of the Granger causal relationship.

In addition, SloVAR modeling posited that inhibitory influences of DLPFC on BA24 are reduced in the depressed sample, suggesting that increased affective reactivity in this group could be due to decreased cortico-cortical inhibitory influences rather than simple hyper-limbic reactivity or abnormal cortico-limbic connectivity. This observation could have high clinical importance as common interventions for depression, such as selective serotonin reuptake inhibitors, appear to directly target limbic (amygdala) activity. Rather, interventions more directly associated with increasing cortico-cortical influences may help to intervene on the underlying mechanisms of sustained emotional-information processing in depression.

Finally, SloVAR modeling suggested that within the depressed group early inhibition of BA24 by DLPFC is predictive of worse depression severity (as measured by the BDI), whereas early excitation of BA24 by DLPFC was predictive of lower severity. To the extent that these data can be interpreted causally, they suggest that depressed individuals in whom the DLPFC does have an excitatory relationship on BA24 are characterized by more "normal" cortico-cortical regulatory information processing. This response style would likely allow BA24 to inhibit the amygdala, yielding decreased sustained reactivity. In contrast, the more severe depressed participants were characterized by an inhibitory relationship, in which top-down cortical influences were associated with inhibiting emotion-regulatory mechanisms. Such a situation is consistent with clinical presentations in which depressed individuals have been shown to value their rumination (e.g., Papageorgiou and Wells, 2001), and thus could work to allow sustained emotional information processing to occur, potentially to their ultimate detriment. Interventions targeted at these mechanisms (e.g., Wells, 2000) may be particularly warranted in such cases.

Note that these effective connectivity relationships would not have been detected in the usual VAR modeling framework. The usual VAR model assumes stationarity, so that effective connectivity relationships at various lags are not allowed to vary in a systematic fashion as a function of time. This does not appear to be a valid model, for example, for the BA24 to amygdala connectivity observed in this experiment. Moreover, early excitation of BA24 by DLPFC being predictive of lower depression severity would have been missed as well, again since typical VAR models do not allow for systematic variation in connectivity estimates over time. Of course, the clinical ties we have just described are speculative and must be more formally tested. Yet without SloVAR modeling, we would not have thought to consider these relationships and could have missed potentially crucial elements of the nature of emotional information processing in depression.

One important issue in implementing the SloVAR model is the choice of number and placement of the knots. Mirroring the approach in the statistics literature on reduced-rank spline smoothing (e.g., Crainiceanu et al 2005), H_1 and H_2 should be chosen to provide a large-enough pool of basis functions without becoming computationally infeasible. By "large-enough," it is meant as many as is necessary to provide a flexible, data-driven fit. For example, our choice of $H_1 = H_2 = 5$ in the data analysis section results in 25 basis functions, which we decided was more than sufficient to obtain a flexible fit for a surface consisting of 21 points; the roughness penalty method prevents overfitting the data. However, we also tried sensitivity analyses using different spacing of the knots, which in this case resulted in little difference in the estimated connectivity surfaces. Other, more formal methods could involve reversible-jump MCMC (as suggested by a referee) or estimation of the posterior probabilities of knot-inclusion indicators (Thompson

and Rosen, 2008). Such methods would add another layer of complexity to the algorithm and are left for future research.

Another useful feature for future research would be a formal test for bivariate Granger causality (whether a given connectivity surface consists of all zeros). This might be done by fitting multiple models and using a model selection statistic such as Bayes factors or as an integral part of the MCMC algorithm through the use of indicators of pairwise connectivity surface inclusion.

An important methodological question not addresssed in this paper is the impact of pre-processing steps on the estimated effective connectivity coefficients. Of particular importance is the choice of smoothing method used to obtain the underlying BOLD responses from noisy fMRI signals on connectivity estimates. This is not just an issue for SloVAR but for any connecitivity model which uses temporal dynamics to determine connectivity among regions or voxels. In this paper, we employed a data-driven approach similar to de Pasquale et al (2008) which accounts for autocorrelation in the fMRI noise when smoothing the data. We intend to examine this issue with Monte Carlo simulations of event-related data in a forthcoming paper.

Other further developments include devising procedures for utilizing SloVAR to build parametric models of effective connectivity. Toward this end, it will be important to develop metrics for model fit and selection among multiple competing effective connectivity models. One way to accomplish this is through Bayesian predictive model checking methods and Bayes factors (Gelman et al 2003); we intend to implement approaches such as these and to more generally investigate issues related to posterior inference in future work.

Acknowledgements

Appendix

The parameters of the connectivity model are denoted by $\Phi = \{\boldsymbol{\eta}_{ip1p2}, \bar{\boldsymbol{\eta}}_{qp1p2}, \boldsymbol{\delta}_{qp1p2}, \Sigma_\omega, \Sigma_\eta, \sigma^2_{\delta_1}, \sigma^2_{\delta_2}\}$, where indices $1 \le i \le N$, $1 \le p_1, p_2 \le P$, and $1 \le q \le Q$ range over subjects, regions, and subgroups, respectively. Priors are specified as in Section (3.3). Let

$\eta_i = (\eta^T_{i11}, \ldots, \eta^T_{i1P}, \ldots, \eta^T_{iPP})^T$ be the $4P^2$-dimensional vector of the concatenated random effects $\boldsymbol{\eta}_{ip1p2}$, and likewise let $\bar{\boldsymbol{\eta}}_q$ and $\boldsymbol{\delta}_q$ be the $4P^2$- and $(H-4)P^2$-dimensional vectors of concatenated fixed effect connectivity parameters. Furthermore, define the $P \times P^2$ matrix

$F^D_{ij}(t) = I_P \otimes f_{ij}(t)^T$ and the $P^2 \times 4P^2$ matrix $\Phi^D_F(s,t) = I_{p^2} \otimes \varphi_F(s,t)^T$, where $\otimes$ denotes the Kronecker product and I_P and I_{P^2} are the P- and P^2-dimensional identity matrices, respectively. Finally, let $\Phi^D_R(s,t) = I_{p^2} \otimes \varphi_R(s,t)^T$, a $P^2 \times (H-4)P^2$ matrix. From Equation (5) we can express the effective connectivity relationships in (1) and (2) as

$$
\begin{aligned}
f_{ij}(t) &= \sum_{s=1}^{t-1} \Gamma_i(s,t) f_{ij}(t-s) + \omega_{ij}(t) \\
&= \sum_{s=1}^{t-1} F^D_{ij}(t-s) \Phi^D_F(s,t) \eta_i + \sum_{s=1}^{t-1} F^D_{ij}(t-s) \Phi^D_R(s,t) \delta_{q_i} + \omega_{ij}(t) \\
&= X_{ij}(t)\eta_i + Z_{ij}(t)\delta_{q_i} + \omega_{ij}(t),
\end{aligned}
$$

where $X_{ij}(t)$ and $Z_{ij}(t)$ are $P \times 4P^2$ and $P \times (H-4)P^2$ dimensional matrices, respectively. The sampling scheme and conditional posteriors for the MCMC algorithm are as follows.

Generating $\boldsymbol{\eta}_i$: The posterior of $\boldsymbol{\eta}_i$ conditional on the data and the other parameters, denoted by $\boldsymbol{\eta}_i | \dots$, is multivariate normal $\mathrm{MVN}_{4P^2}(\boldsymbol{\mu}_{\eta_i|\cdot}, \Sigma_{\eta_i|\cdot})$, where

$$\Sigma_{\eta_i}|\cdot = \left((I_{p^2} \otimes \Sigma_\eta)^{-1} + \sum_{j=1}^{M_i} \sum_{t=2}^{T} X_{ij}(t)^T \Sigma_\omega^{-1} X_{ij}(t) \right)^{-1}$$

$$\mu_{\eta_i|\cdot} = \Sigma_{\eta_i|\cdot} \left((I_{p^2} \otimes \Sigma_\eta)^{-1} \bar{\eta}_{q_i} + \sum_{j=1}^{M_i} \sum_{t=2}^{T} X_{ij}(t)^T \Sigma_\omega^{-1} (f_{ij}(t) - Z_{ij}(t)\delta_{q_i}) \right).$$

Generating $\bar{\boldsymbol{\eta}}_q$: The conditional posterior of $\bar{\boldsymbol{\eta}}_q | \dots \sim \mathrm{MVN}_{4P^2}(\boldsymbol{\mu}_{\bar{\eta}_q|\cdot}, \Sigma_{\bar{\eta}_q|\cdot})$, where

$$\Sigma_{\bar{\eta}_q|\cdot} = \left(\tfrac{1}{c} I + N_q (I_{p^2} \otimes \Sigma_\eta)^{-1} \right)^{-1}$$

$$\mu_{\bar{\eta}_q|\cdot} = \Sigma_{\bar{\eta}_q|\cdot} \left(\sum_{i=1, q_i=q}^{N} (I_{p^2} \otimes \Sigma_\eta)^{-1} \eta_i \right),$$

and N_q is the number of subjects in group q.

Generating $\boldsymbol{\delta}_q$: The conditional posterior of $\boldsymbol{\delta}_q | \dots \sim \mathrm{MVN}_{(H-4)P^2}(\boldsymbol{\mu}_{\delta_q|\cdot}, \Sigma_{\delta_q|\cdot})$, where

$$\Sigma_{\delta_q|\cdot} = \left(\left(I_{p^2} \otimes (\tfrac{1}{\sigma_{\delta_1}^2} S_{\varphi_1} + \tfrac{1}{\sigma_{\delta_2}^2} S_{\varphi_2}) \right)^{-1} + \sum_{i=1, q_i=q}^{N} \sum_{j=1}^{M_i} \sum_{t=2}^{T} Z_{ij}(t)^T \Sigma_\omega^{-1} Z_{ij}(t) \right)^{-1}$$

$$\mu_{\delta_q|\cdot} = \Sigma_{\delta_q|\cdot} \left(\sum_{i=1, q_i=q}^{n} \sum_{j=1}^{M_i} \sum_{t=2}^{T} Z_{ij}(t)^T \Sigma_\omega^{-1} (f_{ij}(t) - X_{ij}(t)\eta_i) \right).$$

Generating Σ_ω: The conditional posterior of $\Sigma_\omega | \dots \sim \mathrm{IW}(S_{\omega|\cdot}, d_{\omega|\cdot})$, where

$$d_{\omega|\cdot} = d_\omega + MT$$

$$S_{\omega|\cdot} = \tfrac{1}{d_{\omega|\cdot}} \left(d_\omega S_\omega + \sum_{i=1}^{N} \sum_{j=1}^{M_i} \sum_{t=1}^{T} (f_{ij}(t) - X_{ij}(t)\eta_i - Z_{ij}(t)\delta_{q_i})(f_{ij}(t) - X_{ij}(t)\eta_i - Z_{ij}(t)\delta_{q_i})^T \right),$$

where M is the total number of trials across all subjects.

Generating Σ_η: The conditional posterior of $\Sigma_\eta | \dots \sim \mathrm{IW}(S_{\eta|\cdot}, d_{\eta|\cdot})$, where

$$d_{\eta|\cdot} = d_\eta + NP^2$$

$$S_{\eta|\cdot} = \tfrac{1}{d_{\eta|\cdot}} \left(d_\eta S_\eta + \sum_{i=1}^{N} \sum_{p_1, p_2=1}^{P} (\eta_{i p_1 p_2} - \bar{\eta}_{q_i})(\eta_{i p_1 p_2} - \bar{\eta}_{q_i})^T \right).$$

Generating $\{\sigma_{\delta_1}^2, \sigma_{\delta_2}^2\}$: The joint conditional posterior of $\{\sigma_{\delta_1}^2, \sigma_{\delta_2}^2\}$ is nonstandard. It is convenient to first transform the random variables so that at least one has a standard posterior.

As in Section (3.3), $\{\sigma_{\delta_1}^2, \sigma_{\delta_2}^2\}$ are *a priori* independent IG(π_1, π_2) random variables. Let $\tau_1 = \sigma_{\delta_1}^2$ and $\tau_2 = \sigma_{\delta_1}^2 / \sigma_{\delta_2}^2$. After this transformation, the conditional posterior for $\tau_1 | \tau_2, \ldots \sim$ IG ($\pi_{1,\tau_1|\cdot}, \pi_{2,\tau_1|\cdot}$), where

$$\pi_{1,\tau_1|\cdot} = 2\pi_1 + .5Q(H-4)P^2$$

$$\pi_{2,\tau_1|\cdot} = \pi_2(\tau_2+1) + .5\sum_{q=1}^{Q}\sum_{p_1,p_2=1}^{P}\delta_{qp_1 p_2}^T(S_{\varphi_1} + \tau_2 S_{\varphi_2})\delta_{qp_1 p_2}.$$

The random variable $\tau_2 | \tau_2, \ldots$ has a nonstandard conditional posterior proportional to

$$p(\tau_2 | \tau_2, \ldots) \propto \tau_2^{(\pi_1 - 1)} \exp\left\{-\tau_2\left(\frac{\pi_2}{\tau_1} + .5\sum_{q=1}^{Q}\sum_{p_1,p_2=1}^{P}\delta_{qp_1 p_2}^T S_{\varphi_2}\delta_{qp_1 p_2}\right)\right\}|S_{\varphi_1} + \tau_2 S_{\varphi_2}|^{-QP^2/2}.$$

With the exception of the last multiplicative term in this expression, this is a Gamma distribution. We sample from this conditional posterior by evaluating this expression on a fine grid of values of τ_2 on a nonnegative interval containing almost all of the probability and normalizing by numerically integrating (Gelman et al 2003). A random variable $U \sim \mathcal{U}(0, 1)$ is drawn and the inverse cdf transformation is applied to obtain τ_2.

References

Aertson, A.; Preissl, H. Non Linear Dynamics and Neuronal Networks, chapter Dynamics of activity and connectivity in physiological neuronal networks. VCH Publishers; New York: 1991. p. 281-302.

Amaral, D.; Price, J.; Pitkanen, A.; Carmichael, S. The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction, chapter Anatomical organization of the primate amygdaloid complex. Wilely-Liss; New York: 1992. p. 1-66.

Anand A, Li Y, Wang Y, Wu J, Gao S, Bukhari L, Matthews V, Kalnin A, Lowe M. Activity and connectivity of brain mood regulating circuit in depression: A functional magnetic resonance study. Biological Psychiatry 2005;57:1079–1088. [PubMed: 15866546]

Bhattacharya S, Ringo Ho MH, Purkayastha S. A bayesian approach to modeling dynamic effective connectivity with fmri data. NeuroImage 2006;30:794–812. [PubMed: 16364661]

Bullmore E, Horwitz B, Honey G, Brammer M, Williams S, Sharma T. How good is good enough in path analysis of fmri data? NeuroImage 2000;11:289–301. [PubMed: 10725185]

Crainiceanu C, Ruppert D, Wand M. Bayesian analysis for penalized spline regression using winbugs. Journal of Statistical Software 2005;14

Davidson R, Jackson D, Kalin N. Emotion, plasticity, context, and regulation: Perspectives from affective neuroscience. Psychol Bull 2000;126:890–909. [PubMed: 11107881]

de Pasquale F, Del Gratta C, Romani G. Empircal markov chain monte carlo bayesian analysis of fmri data. NeuroImage 2008;42:99–111. [PubMed: 18538586]

Eichler M. A graphical approach for evaluating effective connectivity in neural systems. Phil Trans R Soc B 2005;360:953–967. [PubMed: 16087440]

Friston K. Functional and effective connectivity in neuroimaging: a synthesis. Hum Brain Mapping 1994a;2:56–78.

Friston K. Functional and effective connectivity in neuroimaging: A synthesis. Human Brain Mapping 1994b;2:56–78.

Friston K, Harrison L, Penny W. Dynamic causal modeling. NeuroImage 2003;4:1273–1302. [PubMed: 12948688]

Gelman, A.; Carlin, J.; Stern, H.; Rubin, D. Bayesian Data Analysis, Second Edition. Chapman and Hall/ CRC; Boca Raton, FL: 2003.

Ghashghaei H, Barbas H. Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. Neuroscience 2002;115:1261–1279. [PubMed: 12453496]

Glover G. Deconvolution of the impulse response in event-related bold fmri. NeuroImage 1999;9:416–429. [PubMed: 10191170]

Goebel R, Roebroeck A, Kim D, Formisano E. Investigating directed cortical interactions in time resolved fmri data using vector autoregressive modeling and granger causality. Magn Res Imaging 2003;21:1251–1261.

Harrison L, Penny W, Firston K. Multivariate autoregressive modeling of fmri time series. NeuroImage 2003;4:1477–1491. [PubMed: 12948704]

Horwitz B. The elusive concept of brain connectivity. NeuroImage 2003;19:466–470. [PubMed: 12814595]

Kass R, Natarajan R. A default conjugate prior for variance components in generalized linear mixed models (comment on article by browne and draper). Bayesian Analysis 2006;1:535–542.

LeDoux, J. The Emotional Brain. Simon and Shuster; New York: 1996.

Mayberg H, Goldapple K, MacIntosh A. Is there a final common pathway mediating depression remission? the functional neuroimaging evidence. Biological Psychiatry. 2002

Mayberg H, Liotti M, Brannan S, McGinnis S, Mahurin R, Jerabek P, Silva J, Tekell J, Martin C, Lancaster J, Fox P. Reciprocal limbic cortical function and negative mood converging pet findings in depression and normal sadness. American Journal of Psychiatry 1999;156:675–682. [PubMed: 10327898]

McIntosh A, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. Hum Brain Mapping 1994;2:2–22.

Mechelli A, Penny W, Price C, Gitelman D, Friston K. Effective connectivity and inter-subject variability: using a multi-subject network to test differences and commonalities. NeuroImage 2002;17:1459–1469. [PubMed: 12414285]

Metcalfe J, Mischel W. A hot/cool-system analysis of delay of gratification: Dynamics of willpower. Psychol Rev 1999;106:3–19. [PubMed: 10197361]

Moore H, Grace A. Differential effect of tonic and phasic activation of the basolateral amygdala on prefrontal cortical input to nucleus accumbens neurons. Soc Neurosci Abst 2000;385:8.

Nolen-Hoeksema S, Morrow J, Frederickson B. Response styles and the duration of episodes of depressed mood. J Abnorm Psychol 1993;102:20–28. [PubMed: 8436695]

Papageorgiou C, Wells A. Positive beliefs about depressive rumination: Development and preliminary validation of a self-report scale. Behavior Therapy 2001;32:13–28.

Piccard R, Berk K. Data splitting. The Amercian Statistician 1990;44:140–147.

Ringo Ho MH, Ombao H, Shumway R. A state space approach to modelling brain dynamics. Statistica Sinica 2005;15:407–425.

Roebroeck A, Formisano E, Goebel R. Mapping directed influence over the brain using granger causality and fmri. NeuroImage 2005;25:230–242. [PubMed: 15734358]

Rykhlevskaia E, Gratton G. Lagged covariance structure models for studying functional connectivity in the brain. NeuroImage 2006;30:1203–1218. [PubMed: 16414282]

Sato J, Edson A, Takahashi D, de Maria Felix M, Brammerc M, Morettina P. A method to produce evolving functional connectivity maps during the course of an fmri experiment using wavelet-based time-varying granger causality. NeuroImage 2006;31:187–196. [PubMed: 16434214]

Siegle G, Carter C, Thase M. fmri predicts recovery in cognitive behavior therapy for unipolar depression. Am J Psychiatry 2006;163:735–738. [PubMed: 16585452]

Siegle G, Steinhauer S, Thase M, Carter C. Can't shake that feeling: fmri assessment of sustained amygdala activity in response to emotional information in depressed individuals. Biological Psychiatry 2002;51:693–707. [PubMed: 11983183]

Siegle G, Thompson W, Carter C, Steinhauer S, Thase M. Increased amygdala and decreased prefrontal bold responses in depression: Related and independent features. Biological Psychiatry 2007;61:198–209. [PubMed: 17027931]

Thompson W, Rosen O. A Bayesian model for sparse functional data. Biometrics 2008;64:54–63. [PubMed: 17573864]

Valdes-Sosa P, Sanchez-Bornot J, Lage-Castellanos A, Vega-Hernandez M, Bosh-Bayard J, Melie-Garcia L, Canales-Rodriguez E. Estimating brain functional connectivity with sparse multivariate autoregression. Phil Trans R Soc B 2005;360:969–981. [PubMed: 16087441]

Wagner A, Maril A, Bjork R, Schacter D. Prefrontal contributions to executive control: fmri evidence for functional distinctions within lateral prefrontal cortex. NeuroImage 2001;14:1337–1347. [PubMed: 11707089]

Wand M. Smoothing and mixed models. Computational Statistics 2003;18:223–249.

Wells, A. Emotional disorders and metacognition: Innovative cognitive therapy. Wiley; Chichester, UK: 2000.

Wood S. Low-rank scale-invariant tensor product smooths for generalized additive mixed models. Biometrics 2006;62:1025–1036. [PubMed: 17156276]

Woods R, Mazziotta J, Cherry S. Mri pet registration with automated algorithm. J Comput Assist Tomogr 1993;17:536–546. [PubMed: 8331222]

Zhao Y, Staudenmayer J, Coull B, W MP. General design bayesian generalized linear mixed models. Statistical Science 2006;21:35–51.
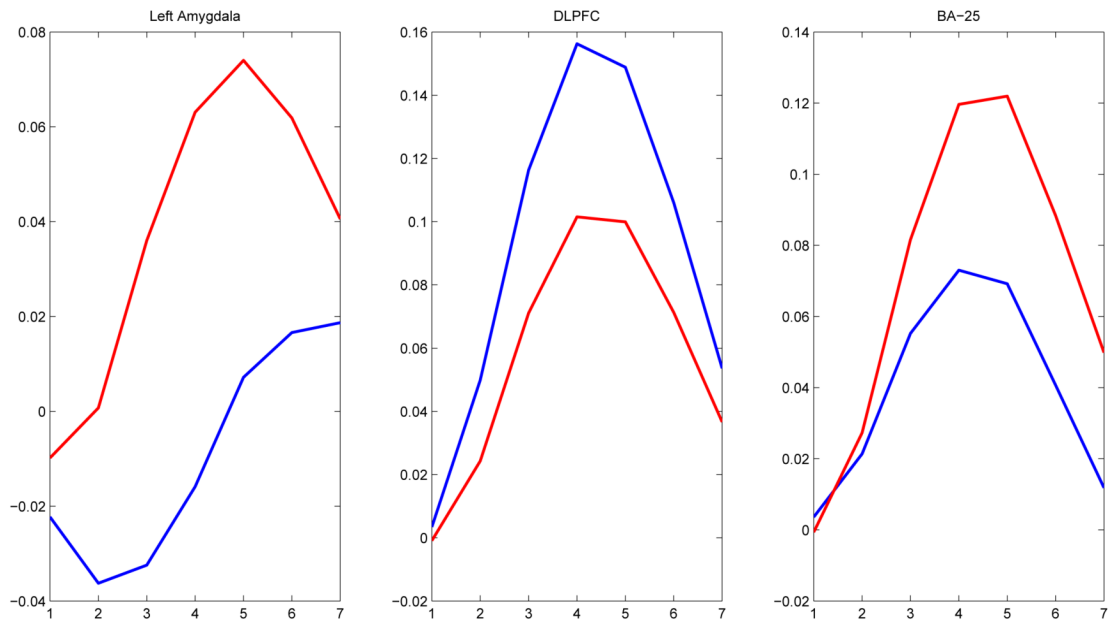
Figure 1.
Group mean stimulus-locked trial activation trajectories in three brain regions in response to 20 negative word stimuli per subject. Regional activation trajectories were determined by averaging activation trajectories over voxels within the region. Blue lines indicate mean stimulus-locked trajectory for 24 mentally-healthy controls and red lines indicate the same for 32 unmedicated depressed subjects.
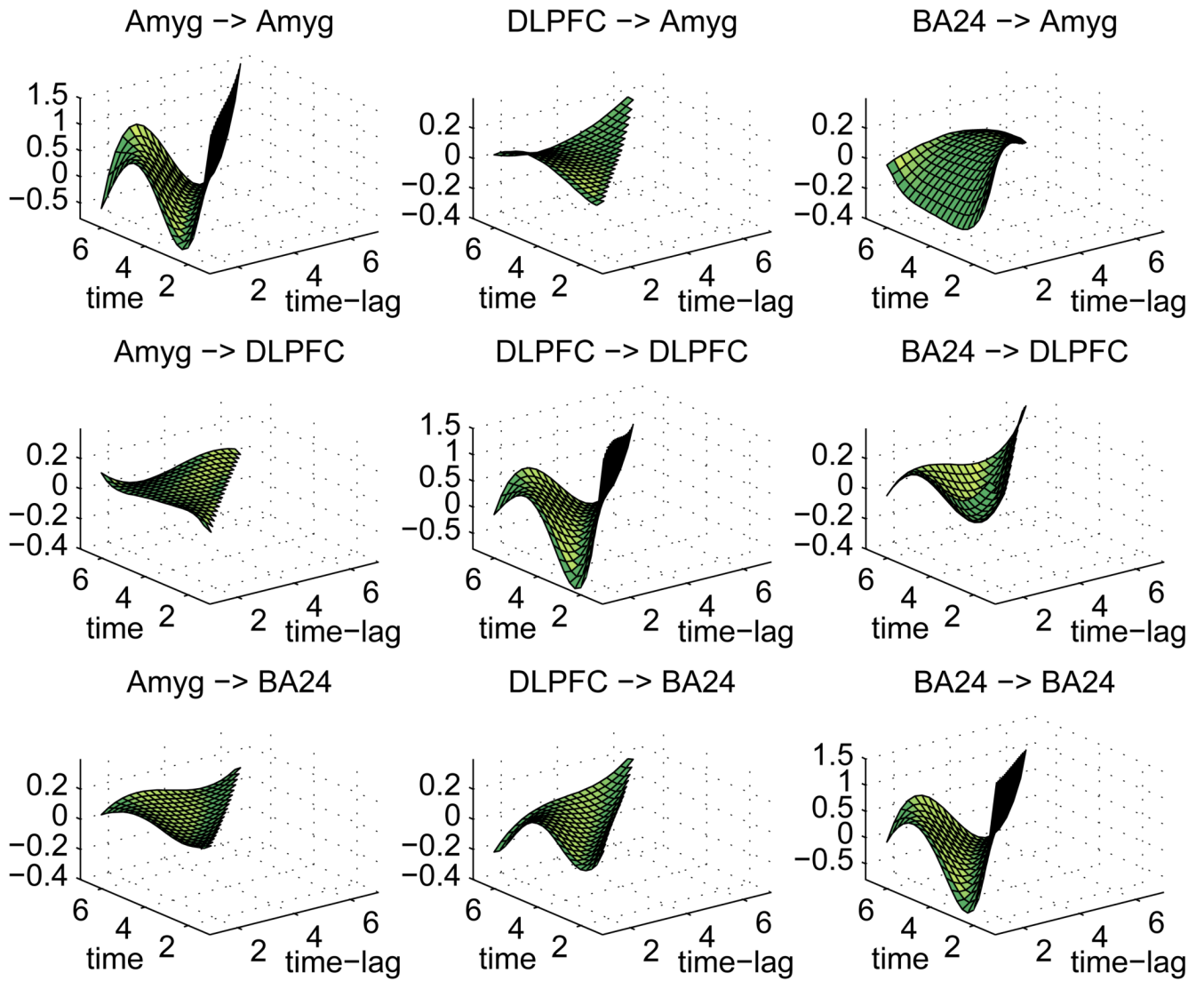
Figure 2.
Posterior connectivity surfaces for control group. Headings indicate influencing region (left of arrow) and influenced region (right of arrow). The x-axis is time-lag ($1 \leq t \leq 6$) and the y-axis is time ($2 \leq t \leq 7$).
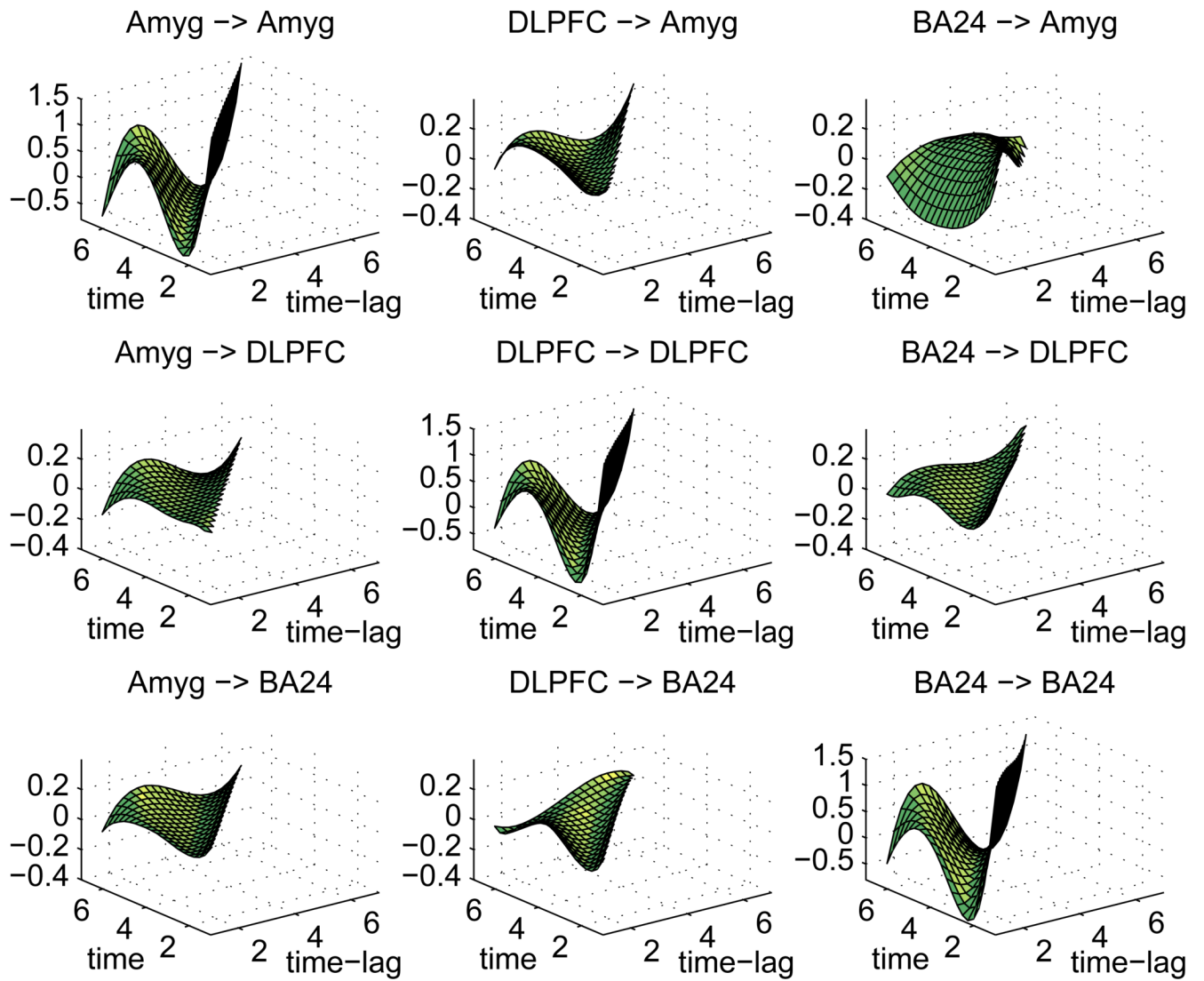
Figure 3.
Posterior connectivity surfaces for depressed group. Headings indicate influencing region (left of arrow) and influenced region (right of arrow). The x-axis is time-lag ($1 \le t \le 6$) and the y-axis is time ($2 \le t \le 7$).

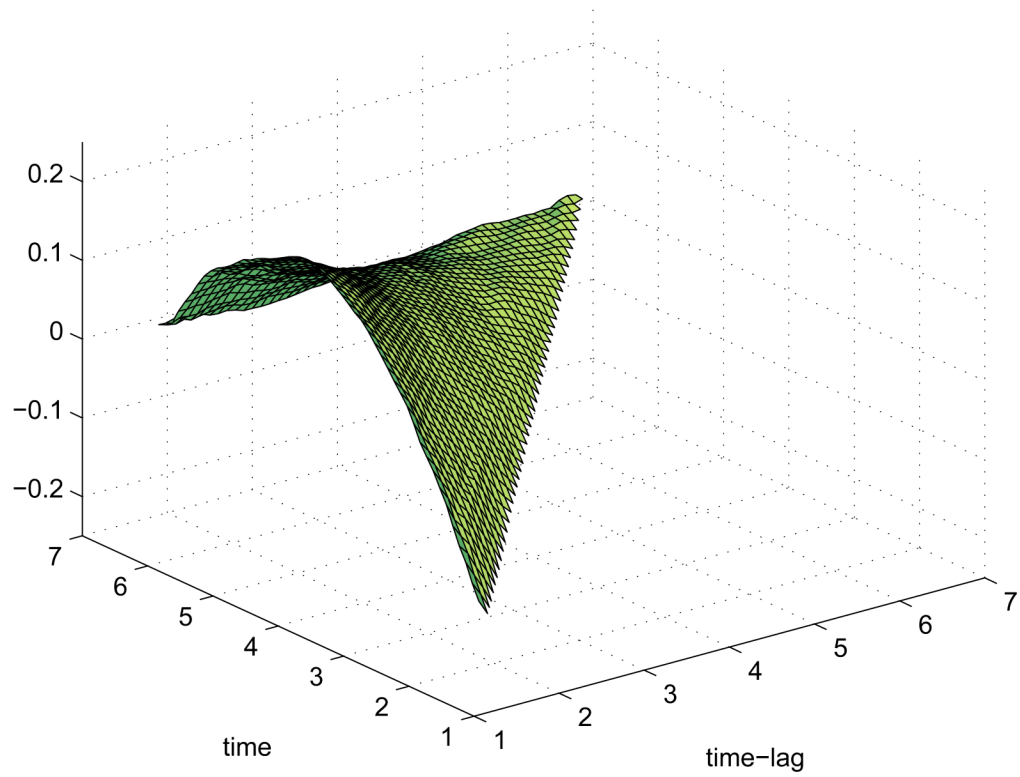**Figure 4.**
Difference in connectivity surfaces (for DLPFC to BA24 connectivity) for 8 depressed subjects with high values of $\eta_{i,\,ba24,\,dlpfc}^{[1]}$ minus 8 depressed subjects with low values of $\eta_{i,\,ba24,\,dlpfc}^{[1]}$.
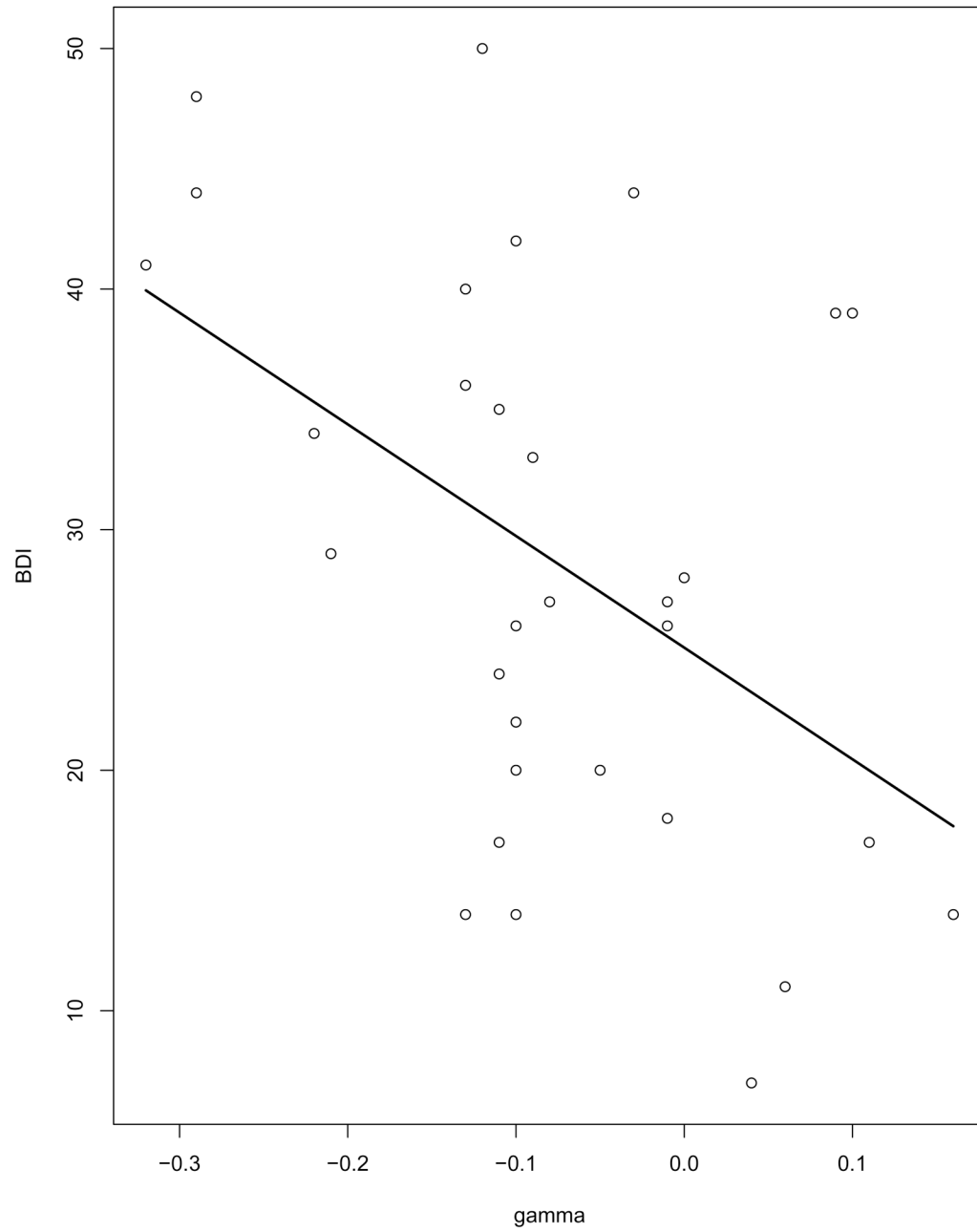
Figure 5.
Scatterplot of pre-treatment BDI in 32 depressed subjects vs. posterior estimates of connectivity coefficients $\gamma_{i\,\text{ba24, dlpfc}}(1, 2)$ determining subject-level variation in DLPFC on BA24 connectivity at time 2 lag 1. Solid line is simple linear regression fit.

Table 1

Estimates of Connectivity Coefficents from VAR and SloVAR Models

time	lag	1 -> 2								2 -> 1	
		true	VAR est	prop	SloVAR est	prop	true	VAR est	prop	SloVAR est	prop
2	1	0.35	0.09	0.00	0.30	0.98	0.00	0.08	0.00	-0.01	0.94
3	1	0.30	0.09	0.00	0.28	0.90	0.00	0.08	0.00	0.00	0.97
3	2	-0.20	-0.10	0.00	-0.16	0.86	0.00	0.05	0.76	0.01	0.91
4	1	0.25	0.09	0.00	0.25	0.96	0.00	0.08	0.00	0.00	0.96
4	2	0.00	-0.10	0.00	-0.01	1.00	0.00	0.05	0.76	0.00	0.98
4	3	-0.20	0.02	0.00	-0.16	0.85	0.00	-0.04	0.06	-0.01	0.99
5	1	0.20	0.09	0.00	0.20	0.97	0.00	0.08	0.00	0.00	0.98
5	2	0.00	-0.10	0.96	0.03	0.96	0.00	0.05	0.76	0.00	1.00
5	3	0.00	0.02	0.00	-0.06	0.74	0.00	-0.04	0.06	0.00	0.98
5	4	-0.20	0.00	1.00	-0.12	0.96	0.00	-0.05	0.12	0.00	0.99
6	1	0.15	0.09	0.04	0.14	0.98	0.00	0.08	0.00	0.00	0.93
6	2	0.00	-0.10	0.00	0.04	0.84	0.00	0.05	0.76	-0.01	0.99
6	3	0.00	0.02	0.96	-0.01	0.99	0.00	-0.04	0.06	0.00	0.99
6	4	0.00	0.00	1.00	-0.05	0.87	0.00	-0.05	0.12	0.00	0.92
6	5	-0.20	-0.01	0.00	-0.08	0.93	0.00	-0.01	1.00	0.01	0.87
7	1	0.10	0.09	0.97	0.09	0.98	0.00	0.08	0.00	0.02	0.91
7	2	0.00	-0.10	0.00	0.02	0.99	0.00	0.05	0.76	-0.01	1.00
7	3	0.00	0.02	0.96	0.02	0.96	0.00	-0.04	0.06	0.00	0.96
7	4	0.00	0.00	1.00	-0.02	0.99	0.00	-0.05	0.12	-0.01	0.98
7	5	0.00	-0.01	0.98	-0.08	0.99	0.00	-0.01	1.00	-0.02	0.98
7	6	-0.20	0.01	0.00	-0.10	0.97	0.00	0.01	1.00	-0.01	0.91

Mean estimates (est) from fitting a standard VAR model and the SloVAR model to 100 simulated datasets with 2 regions. Also indicated are true values (true) of between-region connectivity surfaces and proportion (prop) of times the nominal 95% pointwise posterior credible intervals contained the true parameters. First five columns pertain to the connectivity from region 1 to region 2 and the last five columns are for connectivity from region 2 to region 1.