Attention-dependent Modulation of Cortical Taste Circuits Revealed by Granger Causality with Signal-dependent Noise

Qiang Luo[1,2,3], Tian Ge[3,4], Fabian Grabenhorst[5], Jianfeng Feng[2,3,4,*], Edmund T. Rolls[4,6,*]

[1] Department of Management, School of Information Systems and Management,

National University of Defense Technology, Hunan 410073, P.R. China

[2] Shanghai Center for Mathematical Sciences, Fudan University, Shanghai 200433, P.R. China

[3] Centre for Computational Systems Biology, School of Mathematical Sciences,

Fudan University, Shanghai 200433, P.R. China

[4] Department of Computer Science, University of Warwick, CV4 7AL, UK

[5] Department of Physiology, University of Cambridge, Cambridge, UK

[6] Oxford Centre for Computational Neuroscience, Oxford, UK

[*] Corresponding authors: ER: Edmund.Rolls@oxcns.org, JF: Jianfeng.Feng@warwick.ac.uk

March 27, 2013

Abstract

We show, for the first time, that in cortical areas, for example the insular, orbitofrontal, and lateral prefrontal cortex, there is signal-dependent noise in the fMRI blood-oxygen level dependent (BOLD) signals, with the noise increasing approximately linearly with the square of the signal. Classical Granger causal model is based on autoregressive models with time invariant variance, and thus does not take this signal-dependent noise into account. To address this limitation, here we describe a Granger causal model with signal-dependent noise, and a novel, likelihood ratio test for causal inferences. We applied this approach to the data from an fMRI study to investigate the source of top-down attentional control of taste intensity and taste pleasantness processing. The signal-dependent noise model revealed effects not identified by classical Granger causal analysis. In particular, there was a top-down effect from the posterior lateral prefrontal cortex to the insular taste cortex during attention to intensity but not to pleasantness, and there was a top-down effect from the anterior and posterior lateral prefrontal cortex to the orbitofrontal cortex during attention to pleasantness but not to intensity. In addition, there was stronger forward effective connectivity from the insular taste cortex to the orbitofrontal cortex during attention to pleasantness than during attention to intensity. These findings indicate the importance of explicitly modeling signal-dependent noise in functional neuroimaging, and reveal some of the processes involved in a biased activation theory of selective attention.

1 Introduction

In the past decade, Granger causality (GC) has emerged as a popular method for causal inferences, and has been applied to biological time series obtained from many different modalities, in particular, to the fMRI blood-oxygen level dependent (BOLD) signals to detect effective connectivity between brain areas and understand how the brain works [e.g. Goebel et al., 2003; Hwang et al., 2010; Roebroeck et al., 2005; Wen et al., 2012]. The basic idea of GC can be traced back to [Wiener, 1956], who conceived the notion that if the prediction of one time series can be improved by incorporating the past history of a second one, then the second time series has a causal influence on the first. Granger [1969] later formulated this idea in the context of linear autoregressive (AR) models. GC is completely data-driven and based on time precedence. The interaction discovered by GC may be unidirectional or reciprocal. The simplification of GC is a sward of double edges. It is easy to implement, relies on a small set of straightforward assumptions and does not need any knowledge about how the data are generated. Therefore, it can be applied directly to almost any time series data [Ding et al., 2006]. However, the over-simplification of the model may result in an incorrect use or interpretation of GC and even spurious causal influences in some situations [Friston et al., 2012; Ge et al., 2009].

One possible over-simplification in some scenarios is that the covariance matrix of the noise, conditional on the past history of the time series and the noise process, is assumed to be time invariant. For example, spike trains of neurons are typically close to Poisson processes in their timing, and the variance thus increases linearly with the signal [Gerstein and Mandelbrot, 1964; McAdams and Maunsell, 1999]. Similar conditionally heteroskedas-

tic data have been observed in many physiological recordings, such as the data collected from epilepsy and Parkinson patients. Therefore, it is natural to conjecture that changes in the volatility of one time series may have an impact on the mean activity or volatility of another time series, and vice versa, which indicates that causal influences may happen on the second order statistics or there may be feedbacks between the first and second order moments. Clearly, these causal relationships cannot be captured by classical GC, which is based on a simple AR model, misspecified to time series data with changing volatility. Moreover, although it has been widely observed and investigated that the signal-dependent noise plays important roles in neuronal activities [Harris and Wolpert, 1998; Selen et al., 2009; Todorov, 2005], it is still unclear whether this property carries to fMRI BOLD signals, after the neuronal signals are delayed and blurred by the haemodynamic response.

In this paper, we provide empirical evidence that the variance of the BOLD signal increases linearly with the square of the signal in a number of cortical areas, such as the insula taste, orbitofrontal and lateral prefrontal cortex. We therefore present a Granger causal model with signal-dependent noise to detect GC in both mean and variance from data with time varying volatility. We also propose a likelihood ratio test to infer GC with signal-dependent noise accurately and efficiently. We show, by simulation studies, that this novel method substantially outperforms classical GC when signal-dependent noise is present.

The new method is evaluated with an fMRI investigation [Grabenhorst and Rolls, 2008] to identify the source of the top-down selective attentional control that differentially biases brain systems involved in affective vs sensory analysis [Grabenhorst and Rolls, 2008, 2011; Rolls et al., 2008]. Instructions to pay attention to and later rate the pleasantness of a taste

increased the activations to taste measured with fMRI in the orbitofrontal and pregenual cingulate cortices [Grabenhorst and Rolls, 2008], where the subjective pleasantness of taste is represented [de Araujo et al., 2003b; Grabenhorst et al., 2008, 2010; Kringelbach et al., 2003; Rolls and Grabenhorst, 2008], but not the primary taste cortex in the anterior insula [Grabenhorst and Rolls, 2008], where the subjective intensity and identity of taste are represented [de Araujo et al., 2003b; Grabenhorst et al., 2008; Haase et al., 2009; Kringelbach et al., 2003; Rolls and Grabenhorst, 2008; Small et al., 2003]. Instructions to pay attention to and later rate the intensity of a taste increased the activations to taste in the insular taste cortex but not in the orbitofrontal and pregenual cingulate cortices [Grabenhorst and Rolls, 2008]. Our new method reveal how the effective top-down connectivity changes when attention is paid to the pleasantness vs the intensity of tastes, and helps in the interpretation of the source of the signals that implement top-down attention.

2 Methods

2.1 Granger causality with signal-dependent noise

Classical Granger causality. We start with a brief review of the classical GC. Consider the following zero-mean vector autoregressive model (VAR) of order p:

$$\boldsymbol{x}_t = \sum_{j=1}^{p} \boldsymbol{A}_{x,j} \boldsymbol{x}_{t-j} + \boldsymbol{\epsilon}_{x,t}, \quad t = 1, 2, \cdots, \tag{1}$$

where $\boldsymbol{x}_t$ is a d_x-dimensional column random vector, $\boldsymbol{A}_{x,j}, j = 1, \cdots, p$ are fixed $d_x \times d_x$ coefficient matrices, and $\boldsymbol{\epsilon}_{x,t}$ is a d_x-dimensional independent identically distributed (i.i.d) white noise or innovation process, with a positive definite and time invariant covariance

matrix $\boldsymbol{\Sigma}_x$. We require that this VAR(p) process is stable, that is,

$$\det\left(\boldsymbol{I} - \boldsymbol{A}_{x,1}z - \cdots - \boldsymbol{A}_{x,p}z^p\right) \neq 0, \quad |z| \leqslant 1, \tag{2}$$

where $\det(\cdot)$ is the determinant of a matrix, $\boldsymbol{I}$ is an identity matrix, and z is a complex variable. This stability condition implies that the VAR(p) process is weakly stationary, i.e., its first and second order moments exist and are time invariant [Lütkepohl, 2005].

Now, assume $\boldsymbol{x}_t$ and $\boldsymbol{y}_t$ admit a jointly stable VAR representation. $\boldsymbol{x}_t$ can thus be modeled as

$$\boldsymbol{x}_t = \sum_{j=1}^{p} \boldsymbol{A}_{xx,j}\boldsymbol{x}_{t-j} + \sum_{j=1}^{p} \boldsymbol{A}_{xy,j}\boldsymbol{y}_{t-j} + \boldsymbol{\epsilon}_{xy,t}, \quad t = 1, 2, \cdots, \tag{3}$$

where $\boldsymbol{y}_t$ is a d_y-dimensional column random vector, $\boldsymbol{\epsilon}_{xy,t}$ is a white noise process with a covariance matrix $\boldsymbol{\Sigma}_{xy}$.

The classical GC depends on temporal precedence and predictability. The idea is that a cause cannot come after the effect. Thus, if $\boldsymbol{y}_t$ affects $\boldsymbol{x}_t$, including the past information of $\boldsymbol{y}_t$ should improve the predictions of $\boldsymbol{x}_t$. More formally, if the prediction error of $\boldsymbol{x}_t$ is reduced when the past information of $\boldsymbol{y}_t$ is taken into account, then $\boldsymbol{y}_t$ has a causal influence on $\boldsymbol{x}_t$ in the sense of Granger. Formulating the idea in the context of VAR model, the causal influence from $\boldsymbol{y}_t$ to $\boldsymbol{x}_t$ in the time domain can be quantified as [Geweke, 1982, 1984]

$$\mathcal{F}_{y\to x} = \log\left[\frac{\det(\boldsymbol{\Sigma}_x)}{\det(\boldsymbol{\Sigma}_{xy})}\right]. \tag{4}$$

$\mathcal{F}_{y\to x} > 0$ indicates a causal influence from $\boldsymbol{y}_t$ to $\boldsymbol{x}_t$, and $\mathcal{F}_{y\to x} = 0$ otherwise. Note that (1) is a restricted model of (3), and that $\boldsymbol{y}_t$ does not cause $\boldsymbol{x}_t$ if and only if $\boldsymbol{A}_{xy,j} \equiv 0$ for all $j = 1, \cdots, p$ [Lütkepohl, 2005].

When the white noise is Gaussian distributed, Barnett and Bossomaier [2012] showed that the GC measure in Eq. (4) is equivalent to the likelihood ratio test statistic

$$\mathcal{R}_{y\to x} = \frac{\mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{restricted}} \mid \{\boldsymbol{x}_t\}_{t=1}^T\right)}{\mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right)}, \tag{5}$$

where $\mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{restricted}} \mid \{\boldsymbol{x}_t\}_{t=1}^T\right) \equiv \mathbf{Pr}\left(\{\boldsymbol{x}_t\}_{t=1}^T \mid \widehat{\boldsymbol{\theta}}_{\text{restricted}}\right)$ is the likelihood function, i.e., the probability of the observed time series $\{\boldsymbol{x}_t\}_{t=1}^T$, given the maximum likelihood estimate of the parameters $\widehat{\boldsymbol{\theta}}_{\text{restricted}}$ of the restricted model (1). $\mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right) \equiv \mathbf{Pr}\left(\{\boldsymbol{x}_t\}_{t=1}^T \mid \widehat{\boldsymbol{\theta}}_{\text{full}}, \{\boldsymbol{y}_t\}_{t=1}^T\right)$ is interpreted similarly under the full model (3). Therefore, a likelihood ratio test can be used for causal inference:

$$r_{y\to x} = -2\left[\log \mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{restricted}} \mid \{\boldsymbol{x}_t\}_{t=1}^T\right) - \log \mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right)\right]. \tag{6}$$

The test statistic $r_{y\to x}$ is approximately chi-squared distributed, with degrees of freedom $df_{\text{full}} - df_{\text{restricted}}$, where df_{full} and $df_{\text{restricted}}$ are the number of free parameters of the full model (3) and the restricted model (1), respectively.

Granger causality with signal-dependent noise. To relax the assumption of time invariant covariance matrix in the AR model, Engle [1982] invented the first changing volatility model – the autoregressive conditional heteroskedasticity (ARCH) models, which was then extended to generalized ARCH (GARCH) models [Bollerslev, 1986; Taylor, 1986] as well as multivariate cases [Bollerslev et al., 1988; Diebold and Nerlove, 1989; Engle et al., 1984]. Assume $\boldsymbol{r}_t$ is a d-dimensional zero mean, serially uncorrelated process, which may be the residual process of some dynamic model and can be represented as

$$\boldsymbol{r}_t = \boldsymbol{\Sigma}_{t|t-1}^{1/2}\boldsymbol{\epsilon}_t, \tag{7}$$

where $\boldsymbol{\epsilon}_t$ is a d-dimensional i.i.d white noise process, $\boldsymbol{\Sigma}_{t|t-1}$ is the conditional covariance matrix of $\boldsymbol{r}_t$, given $\Omega_{t-1} = \{\boldsymbol{r}_{t-1}, \boldsymbol{r}_{t-2}, \cdots\}$, and $\boldsymbol{\Sigma}_{t|t-1}^{1/2}$ is the symmetric positive definite square root of $\boldsymbol{\Sigma}_{t|t-1}$. Then a multivariate ARCH process of order q takes the form

$$\text{vech}\left(\boldsymbol{\Sigma}_{t|t-1}\right) = \boldsymbol{\gamma}_0 + \boldsymbol{\Gamma}_1 \text{vech}\left(\boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}^\top\right) + \cdots + \boldsymbol{\Gamma}_q \text{vech}\left(\boldsymbol{r}_{t-q}\boldsymbol{r}_{t-q}^\top\right), \tag{8}$$

where vech denotes the half-vectorization operator which stacks the columns of a square matrix from the diagonal downwards in a vector, $\boldsymbol{\gamma}_0$ is a $\frac{1}{2}d(d+1)$-dimensional column vector of constants and $\boldsymbol{\Gamma}_j, j = 1, \cdots, q$ are $\frac{1}{2}d(d+1) \times \frac{1}{2}d(d+1)$ coefficient matrices. It can be seen that even for a bivariate series with a low order, this general model has a fair number of parameters. Therefore, more restricted models were proposed. For example, Bollerslev et al. [1988] considered diagonal ARCH processes where all the $\boldsymbol{\Gamma}_j$ matrices are diagonal. To guarantee the positive definiteness of the conditional covariance matrix $\boldsymbol{\Sigma}_{t|t-1}$, Baba, Engle, Kraft & Kroner investigated the following variant of a multivariate ARCH model, known as the BEKK model [Baba et al., 1991; Engle and Kroner, 1995]

$$\boldsymbol{\Sigma}_{t|t-1} = \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1^\top \boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}^\top \boldsymbol{\Gamma}_1 + \cdots + + \boldsymbol{\Gamma}_q^\top \boldsymbol{r}_{t-q}\boldsymbol{r}_{t-q}^\top \boldsymbol{\Gamma}_q, \tag{9}$$

where $^\top$ is the matrix transpose, $\boldsymbol{\Gamma}_0 = \boldsymbol{C}_0^\top \boldsymbol{C}_0$ is positive definite, and all $\boldsymbol{\Gamma}_j, j = 1, \cdots, q$ are $d \times d$ matrices. In contrast to the diagonal model, BEKK model produces interactions between second order moments and can generate rich volatility dynamics.

We now present a Granger causal model with signal-dependent noise [Luo et al., 2011]. Consider the following multivariate model with time varying volatility, in particular, signal-dependent noise:

$$\begin{aligned} &\boldsymbol{x}_t = \textstyle\sum_{j=1}^p \boldsymbol{A}_{x,j}\boldsymbol{x}_{t-j} + \boldsymbol{r}_{x,t}, \quad \boldsymbol{r}_{x,t} = \boldsymbol{\Sigma}_{x,t|t-1}^{1/2}\boldsymbol{\epsilon}_{x,t}, \\ &\boldsymbol{\Sigma}_{x,t|t-1} = \boldsymbol{C}_x^\top \boldsymbol{C}_x + \textstyle\sum_{j=1}^q \boldsymbol{B}_{x,j}^\top \boldsymbol{x}_{t-j}\boldsymbol{x}_{t-j}^\top \boldsymbol{B}_{x,j}, \end{aligned} \tag{10}$$

where $\boldsymbol{x}_t$ is a d_x-dimensional column random vector, $\boldsymbol{\epsilon}_{x,t}$ is a d_x-dimensional Gaussian distributed white noise process with zero mean and unit variance, p and q are the model orders, $\boldsymbol{A}_{x,j}, j = 1, \cdots, p$, $\boldsymbol{B}_{x,j}, j = 1, \cdots, q$ and $\boldsymbol{C}_x$ are coefficient matrices. The volatility model is a modification of the BEKK model [Engle and Kroner, 1995] in which the conditional covariance matrix $\boldsymbol{\Sigma}_{x,t|t-1}$ does not regress on the residual process $\boldsymbol{r}_{x,t}$ but only depends on the activity of the signal $\boldsymbol{x}_t$. Hence, the covariance (second order statistics) of the noise process is coupled to the mean (first order statistics). This form also guarantees the positive definiteness of $\boldsymbol{\Sigma}_{x,t|t-1}$. Clearly, when $\boldsymbol{B}_{x,j} \equiv 0$ for all j, the conditional covariance is time invariant and the model reduces to the AR model. In light of these, we term our model (10) the AR-BEKK model.

To define the causal relationship between $\boldsymbol{x}_t$ and another d_y-dimensional time series $\boldsymbol{y}_t$, consider the following joint AR-BEKK model:

$$\boldsymbol{z}_t = \sum_{j=1}^{p} \boldsymbol{A}_j \boldsymbol{z}_{t-j} + \boldsymbol{r}_t, \tag{11}$$

where $\boldsymbol{z}_t = \left(\boldsymbol{x}_t^\top, \boldsymbol{y}_t^\top\right)^\top$, $\boldsymbol{r}_t = \left(\boldsymbol{r}_{xy,t}^\top, \boldsymbol{r}_{yx,t}^\top\right)^\top$, $\boldsymbol{r}_{xy,t} = \boldsymbol{\Sigma}_{xy,t|t-1}^{1/2} \boldsymbol{\epsilon}_{xy,t}$, $\boldsymbol{r}_{yx,t} = \boldsymbol{\Sigma}_{yx,t|t-1}^{1/2} \boldsymbol{\epsilon}_{yx,t}$, $\boldsymbol{\epsilon}_{xy,t}$ and $\boldsymbol{\epsilon}_{yx,t}$ are d_x-dimensional and d_y-dimensional Gaussian distributed white noise process with zero mean and unit variance respectively, and

$$\begin{aligned} \boldsymbol{\Sigma}_{xy,t|t-1} &= \boldsymbol{C}_{xy}^\top \boldsymbol{C}_{xy} + \textstyle\sum_{j=1}^{q} \boldsymbol{B}_{xy,j}^\top \boldsymbol{z}_{t-j} \boldsymbol{z}_{t-j}^\top \boldsymbol{B}_{xy,j}, \\ \boldsymbol{\Sigma}_{yx,t|t-1} &= \boldsymbol{C}_{yx}^\top \boldsymbol{C}_{yx} + \textstyle\sum_{j=1}^{q} \boldsymbol{B}_{yx,j}^\top \boldsymbol{z}_{t-j} \boldsymbol{z}_{t-j}^\top \boldsymbol{B}_{yx,j}. \end{aligned} \tag{12}$$

Here, $\boldsymbol{A}_j, j = 1, \cdots, p$, $\boldsymbol{B}_{xy,j}, \boldsymbol{B}_{yx,j}, j = 1, \cdots, q$ and $\boldsymbol{C}_{xy}$, $\boldsymbol{C}_{yx}$ are all coefficient matrices. The causal influence from $\boldsymbol{y}_t$ to $\boldsymbol{x}_t$ can be defined as [Luo et al., 2011]

$$\mathcal{F}_{y \to x} = \log \left[\frac{\det\left(\boldsymbol{C}_x^\top \boldsymbol{C}_x\right)}{\det\left(\boldsymbol{C}_{xy}^\top \boldsymbol{C}_{xy}\right)} \right]. \tag{13}$$

$\mathcal{F}_{y\to x} > 0$ if $\boldsymbol{y}_t$ has a causal effect on $\boldsymbol{x}_t$, and $\mathcal{F}_{y\to x} = 0$ otherwise. It can been seen that $\boldsymbol{y}_t$ can help improve the prediction of $\boldsymbol{x}_t$ by impacting on either its mean activity through the coefficients in $\boldsymbol{A}_j$, or its variance through the coefficients in $\boldsymbol{B}_{xy,j}$. These two cases correspond to the causality in mean and variance respectively. When the noise in $\boldsymbol{x}_t$ is not signal-dependent, i.e., $\boldsymbol{B}_{x,j} \equiv 0$ and $\boldsymbol{B}_{xy,j} \equiv 0$ for all j, the model reduces to a VAR model and the definition of causality coincides with the classical GC in the time domain (See Eq. (4)).

Stability conditions. To guarantee the stability of the AR-BEKK model, we provide the stability condition for a simple first-order model, i.e., $p = q = 1$ in Eq. (11). This is the model that we usually use for fMRI data analysis, considering the poor temporal resolution of BOLD signals, and the relatively fast signal transmission between groups of neurons. Going forward, we focus on this first-order model unless otherwise specified. We also assume $\boldsymbol{\epsilon}_{xy,t}$ and $\boldsymbol{\epsilon}_{yx,t}$ are uncorrelated. The stability of the model involves both the first and second order stability conditions, i.e., the unconditional mean and covariance of $\boldsymbol{z}_t$ exist and are time invariant. For the first order stability condition, it follows from the theory of the AR model that all the eigenvalues of $\boldsymbol{A}_1$ have modulus less than 1. For the second order stability, note that

$$\operatorname{cov}(\boldsymbol{z}_t \mid \boldsymbol{\Omega}_{t-1}) = \operatorname{cov}(\boldsymbol{u}_t \mid \boldsymbol{\Omega}_{t-1}) = \operatorname{diag}\left\{\boldsymbol{\Sigma}_{xy,t|t-1}, \boldsymbol{\Sigma}_{yx,t|t-1}\right\}, \tag{14}$$

and

$$\operatorname{cov}(\boldsymbol{z}_t \mid \boldsymbol{\Omega}_{t-1}) = \mathbf{E}(\boldsymbol{z}_t \boldsymbol{z}_t^\top \mid \boldsymbol{\Omega}_{t-1}) - \mathbf{E}(\boldsymbol{z}_t \mid \boldsymbol{\Omega}_{t-1})\mathbf{E}(\boldsymbol{z}_t^\top \mid \boldsymbol{\Omega}_{t-1}), \tag{15}$$

where $\boldsymbol{\Omega}_{t-1} = (\boldsymbol{z}_{t-1}, \boldsymbol{z}_{t-2}, \cdots)$. Therefore,

$$\begin{aligned}\mathbf{E}(\boldsymbol{z}_t\boldsymbol{z}_t^\top \mid \boldsymbol{\Omega}_{t-1}) &= \boldsymbol{A}_1\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\boldsymbol{A}_1^\top + \operatorname{diag}\big\{\boldsymbol{C}_{xy}^\top\boldsymbol{C}_{xy} + \boldsymbol{B}_{xy,1}^\top\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\boldsymbol{B}_{xy,1}, \\ &\qquad \boldsymbol{C}_{yx}^\top\boldsymbol{C}_{yx} + \boldsymbol{B}_{yx,1}^\top\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\boldsymbol{B}_{yx,1}\big\}.\end{aligned} \tag{16}$$

Taking expectation of both sides yields

$$\begin{aligned}\mathbf{E}(\boldsymbol{z}_t\boldsymbol{z}_t^\top) &= \boldsymbol{A}_1\mathbf{E}\left(\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\right)\boldsymbol{A}_1^\top + \operatorname{diag}\big\{\boldsymbol{C}_{xy}^\top\boldsymbol{C}_{xy} + \boldsymbol{B}_{xy,1}^\top\mathbf{E}\left(\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\right)\boldsymbol{B}_{xy,1}, \\ &\qquad \boldsymbol{C}_{yx}^\top\boldsymbol{C}_{yx} + \boldsymbol{B}_{yx,1}^\top\mathbf{E}\left(\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\right)\boldsymbol{B}_{yx,1}\big\}.\end{aligned} \tag{17}$$

This can be transformed into the following equation using the vectorization operator, which stacks the columns of a square matrix:

$$\begin{aligned}&\operatorname{vec}\left[\mathbf{E}(\boldsymbol{z}_t\boldsymbol{z}_t^\top)\right] = \boldsymbol{A}_1 \otimes \boldsymbol{A}_1\operatorname{vec}\left[\mathbf{E}\left(\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\right)\right] + \operatorname{vec}\left[\operatorname{diag}\left\{\boldsymbol{C}_{xy}^\top\boldsymbol{C}_{xy}, \boldsymbol{C}_{yx}^\top\boldsymbol{C}_{yx}\right\}\right] \\ &+ \left[\operatorname{diag}\left\{\boldsymbol{B}_{xy,1}^\top, \boldsymbol{B}_{yx,1}^\top\right\} \otimes \operatorname{diag}\left\{\boldsymbol{B}_{xy,1}^\top, \boldsymbol{B}_{yx,1}^\top\right\}\right] \cdot \boldsymbol{T} \cdot \operatorname{vec}\left[\mathbf{E}\left(\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\right)\right] \\ &:= \left(\widetilde{\boldsymbol{A}} + \widetilde{\boldsymbol{B}}\right)\operatorname{vec}\left[\mathbf{E}\left(\boldsymbol{z}_{t-1}\boldsymbol{z}_{t-1}^\top\right)\right] + \widetilde{\boldsymbol{C}},\end{aligned}$$

where $\otimes$ is the Kronecker product,

$$\boldsymbol{T} = \begin{bmatrix}\boldsymbol{I} \\ \\ \boldsymbol{0}\end{bmatrix} \otimes \begin{bmatrix}\boldsymbol{I} \\ \\ \boldsymbol{0}\end{bmatrix} + \begin{bmatrix}\boldsymbol{0} \\ \\ \boldsymbol{I}\end{bmatrix} \otimes \begin{bmatrix}\boldsymbol{0} \\ \\ \boldsymbol{I}\end{bmatrix}.$$

Therefore, it is required that all the eigenvalues of $\widetilde{\boldsymbol{A}} + \widetilde{\boldsymbol{B}}$ have modulus less than 1.

Model estimation. Using Bayes' theorem, the joint density function of $\boldsymbol{r}_1, \cdots, \boldsymbol{r}_T$ is

$$f(\boldsymbol{r}_1, \cdots, \boldsymbol{r}_T) = f(\boldsymbol{r}_1)f(\boldsymbol{r}_2 \mid \boldsymbol{r}_1)\cdots f(\boldsymbol{r}_T \mid \boldsymbol{r}_{T-1}, \cdots, \boldsymbol{r}_1). \tag{18}$$

Thus, the conditional distribution of $\boldsymbol{r}_t$ given $\boldsymbol{\Omega}_{t-1}$ is Gaussian and if the $\boldsymbol{u}_t$ are observed quantities, the log-likelihood function of the AR-BEKK model described by Eq. (11), for

a sample $\boldsymbol{u}_1, \cdots, \boldsymbol{u}_T$ is given by

$$\log \mathcal{L}\left(\boldsymbol{\theta}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right) = \sum_{t=1}^T \log \mathcal{L}_t\left(\boldsymbol{\theta}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right), \tag{19}$$

where $\boldsymbol{\theta}_{\text{full}}$ is a vector of all unknown parameters of the model (11) and

$$\begin{aligned} &\log \mathcal{L}_t\left(\boldsymbol{\theta}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right) \\ &\quad = -\frac{d_x + d_y}{2} \log 2\pi - \frac{1}{2} \log |\boldsymbol{\Sigma}_{xy,t|t-1}| - \frac{1}{2} \boldsymbol{u}_t^\top \boldsymbol{\Sigma}_{xy,t|t-1}^{-1} \boldsymbol{u}_t, \end{aligned} \tag{20}$$

where the required initial values for specifying $\boldsymbol{\Sigma}_{t|t-1}$ are assumed to be available. The likelihood function may be maximized with respect to the parameters $\boldsymbol{\theta}_{\text{full}}$ by using numerical methods. Specifically, the initial values of $\boldsymbol{A}_1$ is given by the least square estimates of $\boldsymbol{z}_t = \boldsymbol{A}_1 \boldsymbol{z}_{t-1} + \boldsymbol{r}_t$, assuming a simple AR model, and $\boldsymbol{B}_{xy,1}, \boldsymbol{B}_{yx,1}$ and $\boldsymbol{C}_{xy}, \boldsymbol{C}_{yx}$ are then initialized to diagonal matrices whose l-th element on the diagonal is the least square estimates of

$$\left[\boldsymbol{r}_{xy,t}^2\right]_l = \left[\boldsymbol{C}_{xy}^2\right]_{ll} + \left[\boldsymbol{B}_{xy,1}^2\right]_{ll} \left[\boldsymbol{z}_{t-1}^2\right]_l, \tag{21}$$

and

$$\left[\boldsymbol{r}_{yx,t}^2\right]_l = \left[\boldsymbol{C}_{yx}^2\right]_{ll} + \left[\boldsymbol{B}_{yx,1}^2\right]_{ll} \left[\boldsymbol{z}_{t-1}^2\right]_l, \tag{22}$$

using the residuals $\boldsymbol{r}_t$ from the AR fitting. The constrained maximum likelihood estimation of the model parameters can be obtained by solving the optimization problem

$$\widehat{\boldsymbol{\theta}}_{\text{full}} = \text{argmax}_{\boldsymbol{\theta}_{\text{full}}} \sum_{t=1}^T \log \mathcal{L}_t(\boldsymbol{\theta}_{\text{full}}), \tag{23}$$

while satisfying the first and second order stability conditions derived above. We use MATLAB function fmincon with the interior-point algorithm to tackle this restricted optimization problem[1]. The parameters of the restricted model (10) can be estimated similarly.

[1]Note that sometimes fmincon with the interior-point algorithm may fail to converge to a reasonable solution. In this case, we use the active-set algorithm as an alternative.

Causal inferences. The nonparametric bootstrap method for causal inferences used in Luo et al. [2011] is time-consuming. Here we develop an analog of the likelihood ratio test for classical GC (See Eq. (6)) to improve computational efficiency. Similarly, the likelihood ratio test statistic takes the form

$$\begin{aligned} r_{y\to x} &= -2\left[\log\mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{restricted}} \mid \{\boldsymbol{x}_t\}_{t=1}^T\right) - \log\mathcal{L}\left(\widehat{\boldsymbol{\theta}}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right)\right] \\ &= -2\sum_{t=1}^T\left[\log\mathcal{L}_t\left(\widehat{\boldsymbol{\theta}}_{\text{restricted}} \mid \{\boldsymbol{x}_t\}_{t=1}^T\right) - \log\mathcal{L}_t\left(\widehat{\boldsymbol{\theta}}_{\text{full}} \mid \{\boldsymbol{x}_t\}_{t=1}^T, \{\boldsymbol{y}_t\}_{t=1}^T\right)\right]. \end{aligned} \tag{24}$$

The test statistic approximately follows a chi-squared distribution, and the degrees of freedom are $2d_xd_y$. Therefore, a parametric chi-squared test can be carried out to test the significance of the causal influence. This likelihood ratio test has also a connection to the transfer entropy between time series [Barnett and Bossomaier, 2012]. However, since the residual process in the AR-BEKK model is not a Gaussian white noise, the likelihood ratio test is not equivalent to the measure defined in Eq. (13).

To test the difference between the causalities in the opposite directions between brain areas, note that the difference of the two causality measures is $r_{\text{diff}} = \frac{1}{2}r_{y\to x} - \frac{1}{2}r_{x\to y}$, where $r_{y\to x}$ and $r_{x\to y}$ are two chi-squared distributed random variables with the same degrees of freedom. Therefore, the distribution function of r_{diff} is

$$T_m(x) = \frac{1}{2^m\sqrt{\pi}\Gamma\left(m+\frac{1}{2}\right)}x^mK_m(x), \tag{25}$$

where K_m is a modified Bessel function, $\Gamma(\cdot)$ is a Gamma function, and $m = d_xd_y - \frac{1}{2}$ [Knepp and Entwisle, 1969]. A table for the two-sided one and five percent quantile of this distribution can be found in Knepp and Entwisle [1969]. For example, in the investigation of a pair of univariate time series, i.e., $d_x = d_y = 1$, a difference measure of 4.61 implies a p-value 0.01.

2.2 Simulation studies

We illustrate the Granger causal model with signal-dependent noise and the likelihood ratio test by a simulation study. Consider the following first-order AR-BEKK model for two univariate time series x_t and y_t.

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1\sqrt{2} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} r_{xy,t} \\ r_{yx,t} \end{bmatrix}, \quad \begin{bmatrix} r_{xy,t} \\ r_{yx,t} \end{bmatrix} = \begin{bmatrix} \mathbf{\Sigma}^{1/2}_{xy,t|t-1}\boldsymbol{\epsilon}_{xy,t} \\ \mathbf{\Sigma}^{1/2}_{yx,t|t-1}\boldsymbol{\epsilon}_{yx,t} \end{bmatrix},$$

$$\begin{aligned} \mathbf{\Sigma}^{1/2}_{xy,t|t-1} &= 1 + [u_1,\ (1-u_1)u_3][x_{t-1},\ y_{t-1}][x_{t-1},\ y_{t-1}]^\top[u_1,\ (1-u_1)u_3]^\top, \\ \mathbf{\Sigma}^{1/2}_{yx,t|t-1} &= 1 + [(1-u_2)u_4,\ u_2][x_{t-1},\ y_{t-1}][x_{t-1},\ y_{t-1}]^\top[(1-u_2)u_4,\ u_2]^\top, \end{aligned}$$

where u_1 and u_2 are random numbers uniformly distributed in $[0, 1]$, and u_2 and u_4 are random numbers with the probability of 0.6 to be 0 and 0.4 to be 1. It is clear that there is a causal influence from y_t to x_t if and only if $(1-u_1)u_3 \neq 0$, and from x_t to y_t if and only if $(1-u_2)u_4 \neq 0$.

We generated 100 models with different $u_i, i = 1, \cdots, 4$, and for each model, we generated time series of 1000 points with 2 replicates. We then fitted both the classical Granger causal model and the signal-dependent noise model to the data. Using different p-value thresholds, the performance of the two models were compared by the ROC (Receiver Operating Characteristic) curve [Fawcett, 2006].

2.3 fMRI experiment

The fMRI data are the same as those used in the previous studies [Ge et al., 2012; Grabenhorst and Rolls, 2008, 2010]. We describe key imaging acquisition, preprocessing and psychophysiological interactions (PPI) analysis for completeness. We refer the readers to

previous publications for full details.

Participants and ethics statement. Twelve healthy volunteers (6 male and 6 female, age range $21 - 35$) participated in the study. Ethical approval (Central Oxford Research Ethics Committee) and written informed consent from all subjects were obtained before the experiment. The subjects had not eaten for three hours before the investigation.

Experimental design. We used the identical taste stimulus, 0.1 M monosodium glutamate (MSG) with 0.005 M inosine monophosphate (see de Araujo et al. [2003a]), referred to throughout this paper for brevity as monosodium glutamate, in two different types of trials. A trial started 5 seconds before the taste delivery with the visual attentional instruction either "Remember and Rate Pleasantness" or "Remember and Rate Intensity", which was shown until the end of the taste period. The 0.75 ml taste stimulus was delivered at $t = 5$ s. The taste period was from $t = 5$ s until $t = 14$ s, and in this period a red cross was also present indicating that swallowing should not occur. The differences between the activations in this period were a measure of the top-down selective attention instructions while the taste was being delivered. (We note that in order to utilize top-down attention, one needs to hold the object of attention in mind, in this case pleasantness or intensity. This requires a short-term memory. Short-term memory is thus a sine qua non of selective attention [Rolls, 2008; Rolls and Deco, 2002], and it is the source of this top-down bias from a short-term memory system in which we are interested in this investigation.) After the end of the taste period the visual instruction and red cross were turned off, and a green cross was shown cueing the subject to swallow. After 2 s a tasteless rinse was delivered with a red cross, and the rinse period was from $t = 16$ s until $t = 23$ s, when the green

cross appeared to cue a swallow. After this the rating of pleasantness or intensity was made using button-press operated visual analog rating scales ranging continuously from $+2$ (very pleasant) to -2 (very unpleasant) for pleasantness, and 4 (intense) to 0 (very weak) for intensity as described previously [Rolls et al., 2003]. These two trail types were interspersed in random permuted sequence with other trials that were part of a different investigation, and each was presented 9 times. As different trial types were being run in the scanner at the same time, and included different stimuli, and no instructions were given about the number of stimuli being used, or that the stimuli were the same on the "Remember and Rate Intensity" and "Remember and Rate Pleasantness" trials, the participants simply had to concentrate on following the instructions about what aspect of the taste stimulus, intensity or pleasantness, had to be rated on that trial. The general protocol and design were described in Grabenhorst et al. [2008], and have been used successfully in previous studies to investigate taste cortical areas [de Araujo et al., 2003c; McCabe and Rolls, 2007; O'doherty et al., 2001].

fMRI data acquisition. Images were acquired with a 3.0-T VARIAN/SIEMENS whole-body scanner at the Centre for Functional Magnetic Resonance Imaging at Oxford (FMRIB), where 27 $T2^*$ weighted EPI coronal slices with in-plane resolution of 3×3 mm and between plane spacing of 4 mm were acquired every 2 seconds ($\mathrm{TR} = 2$ s). We used the techniques that we have developed over a number of years [de Araujo et al., 2003a; O'doherty et al., 2001], and as described in detail by Wilson et al. [2002] we carefully selected the imaging parameters in order to minimize susceptibility and distortion artefact in the orbitofrontal cortex. The relevant factors include imaging in the coronal plane, mini-

mizing voxel size in the plane of the imaging, as high a gradient switching frequency as possible (960 Hz), a short echo time of 28 ms, and local shimming for the inferior frontal area. The matrix size was 64×64 and the field of view was 192×192 mm. Continuous coverage was obtained from $+62$ (A/P) to -46 (A/P). A whole brain $T2^*$ weighted EPI volume of the above dimensions, and an anatomical $T1$ volume with coronal plane slice thickness 3 mm and in-plane resolution of 1×1 mm was also acquired.

fMRI data preprocessing. The imaging data were analyzed using SPM5 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London. `http://www.fil.ion.ucl.ac.uk/spm/`). Preprocessing of the data used SPM5 realignment, reslicing with sinc interpolation, normalization to the Montreal Neurological Institute (MNI) coordinate system [Collins et al., 1994], and spatial smoothing with a 6 mm full width at half maximum (FWHM) isotropic Gaussian kernel. Time series non-sphericity at each voxel was estimated and corrected for [Friston et al., 2002], and a high-pass filter with a cut-off period of 128 seconds was applied.

fMRI data analysis. To investigate task dependent activations of brain areas during the taste period, a Finite Impulse Response (FIR) analysis was performed as implemented in SPM, in order to make no assumption about the time course based on the temporal filtering property of the haemodynamic response function [Gottfried et al., 2006; Yacubian et al., 2006]. The *a priori* defined areas of interest (ROI) for which we reported results included brain areas where activations to taste stimuli have been found in previous studies including the medial and lateral orbitofrontal cortex, the pregenual part of the cingulate cortex, and the taste and oral somatosensory parts of the insular cortex [de Araujo et al., 2003a,c;

Grabenhorst et al., 2008; McCabe and Rolls, 2007; Nitschke et al., 2006; O'doherty et al., 2001; Schoenfeld et al., 2004]; and areas of the lateral prefrontal cortex where activations related to task set, attentional instructions, and remembering rules that guide task performance have been found, including specifically parts of the middle and inferior frontal gyrus [Beck and Kastner, 2009; Bengtsson et al., 2009; Deco and Rolls, 2005a; Kouneiher et al., 2009; Rossi et al., 2009; Sakai and Passingham, 2002, 2006; Veldhuizen et al., 2007]. A contrast of trials where attention was being paid to taste pleasantness with trials where attention was to intensity revealed significant effects in the orbitofrontal cortex [-6, 14, -20]. The reverse contrast of trials where attention was to intensity vs trials where attention was to pleasantness revealed significant effects in the right anterior insular taste cortex [42, 18, -14] [Grabenhorst and Rolls, 2008].

We then performed PPI analyses [Friston et al., 1997; Gitelman et al., 2003], using the above two brain areas as seed regions, to investigate task-dependent functional connectivity of these areas with other brain areas, that might provide the source of the top-down modulation [Grabenhorst and Rolls, 2010]. We identified an anterior lateral prefrontal cortex (AntLPFC) region at $Y = 53$ mm in which the correlation with activity in the orbitofrontal cortex (OFC) seed region was greater when attention was to pleasantness than to intensity [Grabenhorst and Rolls, 2010]. Conversely, in a more posterior region of lateral prefrontal cortex (PostLPFC) at $Y = 34$ mm the correlation with activity in the anterior insula (AntINS) seed region was greater when attention was to intensity than to pleasantness [Grabenhorst and Rolls, 2010]. See Figure 1 for the locations of the seed regions and the identified foci in AntLPFC and PostLPFC.

Granger causal analysis. The PPI analyses described above do not show the directionality of the influences, as they are based on correlations, and for that reason we performed Granger causal analyses to each pair of the four brain areas (OFC, AntINS, AntLPFC and PostLPFC). We extracted the mean BOLD signals from 33 voxels within a sphere of radius 2 voxels centered at the seed voxels in OFC and AntINS, and the peak voxels identified with the largest PPI effect in AntLPFC and PostLPFC, for Granger causal analyses. For each of the two experimental conditions (attention to intensity vs attention to pleasantness), the time series for a single subject consisted of 9 trials, each with 18 BOLD signal data points (2 s apart), starting on each trial at the onset of the instruction to pay attention to the pleasantness or to the intensity of the taste.

For each experimental condition and each pair of the four brain areas, we pooled data from all subjects (12×9 trails) to fit the signal-dependent noise model. We detected unidirectional causal influences as well as significant difference of the causalities in opposite directions [Roebroeck et al., 2005] to identify the dominant causal influences in a particular direction. We also applied classical GC to the same data set as a comparison.

2.4 Empirical analysis of BOLD signals

A model-free analysis is carried out to provide empirical evidence of the signal-dependent noise in BOLD signals. Assume that for each subject, the activation profiles are the same for different trials under the same experimental condition, while the observed intra-subject inter-trail variation is a consequence of noise. Then if the variance of the signal across the trails is large where the mean magnitude of the signal is large (e.g. at the peak), the

signal exhibits signal-dependent noise (See Figure 2A for a cartoon illustration). Therefore, for each of the four brain regions and each subject, we check the correlation between the variance of the BOLD signal at time t across the 9 trails and the squared mean magnitude of the signal at time $t-1$, i.e., the correlation between $\mathrm{var}(x_t)$ and $(\mathbf{E}x_{t-1})^2$. Any nonrandom relationship, in particular, a significant correlation, between $\mathrm{var}(x_t)$ and $(\mathbf{E}x_{t-1})^2$, indicates the presence of the signal-dependent noise.

3 Results

3.1 Empirical evidence of signal-dependent noise in BOLD signals

Figure 2B shows the variance of the BOLD signal at time t across the 9 trails against the squared mean magnitude of the signal at time $t-1$. Significant correlations are observed for both experimental conditions by pooling data from the four brain regions, (attention to intensity, $C = 0.60$, $p < 10^{-6}$, attention to pleasant, $C = 0.44$, $p = 0.0002$), which clearly indicates the presence of signal-dependent noise in BOLD, and, in particular, the variance of the BOLD signal is approximately linearly related to the squared mean magnitude of the signal. Similar effect was also found separately for each brain region.

3.2 Simulation results

Figure 3 shows the comparison of performance for the classical Granger causal model and the Granger causal model with signal-dependent noise by ROC (receiver operating characteristic) analysis. Clearly, classical GC cannot capture the causal influences with the

presence of signal-dependent noise, while the signal-dependent noise model substantially outperforms the classical GC, and shows a good sensitivity and specificity.

3.3 fMRI data investigation

Table 1 shows the causal influences between the four brain areas (OFC, AntINS, AntLPFC, PostLPFC) detected by the signal-dependent noise model. First, we consider attention to intensity. There are significant (top-down) causal influence from both the AntLPFC and PostLPFC to the insular taste cortex (AntINS). Second, we consider attention to pleasantness. There are significant (top-down) causal influence from both the AntLPFC and PostLPFC to the OFC, and a significant effect from the OFC to the antLPFC. There is also a (top-down) effect of the PostLPFC on the taste insula (AntINS). Very interestingly too, during attention to pleasantness, there is increased effective connectivity from the insular taste cortex to the OFC, suggesting that information is routed especially to the OFC during attention to pleasantness.

As a comparison, Table 2 shows the causal influences between the four brain areas detected by the classical Granger causal model. Only one connectivity (PostLPFC to AntLPFC, when paying attention to intensity) is identified as significant. The greater power of the signal-dependent noise model can be clearly observed.

Table 3 shows the difference of the causalities in opposite directions by the signal-dependent noise model. In the pleasantness condition, consistent with the hypothesis that the lateral prefrontal cortex is the source of the top-down modulation of activations in the OFC, there are significantly stronger effects from both the AntLPFC and the PostLPFC to

the OFC than vice versa. It is also of interest that in the pleasantness condition, a significantly stronger forward influence is detected from the antINS to the OFC. Only one significant difference is detected for the intensity condition, that is the effect from the PostLPFC to AntINS is greater than in the reverse direction. This is consistent with the hypothesis that the major top-down effect on the taste insula during attention to intensity is from the PostLPFC. The bi-directional interaction in the pleasantness condition between AntLPFC and OFC (Table 1) may be interpreted in the context that there is a significant difference of the causality with AntLPFC to OFC greater than OFC to AntLPFC, thus indicating a stronger influence of AntLPFC on OFC than vice versa (Table 3).

These analyses provide evidence for the effective connectivities in the intensity and pleasantness conditions that are summarized in Figure 4.

4 Discussion

In this paper, we for the first time provided empirical evidence of signal-dependent noise in fMRI BOLD signals in several cortical areas, such as the insular, orbitofrontal, and lateral prefrontal cortex. We then developed a Granger causal model with signal-dependent noise that can appropriately model BOLD signals and detect causal influences in both mean and variance. By simulation studies, we showed that our model substantially outperforms classical Granger causal model, when signal-dependent noise is present in the time series. We applied our model to the data from an fMRI study to investigate the source of top-down attentional control of taste intensity and taste pleasantness processing. We found a top-down effect from the PostLPFC to the insular taste cortex during attention to intensity but not to

pleasantness, and a top-down effect from the AntLPFC and PostLPFC to OFC during attention to pleasantness but not to intensity. In addition, there was stronger forward effective connectivity from the insular taste cortex to the OFC during attention to pleasantness than during attention to intensity.

4.1 A theoretical explanation of signal-dependent noise in BOLD

In addition to the empirical analysis of BOLD signals in the main text, here we provide a theoretical explanation of the presence of signal-dependent noise in BOLD.

Assume that the activity of each single neuron is governed by a leaky integrate-and-fire (LIF) model [Feng, 2003]. Then given two quantities $V_{\text{thre}} > V_{\text{rest}}$ and when $V_i < V_{\text{thre}}$, the membrane potential of the i-th neuron V_i satisfies the following dynamics

$$\mathrm{d}V_i(t) = -\frac{V_i(t) - V_{\text{rest}}}{\tau}\mathrm{d}t + \mathrm{d}I_{\text{syn},i}, \quad t > 0, \tag{26}$$

where τ is the decay time constant, and the synaptic input $I_{\text{syn},i}$, which is a Poisson point process, is given by the diffusion approximation

$$\mathrm{d}I_{\text{syn},i} = \mu_i \mathrm{d}t + \sigma_i \mathrm{d}B_i(t), \tag{27}$$

with constants $\mu_i > 0$, $\sigma_i > 0$ and the standard Brownian motion $B_i(t)$ [Tuckwell, 1988]. Once V_i is greater than V_{thre}, it is reset to the resting potential V_{rest}. More specifically, we can define

$$\mu_i = a_i(1 - r_i)\lambda_i, \quad \sigma_i^2 = a_i^2(1 + r_i)\lambda_i, \tag{28}$$

where a_i is the magnitude of postsynaptic inputs, λ_i is the input rate, r_i is the ratio between the inhibitory and excitatory inputs. In particular, when $r_i = 0$ the neuron receives

exclusively excitatory inputs; when $r_i = 1$ the inhibitory and excitatory input is balanced. We further assume that the local field potential (LFP) can be modeled by averaging the membrane potentials of neurons in a small neighborhood:

$$\mathrm{d}V(t) = \frac{1}{M}\sum_{i=1}^{M}\mathrm{d}V_i(t), \tag{29}$$

where M is the number of neurons. When M is sufficiently large, it follows from the law of large numbers that

$$\mathrm{d}V(t) = \mu\mathrm{d}t + \sigma\mathrm{d}B(t), \tag{30}$$

where $\mu = \mathbf{E}\left[-(V_i - V_{\mathrm{rest}})/\tau\right] + \mathbf{E}\left[a_i(1-r_i)\lambda_i\right]$, $\sigma = \mathbf{E}\left[\sqrt{a_i^2(1+r_i)\lambda_i}\right]$. Clearly, the variance of the noise σ^2 is coupled to the mean μ. LFP has been demonstrated to be a good predictor of the BOLD signal in many experimental investigations [Goense and Logothetis, 2008; Logothetis et al., 2001; Nir et al., 2007], and it has been widely accepted that the BOLD signal can be generated by convolving the LFP with the haemodynamic response function [Deshpande et al., 2010a; Schippers and Keysers, 2011]. By convolving both sides of Eq. (30) with the haemodynamic response function, it is then clear that BOLD signals also show signal-dependent noise phenomena. Therefore, we envisage that signal-dependent noise is ubiquitous in BOLD signals. Future studies are needed to provide more evidence of the signal-dependent noise in different brain areas and different data sets.

4.2 Methodology assessment

Conditionally heteroskedastic data often show volatility clustering and outliers. In particular, the unconditional distribution of the data is leptokurtic, which means it has more mass around zero and in the tails than the normal distribution and, hence, it can produce

occasional outliers [Lütkepohl, 2005]. Therefore, models with time varying volatility can better capture the nature of the data and it is expected that more reliable causal inferences can be made. Comparing to the earlier approaches of causal inferences in data with time varying volatility, including Luo et al. [2011], the model presented in this paper provides an accurate, efficient and unified method to detect causality in both the mean and variance. The model has a corresponding frequency domain representation [Luo et al., 2011], which may further shed light on frequency specific interactions.

Although we only applied our model to fMRI time series, it is clear that the model can be applied to very many types of data that might exhibit signal-dependent noise, including neurophysiological data such as single or multi-neuronal recordings, magnetoencephalography, local field potentials, and beyond neuroscience also to any possibly causal system where there are time series of data from two or many sources. Indeed, the significance of detecting causality from data with time varying volatility might be partly demonstrated in the 2003 Nobel Prize in Economics shared by Granger, who set up the foundation of Granger causal analysis [Granger, 1969], and Engle, who invented the first changing volatility model [Engle, 1982].

Although our initial implementation of the signal-dependent noise model appears to be successful, due to the highly nonlinear form of the log-likelihood function and optimization problem, fast and robust optimization algorithms deserve future investigation. Also, although a low-order low-dimensional AR-BEKK model is a relatively parsimonious representation of the conditional covariance structure of a process, number of parameters still grows quickly with the dimension of the underlying system. This impedes the application of the model to a modest number of time series. Future studies are needed to find more

restricted models which ensure uniqueness of the parameterization, guarantee the positive definiteness of the conditional covariance, while at the same time still produces rich dynamics.

In spite of the wide and successful applications in neurophysiological data, there is still an ongoing debate on applying GC to fMRI data [David, 2011; David et al., 2008; Friston, 2009, 2011; Roebroeck et al., 2011a,b; Valdes-Sosa et al., 2011]. Inferring causality from fMRI time series – an indirect measure of neuronal activities – imposes many more challenges than those direct electrophysiological recordings. Granger causal model uses the observed fMRI data as a surrogate for the underlying neuronal activity, which is a potential flaw of the method and the main controversy against the application of GC to fMRI data, since BOLD signal is a blurred and delayed representation of the original neuronal signal, and it is now widely recognized that there are intra- and inter-subject variability of haemodynamic responses [Aguirre et al., 1998; Handwerker et al., 2004; Kruggel and Von Cramon, 1999; Rajapakse et al., 1998]. However, there have been a series of numerical and theoretical works showing that GC is quite robust to the difference in haemodynamic delays [Barnett and Seth, 2011; Deshpande et al., 2010b; Schippers et al., 2011]. Moreover, as in Wen et al. [2012], we calculated the cross-correlation function for each pair of time series used in our Granger causal analysis, and most of the cross-correlation peaks appeared at zero lag, indicating that differences in the regional haemodynamic responses may not be a significant factor in this study. We therefore feel that the application of Granger type causal inferences in the analysis of this particular fMRI data set is justified. However, given the complexity of the brain, much work remains to do to provide reliable and accurate causal analysis for neuroscience.

4.3 Biological interpretations

The interpretation of the effective connectivity revealed with our signal-dependent noise model is that during attention to pleasantness, the AntLPFC and PostLPFC regions identified by PPI analysis exert a top-down control of the responsiveness of the OFC to its taste-related inputs, and indeed to how strongly information is routed to the OFC from its preceding area, the AntINS taste cortex. In contrast, during attention to intensity, the PostLPFC identified by PPI analysis exerts a top-down control of the responsiveness of the insular taste cortex to its taste-related inputs. This interpretation is strengthened by the findings with our previous analysis [Ge et al., 2012], which provides evidence that the top-down effects depend on the level of activity in the areas on which there is a top-down effect.

The way that we think of top-down biased competition as operating normally in, for example, visual selective attention [Desimone and Duncan, 1995] is that within an area, e.g. a cortical region, some neurons receive a weak top-down input that increases their response to the bottom-up stimuli [Desimone and Duncan, 1995], potentially supra-linearly if the bottom-up stimuli are weak [Deco and Rolls, 2005a; Rolls, 2008; Rolls and Deco, 2002]. The enhanced firing of the biased neurons then, via the local inhibitory neurons, inhibits the other neurons in the local area from responding to the bottom-up stimuli. This is a local mechanism, in that the inhibition in the neocortex is primarily local, being implemented by cortical inhibitory neurons that typically have inputs and outputs over no more than a few mm [Rolls, 2008; Rolls and Deco, 2002; Shepherd, 2003]. This model of biased competition is illustrated in [Grabenhorst and Rolls, 2010]. That locally implemented

biased competition situation may not apply in the present case, where we have facilitation of processing in a whole cortical area (e.g. orbitofrontal cortex) or even cortical processing stream (e.g. the linked orbitofrontal and pregenual cingulate cortex [Grabenhorst and Rolls, 2010]) in which any taste neurons may reflect pleasantness and not intensity. So the attentional effect might more accurately be described in the present case as biased activation, without local competition being part of the effect. This biased activation theory and model of attention, illustrated in Figure 5, is a rather different way to implement attention in the brain than biased competition, and each mechanism may apply in different cases, or both mechanisms in some cases [Grabenhorst and Rolls, 2010, 2011].

Acknowledgments

JF is a Royal Society Wolfson Research Merit Award holder, partially supported by National Centre for Mathematics and Interdisciplinary Sciences (NCMIS) of the Chinese Academy of Sciences and Key Program of National Natural Science Foundation of China (No. 91230201). QL is partly supported by grants from the National Natural Sciences Foundation of China (No. 11101429, No. 11271121, No. 71171195), Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20114307120019), and National Basic Research Program of China (No. 2011CB707802).

References

G.K. Aguirre, E. Zarahn, and M. D'esposito. The variability of human, BOLD hemodynamic responses. *NeuroImage*, 8(4):360–369, 1998.

Y. Baba, R.F. Engle, D.F. Kraft, and K.F. Kroner. Multivariate simultaneous generalized ARCH. *Discussion paper 89-57, University of California, San Diego, Department of Economics*, 1991.

L. Barnett and T. Bossomaier. Transfer entropy as a log-likelihood ratio. *Physical Review Letters*, 109:138105, Sep 2012.

L. Barnett and A.K. Seth. Behaviour of Granger causality under filtering: Theoretical invariance and practical application. *Journal of Neuroscience Methods*, 201(2):404–419, 2011.

D.M. Beck and S. Kastner. Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research*, 49(10):1154–1165, 2009.

S.L. Bengtsson, J.D. Haynes, K. Sakai, M.J. Buckley, and R.E. Passingham. The representation of abstract task rules in the human prefrontal cortex. *Cerebral Cortex*, 19(8): 1929–1936, 2009.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

T. Bollerslev, R.F. Engle, and J.M. Wooldridge. A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, pages 116–131, 1988.

D.L. Collins, P. Neelin, T.M. Peters, and A.C. Evans. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18(2):192, 1994.

O. David. fMRI connectivity, meaning and empiricism. Comments on: Roebroeck et al. The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage*, 58(2), 2011.

O. David, I. Guillemain, S. Saillet, S. Reyt, C. Deransart, et al. Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS Biology*, 6(12):e315, 2008.

I.E.T. de Araujo, M.L. Kringelbach, E.T. Rolls, and P. Hobden. Representation of umami taste in the human brain. *Journal of Neurophysiology*, 90(1):313–319, 2003a.

I.E.T. de Araujo, M.L. Kringelbach, E.T. Rolls, and F. McGlone. Human cortical responses to water in the mouth, and the effects of thirst. *Journal of Neurophysiology*, 90(3):1865–1876, 2003b.

I.E.T. de Araujo, E.T. Rolls, M.L. Kringelbach, F. McGlone, and N. Phillips. Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *European Journal of Neuroscience*, 18(7):2059–2068, 2003c.

G. Deco and E.T. Rolls. Attention, short-term memory, and action selection: a unifying theory. *Progress in Neurobiology*, 76(4):236–256, 2005a.

G. Deco and E.T. Rolls. Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons. *Journal of Neurophysiology*, 94(1):295–313, 2005b.

G. Deshpande, X. Hu, S. Lacey, R. Stilla, and K. Sathian. Object familiarity modulates effective connectivity during haptic shape perception. *NeuroImage*, 49(3):1991, 2010a.

G. Deshpande, K. Sathian, and X. Hu. Effect of hemodynamic variability on Granger causality analysis of fMRI. *NeuroImage*, 52(3):884–896, 2010b.

R. Desimone and J. Duncan. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 18(1):193–222, 1995.

F.X. Diebold and M. Nerlove. The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. *Journal of Applied Econometrics*, 4(1):1–21, 1989.

M. Ding, Y. Chen, and S.L. Bressler. Granger causality: Basic theory and application to neuroscience. In B. Schelter, M. Winterhalder, and J. Timmer, editors, *Handbook of Time Series Analysis*. Wiley-VCH, Weinheim, 2006.

R.F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

R.F. Engle and K.F. Kroner. Multivariate simultaneous generalized ARCH. *Econometric theory*, 11(01):122–150, 1995.

R.F. Engle, C.W.J. Granger, and D. Kraft. Combining competing forecasts of inflation using a bivariate ARCH model. *Journal of Economic Dynamics and Control*, 8(2):151–165, 1984.

T. Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.

J. Feng. *Computational neuroscience: a comprehensive approach*. Chapman & Hall/CRC, 2003.

K. Friston, R. Moran, and A.K. Seth. Analysing connectivity with Granger causality and dynamic causal modelling. *Current Opinion in Neurobiology*, 23:1–7, 2012.

K.J. Friston. Causal modelling and brain connectivity in functional magnetic resonance imaging. *PLoS Biology*, 7(2):e1000033, 2009.

K.J. Friston. Dynamic causal modeling and granger causality. Comments on: the identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage*, 58(2-2):303, 2011.

K.J. Friston, C. Buechel, G.R. Fink, J. Morris, E.T. Rolls, et al. Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, 6(3):218–229, 1997.

K.J. Friston, D.E. Glaser, R.N.A. Henson, S. Kiebel, C. Phillips, et al. Classical and Bayesian inference in neuroimaging: applications. *NeuroImage*, 16(2):484–512, 2002.

T. Ge, K.M. Kendrick, and J. Feng. A novel extended Granger causal model approach demonstrates brain hemispheric differences during face recognition learning. *PLoS Computational Biology*, 5(11):e1000570, 2009.

T. Ge, J. Feng, F. Grabenhorst, and E.T. Rolls. Componential granger causality, and its application to identifying the source and mechanisms of the top–down biased activation

that controls attention to affective vs sensory processing. *NeuroImage*, 59(2):1846–1858, 2012.

G.L. Gerstein and B. Mandelbrot. Random walk models for the spike activity of a single neuron. *Biophysical Journal*, 4(1):41–68, 1964.

J.F. Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313, 1982.

J.F. Geweke. Measures of conditional linear dependence and feedback between time series. *Journal of the American Statistical Association*, 79(388):907–915, 1984.

D.R. Gitelman, W.D. Penny, J. Ashburner, and K.J. Friston. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *NeuroImage*, 19(1):200–207, 2003.

R. Goebel, A. Roebroeck, D.S. Kim, and E. Formisano. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magnetic Resonance Imaging*, 21(10):1251–1261, 2003.

J.B.M. Goense and N.K. Logothetis. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biology*, 18(9):631–640, 2008.

J.A. Gottfried, J.S. Winston, and R.J. Dolan. Dissociable codes of odor quality and odorant structure in human piriform cortex. *Neuron*, 49(3):467–479, 2006.

F. Grabenhorst and E.T. Rolls. Selective attention to affective value alters how the brain processes taste stimuli. *European Journal of Neuroscience*, 27(3):723–729, 2008.

F. Grabenhorst and E.T. Rolls. Attentional modulation of affective versus sensory processing: functional connectivity and a top-down biased activation theory of selective attention. *Journal of Neurophysiology*, 104(3):1649–1660, 2010.

F. Grabenhorst and E.T. Rolls. Value, pleasure and choice in the ventral prefrontal cortex. *Trends in Cognitive Sciences*, 15(2):56–67, 2011.

F. Grabenhorst, E.T. Rolls, and A. Bilderbeck. How cognition modulates affective responses to taste and flavor: top-down influences on the orbitofrontal and pregenual cingulate cortices. *Cerebral Cortex*, 18(7):1549–1559, 2008.

F. Grabenhorst, E.T. Rolls, B.A. Parris, and A.A. d'Souza. How the brain represents the reward value of fat in the mouth. *Cerebral Cortex*, 20(5):1082–1091, 2010.

C.W.J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969.

L. Haase, B. Cerf-Ducastel, and C. Murphy. Cortical activation in response to pure taste stimuli during the physiological states of hunger and satiety. *NeuroImage*, 44(3):1008–1021, 2009.

D.A. Handwerker, J.M. Ollinger, and M. D'Esposito. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *NeuroImage*, 21(4):1639–1651, 2004.

C.M. Harris and D.M. Wolpert. Signal-dependent noise determines motor planning. *Nature*, 394(6695):780–784, 1998.

K. Hwang, K. Velanova, and B. Luna. Strengthening of top-down frontal cognitive control networks underlying the development of inhibitory control: a functional magnetic resonance imaging effective connectivity study. *The Journal of Neuroscience*, 30(46): 15535–15545, 2010.

D.L. Knepp and D.R. Entwisle. Testing significance of differences between two chi-squares. *Psychometrika*, 34(3):331–333, 1969.

F. Kouneiher, S. Charron, and E. Koechlin. Motivation and cognitive control in the human prefrontal cortex. *Nature Neuroscience*, 12(7):939–945, 2009.

M.L. Kringelbach, J. O'Doherty, E.T. Rolls, and C. Andrews. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cerebral Cortex*, 13(10):1064–1071, 2003.

F. Kruggel and D.Y. Von Cramon. Temporal properties of the hemodynamic response in functional MRI. *Human Brain Mapping*, 8(4):259–271, 1999.

N.K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann. Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412(6843):150–157, 2001.

Q. Luo, T. Ge, and J. Feng. Granger causality with signal-dependent noise. *NeuroImage*, 57(4):1422–1429, 2011.

H. Lütkepohl. *New introduction to multiple time series analysis*. Cambridge Univ Press, 2005.

C.J. McAdams and J.H.R. Maunsell. Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron*, 23:765–773, 1999.

C. McCabe and E.T. Rolls. Umami: a delicious flavor formed by convergence of taste and olfactory pathways in the human brain. *European Journal of Neuroscience*, 25(6): 1855–1864, 2007.

Y. Nir, L. Fisch, R. Mukamel, H. Gelbard-Sagiv, A. Arieli, et al. Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Current Biology*, 17(15):1275–1285, 2007.

J.B. Nitschke, G.E. Dixon, I. Sarinopoulos, S.J. Short, J.D. Cohen, et al. Altering expectancy dampens neural response to aversive taste in primary taste cortex. *Nature Neuroscience*, 9(3):435–442, 2006.

J. O'doherty, E.T. Rolls, S. Francis, R. Bowtell, and F. McGlone. Representation of pleasant and aversive taste in the human brain. *Journal of Neurophysiology*, 85(3):1315–1321, 2001.

J.C. Rajapakse, F. Kruggel, J.M. Maisog, and D.Y. Von Cramon. Modeling hemodynamic response for analysis of functional MRI time-series. *Human Brain Mapping*, 6(4):283–300, 1998.

A. Roebroeck, E. Formisano, R. Goebel, et al. Mapping directed influence over the brain using Granger causality and fMRI. *NeuroImage*, 25(1):230–242, 2005.

A. Roebroeck, E. Formisano, and R. Goebel. The identification of interacting networks in

the brain using fMRI: model selection, causality and deconvolution. *NeuroImage*, 58(2): 296–302, 2011a.

A. Roebroeck, E. Formisano, and R. Goebel. Reply to Friston and David: After comments on: The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution. *NeuroImage*, 58(2):310–311, 2011b.

E.T. Rolls. *Memory, attention and decision-making*. Oxford University Press New York, 2008.

E.T. Rolls and G. Deco. *Computational neuroscience of vision*. Oxford university press, 2002.

E.T. Rolls and F. Grabenhorst. The orbitofrontal cortex and beyond: from affect to decision-making. *Progress in Neurobiology*, 86(3):216, 2008.

E.T. Rolls, M.L. Kringelbach, and I.E.T. de Araujo. Different representations of pleasant and unpleasant odours in the human brain. *European Journal of Neuroscience*, 18(3): 695–703, 2003.

E.T. Rolls, F. Grabenhorst, C. Margot, M. da Silva, and M.I. Velazco. Selective attention to affective value alters how the brain processes olfactory stimuli. *Journal of Cognitive Neuroscience*, 20(10):1815–1826, 2008.

A.F. Rossi, L. Pessoa, R. Desimone, and L.G. Ungerleider. The prefrontal cortex and the executive control of attention. *Experimental Brain Research*, 192(3):489–497, 2009.

K. Sakai and R.E. Passingham. Prefrontal interactions reflect future task operations. *Nature Neuroscience*, 6(1):75–81, 2002.

K. Sakai and R.E. Passingham. Prefrontal set activity predicts rule-specific neural processing during subsequent cognitive performance. *The Journal of Neuroscience*, 26(4): 1211–1218, 2006.

M.B. Schippers and C. Keysers. Mapping the flow of information within the putative mirror neuron system during gesture observation. *NeuroImage*, 57(1):37–44, 2011.

M.B. Schippers, R. Renken, and C. Keysers. The effect of intra-and inter-subject variability of hemodynamic responses on group level Granger causality analyses. *NeuroImage*, 57 (1):22–36, 2011.

M.A. Schoenfeld, G. Neuer, C. Tempelmann, K. Schüßler, T. Noesselt, et al. Functional magnetic resonance tomography correlates of taste perception in the human primary taste cortex. *Neuroscience*, 127(2):347–353, 2004.

L.P.J. Selen, D.W. Franklin, and D.M. Wolpert. Impedance control reduces instability that arises from motor noise. *The Journal of Neuroscience*, 29(40):12606–12616, 2009.

G.M. Shepherd. *The synaptic organization of the brain*. Oxford University Press, USA, 2003.

D.M. Small, M.D. Gregory, Y.E. Mak, D. Gitelman, M.M. Mesulam, et al. Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*, 39(4):701, 2003.

S. Taylor. *Modelling financial time series*, volume 113. World Scientific, 1986.

E. Todorov. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Computation*, 17(5):1084–1108, 2005.

H.C. Tuckwell. *Introduction to theoretical neurobiology. Volume 2: Nonlinear and stochastic theories*. Cambridge University Press, 1988.

P.A. Valdes-Sosa, A. Roebroeck, J. Daunizeau, and K.J. Friston. Effective connectivity: Influence, causality and biophysical modeling. *NeuroImage*, 58(2):339–361, 2011.

M.G. Veldhuizen, G. Bender, R.T. Constable, and D.M. Small. Trying to detect taste in a tasteless solution: modulation of early gustatory cortex by attention to taste. *Chemical Senses*, 32(6):569–581, 2007.

X. Wen, L. Yao, Y. Liu, and M. Ding. Causal interactions in attention networks predict behavioral performance. *The Journal of Neuroscience*, 32(4):1284–1292, 2012.

N. Wiener. The theory of prediction. *Modern mathematics for engineers. New York: McGraw-Hill*, pages 165–190, 1956.

J.L. Wilson, M. Jenkinson, I. de Araujo, M.L. Kringelbach, E.T. Rolls, et al. Fast, fully automated global and local magnetic field optimization for fMRI of the human brain. *NeuroImage*, 17(2):967–976, 2002.

J. Yacubian, J. Gläscher, K. Schroeder, T. Sommer, D.F. Braus, et al. Dissociable systems for gain-and loss-related value predictions and errors of prediction in the human brain. *The Journal of Neuroscience*, 26(37):9530–9537, 2006.

Tables

Intensity				
	OFC	**AntINS**	**AntLPFC**	**PostLPFC**
OFC	–	4.51 (0.10)	0.96 (0.62)	0.68 (0.71)
AntINS	2.15 (0.34)	–	**28.57** ($< 10^{-4}$)	13.59 (1.1×10^{-3})
AntLPFC	0.89 (0.64)	**28.44** ($< 10^{-4}$)	–	5.17 (0.08)
PostLPFC	0.62 (0.73)	**29.62** ($< 10^{-4}$)	4.22 (0.12)	–
Pleasantness				
	OFC	**AntINS**	**AntLPFC**	**PostLPFC**
OFC	–	16.92 (2×10^{-4})	**44.94** ($< 10^{-4}$)	7.12 (0.03)
AntINS	**94.79** ($< 10^{-4}$)	–	2.09 (0.35)	17.56 (2×10^{-4})
AntLPFC	**95.08** ($< 10^{-4}$)	6.72 (0.035)	–	6.93 (0.03)[2]
PostLPFC	**89.16** ($< 10^{-4}$)	**22.81** ($< 10^{-4}$)	5.70 (0.06)	–

Table 1: **The causality results by Granger causal model with signal-dependent noise**. Causal influence is from row to column. The causality is given for each direction and the corresponding p-value is presented in brackets. If the uncorrected p-value is less than 10^{-4} (survived Bonferroni correction), the causal influence is identified as significant and indicated in bold in the table.

[2]The interior-point algorithm failed to converge to a sensible solution. We therefore used the active-set algorithm for this particular link.

Intensity				
	OFC	**AntINS**	**AntLPFC**	**PostLPFC**
OFC	–	0.0039 (0.0061)	0.0006 (0.30)	0.0001 (0.62)
AntINS	0.0015 (0.09)	–	0.0002 (0.56)	0.0003 (0.42)
AntLPFC	0.0000 (0.99)	0.0056 (0.0009)	–	0.0043 (0.0041)
PostLPFC	0.0003 (0.43)	**0.0122 ($< 10^{-4}$)**	0.0013 (0.12)	–
Pleasantness				
	OFC	**AntINS**	**AntLPFC**	**PostLPFC**
OFC	–	0.0000 (0.99)	0.0001 (0.71)	0.0000 (0.98)
AntINS	0.0001 (0.74)	–	0.0023 (0.03)	0.0018 (0.06)
AntLPFC	0.0000 (0.99)	0.0002 (0.51)	–	0.0059 (0.0007)
PostLPFC	0.0008 (0.22)	0.0034 (0.0099)	0.0002 (0.59)	–

Table 2: **The causality results by classical Granger causal model**. Causal influence is from row to column. The causality is given for each direction and the corresponding p-value is presented in brackets. If the uncorrected p-value is less than 10^{-4} (survived Bonferroni correction), the causal influence is identified as significant and indicated in bold in the table.

Direction	Intensity	Pleasantness
PostLPFC $\rightarrow$ AntLPFC $-$ AntLPFC $\rightarrow$ PostLPFC	-0.48	-0.614[3]
PostLPFC $\rightarrow$ AntINS $-$ AntINS $\rightarrow$ PostLPFC	**8.02**	2.62
PostLPFC $\rightarrow$ OFC $-$ OFC $\rightarrow$ PostLPFC	-0.03	**41.03**
AntLPFC $\rightarrow$ AntINS $-$ AntINS $\rightarrow$ AntLPFC	-0.06	2.32
AntLPFC $\rightarrow$ OFC $-$ OFC $\rightarrow$ AntLPFC	-0.03	**25.07**
AntINS $\rightarrow$ OFC $-$ OFC $\rightarrow$ AntINS	-1.18	**38.93**

Table 3: **Difference of the causalities in opposite directions by the Granger causal model with signal-dependent noise**. Difference with a p-value smaller than 0.01 (the difference measure greater than 4.61) is indicated in bold.

[3]We used the active-set algorithm for this particular link.

Figure Legends

Figure 1: **Results of the PPI analysis**. A. The seed areas for the PPI analysis in the orbitofrontal cortex (1) [-6, 14, -20], and insular taste cortex (2) [42, 18, -14]. B. The region of the anterior lateral prefrontal cortex (AntLPFC) [-40, 54, 14] identified by PPI analysis as correlated with the orbitofrontal cortex seed area when attention was to pleasantness ($p < 0.029$). C. The region of the posterior lateral prefrontal cortex [-38, 34, 14] identified by PPI analysis as correlated with the insular taste cortex seed area when attention was to intensity ($p < 0.049$). The full details of the PPI analysis are provided in [Grabenhorst and Rolls, 2010].

Figure 2: **Empirical evidence of signal-dependent noise in BOLD signals**. A. An illustration of signal-dependent noise. For each single subject, the activation profiles are the same for different trials under the same experimental condition (black), while the observed intra-subject inter-trail variation (red) is due to noise. Large variation is seen when the mean amplitude of the signal is large (e.g. at the peak). B. For each of the four brain areas and each subject, the variance of the BOLD signal at time t across the 9 trails is plotted against the squared mean magnitude of the signal at time $t - 1$. Significant correlations are observed for both experimental conditions (attention to intensity, $C = 0.60$, $p < 10^{-6}$, attention to pleasant, $C = 0.44$, $p = 0.0002$).

Figure 3: **Comparison of performance for the classical Granger causal model and the Granger causal model with signal-dependent noise by ROC (receiver operating characteristic) analysis**. The sensitivity of the methods is plotted against 1-specificity for different p-value thresholds. The sensitivity is defined as the proportion of actual causal influences that are correctly identified. The specificity measures the proportion of non-causal influences that are correctly identified. By setting different p-value thresholds for causality, each method gives different sensitivity and specificity. Therefore, the best model is expected to have its performance ROC curve go through the upper left corner, while a random classification algorithm has its performance curve as a diagonal line. The signal-dependent noise model outperforms the classical Granger causal model substantially and consistently.

Figure 4: **Neural circuits revealed by Granger causality with signal-dependent noise**. A. Attention to taste intensity. B. Attention to taste pleasantness. Larger arrows represent a stronger influence. The values of significant likelihood ratio test statistics are indicated.

Figure 5: **A Biased activation theory of selective attention**. The short-term memory systems that provide the source of the top-down activations may be separate (as shown), or could be a single network with different attractor states for the different selective attention conditions. The top-down short-term memory systems hold what is being paid attention to active by continuing firing in an attractor state, and bias separately either cortical processing system 1, or cortical processing system 2. This weak top-down bias interacts with the bottom up input to the cortical stream and produces an increase of activity that can be supralinear [Deco and Rolls, 2005b]. Thus the selective activation of separate cortical processing streams can occur. In the example, stream 1 might process the affective value of a stimulus with the areas involved including the anterior lateral prefrontal cortex with a top-down influence on the orbitofrontal cortex, and stream 2 might process the intensity and physical properties of the stimulus with the areas involved including the posterior lateral prefrontal cortex with a top-down influence on the insular taste cortex. The outputs of these separate processing streams then must enter a competition system, which could be for example a cortical attractor decision-making network that makes choices between the two streams, with the choice biased by the activations in the separate streams [Grabenhorst and Rolls, 2011].