

# On the distinguishability of HRF models in fMRI

Paulo N. Rosa<sup>1\*</sup>, Patricia Figueiredo<sup>2</sup> and Carlos J. Silvestre<sup>3</sup>

<sup>1</sup> Flight Systems Business Unit, Aerospace, Defense & Systems Department, Deimos Engenharia, Lda., Lisboa, Portugal, <sup>2</sup> Institute for Systems and Robotics and Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Portugal, <sup>3</sup> Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Taipa, Macau, China

Modeling the Hemodynamic Response Function (HRF) is a critical step in fMRI studies of brain activity, and it is often desirable to estimate HRF parameters with physiological interpretability. A biophysically informed model of the HRF can be described by a non-linear time-invariant dynamic system. However, the identification of this dynamic system may leave much uncertainty on the exact values of the parameters. Moreover, the high noise levels in the data may hinder the model estimation task. In this context, the estimation of the HRF may be seen as a problem of model falsification or invalidation, where we are interested in distinguishing among a set of eligible models of dynamic systems. Here, we propose a systematic tool to determine the distinguishability among a set of physiologically plausible HRF models. The concept of absolutely input-distinguishable systems is introduced and applied to a biophysically informed HRF model, by exploiting the structure of the underlying non-linear dynamic system. A strategy to model uncertainty in the input time-delay and magnitude is developed and its impact on the distinguishability of two physiologically plausible HRF models is assessed, in terms of the maximum noise amplitude above which it is not possible to guarantee the falsification of one model in relation to another. Finally, a methodology is proposed for the choice of the input sequence, or experimental paradigm, that maximizes the distinguishability of the HRF models under investigation. The proposed approach may be used to evaluate the performance of HRF model estimation techniques from fMRI data.

**Keywords:** HRF, fMRI, BOLD fMRI, distinguishability, model selection, experimental paradigm

## Introduction

The hemodynamic response function (HRF) describes the local changes in cerebral blood flow, volume, and oxygenation associated with neuronal activity, and it is extensively used to model Blood Oxygen Level Dependent (BOLD) signals measured using functional Magnetic Resonance Imaging (fMRI) (Logothetis and Wandell, 2004). In general, fMRI experiments are used to map networks of brain activity that are associated with a specific stimulus or task, or that are functionally correlated during rest. Mapping of stimulus/task-related BOLD changes is most frequently achieved by fitting a general linear model (GLM) to the data, consisting on the stimulus/task time course convolved with a pre-specified HRF model (Friston et al., 1994), assuming a linear time invariant system (Boynton et al., 1996). Although the exact mechanisms underlying the HRF are not yet completely known, the consistency

## OPEN ACCESS

### Edited by:

Tobias Alecio Mattei,  
Brain & Spine Center - InvisionHealth -  
Kenmore Mercy Hospital, USA

### Reviewed by:

Nelson Jesús Trujillo-Barreto,  
University of Manchester, UK  
Daniele Marinazzo,  
University of Ghent, Belgium

### \*Correspondence:

Paulo N. Rosa,  
Flight Systems Business Unit,  
Aerospace, Defense & Systems  
Department, Deimos Engenharia, Av.  
D. Joao II, Lt 1.17.01, 10th floor,  
1998-023 Lisboa, Portugal  
paulo.rosa@deimos.com.pt

**Received:** 22 July 2014

**Accepted:** 24 April 2015

**Published:** 19 May 2015

### Citation:

Rosa PN, Figueiredo P and Silvestre  
CJ (2015) On the distinguishability of  
HRF models in fMRI.  
Front. Comput. Neurosci. 9:54.  
doi: 10.3389/fncom.2015.00054

of its observed shape allowed for canonical (parameterized) HRF models to be derived (Friston et al., 1998). In particular, double-gamma HRF models are commonly employed in fMRI analysis. Nevertheless, extensive HRF variability has been reported across brain regions (Handwerker et al., 2004), scanning sessions (Aguirre et al., 1998), tasks (Cohen and Ugurbil, 2002), physiological modulations (Liu et al., 2004), subjects (Handwerker et al., 2004), and populations (D'Esposito et al., 2003), which may hinder or confound the measurement of BOLD changes associated with brain activity, limiting the interpretability of fMRI studies.

Common approaches attempting to take into account HRF variability allow for greater flexibility in the HRF shape and dynamics by describing it through a set of basis functions in a GLM framework. They include using the partial derivatives with respect to time and dispersion of a canonical HRF (Friston et al., 1998), finite impulse response (FIR) basis sets (Glover, 1999), and specially designed basis functions (Woolrich et al., 2004). An approach that also takes into account the spatial localization of the HRF was very recently proposed in Vincent et al. (2014). While a small number of basis functions cannot accurately model the whole range of HRF shapes and delays, at the other extreme, deconvolution of the BOLD response is a very noisy process. Critically, these approaches do not provide a biophysical foundation for the HRF model, hence limiting the physiological interpretability of the associated parameters. Moreover, they do not explain empirically observed non-linearities in the BOLD responses (Birn et al., 2001).

Biophysically informed non-linear models of the HRF have been proposed, based on the combination of the Balloon model, describing the dynamic changes in deoxyhemoglobin content as a function of blood oxygenation and blood volume (Buxton et al., 1998), with a model of the blood flow dynamics during brain activation, where neuronal activity is approximated by the stimulus/task input scaled by a factor called neural efficiency (Friston et al., 2000). In the original work that proposed this model, the associated parameters were estimated by using a Volterra kernel expansion to characterize the system dynamics (Friston et al., 2000). Later, a Bayesian estimation framework was introduced, allowing for the use of a priori distributions of the parameter values and the production of the respective posterior probability distributions given the data by using Expectation-Maximization methods (Friston, 2002). This HRF model and respective estimation procedure have further been incorporated in Dynamic Causal Models (DCM) developed to study effective connectivity among networks of brain regions from fMRI data (Friston et al., 2003). More recently, the methods of dynamic expectation maximization, variational filtering, and generalized filtering have also been proposed for model inversion (estimation) in this context (Friston et al., 2008).

Several extensions of the Balloon model have since been considered (Buxton et al., 2004), as well as a metabolic/hemodynamic model that takes the metabolic dynamics into account in order to incorporate the separate roles played by excitatory and inhibitory neuronal activities in the generation of the BOLD signal (Sotero and Trujillo-Barreto, 2007). A few alternative approaches for the estimation of these HRF models and related extensions have also been proposed

(Riera et al., 2004). In Riera et al. (2004), a fully stochastic model was presented in order to include physiological noise in the hemodynamic states, in addition to the measurement noise in the observations. A local linearization filter was used for estimating the hemodynamic states as well as the model parameters. In Sotero et al. (2009), a similar approach was used for estimating the metabolic/hemodynamic model proposed by the same group. In contrast to these linearization-based approaches, Johnston et al. (2008) used particle filters so as to truly accommodate the model non-linearities. More recently, Havlicek et al. (2011) proposed non-linear cubature Kalman filtering as a means to invert models of coupled dynamical systems, which furnishes posterior estimates of both the hidden states and the parameters of the system, including any unknown exogenous input.

In fMRI experiments, the system input is given by the stimulus/task time course, which is generally designed as a series of events alternating with baseline periods at specified inter-stimulus intervals (ISIs). A number of studies have addressed the problem of systematically assessing the quality of fMRI experimental designs, both in terms of the ability to detect stimulus/task-related BOLD activation (detection power) and the ability to estimate the HRF model (estimation efficiency) in a given amount of imaging time (Dale, 1999; Liu et al., 2001). Different methodologies have been proposed to determine the optimal design of fMRI experiments for maximal estimation efficiency (Buracas and Boynton, 2002; Wager and Nichols, 2003; Maus et al., 2012), and a few studies have compared different HRF models and the associated estimation efficiency, focusing on specific parameters of interest such as the response latency and duration (Lindquist and Wager, 2007; Lindquist et al., 2009). Importantly, the authors were concerned with the physiological plausibility of the estimated HRF parameters and with their independence, such that differences in one parameter are not confounded with differences in another parameter. However, these studies were based on parameterized HRF models with no direct biophysical groundings, which severely limited the desired physiological interpretability. To our knowledge, no study has so far investigated the effect of experimental design on the estimation of biophysically informed models of the HRF.

When the HRF model is expressed as a dynamic system, the identifiability of this system must be established in order to guarantee that the HRF models inferred from the input/output data are physiologically plausible. It has been shown that the sensitivity of the HRF system input/output behavior to the model parameters is in general small, which means that, when many parameters are estimated together, their values can be varied over a large range with only small changes in the system output (Deneux and Faugeras, 2006). In these cases, the problem of model estimation may be treated as a model falsification (or invalidation) problem, in which we are interested in distinguishing among a set of eligible dynamic systems (Silvestre et al., 2010a). The simplest model falsification problem one can think of is that of stating whether or not a given model is compatible with the current observed input/output data. However, it is important to notice that a model can never be validated in practice. Indeed, the model being compatible with the input/output data up to time  $t$  does not imply that it should be compatible at time  $t + \delta$  where  $\delta > 0$ . Therefore, one can

only say that a given model is not falsified (or invalidated) by the current input/output data. On the other hand, a model is obviously invalidated or falsified once it is not compatible with the observations. Hence, we usually refer to model falsification rather than model validation, since the latter is not achievable in practice. The related problem of model (in)distinguishability arises in a wide range of decision architectures, especially in those that are used in noisy and/or uncertain environments, where more than a single eligible model is compatible with the observed input/output dataset. The distinguishability of two models is in general affected by the input signals, particularly by the uncertainty on the input time-delay and on its magnitude. In fact, model invalidation requires a kind of persistence of the excitation condition in the exogenous inputs, so that the magnitude of the system output signal is large enough when compared to the noise level of the data acquisition process—see (Grewal and Glover, 1976; Walter et al., 1984) and references therein.

In this paper, we extend the results in Silvestre et al. (2010b), by first introducing the concept of absolutely input-distinguishable systems and showing that, for systems with forced responses, the distinguishability between two models can be significantly affected by the shape and magnitude of the external input signals. Moreover, several types of uncertainty, such as unknown input time-delays and uncertain magnitudes of the input signal, can also be adverse to model invalidation. We then exploit the concept of absolutely input-distinguishable systems, in order to optimize the estimation efficiency of fMRI experimental designs through the maximization of the distinguishability among a set of physiologically plausible HRF models. It is stressed that one of the main motivations for the work described herein is the development of a technique that helps define an optimal sequence of stimuli, so that the differences between the models in the set of plausible HRFs become apparent. Hence, the methodology proposed in this paper provides a first step to the so-called experimental paradigm design, while also shedding light on the intrinsic limitations of HRF parameter estimation based on fMRI.

## Methods

The Balloon Model proposed by Buxton et al. (1998), and further analyzed and complemented with the flow dynamics by Friston et al. (2000), consists of a non-linear differential equation that describes the dynamics of normalized values of the blood flow  $b_f$ , with  $s$  being the vasodilatory and activity dependent signal that increases the flow  $b_f$ , the veins deoxyhemoglobin content  $q$ , and the blood venous volume  $v$ , which are considered 1 at rest. This non-linear dynamic system can be described by

$$\left. \begin{aligned} \dot{s} &= \varepsilon u - k_s s - k_f (b_f - 1) && \triangleq F_1 \\ \dot{b}_f &= s && \triangleq F_2 \\ \dot{v} &= \frac{1}{\tau} \left( b_f - v^{\frac{1}{\alpha}} \right) && \triangleq F_3 \\ \dot{q} &= \frac{1}{\tau} \left( b_f \frac{1 - (1 - E_o)^{1/b_f}}{E_o} - v^{\left(\frac{1}{\alpha} - 1\right)} q \right) && \triangleq F_4 \\ y &= V_o \left[ k_1 (1 - q) + k_2 \left( 1 - \frac{q}{v} \right) + k_3 (1 - v) \right] \end{aligned} \right\} \triangleq F(x, \theta, u)$$

$$= f(x, \theta, u) \quad (1)$$

where  $x = [x_1, x_2, x_3, x_4]^T = [s, b_f, v, q]^T$ ,  $E_o$  is the resting net oxygen extraction fraction by capillary bed,  $\varepsilon$  is the efficacy with which neuronal activity causes an increase in signal,  $1/k_s$  and  $1/k_f$  are time constants,  $\tau$  is the mean transit time, and  $\alpha$  is a stiffness exponent that specifies the flow-volume relationship of the venous balloon. The output of this model,  $y(t)$ , is the BOLD signal and represents a complex response controlled by different parameters, that range from the blood oxygenation, to the cerebral blood flow, and cerebral blood volume, and reflects the regional increase in metabolism due to enhancing of the neural activity. In the output equation,  $V_o$  is the resting blood volume fraction, and  $k_1, k_2$ , and  $k_3$  are constants.

The response of the system described by Equation (1), with the parameters in **Table 1** and with initial state  $x^T(0) = [0 \ 1 \ 1 \ 1]$ , to a rectangular input signal, is depicted in **Figure 1**, for different integration periods.

The linear approximation of the model of the system leads to pronouncedly different responses, when compared to the non-linear system. An alternative to this, as described in sequel, is to consider a so-called bilinear model, which accurately mimics the non-linear behavior for sufficiently small integration periods.

## Linearization and Discretization of the Model

The model described by Equation (1) is highly non-linear and parameter-dependent, thus barely allowing any systematic analysis of the associated expected behavior. Hence, to make the problem tractable from a mathematical point of view, the (bi)linearization of the HRF is considered in this paper. This approach allows the use of a widely spread framework for analysis, namely that of the linear time-varying systems. **Figure 1** shows that a close match of the HRF can be obtained by using a bilinear approximation (linear on the state, if the input is fixed, and linear on the input, if the state is fixed). Therefore, in this subsection, a (bi)linearization is derived that approximates the non-linear model locally and that is able to describe the state of the system at a given time,  $x(kT_s)$ , as a function of the state several sampling periods before,  $x((k - N)T_s)$ .

In particular, linearizing Equation (1) around  $x(\cdot) = x^*$  and  $u(\cdot) = 0$ , i.e., writing the associated Taylor expansion and truncating it at the linear term, one obtains (omitting the time-dependence of the variables, for the sake of readability):

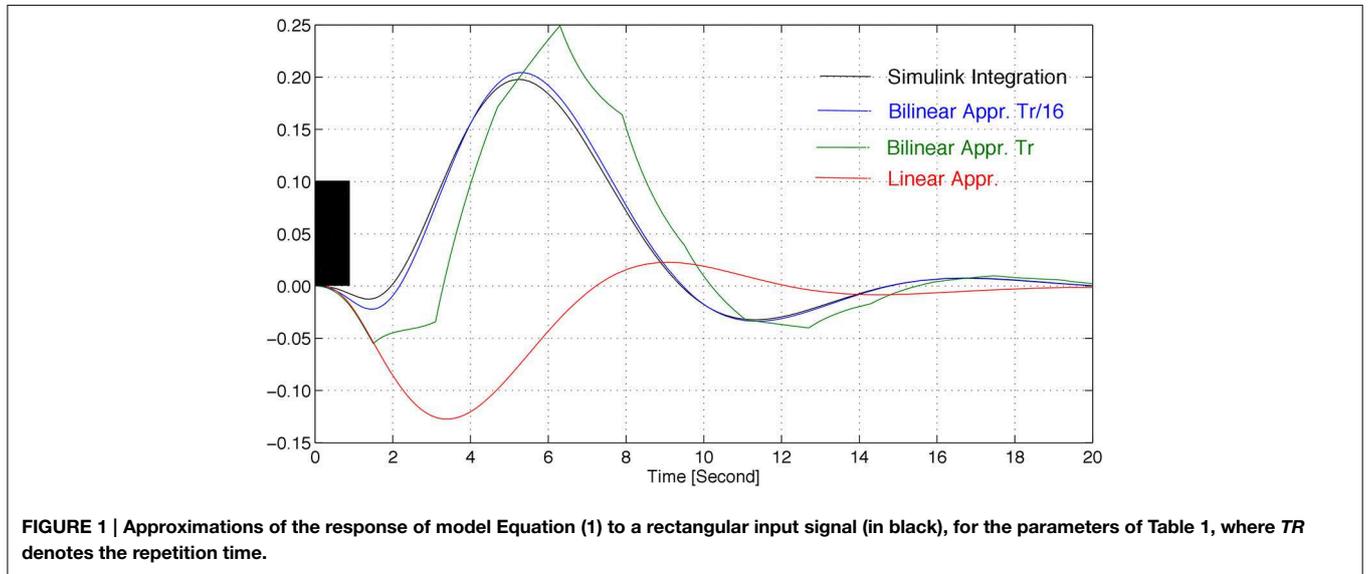
$$\dot{x} \approx F(x^*, \theta, 0) + \left. \frac{\partial F(x, \theta, u)}{\partial x} \right|_{x^*, \theta, 0} (x - x^*) + \sum_i u_i \left( \left. \frac{\partial^2 F(x, \theta, u)}{\partial x \partial u_i} \right|_{x^*, \theta, 0} (x - x^*) + \left. \frac{\partial F(x, \theta, u)}{\partial u_i} \right|_{x^*, \theta, 0} \right),$$

where

$$\frac{\partial F}{\partial x} = \begin{bmatrix} -k_s & -k_f & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\tau} & -\frac{x_3^{\left(\frac{1}{\alpha}-1\right)}}{\alpha\tau} & 0 \\ 0 & \frac{\partial F_4}{\partial x_2} & \frac{\partial F_4}{\partial x_3} & -\frac{x_3^{\left(\frac{1}{\alpha}-1\right)}}{\tau} \end{bmatrix}, \quad (2)$$

**TABLE 1 | Parameters for the non-linear model described by Equation (1).**

Parameter	$\varepsilon$	$k_s[s^{-1}]$	$k_f[s^{-1}]$	$\tau[s]$	$\alpha$	$E_o$	$V_o$	$k_1$	$k_2$	$k_3$
Value	0.065	0.550	0.410	1.280	0.880	0.920	4.88	$7E_o$	2.0	$2E_o - 0.2$



**FIGURE 1 | Approximations of the response of model Equation (1) to a rectangular input signal (in black), for the parameters of Table 1, where  $TR$  denotes the repetition time.**

$$\frac{\partial F_4}{\partial x_2} = \frac{1}{\tau} \left[ \frac{1-(1-E_o)^{\frac{1}{x_2}}}{E_o} + \frac{\log(1-E_o)(1-E_o)^{\frac{1}{x_2}}}{E_o x_2} \right],$$

$$\frac{\partial F_4}{\partial x_3} = \frac{1}{\tau} \left( 1 - \frac{1}{\alpha} \right) x_3^{\left(\frac{1}{\alpha}-2\right)} x_4.$$

and with output equation described by

$$\frac{\partial y}{\partial x} = \left[ 0, \quad 0, \quad -k_3 V_o + k_2 V_o q v^{-2}, \quad -k_1 V_o - k_2 V_o v^{-1} \right].$$

Moreover, given that  $F_1$  depends linearly upon  $u$ , we have that  $\frac{\partial^2 F}{\partial x \partial u_i} = 0$ .

Using the transformation proposed in Friston et al. (2000), one finally obtains the following dynamics:

$$\dot{\tilde{x}} = A\tilde{x} + \sum_i u_i E_i \tilde{x}, \tag{3}$$

where  $\tilde{x} = \left[ 1 \quad x \right]^T$ ,

$$A \triangleq \begin{bmatrix} 0 & 0 \\ \left( F(x^*, \theta, u) - \frac{\partial F(x^*, \theta, u)}{\partial x} x^* \right) & \frac{\partial F(x^*, \theta, 0)}{\partial x} \end{bmatrix},$$

$$E_i \triangleq \begin{bmatrix} 0 & 0 \\ \frac{\partial F(x^*, \theta, 0)}{\partial u_i} & 0 \end{bmatrix},$$

and  $\frac{\partial F(x^*, \theta, 0)}{\partial u_i} = \left[ \varepsilon \quad 0 \quad 0 \quad 0 \right]^T$ .

**Uncertain Dynamic Model Description**

It should be noticed that the dynamics in Equation (3) are bilinear in the state and input variables. This non-linear term hinders the

distinguishability analysis proposed in Rosa and Silvestre (2011) and, thus, a more suitable description is derived in herein.

For the sake of simplicity, we start by redefining  $x(t) \triangleq \tilde{x}(t)$  and  $x^*(t) \triangleq [1 \quad (x^*(t))^T]^T$ . It was previously shown that the continuous-time dynamic model of the HRF, for a single input, can be approximated by

$$\begin{cases} \dot{x}(t) = (A(t) + B_o(t)u(t) + \Delta(t)B_1(t)u(t))x(t), \\ y(t) = h(x(t)), \end{cases} \tag{4}$$

$x(0) = x^*(0)$ ,

with  $t \geq 0$ , and where  $\Delta : \mathbb{R}^+ \rightarrow \mathbb{R}$  was also included to represent an input uncertainty subject to  $|\Delta(t)| \leq 1$  for all  $t \geq 0$ , and where  $B_o = E_1$ . This input uncertainty can be seen as a surrogate for uncertainty in the stimulation signal. The initial state is denoted by  $x(0) \in \mathbb{R}^n$ , and  $n$  is the number of states of the system. Moreover, we assume that

$$B_1(t) = \eta B_o(t),$$

with known  $\eta \in \mathbb{R}$ . We also define  $B(t) = B_o(t) + \Delta(t)B_1(t)$ .

To proceed with the derivation of a discrete-time description of the HRF model in Equation (4), for a given sampling period,  $T_s$ , the following assumptions are posed:

**Assumption 1:** The input signal,  $u(\cdot)$ , is constant during sampling periods, i.e.,  $u(t) = u(kT_s)$ , for all  $t \in [kT_s, (k+1)T_s[$ .

**Assumption 2:** The input uncertainty,  $\Delta(\cdot)$ , is constant during sampling periods, i.e.,  $\Delta(t) = \Delta(kT_s)$ , for all  $t \in [kT_s, (k+1)T_s[$ .

**Assumption 3:** The maps  $A(\cdot)$ ,  $B_o(\cdot)$ , and  $B_1(\cdot)$ , are constant during sampling periods, i.e.,  $A(t) = A(kT_s)$ ,  $B_o(t) = B_o(kT_s)$ , and  $B_1(t) = B_1(kT_s)$ , for all  $t \in [kT_s, (k + 1)T_s[$ .

Under these assumptions, the system in Equation (4) can be rewritten as

$$\begin{cases} \dot{x}(t) = \tilde{A}(k, \Delta(k))x(t), & x(0) = x^*(0), \\ y(t) = g(\tilde{x}(t)), \end{cases} \quad (5)$$

for  $\tilde{x}(t) \in [kT_s, (k + 1)T_s[$ , and where

$$\tilde{A}(k, \Delta(k)) = A_o(k) + \Delta(kT_s)A_1(k),$$

with

$$A_o(k) = A(kT_s) + B_o(kT_s)u(kT_s),$$

and

$$A_1(k) = B_1(kT_s)u(kT_s).$$

In the sequel, we will abbreviate  $x(k) = x(kT_s)$ , for the sake of simplicity. We are now in conditions of stating the following proposition:

**Proposition 1:** Define

$$I^* = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix},$$

and

$$\phi(k) = V(k)\Lambda^*(k)V^{-1}(k)e^{A(kT_s)T_s} - V(k)\Lambda^*(k)V^{-1}(k) - I^*,$$

where  $V(k)\Lambda(k)V^{-1}(k) = A(kT_s)T_s$  is the spectral decomposition of  $A(kT_s)T_s$  with  $\Lambda(k)$  diagonal and  $\Lambda_{11}(k) = 0$ , and

$$\Lambda_{ij}^*(k) = \begin{cases} \frac{1}{\Lambda_{ij}(k)}, & \text{if } i = j \text{ and } \Lambda_{ij} \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, let

$$\begin{aligned} G_o(k) &= e^{A(k)} + B_o(k)u(k) + \phi(k)B_o(k)u(k) \text{ and} \\ G_1(k) &= B_1(k)u(k) + \phi(k)B_1(k)u(k). \end{aligned}$$

Then, the system in Equation (5) is described by

$$\begin{cases} x(k + 1) = G(k, \Delta(k))x(k), & x(0) = x^*(0), \\ y(k) = h(x(k)), \end{cases} \quad (6)$$

where

$$G(k, \Delta(k)) = G_o(k) + \Delta(k)G_1(k),$$

and for  $x(k) = x(kT_s)$ . *Proof:* See Appendix A in Supplementary Material.

Notice that Equation (6), with  $G(k, \Delta(k)) = G_o(k) + \Delta(k)G_1(k)$ , associated with the linearization of the output map,  $g$ , is a full description of the HRF dynamics by means of a linear model with known matrices,  $G_o(k)$  and  $G_1(k)$ , and an uncertain parameter,  $\Delta(k)$ . This description, however, is bilinear in the state,  $x(k)$ , and model uncertainty,  $\Delta(k)$ . This bilinear relationship is tainted once we describe the state  $x(k + 1)$  as a function of  $x(k - 1)$ . Nevertheless, notice that

$(G_o(k + 1) + \Delta G_1(k + 1))(G_o(k) + \Delta G_1(k)) = G_o(k + 1)G_o(k) + \Delta(G_1(k + 1)G_o(k) + G_o(k + 1)G_1(k))$ , since  $G_1(k + 1)G_1(k) = 0$  and where, for the time being, we considered that  $\Delta$  is constant (but unknown), i.e.,  $\Delta(k) = \Delta$  for all  $k$ . To see this, notice that

$$\begin{aligned} G_1(k + 1)G_1(k) &= (B_1(k + 1) + \phi(k + 1)B_1(k + 1))(B_1(k) \\ &\quad + \phi(k)B_1(k)) \\ &= \underbrace{B_1(k + 1)B_1(k)}_{=0} + B_1(k + 1)\phi(k)B_1(k) \\ &\quad + \phi(k + 1)\underbrace{B_1(k + 1)B_1(k)}_{=0} + \\ &\quad + \phi(k + 1)B_1(k + 1)\phi(k)B_1(k), \end{aligned}$$

and that  $B_1(k + 1)\phi(k)B_1(k) = 0$ , due to the fact that the first row of  $\phi$  is zero, and that all but the first column of  $B_1$  are also zero.

By proceeding in a similar manner, we conclude that

$$\begin{aligned} (G_o(k + m) + \Delta G_1(k + m)) \dots (G_o(k) + \Delta G_1(k)) \\ = \Psi_o(k + m) + \Delta \Psi_1(k + m), \end{aligned}$$

where

$$\Psi_o(k + m) = G_o(k + m) \dots G_o(k),$$

and

$$\begin{cases} \Psi_1(k) = G_1(k), \\ \Psi_1(k + m) = G_o(k + m)\Psi_1(k + m - 1) \\ \quad + G_1(k + m)\Psi_o(k + m - 1). \end{cases}$$

Hence, the state  $x(k + m + 1)$  can be written as

$$x(k + m + 1) = (\Psi_o(k + m) + \Psi_1(k + m)\Delta)x(k)$$

Furthermore, the non-linear output Equation of (1) can be linearized as

$$y(x) = y(x^*) + \left. \frac{\partial y}{\partial x} \right|_{x^*} (x - x^*), \quad (8)$$

which, in turn, can alternatively written as:

$$z = y(x) - y(x^*) + \left. \frac{\partial y}{\partial x} \right|_{x^*} x^* = \left. \frac{\partial y}{\partial x} \right|_{x^*} x. \quad (9)$$

where  $z(t)$  can be seen as the measurement for the linear time-varying system obtained by the linearization of Equation (1).

### Absolutely Distinguishable Systems

The problem of indistinguishability typically arises from large amplitudes of the measurement noise, small intensity of the input excitation signals, model uncertainty, and uncertain initial conditions. In particular, if the Signal-to-Noise Ratio (SNR) of the measurements is not sufficiently large, one may be able to explain the observed variables by using more than a single dynamic model, from the set of *eligible* models. A similar conclusion applies if the intensity of the input signal is not sufficient to excite the dynamics of the system.

This section will therefore propose a methodology to systematically derive conditions that guarantee the distinguishability of a set of dynamic models, regardless of the noise sequences and initial states.

### Systems with Uncertain Initial State

We start by analyzing the case where the dynamics of the system are known, although the initial state is uncertain and the measured variables are corrupted by bounded noise. Using Equation (8), we have that

$$y(k) = \underbrace{y(x^*(k)) - C(k)x^*(k)}_{\bar{y}(k)} + C(k)x(k) + n(k), \quad (10)$$

where

$$C(k) = \left. \frac{\partial y}{\partial x} \right|_{x^*(k)},$$

and where  $n(k)$  is the measurement noise. Consider that a given input sequence,  $u(0), \dots, u(N)$ , feeds the inputs of systems  $S_A$  and  $S_B$ , respectively described by

$$S_A : \begin{cases} x_A(k+1) &= G_A(k)x_A(k), \\ y_A(k) &= \bar{y}_A(k) + C_A(k)x_A(k) + n_A(k), \end{cases}$$

$$S_B : \begin{cases} x_B(k+1) &= G_B(k)x_B(k), \\ y_B(k) &= \bar{y}_B(k) + C_B(k)x_B(k) + n_B(k), \end{cases}$$

where  $y_A$  and  $y_B$  are defined as in Equation (10), and  $|n_A(k)| \leq \frac{\bar{n}}{2}$ ,  $|n_B(k)| \leq \frac{\bar{n}}{2}$ . Moreover, we assume that  $x_A(0) \in X_o$  and  $x_B(0) \in X_o$ , where  $X_o \in \mathbb{R}^n$  is a convex polytope. Let  $\phi_i = [n_i^T, u_i^T]^T$ , denote the measurement noise,  $n_i \in W \subseteq \mathbb{R}^{n_n}$ , and input signals,  $u_i \in U \subseteq \mathbb{R}^{n_u}$ , at time instant  $i$ .

**Definition 1:** Systems  $S_A$  and  $S_B$  are said *absolutely*  $(X_o, U, W)$ -input distinguishable in  $N$  sampling times if, for any non-zero

$$(x_A(0), x_B(0), \phi_1, \phi_2, \dots, \phi_N) \in X_o \times X_o \times \overbrace{\Phi \times \dots \times \Phi}^{N \text{ times}},$$

where  $\phi_i \in W \times U =: \Phi \subseteq \mathbb{R}^{n_u+n_d}$  for  $i = 0, 1, \dots, N$ , there exists  $k \in \{0, 1, \dots, N\}$  such that

$$y_A(k) \neq y_B(k).$$

Moreover, two systems are said *absolutely*  $(X_o, U, W)$ -input distinguishable if there exists  $N \geq 0$  such that they are absolutely  $(X_o, U, W)$ -input distinguishable in  $N$  sampling times.

Let  $U = (u(0), u(1), \dots, u(N))$  and

$$W = \left\{ (n(0), n(1), \dots, n(N)) : \forall_{0 \leq k \leq N} |n(k)| \leq \frac{\bar{n}}{2} \right\}.$$

The following proposition can be used to state whether a pair of systems is distinguishable or not.

**Proposition 2:** Systems  $S_A$  and  $S_B$  are absolutely  $(X_o, U, W)$ -input distinguishable in  $N$  sampling times if and only if a solution to the following linear problem does not exist:

$$\begin{bmatrix} C_A(0) & -C_B(0) \\ -C_A(0) & C_B(0) \\ C_A(1)G_A(0) & -C_B(0)G_B(0) \\ -C_A(1)G_A(0) & C_B(0)G_B(0) \\ \vdots & \vdots \\ C_A(N)G_A(N-1) \dots G_A(0) & -C_B(0)G_B(N-1) \dots G_B(0) \\ -C_A(N)G_A(N-1) \dots G_A(0) & C_B(0)G_B(N-1) \dots G_B(0) \\ M_o & 0 \\ 0 & M_o \end{bmatrix} \begin{bmatrix} x_A(0) \\ x_B(0) \end{bmatrix} \leq \begin{bmatrix} \bar{n} - \bar{y}_A(0) + \bar{y}_B(0) \\ \bar{n} + \bar{y}_A(0) - \bar{y}_B(0) \\ \bar{n} - \bar{y}_A(1) + \bar{y}_B(1) \\ \bar{n} + \bar{y}_A(1) - \bar{y}_B(1) \\ \vdots \\ \bar{n} - \bar{y}_A(N) + \bar{y}_B(N) \\ \bar{n} + \bar{y}_A(N) - \bar{y}_B(N) \\ m_o \\ m_o \end{bmatrix}, \quad (11)$$

where  $X_o$  is defined so that  $x \in X_o \Leftrightarrow M_o x \leq m_o$ , which can be written as  $X_o = \text{Set}(M_o, m_o)$ .

*Proof:* See Appendix B in Supplementary Material.

### Systems with Uncertain Model

We now consider the case where the system dynamics are uncertain and described by

$$S_A : \begin{cases} x_A(k+1) &= (G_o^A(k) + \Delta_A G_1^A(k)) x_A(k), \\ y_A(k) &= \bar{y}_A(k) + C_A(k)x_A(k) + n_A(k), \end{cases}$$

$$S_B : \begin{cases} x_B(k+1) &= (G_o^B(k) + \Delta_B G_1^B(k)) x_B(k), \\ y_B(k) &= \bar{y}_B(k) + C_B(k)x_B(k) + n_B(k), \end{cases}$$

where  $y_A$  and  $y_B$  are defined as in Equation (10), and  $|n_A(k)| \leq \frac{\bar{n}}{2}$ ,  $|n_B(k)| \leq \frac{\bar{n}}{2}$ . We also assume that  $|\Delta_A| \leq 1$  and  $|\Delta_B| \leq 1$ . Moreover, for this case we assume that  $X_o$  is a singleton, thus removing the uncertainty in the initial state. In this case,  $S_A$  and  $S_B$  denote families of systems, due to the uncertainties  $\Delta_A$  and  $\Delta_B$ . Therefore, the introduction of the following definition is required.

**Definition 2:** The families of systems  $S_A$  and  $S_B$  are said absolutely  $(X_o, U, W)$ -input distinguishable in  $N$  sampling times if, for any pair of realizations  $(S_1, S_2) \in S_A \times S_B$ , the systems  $S_1$  and  $S_2$  are absolutely  $(X_o, U, W)$ -input distinguishable in  $N$  sampling times.

Hence, we are now in condition to state the following proposition:

**Proposition 3:** The families of systems  $S_A$  and  $S_B$  are absolutely  $(X_o, U, W)$ -input distinguishable in  $N$  sampling times if and only if there does not exist a solution to the following linear problem:

$$\Theta_N \begin{bmatrix} \Delta_A \\ \Delta_B \end{bmatrix} \leq \theta_N, \tag{12}$$

where

$$\Theta_N = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ C_A(1)\Psi_1^A(0)x_A(0) & -C_B(1)\Psi_1^B(0)x_B(0) \\ -C_A(1)\Psi_1^A(0)x_A(0) & C_B(1)\Psi_1^B(0)x_B(0) \\ \vdots & \vdots \\ C_A(N)\Psi_1^A(N-1)x_A(0) & -C_B(N)\Psi_1^B(N-1)x_B(0) \\ -C_A(N)\Psi_1^A(N-1)x_A(0) & C_B(N)\Psi_1^B(N-1)x_B(0) \end{bmatrix}$$

and

$$\theta_N = \begin{bmatrix} \bar{n} - \bar{y}_A(0) + \bar{y}_B(0) - C_A(0)x_A(0) + C_B(0)x_B(0) \\ \bar{n} + \bar{y}_A(0) - \bar{y}_B(0) + C_A(0)x_A(0) - C_B(0)x_B(0) \\ \bar{n} - \bar{y}_A(1) + \bar{y}_B(1) - C_A(1)\Psi_0^A(0)x_A(0) + \Psi_0^B(0)x_B(0) \\ \bar{n} + \bar{y}_A(1) - \bar{y}_B(1) + C_A(1)\Psi_0^A(0)x_A(0) - \Psi_0^B(0)x_B(0) \\ \vdots \\ \bar{n} - \bar{y}_A(N) + \bar{y}_B(N) - C_A(N)\Psi_0^A(N-1)x_A(0) + \Psi_0^B(N-1)x_B(0) \\ \bar{n} + \bar{y}_A(N) - \bar{y}_B(N) + C_A(N)\Psi_0^A(N-1)x_A(0) - \Psi_0^B(N-1)x_B(0) \end{bmatrix}$$

*Proof:* See Appendix C in Supplementary Material.

**Figure 2A** depicts the impulse and step responses of the HRF model with the parameters of **Table 1**, with an uncertainty of 10% in the input signal. It should be noticed that this type of uncertainty mainly affects the amplitude of the responses of the system. Thus, the rise- and fall-times are not significantly influenced by small variations on the amplitude of the input signal.

### Systems with Uncertain Input Time-Delays

In this subsection, a strategy to model uncertain input time-delays is developed. The approach presented in the sequel amounts for rewriting these uncertain input time-delays as model uncertainty.

Consider that the input signal, at sampling time  $k$ , is given by

$$u(k) = \tilde{u}(k - k_d),$$

where  $k_d$  is an integer (the uncertain delay) satisfying  $|k_d| \leq \bar{k}_d$ , with known  $\bar{k}_d$ . The value of  $\tilde{u}(k)$ , for each  $k \geq 0$ , is

also assumed known and bounded. Thus, we have, for each  $k \geq 0$ ,

$$\underline{u}(k) \leq u(k) \leq \bar{u}(k), \tag{13}$$

where  $\bar{u}(k) = \max_{|m| \leq \bar{k}_d} \tilde{u}(k-m)$  and  $\underline{u}(k) = \min_{|m| \leq \bar{k}_d} \tilde{u}(k-m)$ . Therefore, Equation (13) can be rewritten as

$$u(k) = u_o(k) + \Delta_u(k)u_1(k),$$

where  $|\Delta_u(k)| \leq 1$ ,  $u_o(k) = \frac{\bar{u}(k)+\underline{u}(k)}{2}$ , and  $u_1(k) = \frac{\bar{u}(k)-\underline{u}(k)}{2}$ .

Hence, unknown but bounded time-delays on the input can be treated as uncertainty on the  $B$  matrix. The impulse and step responses of the HRF model with the parameters of **Table 1**, with an uncertain input time-delay,  $k_d$ , bounded by  $|k_d| \leq 3$ , are depicted in **Figure 2B**. As seen in the figure, the uncertainty in the input time-delay enlarges the uncertainty in the rise- and fall-times of the output.

### Systems with Uncertain Model and Input Time-Delays

For the sake of completeness, in this subsection we analyze the effects of simultaneous uncertainty on the model and on the input time-delays. The results for this scenario are depicted in **Figure 2C**. As expected, the uncertainty on the model chiefly affects the amplitude of the responses, while the uncertainty on the input time-delay changes the corresponding rise- and fall-times.

### Systems with Uncertain Model and Uncertain Initial State

We now consider the case where both the system dynamics and the initial state are uncertain. The problem is set to that of concluding whether the following two families of systems are distinguishable:

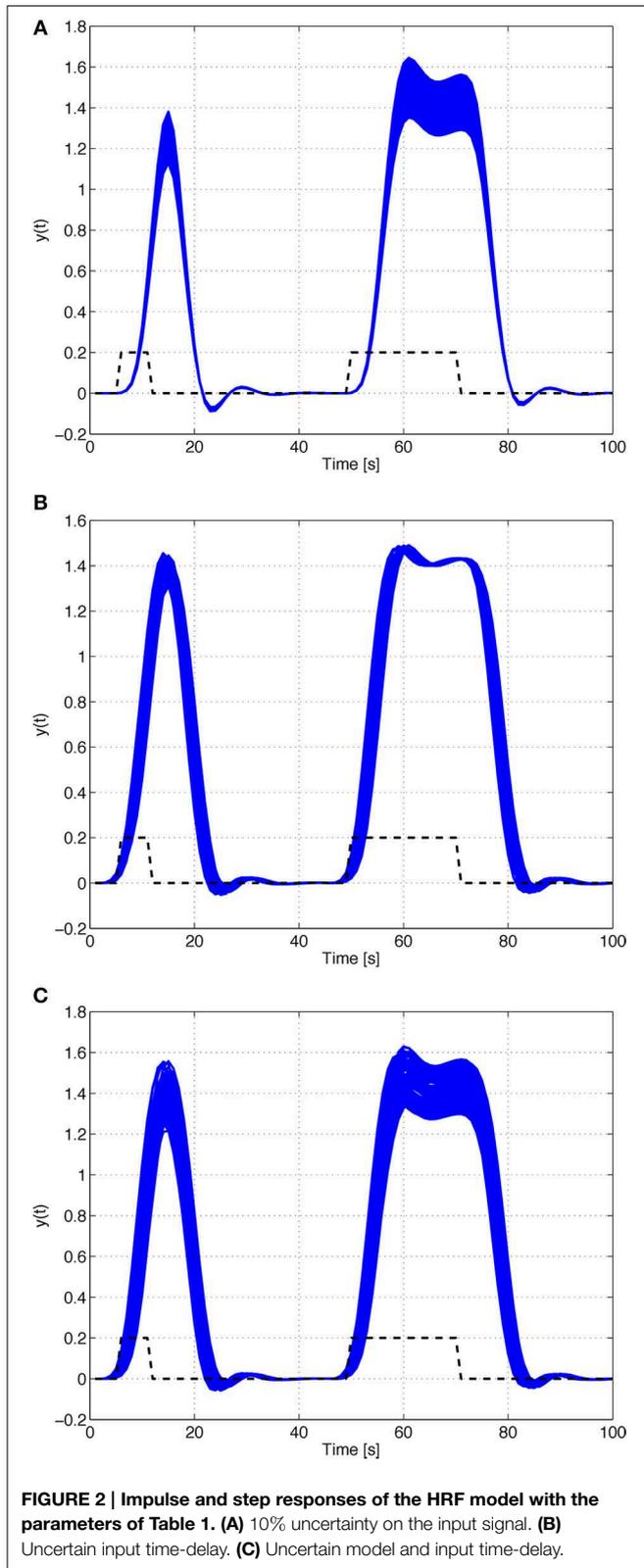
$$S_A : \begin{cases} x_A(k+1) &= (G_o^A(k) + \Delta_A(k)G_1^A(k))x_A(k), \\ y_A(k) &= \bar{y}_A(k) + C_A(k)x_A(k) + n_A(k), \end{cases}$$

$$S_B : \begin{cases} x_B(k+1) &= (G_o^B(k) + \Delta_B(k)G_1^B(k))x_B(k), \\ y_B(k) &= \bar{y}_B(k) + C_B(k)x_B(k) + n_B(k), \end{cases}$$

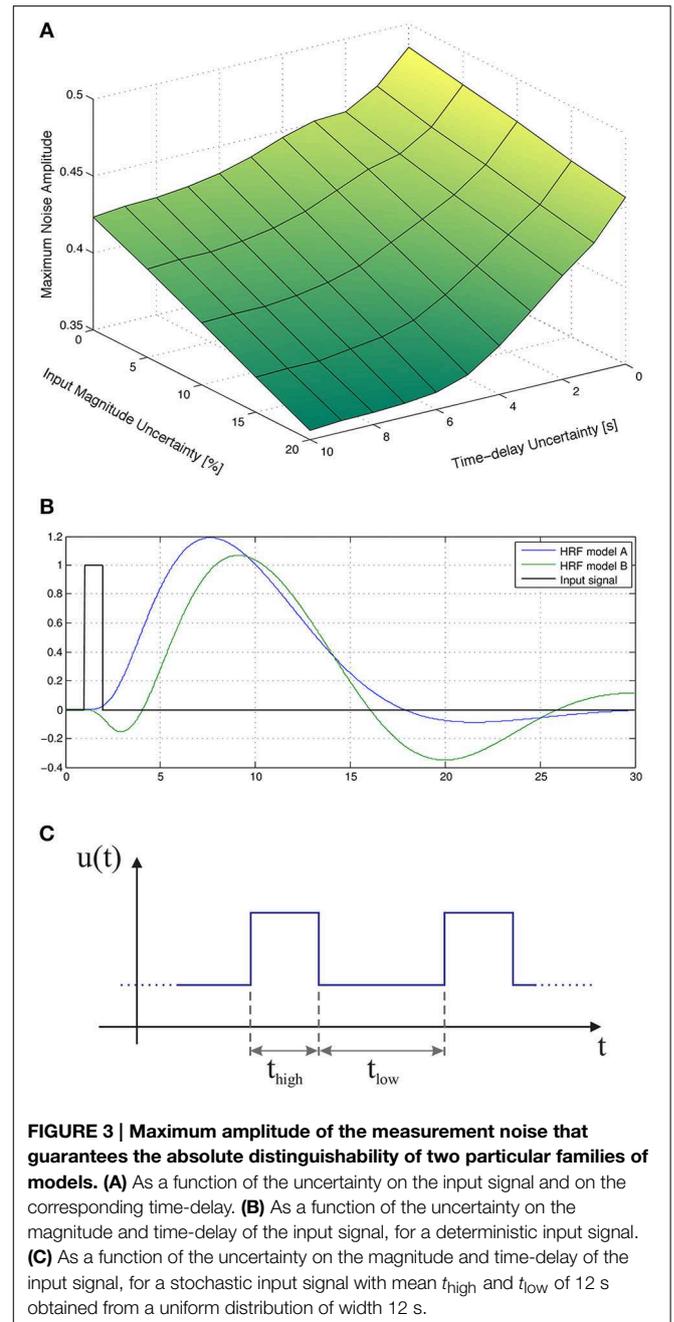
where  $y_A$  and  $y_B$  are defined as in Equation (10), and  $|n_A(k)| \leq \frac{\bar{n}}{2}$ ,  $|n_B(k)| \leq \frac{\bar{n}}{2}$ . We also assume that  $|\Delta_A(k)| \leq 1$  and  $|\Delta_B(k)| \leq 1$ . Moreover, for this case we assume that  $X_o$  is a convex polytope.

**Proposition 4:** Let  $e_1 = [1 \ 0 \ 0 \ 0 \ 0]^T$ . The families of systems  $S_A$  and  $S_B$  are absolutely  $(X_o, U, W)$ -input distinguishable in  $N$  sampling times if and only if there does not exist a solution to the following linear problem:

$$\forall_{k \in \{0, 1, \dots, N\}} : \begin{cases} x_A(0), x_B(0) & \in & X_o, \\ C_A(k)x_A(k) - C_B(k)x_B(k) & \leq & \bar{n} - \bar{y}_A(k) + \bar{y}_B(k), \\ -C_A(k)x_A(k) + C_B(k)x_B(k) & \leq & \bar{n} + \bar{y}_A(k) - \bar{y}_B(k), \\ x_A(k+1) - G_o^A(k)x_A(k) - G_1^A(k)e_1z_A(k) & = & 0, \\ x_B(k+1) - G_o^B(k)x_B(k) - G_1^B(k)e_1z_B(k) & = & 0, \\ -e_1^T x_A(k) \leq z_A(k) & \leq & e_1^T x_A(k), \\ -e_1^T x_B(k) \leq z_B(k) & \leq & e_1^T x_B(k), \end{cases}$$



where the unknown variables are  $x_A(0), \dots, x_A(N)$ ,  $x_B(0), \dots, x_B(N)$ ,  $z_A(0), \dots, z_A(N - 1)$ , and  $z_B(0), \dots, z_B(N - 1)$ .  
*Proof:* See Appendix D in Supplementary Material.



**Figure 3A** depicts the maximum amplitude of the measurements noise that guarantees the absolute distinguishability of two particular families of HRF models, as a function of the uncertainty on the input signal and on the corresponding time-delay. As expected, the maximum level of sensor noise such that the two families of models are absolutely distinguishable, decreases with both types of uncertainty.

### Pre-Processing of fMRI Time Series

We stress that the assumption that the additive noise in the measured signal is bounded is not conservative in practice, since outliers and other *unboundedness* behaviors can, in general, be

tackled during pre-processing, i.e., before performing the main analysis of the signals. This can be done, in particular, by low-pass filtering the signal, so that high-frequency noise is significantly attenuated.

Additionally, the following pre-processing steps are commonly applied to fMRI time series data before submitting them to statistical analysis (Jezzard et al., 2001): (i) normalization of the whole 4D fMRI dataset by scaling each volume by a single (common) scaling factor, so that subsequent analyses are valid; (ii) motion correction by alignment of all fMRI volumes to a reference volume in the time series, usually performed by applying rigid-body transformations, in order to reduce the effect of subject head motion during the experiment; and (iii) high-pass temporal filtering, usually using a local fit of a straight line (Gaussian-weighted within the line to give a smooth response), in order to remove low-frequency artifacts such as signal drifts or physiological fluctuations.

## Results

In this section, we study the influence of the choice of the input signal on the distinguishability of a set of HRF models. A methodology to optimize the fMRI experimental design that takes advantage of this knowledge is also presented.

Throughout the remainder of this paper, we are going to refer to the families of HRF models *A* and *B*, described by the dynamics in Equation (1), with the physiologically plausible parameters presented in **Table 2**. Model family *B* displays a pronounced undershoot and the presence of an initial dip, in stark contrast to model family *A*.

The response of the nominal HRF models, for the parameter configurations of **Table 2**, with initial state  $x^T(0) = [0 \ 1 \ 1 \ 1]$ , to a rectangular input signal of duration 1 s and unit magnitude, is depicted in **Figure 4**.

In general, the input signal is composed of a series of rectangular pulses (events) of duration  $t_{high}$  alternating with baseline periods of duration  $t_{low}$ , with a total duration of 200 s (see **Figure 5**).

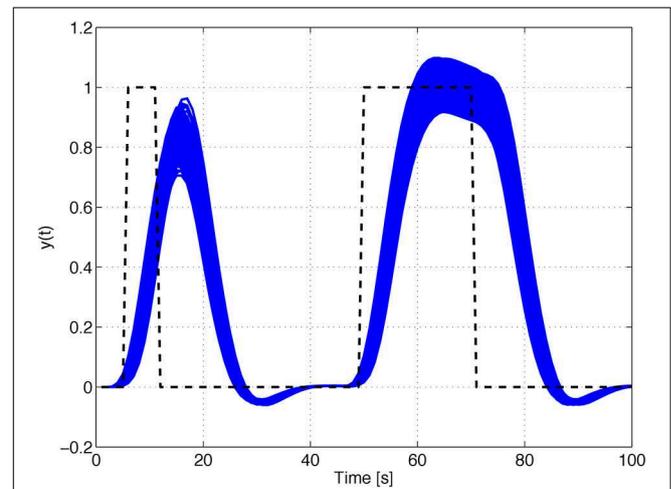
In order to illustrate the characteristic behavior of HRF model family *A*, their responses to rectangular input signals of duration 5 and 20 s and unit magnitude, with an uncertain input time-delay,  $k_d$ , bounded by  $|k_d| \leq 3$  s, and input uncertainty of 10%, are depicted in **Figure 6**. The uncertainty on the input time-delay enlarges the uncertainty in the rise- and fall-times of the output, while the uncertainty in the input mainly affects the amplitude of the responses of the system.

**Figure 3B** depicts the maximum amplitude of the measurements noise that guarantees the absolute distinguishability of the families of models *A* and *B*, for an input signal with  $t_{low} = 12$  s and  $t_{high} = 12$  s, as a function of

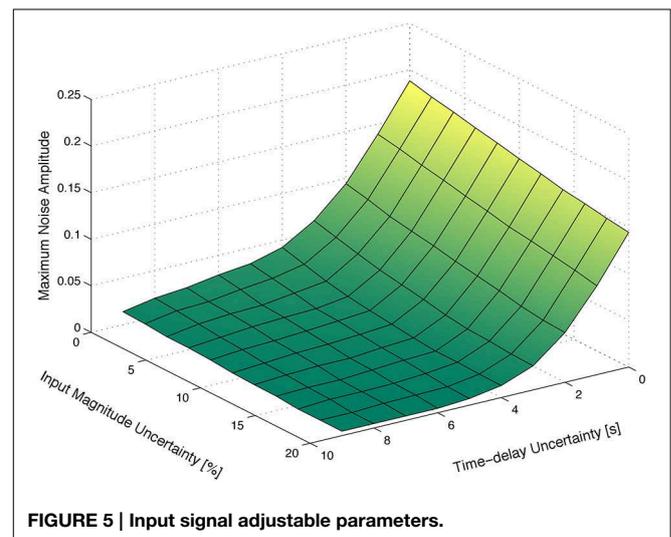
the uncertainty on the magnitude of the input signal and on the corresponding time-delay. As expected, the maximum level of measurement noise such that the families of models *A* and *B* are absolutely distinguishable decreases with both types of uncertainty.

Furthermore, we considered a stochastic input signal, composed of a series of rectangular pulses with mean duration of  $E(t_{high}) = 12$  s, and mean baseline period of  $E(t_{low}) = 12$  s drawn from a uniform distribution of width 12 s. According to the results in the literature (see, for instance, Josephs et al., 1997; Miezin et al., 2000), we observe that, by performing random small variations on  $t_{high}$  and  $t_{low}$ , alternative trajectories of the non-linear model Equation (1) are exploited, which in turn improves the identifiability of the models, as depicted in **Figure 3C**.

We now analyze the effect of different experimental designs on the distinguishability of the families of models at hand. At this point, our goal is to find the combination of values of  $t_{low}$  and  $t_{high}$  such that the absolute distinguishability of two or more



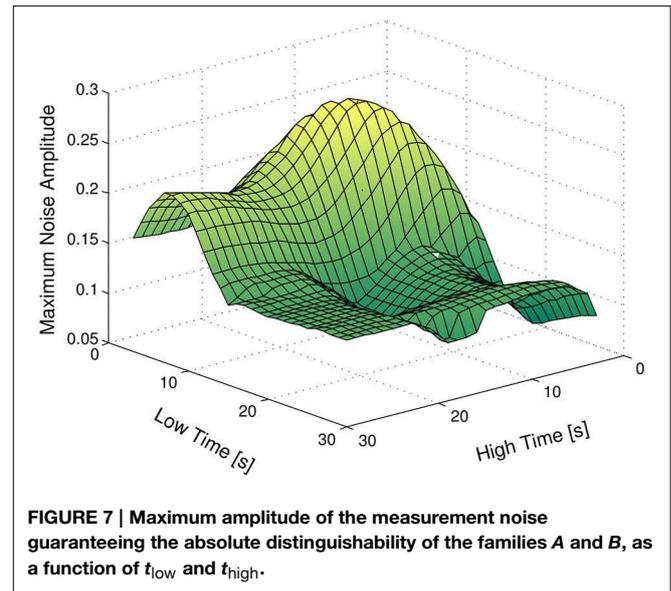
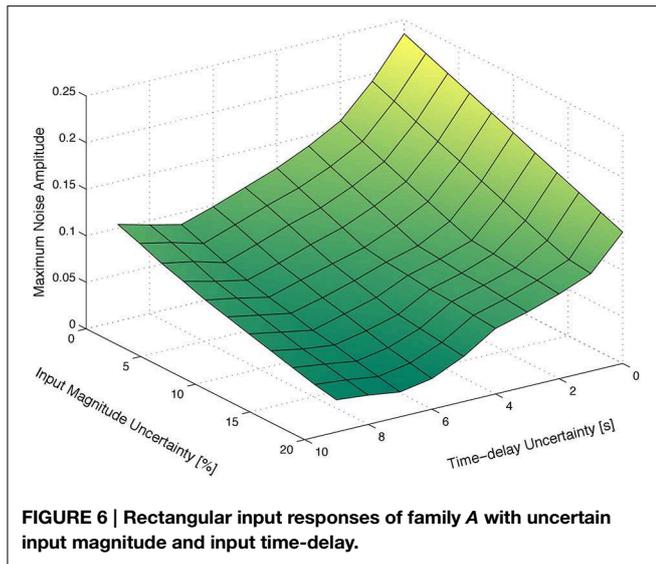
**FIGURE 4 |** Time responses of the nominal models of families *A* and *B*.



**FIGURE 5 |** Input signal adjustable parameters.

**TABLE 2 |** Parameters for the families of non-linear models.

HRF	$k_s$ [ $s^{-1}$ ]	$k_f$ [ $s^{-1}$ ]	$\tau$ [s]	$\alpha$	$E_o$
A	0.400	0.100	2.080	0.320	0.340
B	0.220	0.110	2.180	0.320	0.985



families of models is guaranteed for the highest upper bound on the amplitude of the measurement noise. We denote this optimal combination of values by  $(t_{low}^*, t_{high}^*)$ . The advantage of using an input signal with parameters  $(t_{low}^*, t_{high}^*)$  obviously stems from the fact that we can allow for the highest amplitude on the measurement noise, while guaranteeing the distinguishability of the families.

**Figure 7** depicts the results obtained, considering no time-delay or magnitude uncertainty. As expected, input signals with very small values of  $t_{high}$  and large values of  $t_{low}$  do not have the power required to significantly stimulate the system. On the other hand, input signals with very small values of  $t_{high}$  and  $t_{low}$  are faster than the dynamics of the system, and hence do not produce remarkable changes in the output of the plant. As a final remark, the optimum value for  $t_{low}$  and  $t_{high}$  is 10 s, i.e.,  $t_{low}^* = 10$  s and  $t_{high}^* = 10$  s.

## Discussion

We have addressed the problem of the distinguishability of HRF models in the analysis of fMRI data of brain activation, based on the biophysically informed description of the HRF as a non-linear time-invariant input-state-output dynamic system. We first introduced the concept of absolutely input-distinguishable systems and then showed that the distinguishability between two HRF models, and hence system identification, is significantly affected by the external input (stimulus/task) signals. In particular, the uncertainty in the input time-delays and its magnitude may adversely affect model identification, by reducing the maximum noise level below which model distinguishability is guaranteed. We then applied the concept of absolutely input-distinguishable systems to the development of a methodology for the assessment of the HRF estimation efficiency of fMRI experimental designs, through the maximization of the distinguishability level among a set of physiologically plausible HRF models.

The main contribution of this paper is therefore 2-fold. On the one hand, we show that the distinguishability of two HRF models depends on the level of the measurement noise as well as on the characteristics of the input signal. On the other hand, we develop a methodology to optimize fMRI experimental designs for HRF estimation, which maximizes the allowable noise amplitude that does not impair the distinguishability of a set of a priori admissible dynamic systems.

In this paper, it is assumed that the system inputs can be selected or, at least, measured. This assumption is verified in a straightforward manner when external inputs are present, such as sensory stimuli or cognitive tasks. Although no explicit external inputs exist in resting-state fMRI acquisitions, it has been observed that discrete neuronal events do occur (Deco and Jirsa, 2012). Most interestingly, it has been recently suggested that such events can be identified as peaks of relatively large BOLD signal amplitude (Tagliazucchi et al., 2011), and resting-state fMRI data can then be seen as “spontaneous event-related” data (Wu et al., 2013).

## Significance of HRF Estimation

The importance of estimating the HRF in fMRI experiments is based on the extensively observed variability of its shape and dynamics across brain regions, conditions, subjects, and populations, with critical consequences in the analysis of fMRI data. In fact, one direct consequence of HRF variability is that the deviation of the real HRF from the pre-specified HRF leads to a poorer model of the observed BOLD signal and hence reduces the sensitivity to detect BOLD changes (Handwerker et al., 2004). Another consequence is the potential detection of a group effect due to a systematic HRF difference, which would then be incorrectly interpreted as a neuronal effect. Moreover, when attempting to infer causality within brain networks from BOLD data, differences in HRF latency across brain regions can potentially confound the directionality of information flow (David et al., 2008; Smith et al., 2011; Murta

et al., 2012; Jorge et al., 2014). On the other hand, HRF variability may be an object of interest on its own, potentially reflecting physiological changes associated with the effects of drugs, aging or pathology, for example (Iadecola, 2004). Additionally, there is a growing interest in studying, not only the amplitude of BOLD activation, but also its dynamics, namely its latency and duration, which are reflected in the HRF (Bellgowan et al., 2003). In these cases, it would be desirable to estimate the actual HRF model underlying the BOLD signal measured in each voxel, experiment, subject or population, or otherwise account for its variability.

Despite the acknowledged need for modeling the HRF underlying fMRI BOLD data, and although different approaches have been continuously proposed in the literature for this purpose, our ability to understand HRF variability remains poor (Handwerker et al., 2012). Critically, most studies have focused on parameterized HRF models in a linear framework, while the estimation of physiologically plausible non-linear HRF models with direct biophysical interpretability has been very limited. In particular, no previous study has investigated the optimal fMRI experimental design for the estimation of such biophysical HRF models. We believe that our work therefore makes an important contribution for understanding how a biophysically informed model of the HRF may be inferred from fMRI data, as a function of experimental design and measurement noise.

### Biophysically Informed HRF Modeling

Using a biophysically informed model of the HRF not only allows for a physiologically plausible interpretation of the results, but it also more accurately explains empirical BOLD data, particularly concerning commonly observed non-linearities. Importantly, in contrast to parameterized HRF models, biophysical models described by dynamic systems can account for the detailed dynamics of BOLD responses through a reduced number of parameters, while constraining it to be physiologically plausible. For example, the post-stimulus undershoot and the initial dip are two features of observed BOLD responses that naturally emerge from this dynamic system under slightly different combinations of a limited number of parameters. Although using such dynamic systems represents an additional computational effort compared to the more straightforward linear methods, this may nevertheless become the chosen approach in studies where a detailed characterization of the BOLD temporal dynamics is desirable. In particular, the combination of EEG with fMRI may greatly benefit from such approaches (Riera et al., 2005). On the other hand, important complementary information may be gained for HRF model estimation by combining BOLD recordings with the acquisition of blood flow data using Arterial Spin Labeling (ASL) or near-infrared spectroscopy (NIRS) (Huppert et al., 2006). Despite the potential advantages of such a biophysically informed dynamic system approach to fMRI data analysis, only a few studies have been dedicated to the associated problem of system identification/model estimation (Friston, 2002; Riera et al., 2004). Our study therefore makes a significant contribution to this limited body of literature, by introducing the concept of input-distinguishability of HRF models in order to inform model selection in this context.

### Optimization of the Experimental Design

Previous studies systematically assessing the quality of fMRI experimental designs have again been focused on parameterized HRF models within a linear framework (Dale, 1999; Liu et al., 2001). They found that optimal estimation efficiency is obtained at the cost of reduced detection power by employing randomized rapid event-related designs. In fact, it was shown that, if the ISI is properly jittered or randomized from trial to trial, the efficiency improves monotonically with decreasing mean ISI (Dale, 1999). In general, it is found that a trade-off exists between detection power and estimation efficiency, with block designs being optimal for the former while event-related designs are optimal for the latter (Liu et al., 2001). Nevertheless, a recent report established the feasibility and test-retest reliability of estimating HRF parameters from block design fMRI data (Shan et al., 2014). In our work, we have used a randomized design by introducing uncertainty in the ISI, and we showed that smaller uncertainty leads to better distinguishability for the same noise level. Our results are therefore consistent with the literature.

### Limitations

The framework adopted in this work resorts to deterministic concepts and, therefore, certain assumptions are posed on the signals acting on the system, in particular in terms of maximum amplitudes. Stochastic approaches are more flexible in that sense, but require the knowledge regarding the statistical properties of those signals, which may not be trivial to obtain, or which may be violated in practice. Therefore, a compromise between these two alternative frameworks—deterministic and stochastic—for the distinguishability of HRF models is still a subject of further research.

### Conclusion

In summary, in this paper we proposed a novel approach to assess distinguishability among a set of physiologically plausible biophysically informed HRF models, and to design fMRI experiments for optimal estimation efficiency of such HRF models, with potentially great impact in further understanding HRF variability and its physiological meaning.

### Acknowledgments

We acknowledge financial support by the Portuguese Science Foundation through Projects PTDC/SAU-ENB/112294/2009, PTDC/BBB-IMG/2137/2012 and FCT [UID/EEA/50009/2013], and project MYRG117(Y1-L3)-FST12-MKM of the University of Macau. We also thank the reviewers for their insightful comments and corrections.

### Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fncom.2015.00054/abstract>

## References

- Aguirre, G., Zarahn, E., and D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *Neuroimage* 8, 360–369.
- Bellgowan, P., Saad, Z., and Bandettini, P. (2003). Understanding neural system dynamics through task modulation and measurement of functional mri amplitude, latency, and width. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1415–1419. doi: 10.1073/pnas.0337747100
- Birn, R., Saad, Z., and Bandettini, P. (2001). Spatial heterogeneity of the nonlinear dynamics in the fMRI BOLD response. *Neuroimage* 14, 817–826. doi: 10.1006/nimg.2001.0873
- Boynton, G., Engel, S. A., Glover, G., and Heeger, D. (1996). Linear systems analysis of functional magnetic resonance imaging in human v1. *J. Neurosci.* 16, 4207–4221.
- Buracas, G., and Boynton, G. (2002). Efficient design of event-related fMRI experiments using msequences. *Neuroimage* 16(3 Pt 1), 801–813. doi: 10.1006/nimg.2002.1116
- Buxton, R., Uluda, K., Dubowitz, D., and Liu, T. (2004). Modeling the hemodynamic response to brain activation. *Neuroimage* 23, 220–233. doi: 10.1016/j.neuroimage.2004.07.013
- Buxton, R., Wong, E., and Frank, L. (1998). Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. *Magn. Reson. Med.* 39, 855–864. doi: 10.1002/mrm.1910390602
- Cohen, E., and Ugurbil, K. (2002). Effect of basal conditions on the magnitude and dynamics of the blood oxygenation level-dependent fMRI response. *J. Cereb. Blood Flow Metab.* 22, 1042–1053. doi: 10.1097/00004647-200209000-00002
- D'Esposito, M., Deouell, L., and Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. *Nat. Rev. Neurosci.* 4, 863–872. doi: 10.1038/nrn1246
- Dale, A. (1999). Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.* 8, 109–114.
- David, O., Guillemain, I., Saittel, S., Rey, S., Deransart, C., Segebarth, C., et al. (2008). Identifying neural drivers with functional mri: an electrophysiological validation. *PLoS Biol.* 6:2683–2697. doi: 10.1371/journal.pbio.0060315
- Deco, G., and Jirsa, V. K. (2012). Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. *J. Neurosci.* 32, 3366–3375. doi: 10.1523/JNEUROSCI.2523-11.2012
- Deneux, T., and Faugeras, O. (2006). Using nonlinear models in fMRI data analysis: model selection and activation detection. *Neuroimage* 32, 1669–1689. doi: 10.1016/j.neuroimage.2006.03.006
- Friston, K. (2002). Bayesian estimation of dynamical systems: an application to fMRI. *Neuroimage* 16, 513–530. doi: 10.1006/nimg.2001.1044
- Friston, K., Fletcher, P., Josephs, O., Holmes, A., Rugg, M., and Turner, R. (1998). Event-related fMRI: characterizing differential responses. *Neuroimage* 7, 30–40. doi: 10.1006/nimg.1997.0306
- Friston, K., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage* 19, 1273–1302. doi: 10.1016/S1053-8119(03)00202-7
- Friston, K., Jezzard, P., and Turner, R. (1994). Analysis of functional mri time-series. *Hum. Brain Mapp.* 1, 153–171. doi: 10.1002/hbm.460010207
- Friston, K., Mechelli, A., Turner, R., and Price, C. (2000). Nonlinear responses in fMRI: the Balloon model, Volterra kernels and other hemodynamics. *Neuroimage* 12, 466–477. doi: 10.1006/nimg.2000.0630
- Friston, K., Trujillo-Barreto, N., and Daunizeau, J. (2008). DEM: a variational treatment of dynamic systems. *Neuroimage* 41, 849–885. doi: 10.1016/j.neuroimage.2008.02.054
- Glover, G. H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9, 416–429. doi: 10.1006/nimg.1998.0419
- Grewal, M., and Glover, K. (1976). Identifiability of linear and nonlinear dynamical systems. *IEEE Trans. Autom. Control* 21, 833–837. doi: 10.1109/TAC.1976.1101375
- Handwerker, D. A., Gonzalez-Castillo, J., D'Esposito, M., and Bandettini, P. A. (2012). The continuing challenge of understanding and modeling hemodynamic variation in fMRI. *Neuroimage* 62, 1017–1023. doi: 10.1016/j.neuroimage.2012.02.015
- Handwerker, D. A., Ollinger, J. M., and D'Esposito, M. (2004). Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21, 1639–1651. doi: 10.1016/j.neuroimage.2003.11.029
- Havlicek, M., Friston, K. J., Jan, J., Brazdil, M., and Calhoun, V. D. (2011). Dynamic modeling of neuronal responses in fMRI using cubature kalman filtering. *Neuroimage* 56, 2109–2128. doi: 10.1016/j.neuroimage.2011.03.005
- Huppert, T., Hoge, R., Diamond, S., Franceschini, M. A., and Boas, D. A. (2006). A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *Neuroimage* 29, 368–382. doi: 10.1016/j.neuroimage.2005.08.065
- Iadecola, C. (2004). Neurovascular regulation in the normal brain and in Alzheimer's disease. *Nat. Rev. Neurosci.* 5, 347–360. doi: 10.1038/nrn1387
- Jezzard, P., Matthews, P. M., and Smith, S. M. (Eds.). (2001). *Functional MRI: An Introduction to Methods*, Vol. 61. Oxford: Oxford University Press.
- Johnston, L. A., Duff, E., Mareels, I., and Egan, G. F. (2008). Nonlinear estimation of the BOLD signal. *Neuroimage* 40, 504–514. doi: 10.1016/j.neuroimage.2007.11.024
- Jorge, J. P., van der Zwaag, W., and Figueiredo, P. (2014). EEG-fMRI integration for the study of human brain function. *NeuroImage* 102(Pt1), 24–34. doi: 10.1016/j.neuroimage.2013.05.114
- Josephs, O., Turner, R., and Friston, K. (1997). Event-related fMRI. *Hum. Brain Mapp.* 5, 243–248.
- Lindquist, M. A., Meng Loh, J., Atlas, L. Y., and Wager, T. D. (2009). Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. *Neuroimage* 45, S187–S198. doi: 10.1016/j.neuroimage.2008.10.065
- Lindquist, M. A., and Wager, T. D. (2007). Validity and power in hemodynamic response modeling: a comparison study and a new approach. *Hum. Brain Mapp.* 28, 764–784. doi: 10.1002/hbm.20310
- Liu, T. T., Behzadi, Y., Restom, K., Uludag, K., Lu, K., Buracas, G. T., et al. (2004). Caffeine alters the temporal dynamics of the visual BOLD response. *Neuroimage* 23, 1402–1413. doi: 10.1016/j.neuroimage.2004.07.061
- Liu, T. T., Frank, L. R., Wong, E. C., and Buxton, R. B. (2001). Detection power, estimation efficiency, and predictability in event-related fMRI. *Neuroimage* 13, 759–773. doi: 10.1006/nimg.2000.0728
- Logothetis, N. K., and Wandell, B. A. (2004). Interpreting the BOLD signal. *Annu. Rev. Physiol.* 66, 735–769. doi: 10.1146/annurev.physiol.66.082602.092845
- Maus, B., van Breukelen, G. J., Goebel, R., and Berger, M. P. (2012). Optimal design for nonlinear estimation of the hemodynamic response function. *Hum. Brain Mapp.* 33, 1253–1267. doi: 10.1002/hbm.21289
- Miezin, F. M., MacCotta, L., Ollinger, J. M., Petersen, S. E., and Buckner, R. L. (2000). Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11, 735–759. doi: 10.1006/nimg.2000.0568
- Murta, T., Leal, A., Garrido, M. I., and Figueiredo, P. (2012). Dynamic causal modelling of epileptic seizure propagation pathways: a combined EEG-fMRI study. *Neuroimage* 62, 1634–1642. doi: 10.1016/j.neuroimage.2012.05.053
- Riera, J., Aubert, E., Iwata, K., Kawashima, R., Wan, X., and Ozaki, T. (2005). Fusing EEG and fMRI based on a bottom-up model: inferring activation and effective connectivity in neural masses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 1025–1041. doi: 10.1098/rstb.2005.1646
- Riera, J. J., Watanabe, J., Kazuki, I., Naoki, M., Aubert, E., Ozaki, T., et al. (2004). A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals. *Neuroimage* 21, 547–567. doi: 10.1016/j.neuroimage.2003.09.052
- Rosa, P., and Silvestre, C. (2011). “On the distinguishability of discrete linear time-invariant dynamic systems,” in *Proceedings of the 50th IEEE Conference on Decision and Control*. (Orlando, FL).
- Shan, Z. Y., Wright, M. J., Thompson, P. M., McMahon, K. L., Blokland, G. G., de Zubicaray, G. I., et al. (2014). Modeling of the hemodynamic responses in block design fMRI studies. *J. Cereb. Blood Flow Metab.* 34, 316–324. doi: 10.1038/jcbfm.2013.200
- Silvestre, C., Figueiredo, P., and Rosa, P. (2010a). “Multiple-model set-valued observers: a new tool for HRF model selection in fMRI” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Buenos Aires), 5704–5707.
- Silvestre, C., Figueiredo, P., and Rosa, P. (2010b). “On the distinguishability of HRF models in fMRI” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Buenos Aires), 5677–5680.
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for fMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063

- Sotero, R. C., and Trujillo-Barreto, N. J. (2007). Modelling the role of excitatory and inhibitory neuronal activity in the generation of the BOLD signal. *Neuroimage* 35, 149–165. doi: 10.1016/j.neuroimage.2006.10.027
- Sotero, R. C., Trujillo-Barreto, N. J., Jiménez, J. C., Carbonell, F., and Rodríguez-Rojas, R. (2009). Identification and comparison of stochastic metabolic/hemodynamic models (sMHM) for the generation of the BOLD signal. *J. Comput. Neurosci.* 26, 251–269. doi: 10.1007/s10827-008-0109-3
- Tagliazucchi, E., Balenzuela, P., Fraiman, D., Montoya, P., and Chialvo, D. R. (2011). Spontaneous BOLD event triggered averages for estimating functional connectivity at resting state. *Neurosci. Lett.* 488, 158–163. doi: 10.1016/j.neulet.2010.11.020
- Vincent, T., Badillo, S., Risser, L., Chaari, L., Bakhous, C., Forbes, F., et al. (2014). Flexible multivariate hemodynamics fMRI data analyses and simulations with PyHRF. *Front. Neurosci.* 8:67. doi: 10.3389/fnins.2014.00067
- Wager, T. D., and Nichols, T. E. (2003). Optimization of experimental design in fMRI: a general framework using a genetic algorithm. *Neuroimage* 18, 293–309. doi: 10.1016/S1053-8119(02)00046-0
- Walter, E., Lecourtier, Y., and Happel, J. (1984). On the structural output distinguishability of parametric models, and its relations with structural identifiability. *IEEE Trans. Autom. Control* 29, 56–57. doi: 10.1109/TAC.1984.1103379
- Woolrich, M. W., Behrens, T. E., and Smith, S. M. (2004). Constrained linear basis sets for HRF modelling using variational bayes. *Neuroimage* 21, 1748–1761. doi: 10.1016/j.neuroimage.2003.12.024
- Wu, G. R., Liao, W., Stramaglia, S., Ding, J. R., Chen, H., and Marinazzo, D. (2013). A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. *Med. Image Anal.* 17, 365–374. doi: 10.1016/j.media.2013.01.003
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Rosa, Figueiredo and Silvestre. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.