VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

Ing. MARTIN HAVLÍČEK

EXPLORING BRAIN NETWORK CONNECTIVITY THROUGH HEMODYNAMIC MODELING

ZKOUMÁNÍ KONEKTIVITY MOZKOVÝCH SÍTÍ POMOCÍ HEMODYNAMICKÉHO MODELOVÁNÍ

BIOMEDICÍNSKÁ ELEKTRONIKA A BIOKYBERNETIKA

SHORT VERSION OF PHD THESIS

ŠKOLITEL: Prof. Ing. JIŘÍ JAN, CSc.

OPONENTI:

DATUM OBHAJOBY:

KLÍČOVÁ SLOVA

Efektivní konektivita, fMRI, neuronální, hemodynamické modelování, nelineární kubaturní Kalmanův filtr, vyhlazovač, odhad parametrů, Bayesovský výběr modelu, spojitě-diskrétní systémy, adaptivní filtrace, variační Bayes.

KEYWORDS

Effective connectivity, fMRI, neuronal, hemodynamic modeling, nonlinear cubature Kalman filter, smoother, sequential parameter estimation, Bayesian model selection, pruning, continuous-discrete systems, adaptive filtering, variational Bayes.

Disertační práce je uložena na vědeckém oddělení děkanátu FEKT VUT v Brně, Technická 3058/10, 616 00 Brno.

© 2011 Ing. Martin Havlíček

ISBN 80-214-doplní redakce

ISSN 1213-4198

CONTENTS

1 INTRODUCTION 4
1.1 SCOPE AND CONTRIBUTION OF THE THESIS 7
2 ESTIMATION OF NEURONAL SIGNAL FROM FMRI DATA 10
2.1 PROBABILISTIC INFERENCE 10
2.1.1 Cubature Kalman filter and smoother 13
2.1.2 Sequential join state-parameter estimation 14
2.1.3 Continuous-discrete cubature Kalman filter 16
2.1.4 Adaptive estimation of noise statistics 17
2.2 ALGORITHM FOR ESTIMATION OF NEURONAL SIGNAL 19
3 MODELING BRAIN NETWORK CONNECTIVITY 20
3.1 STOCHASTIC DYNAMIC CAUSAL MODELING 22
3.1.1 Neuronal interaction model 24
3.2 MODEL SELECTION 24
3.2.1 Bayes factor 26
3.2.2 Network pruning 27
4 VALIDATION AND APPLICATION OF THE METHOD 29
4.1 SINGLE TIME COURSE MODEL INVERSION 29
4.2 INVERSION OF STOCHASTIC DCM 32
4.3 APPLICATION TO EMPIRICAL DATA 33
5 CONCLUSIONS 36
BIBLIOGRAPHY 39
CURRICULUM VITAE 43
ABSTRACT 45

1 INTRODUCTION

Functional magnetic resonance imaging (fMRI) utilizing the blood oxygenation level dependent (BOLD) effect as an indicator of local activity is a very useful technique to identify brain regions that are active during perception, cognition, action, but also during rest. The present research interest that dominates in fMRI neuroimaging community can be summarized by quoting Karl Friston [1]: *"A great deal of brain mapping is focused on functional segregation and the localization of function. Functional localization implies that a function can be localized in a cortical area, whereas segregation suggests that a cortical area is specialized with some aspects of perceptual or motor processing, and that this specialization is anatomically segregated within the cortex. The cortical infrastructure supporting a single function may involve many specialized areas whose union is mediated by the functional integration among them. In this view, functional segregation is only meaningful in the context of functional integration and vice versa."*. Since it is generally believed that human cognitive functions emerge from dynamic interactions of brain networks [2], it is not surprising that in the last decade there has been an increasing interest in identifying relationships among brain regions in order to better understand functional integration. This has lead to the formulation of connectivity analysis methods that attempt to identify associated brain regions and their interactions [3-5].

There are two distinct concepts of investigating brain network connectivity (integration) in fMRI data. First, there is a *functional connectivity*, which refers to correlated structures (or any other information theoretic measure) in the data such that brain areas can be grouped into interacting networks. This is usually accessed either by a pair-wise correlation (or a coherence in frequency domain) between a

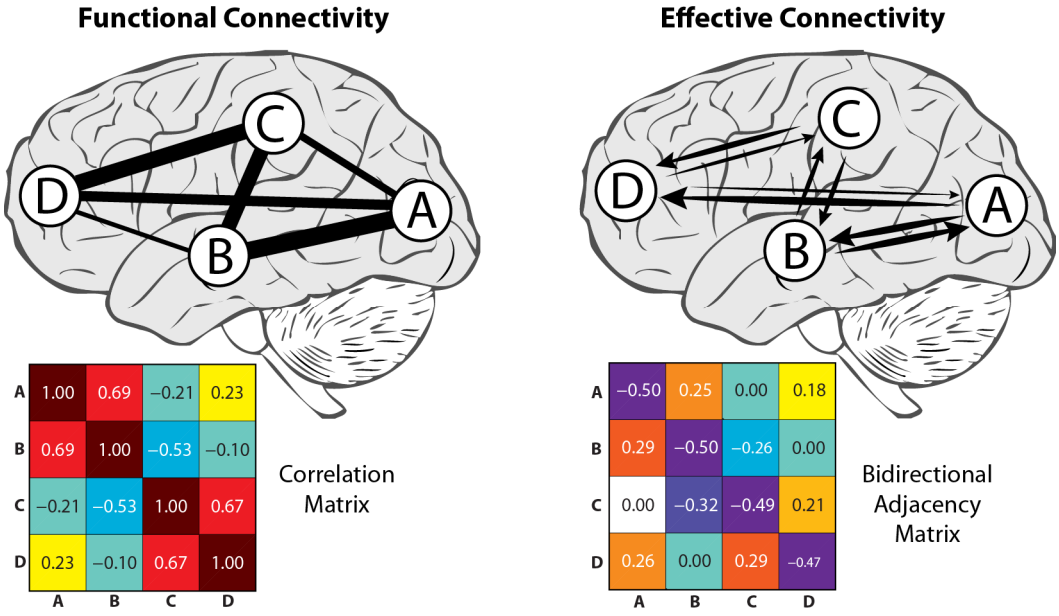

Figure 1.1 *The functional and effective connectivity. The conceptual illustration of functional connectivity (left) and effective connectivity (right) with corresponding connectivity matrix representation.*

region of interest (ROI) and the rest of the brain [6] or by a multivariate approach such as independent component analysis (ICA) [7, 8]. Second, there is *effective connectivity*, which refers to the influence that one neural system exerts over another, either at the synaptic or population level [9]. In other words, effective connectivity moves beyond statistical dependency of functional connectivity, onto measures of directed (causal) influence. This is accessed through models of interactions, which try to explain observed dependences (functional connectivity). In addition, there is a principal difference between these two concepts regarding the questions they are able to address. Critically, effective connectivity enables to distinguish between a correlation and a causation. Just because two events correlate does not mean that one has caused the other.

Evaluation of effective connectivity often requires the definition of a structural model, i.e. an assembly of brain regions (nodes) among which the causal influence is assessed. In this work, the main interest rests upon the effective connectivity framework, where an overview of the most common methods is provided in the introduction to Chapter 3.

Following the above mentioned definition of effective connectivity, it is desirable to detect causal influences among different brain regions at the neuronal (synaptic) level. This desire automatically raises an important question: Considering that the BOLD signal offers only a very indirect measure of neuronal activation, is it possible to evaluate effective connectivity at the neuronal level from fMRI data? In fMRI we measure hemodynamic responses, which reflect changes in blood flow and blood oxygenation that follow neuronal activation. Crucially, the form of this hemodynamic response can vary across subjects and different brain regions [10, 11]. These facts seriously complicate the identification of effective connectivity from fMRI [12]. However, one can reasonably justify that if it is possible to remove the effect of this hemodynamic blurring and variation, we could still achieve the aim of identifying effective connectivity from fMRI data.

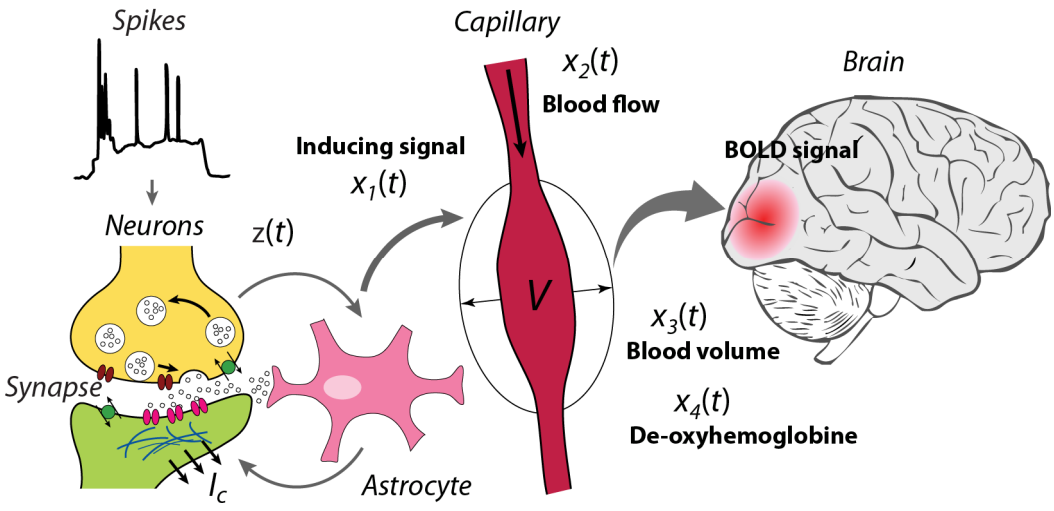

Figure 1.2 *The diagram of physiological process underlying BOLD signal.*

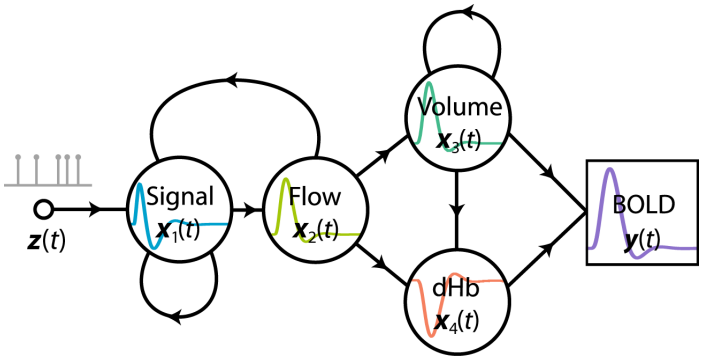

Figure 1.3 *The flowchart of nonlinear hemodynamic model.*

In general, the relationship between initial neuronal activation and our fMRI observations rests on a complex biophysiological dynamic process (see Figure 1.2). If this process is known and well described, it can be approximated by mathematical modeling. Fortunately, the physiological mechanisms mediating the relationship between neuronal activation and vascular/metabolic systems have been studied extensively [13-15] and models of hemodynamic responses have been described at macroscopic level. This hemodynamic model is nonlinear in nature [16-18]. The model flowchart is summarized in Figure 1.3. Here, the neural activity $z(t)$ causes an increase in vasodilatory signal $x_1(t)$ which is subject to auto-regulatory feedback. This flow-inducing signal is artificially designed to subsume many neurogenic and diffusive signal subcomponents. Blood flow $x_2(t)$ responds in proportion to this signal and causes changes in blood volume x_3(t) and deoxyhemoglobin content, $x_4(t)$. The dynamics of these four hemodynamic states are modeled by a set of differential equations. Finally, the observed BOLD signal y_t at discrete times is obtained as a nonlinear combination of blood volume and deoxyhemoglobin content.

Considering this model, there are several ways to perform mapping from observed data to estimated neuronal signals that interact among each other, where this is partly defined by the experimental design conducting the acquisition of the data. For example, in the case of an experiment with specific task (task data), we have prior knowledge of the stimulation paradigm (i.e. any kind of stimulus presented to the subject during a scanning session), which can be used as the definition of a driving exogenous input into the model. It is then possible to model the relationship between neuronal signals and observed responses by considering a deterministic model [3], and simply infer the model parameters to fit the data. This formulation is often unsatisfactory since unexpected contributions in the "real world", which deviate from the model, disturb the considered dynamic phenomena so that the deterministic models have a little explanatory power concerning the dynamics [19]. Therefore, it is usually preferable to consider some additive randomness to the modeled process, which then represents stochastic modeling [20, 21]. This approach is more general and is expected to have much more explanatory power than the deterministic one [22]. In a related context, there are many fMRI

studies, where the data are collected when the subject is at rest (resting-state data). In this case, there is no stimulation paradigm and therefore no exogenous input that can be used for modeling. It means that the neuronal signal, which generates observed hemodynamics, has purely endogenous character. Until very recently, this fact did not allow estimation of effective connectivity in resting-state fMRI data. It is specifically the form of stochastic modeling, which enables the estimation of neuronal signals and their interactions without any prior knowledge of exogenous input that allows evaluation of effective connectivity even in resting-state data [22-25]. Critically, the inversion of such a stochastic model leads to a blind deconvolution[1] problem, which is described as estimating the unknown input to a dynamic system, given output data, when the model of the system contains unknown parameters [26].

1.1 SCOPE AND CONTRIBUTION OF THE THESIS

In the introduction we have emphasized the general problem of estimating the effective connectivity among different brain regions from fMRI data. This problem stems from the fact that the BOLD signal is an indirect measure of the neuronal signal, and the shape of hemodynamic response function varies across different brain regions and also across subjects. In order to enable the identification of effective connectivity from fMRI data that is in the agreement with the true effective connectivity at the neuronal level, one has to solve the inverse ("deconvolution") problem. Moreover, we have also highlighted the methodological enrichment in considering stochastic representation of dynamic modeling as opposed to the limited deterministic one. Finally, we have mentioned the motivation to the inverse problem, where we do not have a prior knowledge of exogenous input, as it can be applied also to the resting-state data.

Although some attempts were already made in this direction as discussed above, there is still considerable room for improvement. Here, the main aim is to build a more accurate but less restrictive estimation approach that can be broadly applied to any fMRI data.

The successful solution to this inverse problem, i.e. successful estimation of the neuronal signal, requires the following:

1) An estimation framework that is able to handle the nonlinear characteristics of a hemodynamic model that couples neuronal activity to BOLD signal.
2) Fully stochastic modeling, since no model is completely able to catch the real world dynamics and that at any modeled physiological level, the likely contribution of the noise has to be taken into account. Moreover, the endogenous neuronal signal can be recovered only by considering a stochastic modeling.

[1] A note on terminology is needed here: although convolution is usually defined as a linear operation, the term deconvolution is generally used in reference to the inversion of nonlinear (generalized) convolution models (i.e. restoration); we adhere to this convention.

3) A robust approach to stochastic continuous-discrete modeling, because the causal chain of hemodynamic model is described in continuous time and, as required above, should also account for randomness.
4) An efficient framework for the estimation of model parameters in order to achieve a good fit of the model to the data and allow for diversity of hemodynamic responses across the brain. Additionally, this framework should preferably enable sequential modeling of conditional dependencies between parameters and modeled states.

These points define the topics that are addressed in this thesis (in the first half of the thesis in particular). When considering a suitable estimation framework that could possibly meet all the above requirements, the preference was to use and further develop new methods from the field of engineering, in the hope that their introduction to the society of computational neuroscience could raise the interest. Another important factor was to consider reasonable computational demands of the employed methods. A great deal of effort has been devoted to the introduction, description and motivation of using these methods. Specifically, we took an advantage and highlighted a recent development of new nonlinear cubature Kalman filter [27]. In this context, special attention is devoted to the joint state-parameter estimation problem. Another relevant part of the thesis describes an accurate discretization of continuous model based on local linearization scheme [28] and online Bayesian learning of measurement noise statistics.

In the latter part of the thesis, we generalize the inverse problem into the multivariate case, where multiple brain regions are involved and where the model of causal interactions at the neuronal level is considered. Critically, this introduces a new concept in evaluation of effective connectivity through stochastic dynamic causal modeling. This is accompanied by a description of the second level inference that is known as model selection. In particular, we discuss a Bayesian approaches to model selection based on different approximations of the marginal likelihood. Consequently, we introduce a simple algorithm for detection of irrelevant parameters in neuronal interaction model based on network pruning.

Finally, in the last part we validate the proposed method from different perspectives, and try to address questions, which presently dominate in the neuroscience community, regarding possible application of methods for analysis of effective connectivity.

This thesis proposes the following:

Contribution to computational neuroscience:

- A novel approach to the nonlinear modeling of hemodynamic signal, where the underlying neuronal signal is estimated from the measured BOLD time series. This approach performs a blind (nonlinear) deconvolution, where the model parameters, physiological states and mainly the endogenous input into the model (neuronal signal) are estimated from measured data.
- A new approach for the evaluation of effective connectivity based on stochastic dynamic causal modeling. This enables inversion of full connectivity models without knowing the driving input or having a hypothesis about the connectivity structure. This means that the *a priori* unknown model of neuronal interactions is learned from the data.

Contribution to engineering methodology:

- Formulation of combined use of cubature Kalman filtering and Rauch-Tung-Striebel smoothing in system identification for joint estimation of the hidden states, model parameters, and endogenous model input. The convergence is supported by an iterative scheme that automatically maximizes the log-likelihood.
- A new algorithm for estimation of continuous-discrete state-space models based on combination of the (square-root) cubature Kalman filter and local linearization scheme, which provides an accurate and stable discretization of a continuous model represented by stochastic differential equations.
- A new nonlinear adaptive Kalman filter for joint estimation problem, where the measurement noise covariance is effectively learned through a variational Bayesian approach, and parameter and state noise covariance are estimated by the Robbins-Monro stochastic approximation scheme.

Note that this short version of dissertation thesis contains only a very limited mathematical description and validation of the developed approach. Any potential reader is strongly encouraged to see the complete version, which can be downloaded at: https://sites.google.com/site/havlicekmartin .

2 ESTIMATION OF NEURONAL SIGNAL FROM FMRI DATA

The nonlinear hemodynamic model naturally forms the state-space model, where the measured data are related to the subset of state-space variables (physiological states) by an observation equation. In this continuous-discrete time dynamic system, which represents a generative model of the BOLD signal, both state and observation equations are nonlinear and polluted by physiological and instrumental noise, respectively. In general, the estimation of the state of a continuous system from noisy discrete observations can be performed using the nonlinear filter theory, which is an extension of the original framework (Kalman filter theory) formulated to provide a sequential and computationally efficient solution to the linear filtering and prediction problems [29]. Finding the optimal nonlinear system identification method (i.e. the estimation of the model parameters and the trajectories of unobservable states) is an active research area.

In this chapter we introduce a new approach to this identification problem. In particular, we will first provide a short introduction to the probabilistic inference based on optimal recursive Bayesian solution. Since this solution is tractable only for linear systems, we will focus on very recent developments in nonlinear Kalman filtering based on efficient cubature integration rules [27]. This numerical tool called cubature Kalman filtering will serve as the cornerstone for further extensions and developments. Specifically, we will describe the cubature Rauch-Tung-Striebel smoother to obtain more accurate estimates of the state, including efficient square-root implementation. Next, we propose the joint estimation framework to simultaneously infer the hidden states and model parameters. We will also introduce a new filtering approach for hybrid continuous-discrete systems based on an accurate discretization scheme of stochastic differential equations called local linearization combined with the above mentioned cubature integration rules. Consequently, in order to make the algorithm easily adaptable to the real data, we will discuss the extension to adaptive filtering through Bayesian estimation of the measurement noise covariance [27], and Robbins-Monro approximation of the parameter and state noise covariance matrices [30].

Finally, all these extensions and developments will be combined into one single algorithm, which will represent a new approach to estimation of neuronal signal from BOLD responses; i.e. the blind (nonlinear) deconvolution approach where all the hemodynamic states, the model parameters and mainly the input (neuronal signal) are estimated from observed BOLD responses.

2.1 PROBABILISTIC INFERENCE

The problem of estimating the hidden states (causing data), parameters (influencing the dynamics of hidden states) and any non-controlled endogenous input to the system, in a situation when only observations are given, requires

probabilistic inference. If we interpret our data through a dynamic state-space model (DSSM), then we are facing the sequential (recursive) probabilistic inference problem.

Assuming the first-order Markov process, a discrete dynamic state-space system is described by a pair of equations:

$$\mathbf{x}_t = \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{u}_{t-1}; \boldsymbol{\theta}) + \mathbf{q}_{t-1} \tag{2.1}$$
$$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t, \mathbf{u}_t; \boldsymbol{\theta}) + \mathbf{r}_t, \tag{2.2}$$

where the first equation represents the system (state) model, describing the evaluation of the states $\mathbf{x}_t$ as a function of time. Here, $\mathbf{q}_{t-1}$ is the process (state) noise that drives the dynamic system through an arbitrary (possibly nonlinear and time-varying) transition function $\mathbf{f}$, and $\mathbf{u}_t$ is the exogenous input to the system that is usually assumed known (though later in this thesis $\mathbf{u}_t$ will be considered unknown). The second equation represents the measurement (observation) model, where the measurement noise $\mathbf{r}_t$ corrupting the observation of the (hidden) states through arbitrary observation function $\mathbf{h}$. Both $\mathbf{f}$ and $\mathbf{h}$ can be parameterized using a set of parameters $\boldsymbol{\theta}$. In a Markovian setting, the current state $\mathbf{x}_t$ depends only on the immediate past state $\mathbf{x}_{t-1}$ through the state-transition distribution $p(\mathbf{x}_t|\mathbf{x}_{t-1})$; i.e. conditional probability density. The observations $\mathbf{y}_t$ are conditionally independent, given the state, and are generated according to the observation likelihood $p(\mathbf{y}_t|\mathbf{x}_t)$ [31]. Therefore, the dynamic state-space model, together with the known statistics of the noise (and the prior distribution of the system states), defines a probabilistic generative model of how system evolves over time and how we (partially or inaccurately) observe this hidden state.

The optimal solution to the above inference problem is given by the recursive Bayesian estimation algorithm, which recursively updates the posterior density of the system state $p(\mathbf{x}_t|\mathbf{y}_t)$ as new observations arrive. Generally in Bayesian framework, the posterior density of the states $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ given all the observations $\mathbf{y}_{1:t}$, embodies the complete solution to the probabilistic inference problem. In other words, $p(\mathbf{x}_t|\mathbf{y}_{1:t})$ contains all information necessary to calculate an optimal estimate

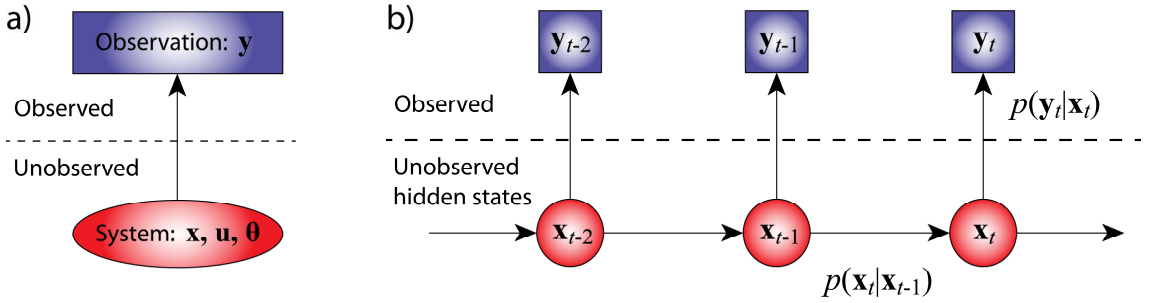

Figure 2.1 *Schematic diagrams of probabilistic inference. (a) Given noisy observation* $\mathbf{y}$*, what can we infer about system state, parameters or input. (b) Graphical model of a probabilistic dynamic state-space model. This representation is also known as a directed acyclic graph (DAG) in the graph theory field.*

of the state, such as the conditional mean:

$$\hat{\mathbf{x}}_{t|t} = \mathbb{E}[\mathbf{x}_t|\mathbf{y}_{1:t}] = \int_{\mathbb{R}^{n_x}} \mathbf{x}_t \, p(\mathbf{x}_t|\mathbf{y}_{1:t}) d\mathbf{x}_t, \tag{2.3}$$

and the covariance matrix, as a measure of accuracy of the estimate $\hat{\mathbf{x}}_{t|t}$:

$$\begin{aligned} \mathbf{P}_{t|t} &= \mathbb{E}\left[\left(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}\right)\left(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}\right)^T\right] \\ &= \int_{\mathbb{R}^{n_x}} \left(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}\right)\left(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}\right)^T p(\mathbf{x}_t|\mathbf{y}_{1:t}) d\mathbf{x}_t. \end{aligned} \tag{2.4}$$

For linear and Gaussian dynamic systems, where $\mathbf{f}$ and $\mathbf{h}$ are linear functions and additive noise and state prior distributions are Gaussian, the solution to the filtering recursion is obtained by well known *Kalman filter* [32]. The Kalman filter provides the posterior *filtering* distribution of the state $\mathbf{x}_t$ at time t given the history of the measurement up to the time step t, $p(\mathbf{x}_t|\mathbf{y}_{1:t})$. However, it is also possible to obtain the *smoothing* distribution $p(\mathbf{x}_t|\mathbf{y}_{1:T})$, where the posterior of the state is computed at the time step t after receiving the measurements up to time step T, where $T > t$. Typically, a smoother is statistically more accurate than filter, as it use more observations. The solution to the forward-backward smoothing is then obtained via *Rauch-Tung-Striebel smoother* (RTS) [33]; i.e. fixed interval Kalman smoother.

Unfortunately, in more realistic environment, which is nonlinear and possibly non-Gaussian, the optimal Bayesian recursion is intractable and an approximate solution must be used. Numerous approximation solutions to the recursive Bayesian estimation problem have been proposed over the last couple of decades. Here we mention only the basic categories: Gaussian approximate methods [34]; direct numerical integration methods [35]; sequential Monte Carlo methods [36]; variational Bayesian methods [37]. More detailed overview can be found in the full version of the thesis.

The selection of suitable sub-optimal approximate solutions to the recursive Bayesian estimation problem represents a trade-off between global optimality on one hand and computational tractability (and robustness) on the other hand. In our case, the best criterion for sub-optimality is formulated as: "Do as best as you can, and not more". Under this criterion, the natural choice is to apply the cubature Kalman filter [27]. The CKF is the closest known direct approximation to the Bayesian filter, which outperforms all other nonlinear filters in any Gaussian setting, including particle filters [27, 38, 39]. The CKF is numerically accurate, can capture true nonlinearity even in highly nonlinear systems, and it is easily extendable to high dimensional problems (the number of sample points grows linearly with the dimension of the state vector). The CKF belongs to the group of so-called Gaussian assumed density filters, which are considered as local approximation methods.

2.1.1 Cubature Kalman filter and smoother

Cubature Kalman filter

The cubature Kalman filter [27] is a recursive, nonlinear and derivative free filtering algorithm, which computes the first two moments (i.e. mean and covariance) of all conditional densities by using the third-degree cubature integration rules to approximate *n*-dimensional Gaussian weighted integrals; i.e. integrals of the form $nonlinear\ function \times Gaussian\ density$. Critically, this cubature rule defines a way how to deterministically select a set of cubature points, and their corresponding weights, so that they completely capture the true mean and covariance of the prior random variable $\mathbf{x}\sim\mathcal{N}(\bar{\mathbf{x}}, \mathbf{P})$:

$$\int_{\mathbb{R}^{n_x}} \mathbf{f}(\mathbf{x})\mathcal{N}(\mathbf{x}; \bar{\mathbf{x}}, \mathbf{P})d\mathbf{x} \approx \sum_{i=1}^{2n} w_i \mathbf{f}(\boldsymbol{\mathcal{X}}_i), \tag{2.5}$$

where the weights are simply $w_i = \frac{1}{2n}$, with n equals the state dimension, and

$$\boldsymbol{\mathcal{X}}_i = \bar{\mathbf{x}} + \boldsymbol{\xi}_i\sqrt{\mathbf{P}}, \quad i = 1, \dots, 2n. \tag{2.6}$$

This involves factorization of error covariance matrix $\mathbf{P} = \sqrt{\mathbf{P}}\sqrt{\mathbf{P}}^T$ and the elementary cubature points are:

$$\boldsymbol{\xi}_i = \begin{cases} \sqrt{n}\mathbf{e}_i, & i = 1, 2, \dots, n \\ -\sqrt{n}\mathbf{e}_i, & i = n+1, n+2, \dots, 2n. \end{cases} \tag{2.7}$$

Here $\mathbf{e}_i$ represents the *i*-th column vector, whose *i*-th entry is a unit and all other entries are zero. From this definition, it can be seen that the cubature points are distributed uniformly on a sphere centered at the origin, and their number increases linearly with the state dimension. Unlike the most common nonlinear filter called the extended Kalman filter, CKF effectively approximates both the Jacobian and Hessian accurately (in statistically average sense) through its cubature point propagation, without the need to perform any analytic differentiation.

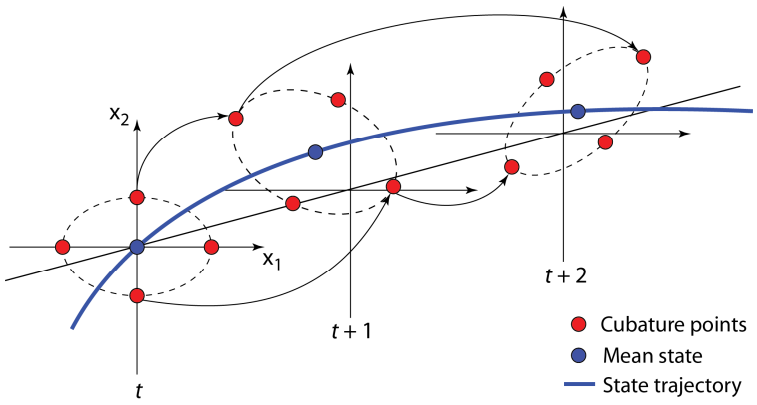

Figure 2.2 *Illustration of cubature points propagation during time update of CKF. The cubature points in the two-dimensional state-space are propagated between time steps. The circles represent cubature points; the new cubature point set at time* $t + 1$ *is computed by simply propagating the old cubature point set at time* t *through the process equation.*

In order to evaluate the dynamic state-space model described by (2.1)-(2.2), the CKF includes two standard Kalman filter steps: a) a time update, after which the predicted density $p(\mathbf{x}_t|\mathbf{y}_{1:t-1}) = \mathcal{N}(\hat{\mathbf{x}}_{t|t-1}, \mathbf{P}_{t|t-1})$ is computed; and b) a measurement update, after which the posterior density $p(\mathbf{x}_t|\mathbf{y}_{1:t}) = \mathcal{N}(\hat{\mathbf{x}}_{t|t}, \mathbf{P}_{t|t})$ is computed.

Cubature Rauch-Tung-Striebel smoother

The same approximation principles that were used in cubature Kalman filter can be applied also during the backward pass of the Rauch-Tung-Striebel (RTS) smoother, yielding the cubature RTS smoother. The backward pass is used for computing suitable corrections to the forward filtering results to obtain the smoothing solution $p(\mathbf{x}_t, \mathbf{y}_{1:T}) = \mathcal{N}\left(\hat{\mathbf{x}}_{t|T}|\hat{\mathbf{x}}^s_{t|T}, \mathbf{P}^s_{t|T}\right)$. Because the filtering and smoothing estimates of the last time step T are the same, we make $\hat{\mathbf{x}}^s_{T|T} = \hat{\mathbf{x}}_{T|T}$, $\mathbf{P}^s_{T|T} = \mathbf{P}_{T|T}$. This means the recursion can be used for computing the smoothing estimates of all time steps by starting from the last step $t = T$ and proceeding backward to the initial step $t = 0$. To accomplish this, all estimates of $\hat{\mathbf{x}}_{0:T}$ and $\mathbf{P}_{0:T}$ from the forward pass have to be stored to be later reused during the backward pass. Note that we will use an abbreviation CKS to refer to the forward run of cubature Kalman filter followed by the backward run of the cubature RTS smoother.

A detail description of CKF and CKS algorithms including their efficient square-root implementations are provided in the full version of the thesis.

2.1.2 Sequential join state-parameter estimation

It is very often the case that both hidden states of dynamic process and model parameters are unknown and have to be inferred from the measured data. Moreover, there might be even a situation, where one wants to estimate also the unknown input into the system. This special case of system identification can be consider as a blind (nonlinear) deconvolution problem, which is described as estimating the unknown input to a dynamic system, given output data, when the model of the system contains unknown parameters. The nonlinear cubature Kalman framework is a well suited approach to robust parameter estimation. What we should add now is the fact that because of so-called natural condition of control [27], it is possible to generate the input $\mathbf{z}_t$ using the state prediction $\hat{\mathbf{x}}_{t|t-1}$. In our case it means that if we augment the state vector by the process representing the input with input noise $\mathbf{v}_t \sim \mathcal{N}(0, \mathbf{V}_t)$ and by the process for parameters (Gaussian random walk model) with process noise $\mathbf{o}_{t-1} \sim \mathcal{N}(0, \mathbf{O}_{t-1})$, we can estimate the system input and parameters together (jointly) with the states. By saying this, we are trying to solve a dual estimation problem, where under consideration of a nonlinear dynamic system, the system states $\mathbf{x}_t$, the parameters $\boldsymbol{\theta}_t$ and the input $\mathbf{z}_t$, are estimated simultaneously from the observed noisy signal $\mathbf{y}_t$.

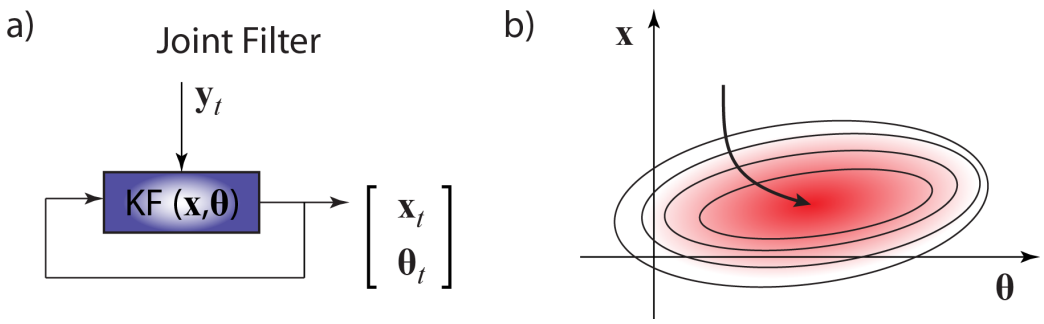

Figure 2.3 *Illustration of joint filtering scheme. (a) Both model states and parameters are estimated simultaneously in the augmented form of the state vector. (b) Joint state-parameter optimization space.*

It should be noted that by the input $\mathbf{z}_t$ we mean an endogenous input (or signal), which might be different from the exogenous input $\mathbf{u}_t$. In the context of fMRI, the input $\mathbf{u}_t$ is presented to the subject, whereas $\mathbf{z}_t$ reflects the actual neuronal response, which might (or might not) reflect the exogenous stimulus. In other words, there is always some endogenous activity present in brain even in the absence of any external stimuli, i.e. at rest.

A general theoretical and algorithmic framework for dual Kalman filter based estimation was presented in [40], [31]. This framework encompasses two main approaches, namely joint estimation and dual estimation. In the dual filtering approach, two Kalman filters are run simultaneously (in an iterative fashion) for a state and a parameter estimation. At every time step, the current estimate of the parameters $\boldsymbol{\theta}_t$ is used in the state filter as a given (known) input and likewise, the current estimate of the state $\hat{\mathbf{x}}_t$ is used in the parameter filter. This results in a step-wise optimization within the joint state-parameter space. On the other hand, in the joint filtering approach, the unknown system state and parameters are concatenated into a single higher-dimensional joint state vector, $\boldsymbol{\mu}_t = [\mathbf{x}_t^T, \mathbf{z}_t^T, \boldsymbol{\theta}_t^T]^T$. This results in a smoothed convergence in the joint state-parameter space (see Figure 2.3). There is a prevalent opinion that the performance of joint estimation scheme is superior to dual estimation scheme [31, 40, 41]. Therefore, the joint estimation framework based on cubature Kalman filtering and smoothing is considered in this work.

The state-space model for joint estimation scheme is then formulated as:

$$\boldsymbol{\mu}_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{z}_t \\ \boldsymbol{\theta}_t \end{bmatrix} = \begin{bmatrix} \mathbf{f}(\mathbf{x}_{t-1}, \boldsymbol{\theta}_{t-1}, \mathbf{z}_{t-1}) \\ \mathbf{z}_{t-1} \\ \boldsymbol{\theta}_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{q}_{t-1} \\ \mathbf{v}_{t-1} \\ \mathbf{o}_{t-1} \end{bmatrix} \tag{2.8}$$

$$\mathbf{y}_t = \mathbf{h}(\boldsymbol{\mu}_t) + \mathbf{r}_t. \tag{2.9}$$

Since the joint filter concatenates the state and parameter variables into a single state vector, it effectively models the cross-covariances between the state, input and parameters estimates:

$$\mathbf{P}_t = \begin{bmatrix} \mathbf{P}_{x,t} & \mathbf{P}_{xz,t} & \mathbf{P}_{x\theta,t} \\ \mathbf{P}_{zx,t} & \mathbf{P}_{z,t} & \mathbf{P}_{z\theta,t} \\ \mathbf{P}_{\theta x,t} & \mathbf{P}_{\theta z,t} & \mathbf{P}_{\theta,t} \end{bmatrix}. \quad (2.10)$$

This full covariance structure allows the joint estimation framework not only to deal with uncertainty about parameter and state estimates (through the cubature-point approach), but also to model the interaction (conditional dependences) between the states and parameters, which generally provides better estimates [31, 40]. Note that since the parameters are estimated simultaneously with the states, the convergence of parameter estimates depends also on the length of the time series. Therefore, we propose a joint estimation scheme, where iterating cubature Kalman filter and smoother steps automatically maximizes the log-likelihood and provides a fast convergence of model parameters and accurate estimates of hidden states including the endogenous input.

2.1.3 Continuous-discrete cubature Kalman filter

In previous sections, we have considered the state-space model to be described in a discrete time, however, in case of hemodynamic model, the process equations of state-space model are derived from underlying physics of a continuous dynamic system, and are expressed in the form of a set of differential equations. But still, the measurements $\mathbf{y}_t$ are acquired by digital devices; i.e. they are available at discrete time points $(t = 1,2, \dots, T)$. Therefore, we have a model with a continuous process equation and a discrete measurement equation. The stochastic representation of this state-space model, with additive noise, can be formulated as:

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t)dt + \sqrt{\mathbf{Q}}d\mathbf{w}(t) \quad (2.11)$$

$$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t, t) + \mathbf{r}_t, \quad (2.12)$$

where $\mathbf{x}(t)$ is the state of the system at time t; $\mathbf{f}(.)$ is a known nonlinear drift function; $\mathbf{w}(t)$ denotes standard Brownian motion that is independent of $\mathbf{x}(t)$; and $\mathbf{Q}$ is a known diffusion matrix. The process equation is the simplest form of Itô's stochastic differential equation [42]. The system is observed through the noisy measurements in discrete time intervals (discrete times are denoted as subscripts).

The recursive Bayesian solution to the above continuous-discrete model is very similar to the standard Kalman filter. The only difference appears during the time update step of Kalman filter, where the old posterior density is propagated trough the process equation (2.11). In this case, the probability density of the state at time t obeys the *Fokker-Plank equation* (FPE), which describes the evolution of probability density between the measurement time instants. The exact solution to FPE is available only for linear Gaussian system represented by the time update of Kalman-Bucy filter [43]. In other cases, the FPE has to be approximated. Here we consider methods, which compute a finite number of summary statistics in terms of conditional moments after discretizing the continuous time process equation. However, the aim will not be to develop an approximate nonlinear filter by

approximating continuous time filter equations (as it is often the case), but rather to use the standard discrete time filtering equations for an approximate discrete time model of the original continuous time dynamical system [28]. This automatically puts high demand on accuracy of the method that discretizes the model represented by stochastic differential equations. In this sense we are seeking a good discrete time approximation of the continuous stochastic dynamical model for nonlinear and Gaussian multivariate process $\mathbf{x}(t)$ that is consistent and stable.

We introduce a new approach to continuous-discrete filtering that combines statistical linearization using cubature rules with local linearization (LL) scheme for accurate and stable discretization of continuous model. The LL scheme intuitively assumes the nonlinear function $\mathbf{f}(.)$ to be locally linear with respect to the process $\mathbf{x}(t)$, where its Jacobian $\mathbf{J}_x$ is constant:

$$\mathbf{x}_{t+\delta} = \mathbf{x}_t + \mathbf{J}_x^{-1}[\exp(\mathbf{J}_x\delta) - \mathbf{I}]\mathbf{f}(\mathbf{x}_t). \tag{2.13}$$

This local linearization equation corresponding to the first order linear approximation [28]. Additionally, it is possible to apply the LL scheme also to the random term of the state equation in (2.11) that follows multivariate normal distribution with zero mean vector and covariance matrix [19]:

$$\mathbf{Q}_{t+\delta} = \int_t^{t+\delta} \exp(\mathbf{J}_x\delta)\sqrt{\mathbf{Q}_t}\sqrt{\mathbf{Q}_t}^T \exp(\mathbf{J}_x\delta)^T\, dt\,. \tag{2.14}$$

The combination of cubature Kalman filter with local linearization provides a very accurate estimates of hidden states even in case when a larger integration step δ is applied. Moreover, this scheme is well suited to the joint estimation framework that was discussed earlier. More detailed derivation of this approach is provided in the full version of the thesis.

2.1.4 Adaptive estimation of noise statistics

The Kalman formulation of filtering problem assumes complete *a priori* knowledge of the process and measurement noise statistics. In most practical situations, these noise statistics are unknown or not known perfectly. When incorrect prior statistics are used to implement sequential filtering algorithm, it might result in suboptimal performance and possibly in filter divergence. Therefore, in the lack of system statistics knowledge, it is desirable to adaptively estimate the process noise and measurement noise statistics simultaneously with the system state. Since we already have the preference to perform joint estimation of the states and parameters, it is a logical choice to adaptively estimate the noise statistics as well.

In standard Kalman filtering framework, all noise statistics are described by the first two statistic moments, i.e. by the mean and the covariance, where the mean is usually assumed to be zero. Therefore, the goal of adaptive filtering is to estimate the covariances of process noise and measurement noise.

In the case of parameter estimation, the amount of oscillations in the model prediction clearly depends on the value of the (parameter) process noise covariance. As a result, this covariance can be used as a regularization mechanism to control the smoothness of the prediction.

Relatively robust approach to recursive estimation of noise covariance is the Robbins-Monro (RM) stochastic approximation [44]:

$$\mathbf{O}_t = \lambda_\theta \mathbf{O}_{t-1} + \left(\lambda_\theta^{-1} - 1\right)\mathbf{K}_t\left[\mathbf{y}_t - \mathbf{g}\left(\mathbf{x}_t; \widehat{\boldsymbol{\theta}}_t\right)\right]\left[\mathbf{y}_t - \mathbf{g}\left(\mathbf{x}_t; \widehat{\boldsymbol{\theta}}_t\right)\right]^T \mathbf{K}_t^T \quad (2.15)$$

The method assumes that the covariance of the Kalman update model should be consistent with the actual update model. Here, $\lambda_\theta \in (0,1]$ is called the forgetting factor, which allows to put exponentially decaying weigh on past data, and $\mathbf{K}_t$ is the Kalman gain. It is known that an RM approximation provides a very fast convergence and a low final MMSE [31].

The RM helps to escape poor local minima of the error surface. We made a choice to apply the RM approximation to the recursive estimation of parameter noise covariance $\mathbf{O}_t$ and also to approximate the process noise covariance matrix $\mathbf{Q}_t$, since it proved to have better convergence properties [31]. We should also note that we do not expect to estimate the exact process noise covariance of the dynamic model with RM approach. The aim here is to maintain some artificial level of randomness, which supports the convergence of the algorithm, and prevents the filter from becoming overconfident with the estimate (i.e. it avoids overfitting). By saying this, we consider the noise covariance of the input to be fixed. In this case, any attempt to adaptively estimate the input noise covariance led to the quick divergence of the filter.

The most important part in adaptive filtering is accurate estimation of measurement noise statistics. For this particular task, we consider a recently introduced variational Bayesian (VB) approach [45] to recursive estimation of measurement noise covariance, which is suitable also to nonlinear filtering and is able to take the advantage of assumed Gaussian density filter such as CKF.

If the measurement noise covariance $\mathbf{R}_t$ is unknown then the goal of Bayesian filtering is to compute the joint posterior distribution $p(\mathbf{x}_t, \mathbf{R}_t|\mathbf{y}_{1:t})$ of the state $\mathbf{x}_t$ and of the covariance $\mathbf{R}_t$. The posterior is given by a product of observation likelihood and predictive distribution, which now takes a form:

$$p(\mathbf{x}_t, \mathbf{R}_t|\mathbf{y}_{1:t}) \approx p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{R}_t)p(\mathbf{x}_t, \mathbf{R}_t|\mathbf{y}_{1:t-1}). \quad (2.16)$$

At this point, in order to make the computation of posterior (2.16) tractable, it is possible to apply the VB approach. The VB usually applies the mean field approximation [37] that factorizes the posterior distribution as follows:

$$p(\mathbf{x}_t, \mathbf{R}_t|\mathbf{y}_{1:t}) \approx Q_x(\mathbf{x}_t)Q_R(\mathbf{R}_t), \quad (2.17)$$

where $Q_x(\mathbf{x}_t)$ and $Q_R(\mathbf{R}_t)$ are in our case the approximations of normal and inverse-Gamma densities, respectively. It is further assumed that the measurement noise covariance has a form of diagonal matrix $\mathbf{R}_t = \mathrm{diag}(R_{1,t}, \dots, R_{d,t})$. The inverse-Gamma distribution is chosen because it represents the conjugate prior distribution

for variance of Gaussian distribution. The VB approximation can now be formed by minimizing the Kullback-Leibler (KL) divergence between the separable approximations and the true posterior distribution [45], where the solution can be obtained iteratively along with the measurement update of Kalman filter.

We have extended this approach also to the nonlinear models, where we use a cubature integration rules to approximate the coefficients of inverse-Gamma distribution, which then provide the estimate of measurement noise covariance matrix. The complete description is provided in the full version of the thesis.

2.2 ALGORITHM FOR ESTIMATION OF NEURONAL SIGNAL

Based on the developments and extensions described in the previous sections, we are able to build the complete algorithm for estimation of neuronal signal from fMRI data. In this sense, we are introducing a novel algorithm that is able to solve a triple estimation problem, i.e. we jointly estimate not only the model states (including the endogenous input), the model parameters, but also hyperparameters that represent noise statistics.

In particular, we have proposed an approach based on nonlinaer cubature Kalman filtering [27] to the joint estimation problem where the hidden states and parameters are concatenated into a single joint state vector and estimated simultaneously, yielding joint maximum *a posteriori* estimates. This form allows us not to only accurately treat the dual uncertainty of the parameter and state estimates (by using cubature point approach), but also to accurately model the interaction (conditional dependence) between the states and parameters. To obtain more accurate estimates of hemodynamic states and mainly of neuronal signal, we have employed forward-backward smoothing, encompassing also the cubature points formulation of Rauch-Tung-Striebel smoother. The overall estimator performance is further enhanced by considering a square-root formulation that ensures a numerical stability during the recursion.

Next, because the states of hemodynamic model are represented by ordinary differential equations, we have introduced a novel continuous-discrete time representation of CKF that combines a statistical linearization with the local linearization approach for accurate and stable discretization of the process model. This new algorithm is also suitable for joint estimation, where the states are propagated through the continuous model and the parameters are propagated through the discrete model.

The estimation framework would not be complete if we were restricted to the informed model inversion, where one assumes the noise statistic to be known, which is typically not the case. Therefore, we have proposed an adaptive estimation of state and parameter noise covariance matrix based on Robbins-Monro stochastic approximation scheme, and further adopted the recently introduced variational Bayesian approach for estimation of measurement noise variance [45] to the cubature point Kalman filter.

All these developments and extensions were combined to create an iterative optimization method, which maximizes the log-likelihood with each iteration and achieves a fast convergence. As a result, we have obtained a novel advanced approach to the estimation of the neuronal signal from the observed BOLD signal superior to what has been so far introduced in the field of neuroscience (comparison is provided in [26]). Additional demonstration of the algorithm can be found in the full version of the thesis, specifically in the Chapter 4 and in the Appendix A.

3 MODELING BRAIN NETWORK CONNECTIVITY

In this chapter, we focus on modeling coupling among different brain regions (nodes) in terms of effective connectivity. In particular, we will introduce a direct generalization of the estimation framework described in the previous chapter to a multivariate case, where the main goal will be to infer the directional influence among different brain regions at the neuronal level. Before we do so, it will be useful to provide a short overview and motivation on methods that attempt to assess effective connectivity.

In effective connectivity, the neuronal states describe the activity of set of nodes that comprise a graph. The aim of analysis is to identify the directional (causal) influence of activated links in the graph. Importantly, these nodes are in fMRI defined as neural populations at macroscopic level, i.e. whole brain areas, whose activity is summarized by a time varying state vector.

In general, there are two streams of statistical causal modeling: one based on Bayesian dependency graphs or graphical models called structural causal modeling and the other based on causal influence over time [1]. Structural causal modeling is related to structural equation modeling (SEM) [46, 47] and uses graphical models in which direct causal links are encoded by directed edges. However, this approach has two limitations. First, it is restricted to discovering conditional independencies in directed acyclic graphs (DAG), i.e. it cannot deal with (cyclic) feedback loops. This is a serious drawback because the brain works as a directed cyclic graph, where every brain region is connected reciprocally (at least polysynaptically) [22]. Second, the estimation is completely based on the sample covariance matrix, i.e. it ignores time dynamics. Fortunately, the DAG restriction can be finessed by considering dynamics and temporal precedence within structural causal modeling. This is because the arrow of time can be used to convert a directed acyclic graph into a cyclic graph when the nodes are deployed over successive time points. This leads to SEM with time-lagged data described by autoregressive (AR) models, which are the ground for Granger causality modeling (GCM). The GCM approach is based on temporal precedence, i.e. A causes B if one reduces uncertainty about the future of B given the past of A. It is formulated in discrete time analysis framework, where the directionality is usually inferred directly from measured signals. Although GCM has

become quite popular in the neuroimaging community during the last several years, there is an ongoing discussion to determine if the concept of temporal precedence is suitable for application to fMRI time series analysis. Main concerns are that GCM does not account for variability in hemodynamic response function across different brain regions [12, 48]; the measurement noise can reverse the estimation of causality direction [49]; and the coupling strengths are parameterized in terms of regression coefficients, which are not the true coupling parameters of effective connectivity. Additionally, the reliability of GCM degrades with the increase of sampling interval [5], which is important for fMRI because the sampling interval is quite large with respect to the time scale of neuronal events.

As a result of these recent discussions [50-53], it is now clear that discovering effective connectivity should be based on state-space models of controllable (causal in the control theory sense) biophysical processes that have hidden neuronal states and possibly exogenous input [54]. Further, the optimal statistical procedure is to invert the complete generative model described by a set of state equations that quantify how the observed data are affected by the presence of causal links [55]. This possibly allows to accommodate the conditional dependencies between parameters of the state equations, which are mapped to the observations [22]. If we now recall the definition of the effective connectivity as stated in Chapter 1, i.e. that effective connectivity refers to the influence that one neural system exerts over another, either at synaptic or population level, one realizes that the procedure mentioned above is a sensible choice and, at the moment, probably also the only choice suitable for application to fMRI data.

This reasoning has led to the development of dynamic causal modeling (DCM), which employs biophysically motivated generative model that relates the observed BOLD data to neuronal signal [3]. Here, the causal influence is defined as a physical influence, where changing influences causes changes in their consequences [55], and it is modeled by a continuous time dynamic state-space system. The original formulation of DCM requires knowledge of known exogenous input, assigned to some of the network nodes, which drives the dynamics of the system. In this case, all hidden states are treated deterministically and the random term is considered only at the level of observation equation. The coupling and hemodynamic parameters are inferred through variational Bayesian formulation of EM algorithm, which maximizes the model evidence [3]. In this scenario, DCM is seen as a hypothesis-based approach to understanding distributed neuronal architectures underlying observed brain responses. Then, different hypotheses (model candidates) represented by different networks (or graphs) are compared based on the model fit reflected in evidence, via Bayesian model selection (BMS) [56]. However, as we emphasized in Chapter 1, one can do better if the model accounts for randomness at all levels, including hidden states, i.e. it is formulated as a fully stochastic system.

Since we have introduced a fully stochastic scheme in the previous chapter for simultaneous estimation of neuronal signal (i.e. endogenous input) and model parameters for single time course, it is reasonable to think that the same scheme can

be extended to multivariate case, which includes the modeling of neuronal interactions among different brain regions. In other words, we do not utilize any prior knowledge about the experimental causes of observed responses as required by deterministic DCM and introduce a stochastic DCM, which can be completely data-driven. This enables network discovery using both observed and unobserved responses during both activation based studies and (task-free) studies of autonomous or endogenous activity during the resting state [22]. In addition, because we jointly estimate both model parameters and neuronal signals in temporally sequential sense, i.e. we estimate the hidden states generating observed data, while properly accommodating endogenous inputs and model parameters, we implicitly assume that the uncertainty about the parameter estimates depends on uncertainty of hidden states (including endogenous inputs). This is more proper assumption compared to the deterministic DCM, which assumes that the uncertainty about parameters (after seeing data) does not depend on uncertainty about the states [22].

3.1 STOCHASTIC DYNAMIC CAUSAL MODELING

In this section we propose method for evaluation of stochastic DCM (sDCM). This method represents a straightforward extension to the deterministic DCM (dDCM), when it has the following properties: (i) it releases the need of known exogenous input; (ii) accounts for random process at hidden states level; (iii) and provides conditionally dependent estimates of the states and parameters.

Similarly to dDCM, sDCM is formulated as a multiple-input multiple-output (MIMO) system that comprises *m* inputs and *l* outputs with one output per region. Unlike in dDCM, where inputs must correspond to causes, $\mathbf{u}_t$ (i.e. designed exogenous inputs), in sDCM the inputs can be treated as endogenous $\tilde{\mathbf{z}}_t$, i.e. they can be generated by the fMRI data, which makes sDCM data-driven approach. However, it does not mean that sDCM is limited only to this scenario. Importantly, as we will see in the following section, one can still define any exogenous input as in the case of dDCM and use the sDCM for testing different hypotheses that motivated the experimental design but with the fully stochastic treatment of the model. Also, the stochastic formulation of DCM can be always easily converted to deterministic one, by setting the process noise variances of hidden states to zero (or to very small values). In either case, DCM rests on a choice of neuronal model $\mathbf{f}^n(.)$ of interacting cortical regions, which is defined in continuous time. This neuronal model is further supplemented with a forward hemodynamic model (summarized by $\mathbf{f}^h(.)$ and $\mathbf{h}(.)$), which describes how neuronal or synaptic activity is transformed into a measured response $\mathbf{y}_t$. This complete generative model allows to estimate the neuronal model parameters $\boldsymbol{\theta}_t^n$ (i.e. effective connectivity) from observed data, where the parameters represent couplings among unobserved brain states (i.e. neuronal activity in different brain regions), but it also accounts for parameterization of the hemodynamic response, $\boldsymbol{\theta}_t^h$. Then, the state-space model can have the following joint form:

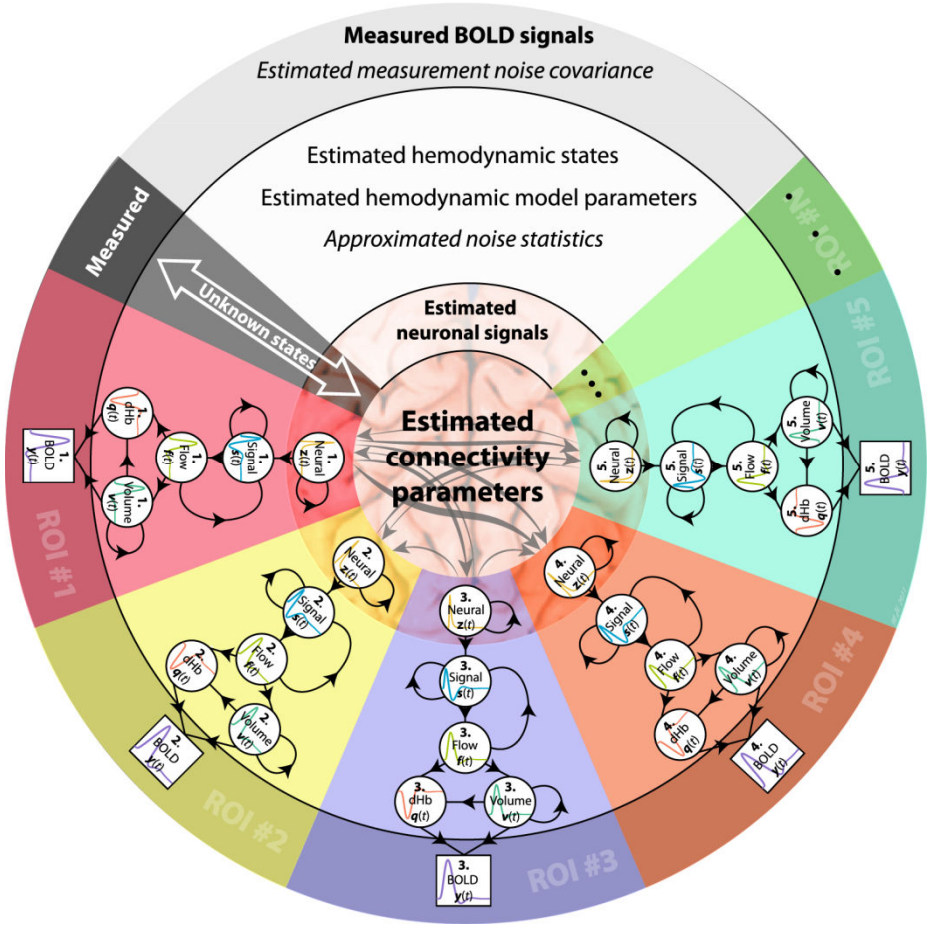

Figure 3.1 *Schematic illustration of stochastic DCM. From measured BOLD signals associated with different brain regions we perform model inversion assuming hemodynamic model and neuronal model of interactions.*

$$\boldsymbol{\mu}_t = \begin{bmatrix} \tilde{\mathbf{z}}_t \\ \tilde{\mathbf{x}}_t \\ \boldsymbol{\theta}_t^{n,h} \end{bmatrix} = \begin{bmatrix} \mathbf{f}^n(\tilde{\mathbf{z}}_{t-1}, \boldsymbol{\theta}_{t-1}^n, \mathbf{u}_{t-1}) \\ \mathbf{f}^h\left(\tilde{\mathbf{x}}_{t-1}, \boldsymbol{\theta}_{t-1}^h, \tilde{\mathbf{z}}_{t-1}\right) \\ \boldsymbol{\theta}_{t-1}^{n,h} \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{q}}_{t-1}^n \\ \tilde{\mathbf{q}}_{t-1}^h \\ \mathbf{o}_{t-1}^{n,h} \end{bmatrix} \tag{3.1}$$

$$\mathbf{y}_t = \mathbf{h}(\boldsymbol{\mu}_t) + \mathbf{r}_t, \tag{3.2}$$

where for simplicity we skipped the notation for multiple l regions; e.g. $\tilde{\mathbf{z}} = [\tilde{\mathbf{z}}_1, \dots, \tilde{\mathbf{z}}_l]^T$, $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_l]^T$, $\boldsymbol{\theta}^h = \left[\boldsymbol{\theta}_1^h, \dots, \boldsymbol{\theta}_l^h\right]^T$, etc. Further, we mark the variables that are obtained by discretization of the continuous process, using a local linearization approach, with tilde.

In summary, each of the l regions is described by one neuronal state $\tilde{\mathbf{z}}$, four hemodynamic states $\tilde{\mathbf{x}} = [\tilde{s}, \tilde{f}, \tilde{v}, \tilde{q}]$, and by a set of hemodynamic parameters $\boldsymbol{\theta}^h$. Crucially, all regions are coupled together (with mutual influence) through the neuronal model, where the strength of couplings is encoded by parameters $\boldsymbol{\theta}^n$. The neuronal model represents a bottom of the generative model, where the neuronal activities from different regions talk to each other. It is supposed that there is no influence or interaction between hemodynamic states of different regions; i.e. at the higher level of the generative model. Schematic representation of this model is depicted in Figure 3.1.

3.1.1 Neuronal interaction model

For fMRI data it is reasonable to define the model of neuronal interactions at macroscopic level, where one can study the whole brain dynamics and interactions between large-scale neuronal systems such as cortical regions. In this sense, it is common to consider a simple model of neuronal responses distributed over l nodes, where under the mean field assumption (see [57, 58]) the dynamics of one node are determined by the average activity of another. This is like assuming that each neuron in one node can see a sufficiently large number of neurons in another node to render the effective influence that is the same as the average over all neurons in the source node. As a result, only the slow dynamics are communicated among nodes, which means we can model distributed activity with a small number of macroscopic variables (e.g. one per node), whose time constants are greater than underlying fast fluctuations that are specific to each node. These fluctuations are continuous and can be represented by system noise. Therefore, the neuronal model can be described through simple linear stochastic differential equation:

$$\begin{aligned} d\mathbf{z}(t) &= \mathbf{f}(\mathbf{z}(t), \boldsymbol{\theta}^n(t))dt + \sqrt{\mathbf{Q}^n}d\mathbf{w}(t) \\ &= \mathbf{A}\mathbf{z}(t)dt + \sqrt{\mathbf{Q}^n}d\mathbf{w}(t), \end{aligned} \tag{3.3}$$

where

$$\mathbf{A} = \frac{\partial \mathbf{f}(\mathbf{z}, \boldsymbol{\theta}^n)}{\partial \mathbf{z}}, \tag{3.4}$$

is the connectivity matrix (the Jacobian), also called adjacency matrix, which represents the first-order connectivity among nodes [59]. The elements of this connectivity matrix are function of endogenous neuronal states and represent the unknown parameters which we want to estimate, $\boldsymbol{\theta}^n = \mathbf{A}$. One can also understand these coupling parameters of effective connectivity as a rate constants (with units s^{-1}) of neuronal population responses that have exponential nature (the solution of differential equation (3.3) is exponential function). Additionally, since the parameters $\boldsymbol{\theta}^n$ are estimated sequentially with the proposed model inversion scheme, it means that we are able to obtain time-varying parameters of effective connectivity, where the uncertainty about the parameters might change with time as well. Crucially, this continuous model allows estimating the cyclic directed graphs, i.e. it enables distinction between forward and backward (feedback) connections.

In the full version of the thesis we also define structure priors on coupling parameters of connectivity matrix, which are necessary for successful inversion of the model.

3.2 MODEL SELECTION

Until now, we have always considered only the first level of inference, where we fit the model m_i to the data $\mathbf{y}$. Our model includes free parameters $\boldsymbol{\theta}$ and by fitting the model to the data we are inferring what values those parameters should probably take, given the data. As a result of this inference we have obtained the most probable

posterior parameter estimates and the uncertainties on these estimates described by posterior error variances. Using Bayes' rule, we define the first level inference (model fitting) in the probabilistic sense as:

$$p(\boldsymbol{\theta}|\mathbf{y}, m_i) = \frac{p(\mathbf{y}|\boldsymbol{\theta}, m_i)p(\boldsymbol{\theta}|m_i)}{p(\mathbf{y}|m_i)}. \tag{3.5}$$

The denominator on the right hand side of (3.5) represents normalizing constant known as evidence or marginal likelihood. The normalizing constant $p(\mathbf{y}|m_i)$ is ignored during the first level of inference. In our case, we obtain posterior estimates $p(\boldsymbol{\theta}|\mathbf{y}, m_i)$ of model parameters by using square-root CKS (SCKS) estimation scheme, where the estimates are optimal in both maximum likelihood and maximum *a posteriori* sense.

However, in the task where one wants to identify the connectivity couplings between nodes, which possibly involves many free parameters (depending on the size of network), we cannot be sure that the model m_i we have inverted is the best one. In other words, there are many possible models $m_i \in M$, where each model is defined by its unique structure (or adjacency matrix) of allowed connections between nodes. Therefore, after model fitting it is common to perform the second level of inference, represented by model comparison. At this level we wish to infer which model is the most plausible, given the data. This is the reason, why one usually considers a set of alternative model candidates, and for each of the model inversion is performed. In this case, the posterior probability of each model is given by:

$$p(m_i|\mathbf{y}) \propto p(\mathbf{y}|m_i)p(m_i), \tag{3.6}$$

where the data-dependent term on the right side of (3.6) is the marginal likelihood (evidence) that appeared already in (3.5), but was ignored. Assuming that there is no *a priori* belief that one model should be better than others, we usually assign equal priors $p(m_i)$ for all model candidates. This means that the models can be uniquely ranked by evaluating the marginal likelihood [60].

Unfortunately, the marginal likelihood is not straightforward to compute, since this computation involves integrating out the dependence on model parameters:

$$p(\mathbf{y}|m_i) = \int p(\mathbf{y}|\boldsymbol{\theta}, m_i)\, p(\boldsymbol{\theta}|m_i) d\boldsymbol{\theta}. \tag{3.7}$$

Therefore, the approximation to marginal likelihood (evidence) is generally considered. Critically, this approximation should represent a balance between the model fit and model complexity:

$$Evidence(m_i) = Best\ fit\ likelihood(m_i) \times Complexity(m_i), \tag{3.8}$$

where the complexity term is usually known as Occam's factor (which is always less or equal to one), which scales with respect to the number of parameters (if number of parameters increases, Occam's factor decreases). Thus, the models with more parameters are automatically penalized. Note that this is very simplified

interpretation of Occam's factor, which fits to our framework. In general, Occam's factor can represent much more than that (see [60]).

3.2.1 Bayes factor

By following the Bayes' rule, models can be compared in Bayesian sense as a ratio of posterior model probabilities (i.e. posterior odds ratio). If we have no prior preference for either model, the prior odds ratio will be equal to 1. Then the posterior odds ratio reduces to the ratio of marginal likelihood, which is called Bayes factor (B):

$$B_{ij} = \frac{p(\mathbf{y}|m_i)}{p(\mathbf{y}|m_j)}. \tag{3.9}$$

Now we know how to compare models between each other, but what we do not know yet is how to approximate the marginal likelihood. Fortunately, asymptotic approximation to the logarithm of marginal likelihood can be provided by Bayesian information criterion (BIC):

$$BIC_i = -2\mathcal{L}_i + n_{\theta,i} \log T, \tag{3.10}$$

where $\mathcal{L}_i$ is the log-likelihood obtained during model inversion. The second term on the right side of (3.10) represents the approximation to the model complexity, where $n_{\theta,i}$ is the number of free parameters considered in the model, and T is the number of observation samples.

Although we can apply BIC as a measure for model comparison, it is limited only to very small networks (upto 4 nodes) because it requires individual model inversion for each possible model candidate, i.e. model restricted to specific structure of allowed coupling parameters. Moreover, it starts to be a very common practice to

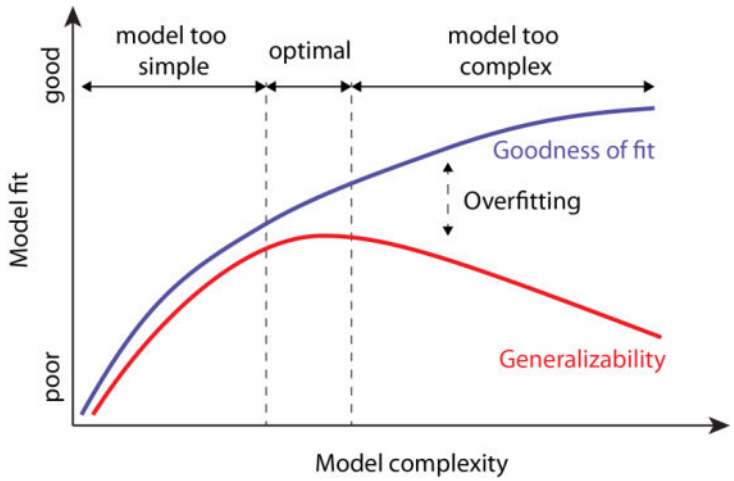

Figure 3.2 *Illustration of optimal model fit. Relationship between goodness of fit (blue line) and generalizability (red line) as a function of model complexity. The y-axis represents any measure of goodness of fit (e.g. log-likelihood), where a larger value represents a better fit. The goal of model selection is to choose the model that generalizes the best across all model candidates.*

search over a larger number of competing models. Therefore, we are interested in model selection strategies, which can compare a large number of models, but do not require to invert each model candidate separately.

To address a problem of selecting the best model among large number of model candidates, in [61] they introduced a Bayesian model selection procedure for *post hoc* inferences about reduced (nested) versions of a full model. This method enables to calculate the marginal likelihood for any reduced model that is nested within a larger (full) model as a function of the posterior density of the full model. Critically, this procedure requires only a single inversion of the full model, where all connections are allowed.

In particular, any reduced model can be created from the full model by collapsing the prior density over one or more parameters; i.e. by setting the corresponding elements of the prior mean $\boldsymbol{\eta}_i$ and precision $\boldsymbol{\Pi}_i^{\theta}$ to zero. Then the free-energy of reduced model m_i, which represents the approximation of marginal likelihood, can be expressed as a simple analytic function of the means and precisions[2] of the prior and posterior of the full model [61]:

$$\mathcal{F}_i = \frac{1}{2}\ln\frac{\left|\boldsymbol{\Pi}_i^{\theta}\right|\left|\boldsymbol{\Upsilon}_F^{\theta}\right|}{\left|\boldsymbol{\Upsilon}_i^{\theta}\right|\left|\boldsymbol{\Pi}_F^{\theta}\right|} - \frac{1}{2}\left(\boldsymbol{\theta}_F^T\boldsymbol{\Upsilon}_F^{\theta}\boldsymbol{\theta}_F + \boldsymbol{\eta}_i^T\boldsymbol{\Pi}_i^{\theta}\boldsymbol{\eta}_i - \boldsymbol{\eta}_F^T\boldsymbol{\Pi}_F^{\theta}\boldsymbol{\eta}_F - \boldsymbol{\theta}_i^T\boldsymbol{\Upsilon}_i^{\theta}\boldsymbol{\theta}_i\right), \quad (3.11)$$

where the reduced posterior precision $\boldsymbol{\Upsilon}_i^{\theta}$ is the posterior precision of the full model $\boldsymbol{\Upsilon}_F^{\theta}$ plus the difference between the reduced and full prior precisions $\boldsymbol{\Pi}_i^{\theta}$ and $\boldsymbol{\Pi}_F^{\theta}$, respectively. Similarly, the reduced posterior mean $\boldsymbol{\theta}_i$ is a mixture of precision-weighted means:

$$\boldsymbol{\Upsilon}_i^{\theta} = \boldsymbol{\Upsilon}_F^{\theta} + \boldsymbol{\Pi}_i^{\theta} - \boldsymbol{\Pi}_F^{\theta} \quad (3.12)$$

$$\boldsymbol{\theta}_i = \mathbf{P}_i^{\theta}\left(\boldsymbol{\Upsilon}_F^{\theta}\boldsymbol{\theta}_F + \boldsymbol{\Pi}_i^{\theta}\boldsymbol{\eta}_i - \boldsymbol{\Pi}_F^{\theta}\boldsymbol{\eta}_F\right). \quad (3.13)$$

The post-hoc estimate of the free-energy based approximation of log marginal likelihood can now be used to compute the Bayes factor, comparing the reduced model m_i with reduced model m_j, as:

$$B_{ij} = \exp\left(\mathcal{F}_i - \mathcal{F}_j\right). \quad (3.14)$$

This model selection approach is suitable also for larger networks. However, for networks consisting of more than six nodes, one has to apply this approach within a greedy search procedure. Detail description of the model selection approaches is provided in the full version of the thesis.

3.2.2 Network pruning

Although the *post hoc* model selection does a great deal of work for us, it is still the first level of inference (model inversion) which must provide a confident estimates of coupling parameters, given the data. However, since there are always some random correlations between time courses (which correspond to particular

[2] Precision is the inverse of variance.

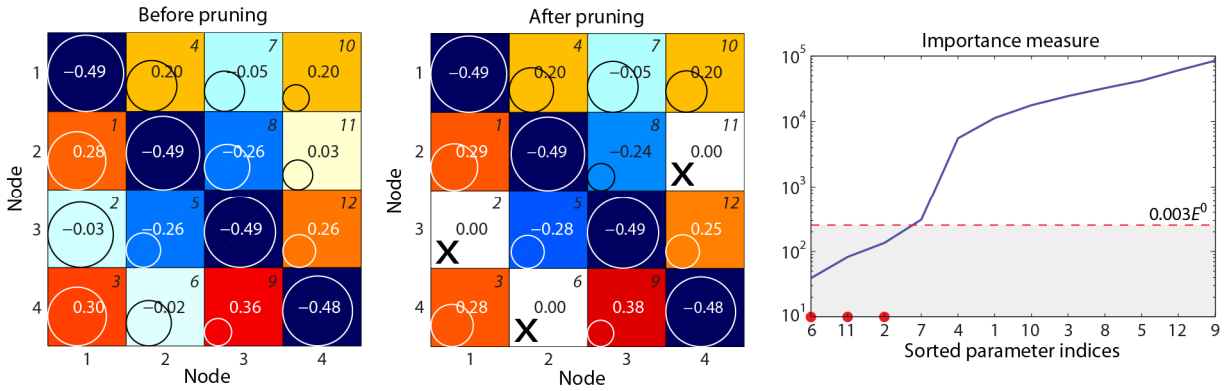

Figure 3.3 *Illustration of the pruning procedure. The upper part shows the estimate of adjacency matrix obtained after* i*-th iteration (e.g.* $i = 5$*) before pruning procedure was applied (left) and after pruning (right). Here the coupling parameters scored with importance measure that is lower than a priori selected threshold (lower plot, y-axis is displayed in log scale to enable visualization of the threshold) are set to zero (black cross).*

nodes in the network), the inversion scheme does not set automatically the irrelevant coupling parameters to zero. Thus these spurious couplings (with non-zero variance) spoil the performance of model inversion. Therefore, we seek a procedure, which can automatically infer the relevant connections and suppress the irrelevant ones.

In principle, there are two ways how to achieve this. The first one supplements a penalty term to the objective function, which causes that the irrelevant couplings tend to zero value. This approach is known as automatic relevance determination (ARD) [60, 62]. The second way involves an estimation of sensitivity of the error function to removal of a coupling parameter (when set to zero), where the connections with the least effect on the error function are subsequently removed. This clearly requires a threshold that has to be specified *a priori*. This approach is mostly known as network pruning.

The estimation framework that is considered in this work does not allow to optimize the priors (required for ARD approach) and include them directly into Kalman filter. Therefore, we will proceed with the parameter relevance determination based on network pruning. The pruning method considered in this work is based on computation of importance function of individual coupling parameters (or subset of them) by determining sensitivity on their removal; i.e. by setting $\boldsymbol{\theta}_i = 0$. Here both the importance function and the sensitivity are derived from a scaled (expected) Fisher information matrix (inverse of parameter error covariance):

$$E = \boldsymbol{\theta}^T \boldsymbol{\Upsilon}^{\theta} \boldsymbol{\theta}, \tag{3.15}$$

where precision $\boldsymbol{\Upsilon}^{\theta}$ is in this case a diagonal matrix (off-diagonal elements are ignored, i.e. effectively set to zero). The complete algorithm for pruning of coupling parameters is described in the full version the thesis.

This pruning procedure enables us to remove at least some of irrelevant parameters already during the optimization process of SCKS, while it also contributes to the reduction of model space that we have to search after the model inversion. Finally, the effect of pruning becomes especially significant when larger networks are estimated.

4 VALIDATION AND APPLICATION OF THE METHOD

The first three chapters described the theoretical background and introduced the new method for the estimation of neuronal signal and evaluation of effective connectivity from fMRI data. Until this point, we have not shown or discussed any results that can be obtained by applying these algorithms. Therefore, it will be the focus of this chapter to provide a demonstration and validation of the proposed method. However, the validation is too extensive to be covered in the short version of the thesis. The reader is referred to the full version of the thesis, where detailed description and complete results are provided. The full version is possible to download at *https://sites.google.com/site/havlicekmartin*. The short version provides only a small fraction of the results compared to the full version.

4.1 SINGLE TIME COURSE MODEL INVERSION

In this section we provide a simple demonstration of neuronal signal estimation from single fMRI time course, where we also focus on the identification problem of hemodynamic model parameters. Note that all details regarding data simulation and model inversion are described in the full version of the thesis.

We considered simulated data having the character of resting-state fMRI time courses of the length 256 time points with temporal resolution 2 s and signal to noise ratio (SNR) equals 2. After obtaining the data, we performed model inversion by the SCKS algorithm, where we allowed maximum of 20 iterations, and considered the discretization of continuous model by using local linearization scheme with the time step 1 s. It means we linearly interpolated the observation sequence by factor 2.

In Figure 4.1 we can see a prediction of BOLD responses compared to the noisy and to the clean BOLD signal. Due to employed estimation of measurement noisy we are not overfitting and the prediction corresponds well to the clean signal. In this case we have reached the convergence with 11 iterations, as it can be seen from the plots of log-likelihood and convergence rate. In Figure 4.2 we display the estimates of hemodynamic states as they are delivered by forward run of the filter and backward run of the smoother, respectively. Clearly, by performing the smoothing, the estimates are more correct and confident (narrow confidence interval around the posterior means). More importantly, when looking at the results of estimated of neuronal signal, we can see that only forward run is unable to recover the true neuronal signal correctly. Therefore, it is now obvious why one has to employ also

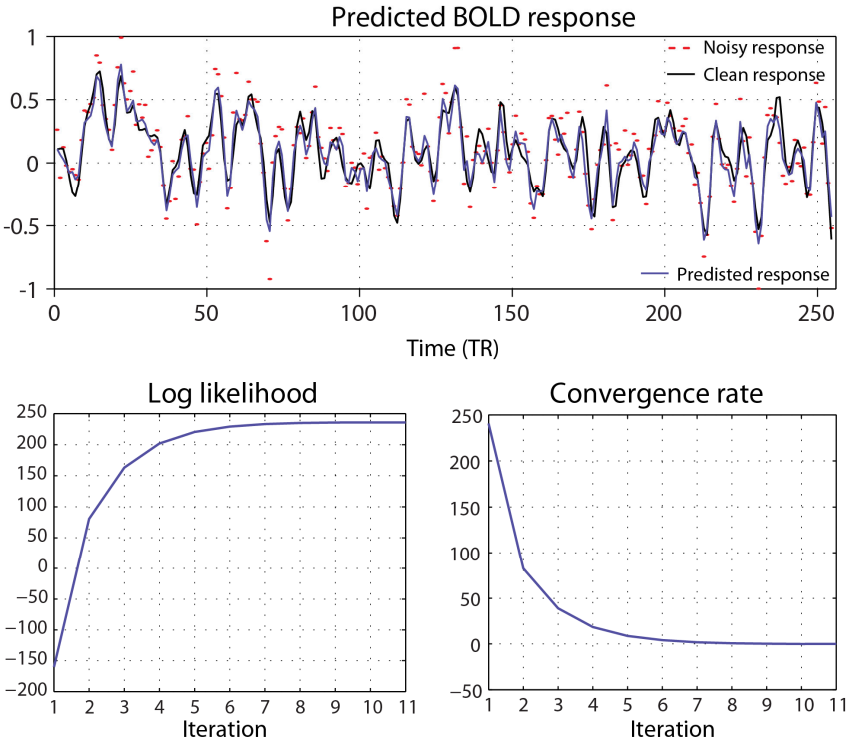

Figure 4.1 *Results of single fMRI time course model inversion (part 1.). The upper plot shows the predicted BOLD responses by SCKS algorithm and provides the comparison with the noisy observed responses and the original noiseless signal. The lower plots show the increase of log-likelihood and the decrease of convergence rate, which indicate that algorithm converged after 11 iteration.*

backward run. By applying it we receive a correct estimate with much narrower confidence intervals.

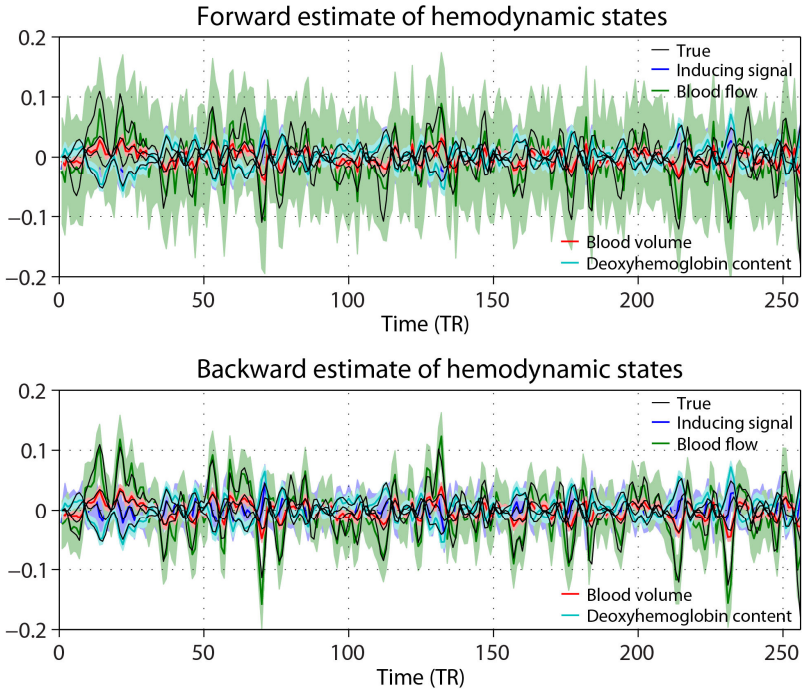

Figure 4.2 *Results of single fMRI time course model inversion (part 2.). The upper plot shows the estimates of hidden hemodynamic states as they are provided by forward run of SCKS algorithm. The shaded area represents the 95 % posterior confidence intervals. The lower plot displays the estimated provided by the backward run (smoother). Note that in the later case the confidence intervals are already much smaller.*

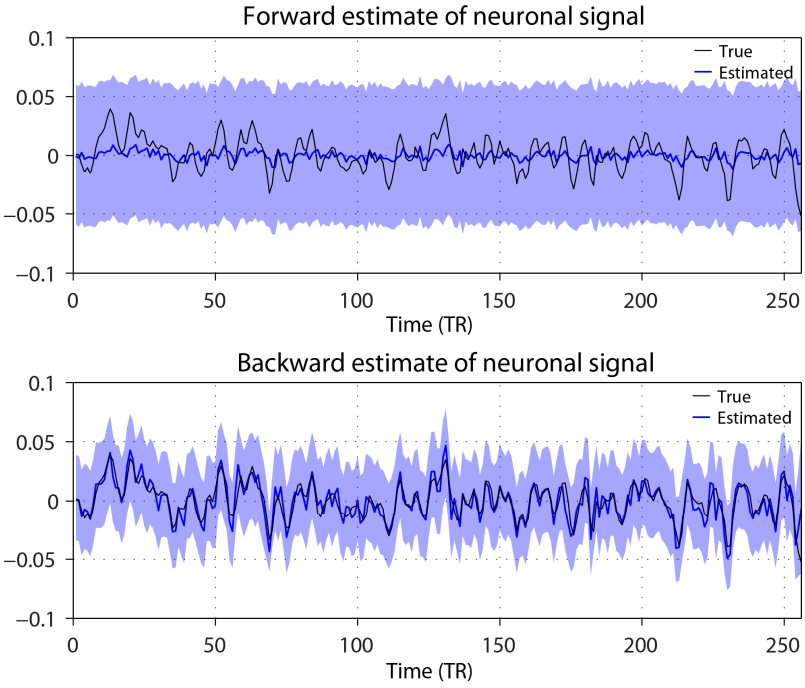

Figure 4.3 *Results of single fMRI time course model inversion (part 3.). The upper plot shows the estimates of neuronal signal provided by forward run of SCKS algorithm. In this case there is significant difference between the true signal and the estimated signal. The estimated signal has much lower amplitude is also delayed. The lower plot shows the estimate of neuronal signal provided by the backward run of SCKS, which well much the true neuronal signal.*

4.2 INVERSION OF STOCHASTIC DCM

Primary method validation was carried out for the multivariate case using Monte Carlo simulations, where we tested the performance of SCKS accompanied by a *post hoc* Bayesian model selection as an approach to stochastic DCM. Details can be found in the full version of the thesis.

In particular, we focused on the estimation of coupling parameters in neuronal interaction model, where besides the general validation of model inversion and model selection we tried to address the main concerns that are very often associated with the methods designed for evaluation of effective connectivity or connectivity analysis in general. These questions and the main results can be summarized as follows:

- *Effect of noisy data:* The method seems to be robust enough in situations of very noisy data with relatively large sampling period. The accuracy of coupling parameters estimates increases with higher SNR. Also the adaptive estimation of measurement noise variance has very positive influence on the obtained estimates, providing more confident estimates (compared to fixed measurement noise variance).
- *Effect of sampling period:* The method is not sensitive to sampling period. Even in case of very poor temporal resolution (5 s), the method is still able to provide a correct estimates of the connectivity structure. Results also suggest that the accuracy of coupling estimates can be improved by increasing data SNR and temporal resolution.
- *Variability of hemodynamic response function among different brain regions:* The method is able to account for hemodynamic response variability across different brain regions. In other words, our method is not sensitive to this sort of variability, providing correct estimates. This feature makes our approach superior to other approaches which do not consider the generative model.
- *Possible confusion of causality by an influence of the third (missing) region:* The results obtained by testing for the robustness against the missing region problem are slightly less promising. In this case we observed a decrease in posterior model probability estimates and an increase of false positive identifications. Although generally these results are still satisfactory, in future work we expect to enhance regularization techniques for parameter estimation that could improve the performance in this particular case.
- *Application to the larger networks:* The results suggest that it should be possible to apply the method to the networks consisting of larger number of nodes. Although it is still not possible to think about the application to the whole data consisting of thousands of voxels, one might imagine its application to the brain activity summarized by spatially distributed modes (often around 20), such as those obtained from the independent component analysis.

4.3 APPLICATION TO EMPIRICAL DATA

We also attempted to apply the SCKS algorithm to empirical data. Although, we mostly emphasized on application of stochastic DCM to resting-state data, there is no reason why this approach should not work also with a task data. Further, since there is very little known about effective connectivity in resting state data, we choose to demonstrate the algorithm performance on task data, but under the assumption that we do not know the exogenous (experimental) input. In particular, we apply SCKS to empirical data-set of visual attention study, which has been used previously to describe developments in causal modeling and related analysis [3, 22, 23, 46, 63, 64]. Detailed description of this data-set and selected region can be found in the full version of the thesis.

We have selected 6 relevant regions and performed model inversion by SCKS algorithm. Results were obtained after 24 iterations of SCKS algorithm (using standard initialization), when it reached the convergence criteria. Then the best model was selected among 32,768 possible models using *post hoc* BMS with posterior model probability about 80 %. The result after model selection are depicted in Figure 4.4. We can see that the final result has relatively sparse structure of connectivity matrix with four bidirectional connections switched off. The architecture of this identified network is then shown also in an anatomical space, where the color of the arrow reports the source of the strongest bidirectional connection and the width represents its absolute (positive or negative) strength. This visualization refers to undirected graphs, although our scheme provides separate estimates for both directions of reciprocal connections. As maybe expected, there are stronger forward connections coming from the visual cortex, which can be considered as a bottom of functional hierarchy, to posterior parietal cortex and prefrontal cortex. Interestingly, there are also many backward connections that are stronger than the forward ones. For example from frontal eye fields, which could be considered as the top of the functional hierarchy, to the visual cortex, prefrontal cortex and to superior temporal sulcus. This is quite sensible given the greater amount of backward connections (neuronal pathways) anatomically, both within the cortical hierarchy and from cortex to subcortical structures [22, 65]. Finally, most of identified connections (but not quite all) are in agreement with a previous results obtained by analyzing effective connectivity using the same data-set [22, 46].

Besides identification of effective connectivity we can be also interested in estimated time courses of neuronal signals (bottom of Figure 4.5). Specifically, because we know the experimental paradigm of visual stimulation (grey filled areas - high for attention and low for no attention), we can compare it to our estimates. In

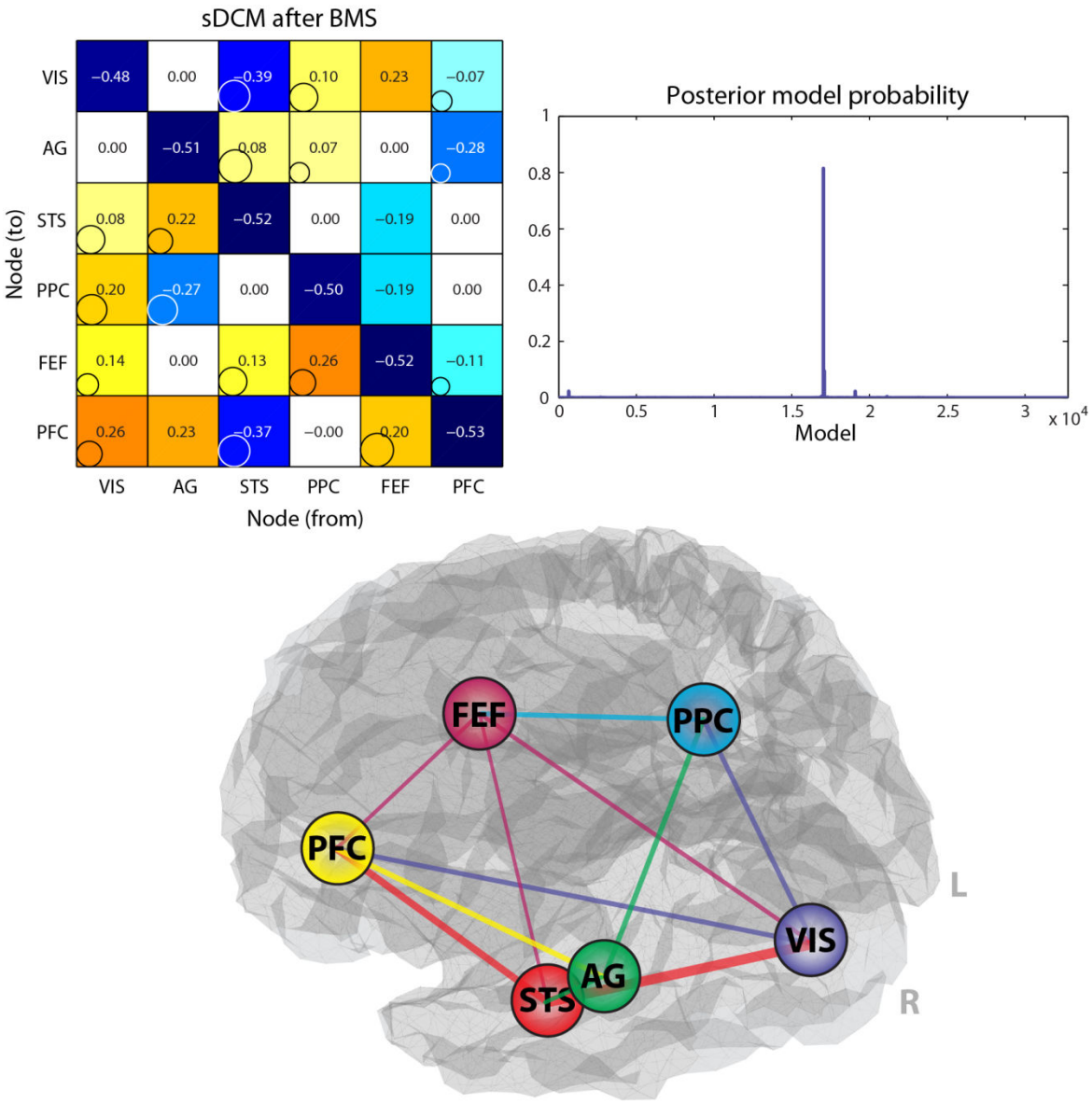

Figure 4.4 *Results of empirical data analysis (part 1.). The upper row shows the results of connectivity structures obtained by model inversion using SCKS algorithm after post hoc model selection (left), where this model was selected with 80 % evidence as it can be seen from the plot of posterior model probability (right). Visualization of identified connectivity structure in anatomical space is depicted in the lower row. The color of the lines reports the source of the strongest bidirectional connection and the width represents its absolute (positive or negative) strength.*

particular, looking at the estimate of neuronal signal associated with the visual cortex (highlighted in blue) we can clearly see that the activation and deactivation phase of this region match exactly to the start and end phase of the stimulus. This confirms that model inversion has effectively estimated neuronal activity from observed BOLD signal and that this estimate is veridical in relation to the experimental manipulation. For comparison to neuronal estimates we also show the original observed BOLD signals and our predictions (bottom of Figure 4.25), where

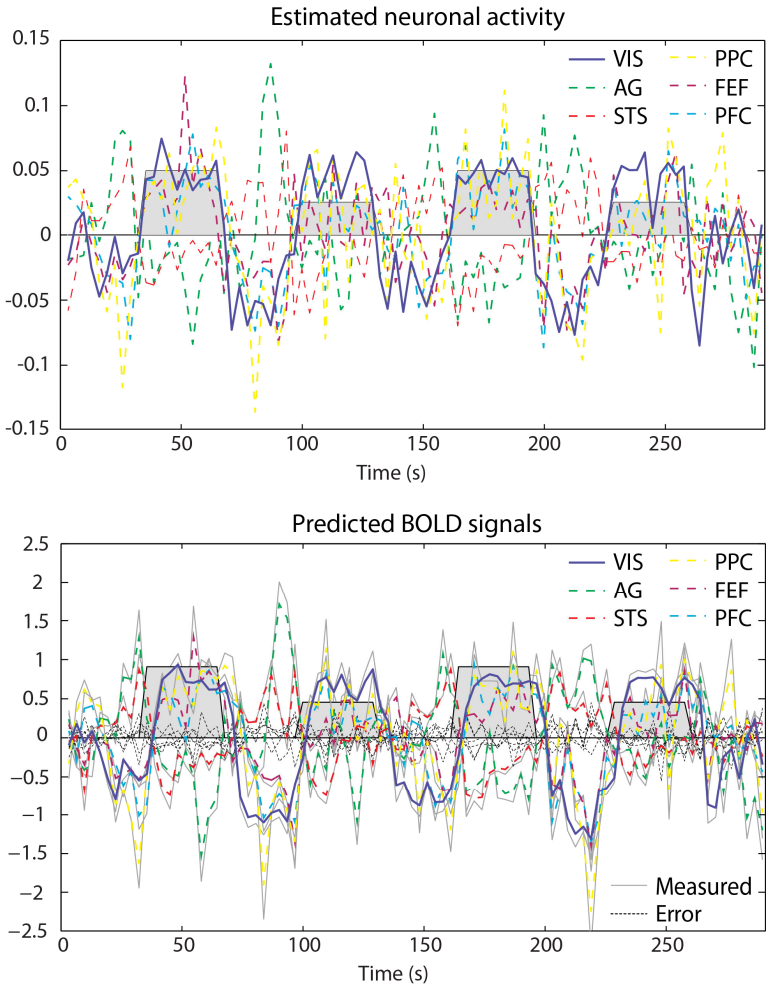

Figure 4.5 *Results of empirical data analysis (part 2.). The upper plot displays the estimated neuronal signals, where we highlight the neuronal responds of the visual cortex. The shaded represents the paradigm of the visual stimulation. Similarly, the lower plot displays predicted BOLD responses.*

there is a delay of several seconds (about 6 s) between the start of stimulation and actual increase in BOLD signal.

Before closing we should again emphasize that the model inversion was not informed by a known stimulation function but can still recover the evoked responses.

5 CONCLUSIONS

In order to evaluate effective connectivity among different brain regions we need to model interactions at the neuronal level. In the case of fMRI data this is complicated by the fact that the measured BOLD signal is only an indirect representation of neuronal activations. The chain of physiological processes that connect the neuronal activation to the BOLD signal can be described by a continuous nonlinear hemodynamic model. Clearly, no model is perfect, which means that it is very important to allow for random fluctuations in unobserved (hidden) neuronal and physiological states by assuming a stochastic representation. Moreover, if we are not restricted to a deterministic model, we are able to account also for (endogenous) autonomous dynamics that cannot be explained by known (exogenous) experimental inputs. We can even throw away the prior knowledge about experimental causes of observed responses and make the evaluation of effective connectivity completely data-driven. Crucially, this enables us to assess causal influence at the neuronal level even from the resting-state fMRI data.

To allow evaluation of this stochastic model we consider the brain as a learning machine that infers information about states and parameters from the observed data. This inference requires representation of uncertainty. Probability theory provides a language for representing the uncertainty beliefs and a framework for maintaining these beliefs in consistent manner. Utilizing probability theory and the general descriptive power of dynamic state-space models, recursive Bayesian estimation provides a theoretically well founded and mathematically robust framework to facilitate sequential probabilistic inference in systems where reasoning under uncertainty is essential. However, because the hemodynamic model we employ is nonlinear, the optimal Bayesian solution to the probabilistic inference problem under consideration is intractable. Therefore, we have to consider only an approximate solution.

In this thesis, we have focused on an approximate solution provided by the Gaussian integration method based on cubature integration rules which was recently introduced to nonlinear Kalman filtering [27]. Specifically, we have proposed a new approach based on cubature Kalman filtering and Rauch-Tung-Striebel smoothing to a joint estimation problem, where both model states and parameters are estimated sequentially, which also models the interaction (conditional dependences) between them. This framework was further extended to meet all requirements given by the model and the fMRI data we work with. First, we have introduced an extension of this approach to the continuous-discrete time systems, where the accurate and stable discretization of the process model was achieved by a local-linearization scheme [28]. Second, to allow the model inversion in a situation, where we *a priori* do not know the noise statistics of the observed BOLD signal, we have adopted an iterative variational Bayesian approach [66] to sequential estimation of measurement noise variance. Third, to improve the convergence of joint state and parameter estimation, we have proposed an adaptive scheme for the estimation of the parameters and state

process noise covariance by efficient Robbins-Monro stochastic approximation scheme. Fourth, since we deal with observed data of a limited length, the forward cubature Kalman filter pass and the backward cubature RTS smoother pass, were wrapped into a simple iterative scheme that maximizes the log-likelihood with each iteration and provides fast convergence. Finally, to further improve the numerical stability of the filter the entire scheme was considered in its square-root form.

All these developments and extensions had one common aim: to enable the estimation of the neuronal signal from a noisy BOLD signal, while considering a realistic nonlinear generation model of the observed signal which also includes stochastic fluctuations contributing to the hidden hemodynamic and neuronal states. In addition, the proposed approach to inversion of the model has a character of (generalized nonlinear) blind deconvolution, because the unknown endogenous neuronal signal (input) to hemodynamic model, which contains unknown parameters is estimated (only) from observed BOLD signal.

Although a very advanced and efficient method was proposed in Chapter 2, it fulfils only a first part of the goal that was described in this thesis. The second (main) goal was to enable evaluation of effective connectivity at the level of neuronal signals given the observed hemodynamic responses. Clearly, the methodological framework described in Chapter 2 solves the more difficult part, i.e. provides estimates of (endogenous) neuronal activity.

In Chapter 3 we have extended this framework to multivariate case, where the main interest is estimation of (effective) coupling parameters that inform the neuronal interaction model. In this case we have considered a neuronal interaction model in a form of linear stochastic differential equations, which define interactions as the communication of slow dynamics among macroscopic variables; i.e. brain regions (nodes). By connecting this neuronal model to region-specific hemodynamic models that link the neuronal activation to observations, we enabled full model inversion, which provides conditional estimates of coupling parameters, region-specific neuronal signals, hemodynamic states, and associated hemodynamic model parameters. All that is possible by applying the approach (iterated square-root cubature RTS smoother) developed in Chapter 2. Importantly, this neuronal interaction model allows one to estimate bidirectional connectivity (causal influence) between different nodes. Further, as an extension to the estimation framework, we have introduced an automatic detection of irrelevant coupling parameters using a network pruning algorithm based on calculation of scaled Fisher information matrix. This addition was necessary to improve the performance of estimating coupling (connectivity) parameters, especially in cases when the spurious correlations between observed signals are present. A complete form of this model inversion represents a stochastic treatment of dynamic causal modeling that makes it possible to estimate effective connectivity even in case of unknown model input; i.e. in case of resting-state data, where the neuronal signals causing the hemodynamic responses have purely endogenous character. This is an important departure from the

original dynamic causal modeling [3], which was limited to a discrete model of hemodynamic states and assumed *a priori* knowledge of the model input.

This novel approach represents the first level of inference where we are especially interested in conditional estimates of coupling parameters and in the associated error covariance matrix. However, it is still necessary to perform the second level of inference, where we identify (select) the most likely model candidate, which in the case of effective connectivity corresponds to the most likely connectivity matrix. In Chapter 3 we considered model comparison based on calculation of Bayes factor (defined as a ratio of marginal likelihoods of restricted models) and discussed two different approximations to marginal likelihood through common Bayesian information criteria and through recently introduced concept of reduced free energy [61]. We have emphasized the convenience of the later approximation, because it requires only a single inversion of the full model, where all connections all between nodes are allowed. Moreover, this approach is also well suited for model selection in larger networks.

In Chapter 4, we demonstrated the performance of the proposed approach, first focusing on estimation of the neuronal signal from a single fMRI time course. We also addressed the principal questions one might have regarding the performance of the introduced method to correctly infer the coupling parameters of neuronal interaction model. In this case we were able to show that the method is robust even when applied to data with lower SNR and larger sampling period. Also it is not sensitive to variability of hemodynamic response function across different brain regions. These are important properties, which make the approach superior to other approaches for the evaluation of effective connectivity that are not based on generative models and are not formulated in continuous time. We also showed that there is a good perspective for this approach to be applied to larger networks, where possibly all relevant brain regions are included. As a relative weakness we found that the method is only partly immune to the strong spurious correlations caused by exclusion of relevant region (node) from the analysis.

BIBLIOGRAPHY

[1] K. FRISTON, "Functional and effective connectivity: a review," *Brain Connectivity,* vol. 1, 2011.

[2] A. R. MCINTOSH, "Towards a network theory of cognition," *Neural Networks,* vol. 13, pp. 861-870, 2000.

[3] K. J. FRISTON, L. HARRISON, W. PENNY, "Dynamic causal modelling," *Neuroimage,* vol. 19, pp. 1273-1302, 2003.

[4] R. GOEBEL, A. ROEBROECK, D. S. KIM, E. FORMISANO, "Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping," *Magnetic resonance imaging,* vol. 21, pp. 1251-1261, 2003.

[5] A. ROEBROECK, E. FORMISANO, R. GOEBEL, "Mapping directed influence over the brain using Granger causality and fMRI," *Neuroimage,* vol. 25, pp. 230-242, 2005.

[6] K. J. FRISTON, "Functional and effective connectivity in neuroimaging: a synthesis," *Human Brain Mapping,* vol. 2, pp. 56-78, 1994.

[7] M. MCKEOWN, T. SEJNOWSKI, "Independent component analysis of fMRI data: examining the assumptions," *Human Brain Mapping,* vol. 6, 1998.

[8] V. D. CALHOUN, T. ADALI, G. D. PEARLSON, J. J. PEKAR, "A method for making group inferences from functional MRI data using independent component analysis," *Human Brain Mapping,* vol. 14, pp. 140-151, 2001.

[9] K. FRISTON, C. BUECHEL, G. FINK, J. MORRIS, E. ROLLS, R. DOLAN, "Psychophysiological and modulatory interactions in neuroimaging," *Neuroimage,* vol. 6, pp. 218-229, 1997.

[10] G. K. AGUIRRE, E. ZARAHN, M. D'ESPOSITO, "The variability of human, BOLD hemodynamic responses," *Neuroimage,* vol. 8, pp. 360-369, 1998.

[11] D. HANDWERKER, J. OLLINGER, M. D'ESPOSITO, "Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses," *Neuroimage,* vol. 21, pp. 1639-1651, 2004.

[12] O. DAVID, I. GUILLEMAIN, S. SAILLET, S. REYT, C. DERANSART, C. SEGEBARTH, A. DEPAULIS, "Identifying neural drivers with functional MRI: an electrophysiological validation," *PLoS Biology,* vol. 6, pp. 2683-2697, 2008.

[13] C. IADECOLA, "CC commentary: intrinsic signals and functional brain mapping: caution, blood vessels at work," *Cerebral Cortex,* vol. 12, pp. 223-224, 2002.

[14] P. MAGISTRETTI, L. PELLERIN, "Cellular Mechanisms of Brain Energy Metabolism and their Relevance to Functional Brain Imaging," *Philos Trans R Soc Lond B Biol Sci,* vol. 354, pp. 1155-1163, 1999.

[15] D. ATTWELL, A. M. BUCHAN, S. CHARPAK, M. LAURITZEN, B. A. MACVICAR, E. A. NEWMAN, "Glial and neuronal control of brain blood flow," *Nature,* vol. 468, pp. 232-243, 2010.

[16] G. S. BERNS, A. W. SONG, H. MAO, "Continuous functional magnetic resonance imaging reveals dynamic nonlinearities of "dose-response" curves for finger opposition," *Journal of Neuroscience,* vol. 19, pp. 1-6, 1999.

[17] A. MECHELLI, C. PRICE, K. FRISTON, "Nonlinear coupling between evoked rCBF and BOLD signals: a simulation study of hemodynamic responses," *NeuroImage,* vol. 14, pp. 862-872, 2001.
[18] K. L. MILLER, W. M. LUH, T. T. LIU, A. MARTINEZ, T. OBATA, E. C. WONG, L. R. FRANK, R. B. BUXTON, "Nonlinear temporal dynamics of the cerebral blood flow response," *Human Brain Mapping,* vol. 13, pp. 1-12, 2001.
[19] I. SHOJI, "A comparative study of maximum likelihood estimators for nonlinear dynamical system models," *International Journal of Control,* vol. 71, pp. 391-404, 1998.
[20] J. RIERA, J. WATANABE, I. KAZUKI, M. NAOKI, E. AUBERT, T. OZAKI, R. KAWASHIMA, "A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals," *NeuroImage,* vol. 21, pp. 547-567, 2004.
[21] K. J. FRISTON, N. TRUJILLO-BARRETO, J. DAUNIZEAU, "DEM: a variational treatment of dynamic systems," *Neuroimage,* vol. 41, pp. 849-885, 2008.
[22] K. J. FRISTON, B. LI, J. DAUNIZEAU, K. E. STEPHAN, "Network discovery with DCM," *Neuroimage,* pp. 1202-1221, 2010.
[23] B. LI, J. DAUNIZEAU, K. E. STEPHAN, W. PENNY, D. HU, K. FRISTON, "Generalised filtering and stochastic DCM for fMRI," *NeuroImage,* pp. 442-457, 2011.
[24] T. OBATA, T. T. LIU, K. L. MILLER, W. M. LUH, E. C. WONG, L. R. FRANK, R. B. BUXTON, "Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas: application of the balloon model to the interpretation of BOLD transients," *NeuroImage,* vol. 21, pp. 144-153, 2004.
[25] M. HAVLICEK, K. J. FRISTON, J. JAN, M. BRAZDIL, V. D. CALHOUN, "Effective connectivity in fMRI resting state data via blind deconvolution," in *17th Annual Meeting of the Organization for Human Brain Mapping*, Quebec, 2011, pp. 1-2.
[26] M. HAVLICEK, K. J. FRISTON, J. JAN, M. BRAZDIL, V. D. CALHOUN, "Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering," *NeuroImage,* vol. 56, pp. 2109-2128, 2011.
[27] I. ARASARATNAM, S. HAYKIN, "Cubature Kalman Filters," *IEEE Transactions on Automatic Control,* vol. 54, pp. 1254-1269, 2009.
[28] T. OZAKI, "A bridge between nonlinear time series models and nonlinear stochastic dynamical systems: a local linearization approach," *Statistica Sinica,* vol. 2, pp. 113–-135, 1992.
[29] J. J. RIERA, J. WATANABE, I. KAZUKI, M. NAOKI, E. AUBERT, T. OZAKI, R. KAWASHIMA, "A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals," *NeuroImage,* vol. 21, pp. 547-567, 2004.
[30] H. ROBBINS, S. MONRO, "A stochastic approximation method," *The Annals of Mathematical Statistics,* vol. 22, pp. 400-407, 1951.
[31] R. VAN DER MERWE, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," Ph.D.thesis, Oregon Graduate Institute of Science and Technology, 2004.
[32] R. E. KALMAN, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering,* vol. 82, pp. 35-45, 1960.
[33] H. E. RAUCH, F. TUNG, C. T. STRIEBEL, "Maximum likelihood estimates of linear dynamic systems," *AIAA Journal,* vol. 3, pp. 1445-1450, 1965.

[34] S. JULIER, J. UHLMANN, H. F. DURRANT-WHYTE, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *Automatic Control, IEEE Transactions on,* vol. 45, pp. 477-482, 2002.
[35] R. S. BUCY, K. D. SENNE, "Digital synthesis of non-linear filters," *Automatica,* vol. 7, pp. 287-298, 1971.
[36] A. DOUCET, N. DE FREITAS, N. GORDON, *Sequential Monte Carlo methods in practice*: Springer Verlag, 2001.
[37] T. S. JAAKKOLA, "Tutorial on variational approximation methods," *Advanced mean field methods: theory and practice,* pp. 129-159, 2000.
[38] C. FERNANDEZ-PRADES, J. VILA-VALLS, "Bayesian Nonlinear Filtering Using Quadrature and Cubature Rules Applied to Sensor Data Fusion for Positioning," in *In proceeding of IEEE International Conference on Communications*, 2010, pp. 1-5.
[39] P. LI, J. YU, M. WAN, J. HUANG, "The augmented form of cubature Kalman filter and quadrature Kalman filter for additive noise," in *IEEE Youth Conference on Information, Computing and Telecommunication, YC-ICT '09.*, 2009, pp. 295-298.
[40] A. T. NELSON, "Nonlinear estimation and modeling of noisy time-series by dual Kalman filtering methods," Ph.D thesis, Oregon Graduate Institute of Science and Technology, 2000.
[41] Z. JI, M. BROWN, "Joint State and Parameter Estimation For Biochemical Dynamic Pathways With Iterative Extended Kalman Filter: Comparison With Dual State and Parameter Estimation," *Open Automation and Control Systems Journal,* vol. 2, pp. 69-77, 2009.
[42] P. E. KLOEDEN, E. PLATEN, *Numerical Solution of Stochastic Differential Equations, Stochastic Modeling and Applied Probability, 3rd editon*: Springer, 1999.
[43] R. E. KALMAN, R. S. BUCY, "New results in linear filtering and prediction theory," *Transactions of the ASME. Series D, Journal of Basic Engineering,* vol. 83, pp. 95-107, 1961.
[44] L. LJUNG, T. SÖDERSTRÖM, *Theory and practice of recursive identification*. Cambridge, MA: MIT Press, 1983.
[45] S. SÄRKKÄ, A. NUMMENMAA, "Recursive noise adaptive Kalman filtering by variational Bayesian approximations," *Automatic Control, IEEE Transactions on,* vol. 54, pp. 596-600, 2009.
[46] C. BUCHEL, "Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI," *Cerebral Cortex,* vol. 7, pp. 768-778, 1997.
[47] A. R. MCLNTOSH, F. GONZALEZ-LIMA, "Structural equation modeling and its application to network analysis in functional brain imaging," *Human Brain Mapping,* vol. 2, pp. 2-22, 1994.
[48] A. ROEBROECK, E. FORMISANO, R. GOEBEL, "The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution," *Neuroimage,* vol. In Press, 2009.
[49] S. M. SMITH, K. L. MILLER, G. SALIMI-KHORSHIDI, M. WEBSTER, C. F. BECKMANN, T. E. NICHOLS, J. D. RAMSEY, M. W. WOOLRICH, "Network modelling methods for FMRI," *Neuroimage,* vol. 54, pp. 875-891, 2010.

[50] A. ROEBROECK, E. FORMISANO, 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,* 2009.
[51] A. ROEBROECK, E. FORMISANO, R. GOEBEL, "The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution," *Neuroimage,* p. In Press, 2009.
[52] 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,* p. In Press, 2009.
[53] K. FRISTON, "Dynamic casual modeling and Granger causality Comments on: The identification of interacting networks in the brain using fMRI: Model selection, causality and deconvolution," *Neuroimage,* p. In Press, 2009.
[54] A. ROEBROECK, A. K. SETH, P. VALDES-SOSA, "Causal Time Series Analysis of functional Magnetic Resonance Imaging Data," in *JMLR: Workshop and Conference Proceedings*, 2011, pp. 65-94.
[55] P. A. VALDES-SOSA, A. ROEBROECK, J. DAUNIZEAU, K. FRISTON, "Effective connectivity: Influence, causality and biophysical modeling," *NeuroImage,* vol. 58, 2011.
[56] W. PENNY, K. STEPHAN, A. MECHELLI, K. FRISTON, "Comparing dynamic causal models," *NeuroImage,* vol. 22, pp. 1157-1172, 2004.
[57] P. DAYAN, L. F. ABBOTT, *Theoretical neuroscience: Computational and mathematical modeling of neural systems*: MIT Press, 2001.
[58] G. DECO, V. K. JIRSA, P. A. ROBINSON, M. BREAKSPEAR, K. FRISTON, "The dynamic brain: from spiking neurons to neural masses and cortical fields," *PLoS computational biology,* vol. 4, pp. 1-35, 2008.
[59] J. M. KILNER, J. MATTOUT, R. HENSON, K. J. FRISTON, "Hemodynamic correlates of EEG: a heuristic," *Neuroimage,* vol. 28, pp. 280-286, 2005.
[60] D. J. C. MACKAY, "Bayesian interpolation," *Neural computation,* vol. 4, pp. 415-447, 1992.
[61] K. FRISTON, W. PENNY, "Post hoc Bayesian model selection," *NeuroImage,* pp. 2089-2098, 2011.
[62] D. J. C. MACKAY, "A practical Bayesian framework for backpropagation networks," *Neural computation,* vol. 4, pp. 448-472, 1992.
[63] L. HARRISON, W. PENNY, K. FRISTON, "Multivariate autoregressive modeling of fMRI time series," *NeuroImage,* vol. 19, pp. 1477-1491, 2003.
[64] K. E. STEPHAN, L. KASPER, L. M. HARRISON, J. DAUNIZEAU, H. E. M. DEN OUDEN, M. BREAKSPEAR, K. J. FRISTON, "Nonlinear dynamic causal models for fMRI," *NeuroImage,* vol. 42, pp. 649-662, 2008.
[65] A. M. SILLITO, H. E. JONES, "Corticothalamic interactions in the transfer of visual information," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences,* vol. 357, p. 17391752, 2002.
[66] S. SÄRKKÄ, "Recursive Bayesian inference on stochastic differential equations," Ph.D. Thesis: Helsinki University of Technology Laboratory of Computational Engineering, 2006.

MARTIN HAVLÍČEK CURRICULUM VITAE

CONTACT

Address:
Ostrožná 13
74601 Opava
Czech Republic

Mobile: +420 777 005 683
E-mail: havlicekmartin@gmail.com

EDUCATION

Brno University of Technology, Czech Republic: 2002-present

Ph.D. in Biomedical Engineering: Expected 2011

- *Thesis:* Exploring brain network connectivity through hemodynamic modeling.
- *Advisors:* Prof. Jiří Jan, Prof. Milan Brázdil, Prof. Vince D. Calhoun.

M.S. in Biomedical Engineering: 2007

- *Thesis:* Modeling of SPECT Imaging.
- *Advisor:* Ass. Prof. Aleš Drastich.

B.S. in Electrical Engineering: 2005

- *Thesis:* Metrology of Positron Emission Tomography.
- *Advisor:* Ass. Prof. Aleš Drastich.

University of Kuopio, Finland: 2006

- Exchange student: Dept. of Medical Physics
- *Advisor:* Prof. Pasi Karjalainen.

SKILLS

Programming: Matlab (expert), R, Python, LaTEX, HTML, C/C++ (partly).
Software: SPM8, GIFT, AFNI, EEGlab, Windows (95-2007), Unix.
Languages: Czech (mother tongue), English (fluent), German (basics).

RESEARCH POSITIONS

Research Scholar, University of New Mexico, Albuquerque, USA 2009-2011

- Mind Research Network Institute,MIALab
- Advisor: Prof. Vince D. Calhoun.

Graduate Student Researcher, Brno University of Technology, Czech Republic 2008-2009

- Research Centre Data-Algorithms-Decision Making
- Advisor: Prof. Jiří Jan.

JOURNAL PUBLICATIONS

HAVLICEK, M., FRISTON, K.J., JAN, J., BRAZDIL, M., CALHOUN, V.D. *Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering*. NeuroImage, 2011, vol. 56, no. 4, pp. 2109-2128.

HAVLICEK, M., JAN, J., BRAZDIL, M., CALHOUN, V.D. *Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data*. NeuroImage, 2010, vol. 53, no. 1, pp. 65-77.

SELECTED CONFERENCE PRESENTATIONS

HAVLICEK, M., FRISTON, K.J, JAN, J., BRAZDIL, M., CALHOUN, V.D. *Effective connectivity in fMRI resting state data via blind deconvolution.* In 17th Annual Meeting of the Organization for Human Brain Mapping. Quebec: Organization for Human Brain Mapping, 2011.

HAVLICEK, M., JAN, J., BRAZDIL, M., CALHOUN, V.D. Estimation of neuronal signal from hemodynamic responses. In Proc. of 33st IEEE EMBS conf. 2011. Boston: IEEE EMBS, 2011, pp. 8122-8125.

HAVLICEK, M., JAN, J., BRAZDIL, M., CALHOUN, V.D. *Deconvolution of neuronal signal from hemodynamic response.* In Proceedings of 2011 International Conference on Acoustics, Speech, and Signal Processing. Prague: IEEE ICASSP, 2011. pp. 617-620.

HAVLICEK, M., JAN, J., PLIS, S., BRAZDIL, M., CALHOUN, V.D. *Dynamic Granger causality with embedded hemodynamic model*. In 16th Annual Meeting of the Organization for Human Brain Mapping. Barcelona: Organization for Human Brain Mapping, 2010. (Talk)

HAVLICEK, M., JAN, J., BRAZDIL, M., CALHOUN, V.D. *Nonlinear estimation of BOLD signal based on cubature particle filter*. In Proceedings of Biosignal 2010: Analysis of Biomedical Signals and Images. Brno, Czech Republic: Brno University of Technology, 2010. pp. 328-332.

HAVLICEK, M., JAN, J., CALHOUN, V.D., BRAZDIL, M. *Extended time-frequency Granger causality for evaluation of functional network connectivity in event-related FMRI data.* In Proc. of 31st IEEE EMBS conf. 2009. Minneapolis: IEEE EMBS, 2009. pp. 4440-4443. (Talk)

HAVLICEK, M., JAN, J., CALHOUN, V., BRAZDIL, M., MIKL, M. *Evaluation of functional network connectivity in event-related fMRI data based on ICA*. In Proc. WC 2009. Munich: IFMBE, 2009. pp. 716-719.

AWARDS

- Organization of Human Brain Mapping - Abstract award, 2010
- Dean's award for Master thesis, 2007
- Dean's award for Bachelor thesis, 2005

JOURNAL REVIEWER

Human Brain Mapping, NeuroImage, Psychiatry Research: Neuroimaging.
Mentored reviews under Dr. Calhoun and Dr. Jan: Frontiers in Neuroscience, Journal of Neuroscience, Journal of Neurophysiology, Computer Methods and Programs in Biomedicine, Medical and Biological Engineering and Computing.

ABSTRACT

Functional magnetic resonance imaging (fMRI) utilizing the blood-oxygen-level-dependent (BOLD) effect as an indicator of local activity is a very useful technique to identify brain regions that are active during perception, cognition, action, and also during rest. Currently, there is a growing interest to study connectivity between different brain regions, particularly in the resting-state.

This thesis introduces a new and original approach to problem of indirect relationship between observed hemodynamic response and its cause represented by neuronal signal, as this indirect relationship complicates the estimation of effective connectivity (causal influence) between different brain regions from fMRI data. The novelty of this approach is in (generalized nonlinear) blind-deconvolution technique that allows estimation of the endogenous neuronal signals (system inputs) from measured hemodynamic responses (system outputs). Thus, it enables a fully data-driven evaluation of effective connectivity on neuronal level, even though only fMRI hemodynamic responses are observed. The solution to this difficult deconvolution (model inversion) problem is obtained through a nonlinear recursive Bayesian estimation framework for joint estimation of hidden model states and parameters.

This thesis is divided into three main parts. The first part proposes a method to solve the above mentioned inversion problem. The method uses a square-root form of a nonlinear cubature Kalman filtering and cubature Rauch-Tung-Striebel smoothing extended to a joint estimation problem defined as a simultaneous estimation of states and parameters in a sequential manner. The method is designed particularly for continuous-discrete systems and obtains an accurate and stable solution to model discretization by combining nonlinear (cubature) filtering with local linearization. Moreover, the inversion method is equipped with the adaptive estimation of measurement, state, and parameter noise statistics. The first part of the thesis is focused only on the single time course model inversion; i.e. estimation of neuronal signal from fMRI signal.

The second part generalizes the proposed approach and applies it to multiple fMRI time courses in order to enable the estimation of coupling parameters of a neuronal interaction model; i.e. estimation of effective connectivity. This method represents a novel stochastic treatment of dynamic causal modeling, which makes it distinct from any previously introduced approach. The second part also deals with methods for Bayesian model selection and proposes a technique for detection of irrelevant connectivity parameters to achieve improved performance of parameter estimation.

Finally, the third part provides a validation of the proposed approach by using both simulated and empirical fMRI data, and demonstrates robust and very good performance.