

1 **Dynamic modeling of neuronal responses in fMRI using**  
2 **cubature Kalman filtering**

3 Martin Havlicek<sup>1,2,4</sup>, Karl J. Friston<sup>3</sup>, Jiri Jan<sup>1</sup>, Milan Brazdil<sup>5,6</sup>, and Vince D. Calhoun<sup>2,4</sup>

4 <sup>1</sup> *Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic,*

5 <sup>2</sup> *The Mind Research Network, Albuquerque, New Mexico, USA,*

6 <sup>3</sup> *The Wellcome Trust Centre for Neuroimaging, UCL, WC1N 3BG, UK,*

7 <sup>4</sup> *Department of Electrical and Computer Engineering, University of New Mexico, USA,*

8 <sup>5</sup> *Behavioral and Social Neuroscience Research Group, Central European Institute of Technology*  
9 *(CEITEC), Masaryk University, Brno, Czech Republic,*

10 <sup>6</sup> *Department of Neurology, St. Anne's University Hospital and Medical Faculty of Masaryk University,*  
11 *Brno, Czech Republic.*

12

13

14

15

16

17

18

19 **Address for correspondence**

20 Martin Havlicek

21 The Mind Research Network

22 1101 Yale Boulevard NE,

23 Albuquerque,

24 NM 87106, USA.

25 Fax: +1 505 272 8002.

26 E-mail: havlicekmartin@gmail.com (M. Havlicek),

27 vcalhoun@mrn.org (V.D. Calhoun).

28

29 **Keywords:** Neuronal; fMRI; Blind deconvolution; Cubature Kalman filter; Smoother;  
30 Stochastic; Hemodynamic modeling; Dynamic Expectation Maximization; Nonlinear;

## 31 **Abstract**

32       This paper presents a new approach to inverting (fitting) models of coupled dynamical  
33 systems based on state-of-the-art (cubature) Kalman filtering. Crucially, this inversion furnishes  
34 posterior estimates of both the hidden states and parameters of a system, including any unknown  
35 exogenous input. Because the underlying generative model is formulated in continuous time  
36 (with a discrete observation process) it can be applied to a wide variety of models specified with  
37 either ordinary or stochastic differential equations. These are an important class of models that  
38 are particularly appropriate for biological time-series, where the underlying system is specified  
39 in terms of kinetics or dynamics (i.e., dynamic causal models). We provide comparative  
40 evaluations with generalized Bayesian filtering (dynamic expectation maximization) and  
41 demonstrate marked improvements in accuracy and computational efficiency. We compare the  
42 schemes using a series of difficult (nonlinear) toy examples and conclude with a special focus on  
43 hemodynamic models of evoked brain responses in fMRI. Our scheme promises to provide a  
44 significant advance in characterizing the functional architectures of distributed neuronal systems,  
45 even in the absence of known exogenous (experimental) input; e.g., resting state fMRI studies  
46 and spontaneous fluctuations in electrophysiological studies. Importantly, unlike current  
47 Bayesian filters (e.g. DEM), our scheme provides estimates of time-varying parameters, which  
48 we will exploit in future work on the adaptation and enabling of connections in the brain.

49

50 **Content**

51 Introduction..... 5

52 Nonlinear continuous-discrete state-space models ..... 8

53     Probabilistic inference ..... 10

54     Cubature Kalman filter ..... 12

55     Parameters and input estimation ..... 13

56     Square-root cubature Kalman filter..... 15

57     Forward filtering pass ..... 16

58     Backward smoothing pass..... 19

59     Dynamic expectation maximization..... 21

60 Inversion of dynamic models by SCKF and SCKS ..... 24

61     Lorenz attractor ..... 26

62     Double-well..... 29

63     Convolution model..... 31

64     Hemodynamic model ..... 34

65         • Effect of model parameters on hemodynamic response and their estimation..... 40

66         • Beyond the limits of fMRI signal ..... 45

67 Discussion ..... 48

68 Conclusion ..... 52

69 Acknowledgments..... 53

70 References..... 53

71

## 72 **Introduction**

73       The propagation of neuronal activity in the brain is a dynamic process, which mediates the  
74 communication among functional brain areas. Although, recent advances in neuroimaging allow  
75 for greater insights into brain function, all available noninvasive brain mapping techniques  
76 provide only indirect measures of the underlying electrophysiology. For example, we cannot  
77 observe the time-varying neuronal activation in the brain but we can measure the electrical field  
78 it generates on the scalp using electroencephalography (EEG). Similarly, in functional magnetic  
79 resonance imaging (fMRI) we measure hemodynamic responses, which represent changes in  
80 blood flow and blood oxygenation that follow neuronal activation. Crucially, the form of this  
81 hemodynamic response can vary across subjects and different brain regions (Aguirre et al., 1998;  
82 Handwerker et al., 2004). This complicates the estimation of hidden neuronal states and  
83 identification of effective connectivity (i.e. directed influence) between brain regions (David,  
84 2009; David et al., 2008; Friston, 2009; Roebroeck et al., 2009a, b).

85       In general, the relationship between initial neuronal activation and our observations rests on  
86 a complex electro/bio-physiological process. If this process is known and well described, it can  
87 be approximated by mathematical modeling. Inversion of the ensuing model allows us to  
88 estimate hidden states of neuronal systems (e.g., the neuronal activation) from observations. The  
89 resulting estimate will be affected by the accuracy of the inversion (formulated as an  
90 optimization problem) and by the precision of the observation itself (temporal resolution, signal  
91 to noise ratio (SNR), etc.). In signal processing theory, this problem is called blind deconvolution  
92 and is described as estimating the unknown input to a dynamic system, given output data, when  
93 the model of the system contains unknown parameters. A note on terminology is needed here:  
94 although convolution is usually defined as a linear operation, the term deconvolution is generally  
95 used in reference to the inversion of nonlinear (generalized) convolution models (i.e.  
96 restoration); we adhere to this convention.

97       In fMRI, the physiological mechanisms mediating the relationship between neuronal  
98 activation and vascular/metabolic systems have been studied extensively (Attwell et al., 2010;  
99 Iadecola, 2002; Magistretti and Pellerin, 1999) and models of hemodynamic responses have been  
100 described at macroscopic level by systems of differential equations. The hemodynamic model  
101 (Friston et al., 2000) links neuronal activity to flow and subsumes the balloon-windkessel model

102 (Buxton et al., 1998; Mandeville et al., 1999), linking flow to observed fMRI signals. The  
103 hemodynamic model includes model of neurovascular coupling (i.e., how changes in neuronal  
104 activity cause a flow-inducing signal) and hemodynamic processes (i.e. changes in cerebral  
105 blood flow (CBF), cerebral blood volume (CBV), and total de-oxyhemoglobin (dHb)). In this  
106 paper, we will focus on a hemodynamic model of a single region in fMRI, where experimental  
107 studies suggest that the neuronal activity that drives hemodynamic responses corresponds more  
108 afferent synaptic activity (as opposed to efferent spiking activity (Lauritzen, 2001; Logothetis,  
109 2002)). In the future work, we will use exactly the same scheme to model distributed neuronal  
110 activity as observed in multiple regions.

111 The hemodynamic model is nonlinear in nature (Berns et al., 1999; Mechelli et al., 2001).  
112 Therefore, to infer the hidden states and parameters of the underlying system, we require  
113 methods that can handle these nonlinearities. In Friston et al. (2000), the parameters of a  
114 hemodynamic model were estimated using a Volterra kernel expansion to characterize the  
115 hemodynamic response. Later, Friston et al. (2002) introduced a Bayesian estimation framework  
116 to invert (i.e., fit) the hemodynamic model explicitly. This approach accommodated prior  
117 constraints on parameters and avoided the need for Volterra kernels. Subsequently, the approach  
118 was generalized to cover networks of coupled regions and to include parameters controlling the  
119 neuronal coupling (effective connectivity) among brain regions (Friston et al., 2003). The  
120 Bayesian inversion of these models is known as dynamic causal modeling (DCM) and is now  
121 used widely to analyses effective connectivity in fMRI and electrophysiological studies.  
122 However, current approaches to hemodynamic and causal models only account for noise at the  
123 level of the measurement; where this noise includes thermally generated random noise and  
124 physiological fluctuations. This is important because physiological noise represents stochastic  
125 fluctuations due to metabolic and vascular responses, which affect the hidden states of the  
126 system; furthermore, neuronal activity can show pronounced endogenous fluctuations (Biswal et  
127 al., 1995; Krüger and Glover, 2001). Motivated by this observation, Riera et al. (2004) proposed  
128 a technique based on a fully stochastic model (i.e. including physiological noise) that used the  
129 local linearization filter (LLF) (Jimenez and Ozaki, 2003), which can be considered a form of  
130 extended Kalman filtering (EKF) (Haykin, 2001) for continuous dynamic systems. Besides  
131 estimating hemodynamic states and parameters, this approach allows one to estimate the  
132 system's input, i.e. neuronal activity; by its parameterization via radial basis functions (RBFs). In

133 Riera et al. (2004), the number of RBFs was considered fixed *a priori*, which means that the  
134 solution has to lie inside a regularly distributed but sparse space (otherwise, the problem is  
135 underdetermined). Recently, the LLF technique was applied by Sotero et al. (2009) to identify  
136 the states and parameters of a metabolic/hemodynamic model.

137 The hemodynamic response and hidden states of hemodynamic models possess strong  
138 nonlinear characteristics, which are prescient with respect to stimulus duration (Birn et al., 2001;  
139 Miller et al., 2001). This makes one wonder whether a linearization approach such as LLF can  
140 handle such strong nonlinearities. Johnston et al. (2008) proposed particle filtering, a sequential  
141 Monte Carlo method, that accommodates true nonlinearities in the model. The approach of  
142 Johnston et al. was shown to be both accurate and robust, when used to estimate hidden  
143 physiologic and hemodynamic states; and was superior to LLF, though a suboptimal numerical  
144 procedure was used in evaluating LLF. Similarly, two-pass particle filtering, including a  
145 smoothing (backwards pass) procedure, was introduced by Murray et al. (2008). Another attempt  
146 to infer model parameters and hidden states used the unscented Kalman filter (UKF), which is  
147 more suitable for highly nonlinear problems (Hu et al., 2009). Finally, Jacobson et al. (2008)  
148 addressed inference on model parameters, using a Metropolis–Hastings algorithm for sampling  
149 their posterior distribution.

150 None of the methods mentioned above, except (Riera et al., 2004) with its restricted  
151 parameterization of the input, can perform a complete deconvolution of fMRI signals and  
152 estimate both hidden states and input; i.e. the neuronal activation, without knowing the input  
153 (stimulation function). Here, an important exception is the methodology introduced by Friston et  
154 al. (2008) called dynamic expectation maximization (DEM) and its generalizations: variational  
155 filtering (Friston, 2008a) and generalized filtering (Friston et al., 2010). DEM represents a  
156 variational Bayesian technique (Hinton and van Camp, 1993; MacKay, 1995), that is applied to  
157 models formulated in terms of generalized coordinates of motion. This scheme allows one to  
158 estimate not only the states and parameters but also the input and hyperparameters of the system  
159 generating those states. Friston et al. (2008) demonstrated the robustness of DEM compared to  
160 standard Bayesian filtering methods, particularly the extended Kalman filter and particle filter,  
161 on a selection of difficult nonlinear/linear dynamic systems. They concluded that standard

162 methods are unable to perform joint estimation of the system input and states, while inferring the  
163 model parameters.

164 In this paper, we propose an estimation scheme that is based on nonlinear Kalman filtering,  
165 using the recently introduced cubature Kalman filter (CKF) (Arasaratnam and Haykin, 2009),  
166 which is recognized as the closest known approximation to Bayesian filtering. Our procedure  
167 applies a forward pass using the CKF that is finessed by a backward pass of the cubature Rauch-  
168 Tung-Striebel smoother. Moreover, we utilize the efficient square-root formulation of these  
169 algorithms. Crucially, we augment the hidden states with both parameters and inputs, enabling us  
170 to identify hidden states, model parameters and estimate the system input. We will show that we  
171 can obtain accurate estimates of hidden hemodynamic and neuronal states, well beyond the  
172 temporal resolution of fMRI.

173 The paper is structured as follows: First, we review the general concept of nonlinear  
174 continuous-discrete state-space models for simultaneous estimation of the system hidden states,  
175 its input and parameters. We then introduce the forward-backward cubature Kalman estimation  
176 procedure in its stable square-root form, as a suitable method for solving this complex inversion  
177 problem. Second, we provide a comprehensive evaluation of our proposed scheme and compare  
178 it with DEM. For this purpose, we use the same nonlinear/linear dynamic systems that were used  
179 to compare DEM with the EKF and particle filter algorithms (Friston et al., 2008). Third, we  
180 devote special attention to the deconvolution problem, given observed hemodynamic responses;  
181 i.e. to the estimation of neuronal activity and parameter identification of a hemodynamic model.  
182 Again, we provide comparative evaluations with DEM and discuss the advantages and  
183 limitations of each approach, when applied to fMRI data.

## 184 **Nonlinear continuous-discrete state-space models**

185 Nonlinear filtering problems are typically described by state-space models comprising a  
186 process and measurement equation. In many practical problems, the process equation is derived  
187 from the underlying physics of a continuous dynamic system, and is expressed in the form of a  
188 set of differential equations. Since the measurements  $y$  are acquired by digital devices; i.e. they  
189 are available at discrete time points ( $t = 1, 2, \dots, T$ ), we have a model with a continuous process

190 equation and a discrete measurement equation. The stochastic representation of this state-space  
 191 model, with additive noise, can be formulated as:

$$\begin{aligned} dx_t &= \mathbf{h}(x_t, \theta_t, u_t, t)dt + \mathbf{I}(x_t, t)d\beta_t, \\ y_t &= \mathbf{g}(x_t, \theta_t, u_t, t) + r_t, \end{aligned} \tag{1}$$

192 where  $\theta_t$  represents unknown parameters of the equation of motion  $\mathbf{h}$  and the measurement  
 193 function  $\mathbf{g}$ , respectively;  $u_t$  is the exogenous input (the cause) that drives hidden states or the  
 194 response;  $r_t$  is a vector of random Gaussian measurement noise,  $r_t \sim \mathcal{N}(0, R_t)$ ;  $\mathbf{I}(x_t, t)$  can be a  
 195 function of the state and time; and  $\beta_t$  denotes a Wiener process or state noise that is assumed to  
 196 be independent of states and measurement noise.

197 The continuous time formulation of the stochastic differential equations (SDE) in (1) can  
 198 also be expressed using Riemann and Ito integrals (Kloeden and Platen, 1999):

$$x_{t+\Delta t} = x_t + \int_t^{t+\Delta t} \mathbf{h}(x_t, \theta_t, u_t, t)dt + \int_t^{t+\Delta t} \mathbf{I}(x_t, t)d\beta_t, \tag{2}$$

199 where the second integral is stochastic. This equation can be further converted into a discrete-  
 200 time analogue using numerical integration such as Euler-Maruyama method or the local  
 201 linearization (LL) scheme (Biscay et al., 1996; Ozaki, 1992). This leads to the standard form of a  
 202 first order autoregressive process (AR(1)) of nonlinear state-space models:

$$\begin{aligned} x_t &= \mathbf{f}(x_{t-1}, \theta_t, u_t) + q_t \\ y_t &= \mathbf{g}(x_t, \theta_t, u_t) + r_t, \end{aligned} \tag{3}$$

203 where  $q_t$  is a zero-mean Gaussian state noise vector;  $q_t \sim \mathcal{N}(0, Q_t)$ . Our preference is to use LL-  
 204 scheme, which has been demonstrated to improve the order of convergence and stability  
 205 properties of conventional numerical integrators (Jimenez et al., 1999). In this case, the function  
 206  $\mathbf{f}$  is evaluated through:

$$\mathbf{f}(x_{t-1}, \theta_t, u_t) \approx x_{t-1} + f_{x_t}^{-1}[\exp(f_{x_t} \Delta t) - I]\mathbf{h}(x_{t-1}, \theta_t, u_t), \tag{4}$$

207 where  $f_{x_t}$  is a Jacobian of  $\mathbf{h}$  and  $\Delta t$  is the time interval between samples (up to the sampling  
 208 interval). The LL method allows integration of a SDE near discretely and regularly distributed  
 209 time instants, assuming local piecewise linearity. This permits the conversion of a SDE system

210 into a state-space equation with Gaussian noise. A stable reconstruction of the trajectories of the  
211 state-space variables is obtained by a one step prediction. Note that expression in (4) is not  
212 always the most practical; it assumes the Jacobian has full rank. See (Jimenez, 2002) for  
213 alternative forms.

#### 214 *Probabilistic inference*

215 The problem of estimating the hidden states (causing data), parameters (causing the  
216 dynamics of hidden states) and any non-controlled exogenous input to the system, in a situation  
217 when only observations are given, requires probabilistic inference. In Markovian setting, the  
218 optimal solution to this problem is given by the recursive Bayesian estimation algorithm which  
219 recursively updates the posterior density of the system state as new observations arrive. This  
220 posterior density constitutes the complete solution to the probabilistic inference problem, and  
221 allows us to calculate an “optimal” estimate of the state. In particular, the hidden state  $x_t$ , with  
222 initial probability  $p(x_0)$ , evolves over time as an indirect or partially observed first-order  
223 Markov process, according to the conditional probability density  $p(x_t|x_{t-1})$ . The observations  
224  $y_t$  are conditionally independent, given the state, and are generated according to the conditional  
225 posterior probability density  $p(y_t|x_t)$ . In this sense, the discrete-time variant of state-space  
226 model presented in Eq. (3) can also be written in terms of transition densities and a Gaussian  
227 likelihood:

$$\begin{aligned} p(x_t|x_{t-1}) &= \mathcal{N}(x_t|\mathbf{f}(x_{t-1}, u_t, \theta_t), Q) \\ p(y_t|x_t) &= \mathcal{N}(y_t|\mathbf{g}(x_t, \theta_t, u_t), R). \end{aligned} \tag{5}$$

228 The state transition density  $p(x_t|x_{t-1})$  is fully specified by  $\mathbf{f}$  and the state noise distribution  
229  $p(q_t)$ , whereas  $\mathbf{g}$  and the measurement noise distribution  $p(r_t)$  fully specify the observation  
230 likelihood  $p(y_t|x_t)$ . The dynamic state-space model, together with the known statistics of the  
231 noise (and the prior distribution of the system states), defines a probabilistic generative model of  
232 how system evolves over time and of how we (partially or inaccurately) observe this hidden state  
233 (Van der Merwe, 2004).

234 Unfortunately, the optimal Bayesian recursion is usually tractable only for linear, Gaussian  
235 systems, in which case the closed-form recursive solution is given by the classical Kalman filter

236 (Kalman, 1960) that yields the optimal solution in the minimum-mean-square-error (MMSE)  
237 sense, the maximum likelihood (ML) sense, and the maximum *a posteriori* (MAP) sense. For  
238 more general real-world (nonlinear, non-Gaussian) systems the optimal Bayesian recursion is  
239 intractable and an approximate solution must be used.

240 Numerous approximation solutions to the recursive Bayesian estimation problem have been  
241 proposed over the last couple of decades, in a variety of fields. These methods can be loosely  
242 grouped into the following four main categories:

- 243 • *Gaussian approximate methods*: These methods model the pertinent densities by Gaussian  
244 distributions, under assumption that a consistent minimum variance estimator (of the  
245 posterior state density) can be realized through the recursive propagation and updating of the  
246 first and second order moments of the true densities. Nonlinear filters that fall under this  
247 category are (in chronological order): a) the extended Kalman filter (EKF), which linearizes  
248 both the nonlinear process and measurement dynamics with a first-order Taylor expansion  
249 about current state estimate; b) the local linearization filter (LLF) is similar to EKF, but the  
250 approximate discrete time model is obtained from piecewise linear discretization of nonlinear  
251 state equation; c) the unscented Kalman filter (UKF) (Julier et al., 2002) chooses  
252 deterministic sample (sigma) points that capture the mean and covariance of a Gaussian  
253 density. When propagated through the nonlinear function, these points capture the true mean  
254 and covariance up to a second-order of the nonlinear function; d) the divided difference filter  
255 (DDF) (Norgaard et al., 2000) uses Stirling's interpolation formula. As with the UKF, DDF  
256 uses a deterministic sampling approach to propagate Gaussian statistics through the nonlinear  
257 function; e) the Gaussian sum filters (GSF) approximates both the predicted and posterior  
258 densities as sum of Gaussian densities, where the mean and covariance of each Gaussian  
259 density is calculated using separate and parallel instances of EKF or UKF; f) the quadrature  
260 Kalman filter (QKF) (Ito and Xiong, 2002) uses the Gauss-Hermite numerical integration  
261 rule to calculate the recursive Bayesian estimation integrals, under a Gaussian assumption; g)  
262 the cubature Kalman filter (CKF) is similar to UKF, but uses the spherical-radial integration  
263 rule.
- 264 • *Direct numerical integration methods*: these methods, also known as grid-based filters (GBF)  
265 or point-mass method, approximate the optimal Bayesian recursion integrals with large but

266 finite sums over a uniform  $N$ -dimensional grid that covers the complete state-space in the  
267 area of interest. For even moderately high dimensional state-spaces, the computational  
268 complexity can become untenably large, which precludes any practical use of these filters  
269 (Bucy and Senne, 1971).

- 270 • *Sequential Monte-Carlo (SMC) methods*: these methods (called particle filters) use a set of  
271 randomly chosen samples with associated weights to approximate the density (Doucet et al.,  
272 2001). Since the basic sampling dynamics (importance sampling) degenerates over time, the  
273 SMC method includes a re-sampling step. As the number of samples (particles) becomes  
274 larger, the Monte Carlo characterization of the posterior density becomes more accurate.  
275 However, the large number of samples often makes the use of SMC methods computationally  
276 prohibitive.
- 277 • *Variational Bayesian methods*: variational Bayesian methods approximate the true posterior  
278 distribution with a tractable approximate form. A lower bound on the marginal likelihood  
279 (evidence) of the posterior is then maximized with respect to the free parameters of this  
280 approximation (Jaakkola, 2000).

281 The selection of suitable sub-optimal approximate solutions to the recursive Bayesian  
282 estimation problem represents a trade-off between global optimality on one hand and  
283 computational tractability (and robustness) on the other hand. In our case, the best criterion for  
284 sub-optimality is formulated as: “Do as best as you can, and not more”. Under this criterion, the  
285 natural choice is to apply the cubature Kalman filter (Arasaratnam and Haykin, 2009). The CKF  
286 is the closest known direct approximation to the Bayesian filter, which outperforms all other  
287 nonlinear filters in any Gaussian setting, including particle filters (Arasaratnam and Haykin,  
288 2009; Fernandez-Prades and Vila-Valls, 2010; Li et al., 2009). The CKF is numerically accurate,  
289 can capture true nonlinearity even in highly nonlinear systems, and it is easily extendable to high  
290 dimensional problems (the number of sample points grows linearly with the dimension of the  
291 state vector).

### 292 *Cubature Kalman filter*

293 The cubature Kalman filter is a recursive, nonlinear and derivative free filtering algorithm,  
294 developed under Kalman filtering framework. It computes the first two moments (i.e. mean and  
295 covariance) of all conditional densities using a highly efficient numerical integration method

296 (cubature rules). Specifically, it utilizes the third-degree spherical-radial rule to approximate the  
 297 integrals of the form (*nonlinear function*  $\times$  *Gaussian density*) numerically using a set of  $m$   
 298 equally weighted symmetric cubature points  $\{\xi_i, \omega_i\}_{i=1}^m$ :

$$\int_{\mathbb{R}^N} \mathbf{f}(x) \mathcal{N}(x; 0, I_N) dx \approx \sum_{i=1}^m \omega_i \mathbf{f}(\xi_i), \quad (6)$$

$$\xi = \sqrt{\frac{m}{2}} [I_N, -I_N], \quad \omega_i = \frac{1}{m}, \quad i = 1, 2, \dots, m = 2N, \quad (7)$$

299 where  $\xi_i$  is the  $i$ -th column of the cubature points matrix  $\xi$  with weights  $\omega_i$  and  $N$  is dimension  
 300 of the state vector.

301 In order to evaluate the dynamic state-space model described by (3), the CKF includes two  
 302 steps: a) a time update, after which the predicted density  $p(x_t | y_{1:t-1}) = \mathcal{N}(\hat{x}_{t|t-1}, P_{t|t-1})$  is  
 303 computed; and b) a measurement update, after which the posterior density  $p(x_t | y_{1:t}) =$   
 304  $\mathcal{N}(\hat{x}_{t|t}, P_{t|t})$  is computed. For a detailed derivation of the CKF algorithm, the reader is referred  
 305 to (Arasaratnam and Haykin, 2009). We should note that even though CKF represents a  
 306 derivative-free nonlinear filter, our formulation of the continuous-discrete dynamic system  
 307 requires first order partial derivatives implicit in the Jacobian, which is necessary for  
 308 implementation of LL scheme. Although, one could use simple Euler's methods to approximate  
 309 the numerical solution of the system (Sitz et al., 2002), local linearization generally provides  
 310 more accurate solutions (Valdes Sosa et al., 2009). Note that since the Jacobian is only needed to  
 311 discretise continuous state variables in the LL approach (but for each cubature point), the main  
 312 CKF algorithm remains discrete and derivative-free.

### 313 *Parameters and input estimation*

314 Parameter estimation sometimes referred to as system identification can be regarded as a  
 315 special case of general state estimation in which the parameters are absorbed into the state  
 316 vector. Parameter estimation involves determining the nonlinear mapping:

$$y_t = \mathcal{D}(x_t; \theta_t), \quad (8)$$

317 where the nonlinear map  $\mathcal{D}(\cdot)$  is, in our case, the dynamic model  $\mathbf{f}(\cdot)$  parameterized by the  
 318 vector  $\theta_t$ . The parameters  $\theta_t$  correspond to a stationary process with an identity state-transition  
 319 matrix, driven by an "artificial" process noise  $w_t \sim \mathcal{N}(0, W_t)$  (the choice of variance  $W_t$

320 determines convergence and tracking performance and is generally small). The input or cause of  
321 motion on hidden states  $u_t$  can also be treated in this way, with input noise  $v_t \sim \mathcal{N}(0, V_t)$ . This is  
322 possible because of the so-called natural condition of control (Arasaratnam and Haykin, 2009),  
323 which says that the input  $u_t$  can be generated using the state prediction  $\hat{x}_{t|t-1}$ .

324 A special case of system identification arises when the input to the nonlinear mapping  
325 function  $\mathcal{D}(\cdot)$ , i.e. our hidden states  $x_t$ , cannot be observed. This then requires both state  
326 estimation and parameter estimation. For this dual estimation problem, we consider a discrete-  
327 time nonlinear dynamic system, where the system state  $x_t$ , the parameters  $\theta_t$  and the input  $u_t$ ,  
328 must be estimated simultaneously from the observed noisy signal  $y_t$ . A general theoretical and  
329 algorithmic framework for dual Kalman filter based estimation was presented by Nelson (2000)  
330 and Van der Merwe (2004). This framework encompasses two main approaches, namely joint  
331 estimation and dual estimation. In the dual filtering approach, two Kalman filters are run  
332 simultaneously (in an iterative fashion) for state and parameter estimation. At every time step,  
333 the current estimate of the parameters  $\theta_t$  is used in the state filter as a given (known) input and  
334 likewise, the current estimate of the state  $\hat{x}_t$  is used in the parameter filter. This results in a step-  
335 wise optimization within the joint state-parameter space. On the other hand, in the joint filtering  
336 approach, the unknown system state and parameters are concatenated into a single higher-  
337 dimensional joint state vector,  $\tilde{x}_t = [x_t, u_t, \theta_t]^T$ . It was shown in (Van der Merwe, 2004) that  
338 parameter estimation based on nonlinear Kalman filtering represents an efficient online 2<sup>nd</sup> order  
339 optimization method that can be also interpreted as a recursive Newton-Gauss optimization  
340 method. They also showed that nonlinear filters like UKF and CKF are robust in obtaining  
341 globally optimal estimates, whereas EKF is very likely to get stuck in a non-optimal local  
342 minimum.

343 There is a prevalent opinion that the performance of joint estimation scheme is superior to  
344 dual estimation scheme (Ji and Brown, 2009; Nelson, 2000; Van der Merwe, 2004). Therefore,  
345 the joint CKF is used below to estimate states, input, and parameters. Note that since the  
346 parameters are estimated online with the states, the convergence of parameter estimates depends  
347 also on the length of the time series.

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

$$\tilde{x}_t = \begin{bmatrix} x_t \\ u_t \\ \theta_t \end{bmatrix} = \begin{bmatrix} \mathbf{f}(x_{t-1}, \theta_{t-1}, u_{t-1}) \\ u_{t-1} \\ \theta_{t-1} \end{bmatrix} + \begin{bmatrix} q_{t-1} \\ v_{t-1} \\ w_{t-1} \end{bmatrix} \quad (9)$$

$$y_t = \mathbf{g}(\tilde{x}_t) + r_{t-1}$$

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

$$P_t = \begin{bmatrix} P_{x_t} & P_{x_t u_t} & P_{x_t \theta_t} \\ P_{u_t x_t} & P_{u_t} & P_{u_t \theta_t} \\ P_{\theta_t x_t} & P_{\theta_t u_t} & P_{\theta_t} \end{bmatrix} \quad (10)$$

351 This full covariance structure allows the joint estimation framework not only to deal with  
 352 uncertainty about parameter and state estimates (through the cubature-point approach), but also  
 353 to model the interaction (conditional dependences) between the states and parameters, which  
 354 generally provides better estimates.

355 Finally, the accuracy of the CKF can be further improved by augmenting the state vector  
 356 with all the noise components (Li et al., 2009; Wu et al., 2005), so that the effects of process  
 357 noise, measurement noise and parameter noise are explicitly available to the scheme model. By  
 358 augmenting the state vector with the noise variables (Eqs. 11 and 12), we account for uncertainty  
 359 in the noise variables in the same manner as we do for the states during the propagation of  
 360 cubature-points. This allows for the effect of the noise on the system dynamics and observations  
 361 to be treated with the same level of accuracy as state variables (Van der Merwe, 2004). It also  
 362 means that we can model noise that is not purely additive. Because this augmentation increases  
 363 the number of cubature points (by the number of noise components), it may also capture high  
 364 order moment information (like skew and kurtosis). However, if the problem does not require  
 365 more than first two moments, augmented CKF furnishes the same results as non-augmented  
 366 CKF.

### 367 *Square-root cubature Kalman filter*

368 In practice, Kalman filters are known to be susceptible to numerical errors due to limited  
 369 word-length arithmetic. Numerical errors can lead to propagation of an asymmetric, non-  
 370 positive-definite covariance, causing the filter to diverge (Kaminski et al., 1971). As a robust  
 371 solution to this, a square-root Kalman filter is recommended. This avoids the matrix square-

372 rooting operations  $P = SS^T$  that are necessary in the regular CKF algorithm by propagating the  
 373 square-root covariance matrix  $S$  directly. This has important benefits: preservation of symmetry  
 374 and positive (semi)definiteness of the covariance matrix, improved numerical accuracy, double  
 375 order precision, and reduced computational load. Therefore, we will consider the square-root  
 376 version of CKF (SCKF), where the square-root factors of the predictive posterior covariance  
 377 matrix are propagated (Arasaratnam and Haykin, 2009).

378 Bellow, we summarize the steps of SCKF algorithm. First, we describe the forward pass of a  
 379 joint SCKF for the simultaneous estimation of states, parameters, and of the input, where we  
 380 consider the state-space model in (9). Second, we describe the backward pass of the Rauch-  
 381 Tung-Striebel (RTS) smoother. This can be derived easily for SCKF due to its similarity with the  
 382 RTS smoother for square-root UKF (Simandl and Dunik, 2006). Finally, we will use the  
 383 abbreviation SCKS to refer to the combination of SCKF and our RTS square-root cubature  
 384 Kalman smoother. In other words, SCKF refers to the forward pass, which is supplemented with  
 385 a backward pass in SCKS.

### 386 *Forward filtering pass*

- 387 • *Filter initialization*

388 During initialization step of the filter we build the augmented form of state variable:

$$\hat{x}_0^a = E[x_0^a] = [\tilde{x}_0^T, 0, 0, 0, 0]^T = [x_0, u_0, \theta_0, 0, 0, 0, 0]^T. \quad (11)$$

389 The effective dimension of this augmented state is  $N = n_x + n_u + n_\theta + n_q + n_v + n_w + n_r$ ,  
 390 where  $n_x$  is the original state dimension,  $n_u$  is dimension of the input,  $n_\theta$  is dimension of the  
 391 parameter vector,  $\{n_q, n_v, n_w\}$  are dimensions of the noise components (equal to  $n_x, n_u, n_\theta$ ,  
 392 respectively), and  $n_r$  is the observation noise dimension (equal to the number of observed  
 393 variables). In a similar manner, the augmented state square-root covariance matrix is assembled  
 394 from the individual (square-roots) covariance matrices of  $x$ ,  $u$ ,  $\theta$ ,  $q$ ,  $v$ ,  $w$ , and  $r$ :

$$S_0^a = \text{chol}(E[(x_0^a - \hat{x}_0^a)(x_0^a - \hat{x}_0^a)^T]) = \text{diag}(S_0, S_q, S_v, S_w, S_r), \quad (12)$$

$$S_0 = \begin{bmatrix} \sqrt{P_x} & 0 & 0 \\ 0 & \sqrt{P_u} & 0 \\ 0 & 0 & \sqrt{P_\theta} \end{bmatrix}, S_q = \sqrt{Q}, S_v = \sqrt{V}, S_w = \sqrt{W}, S_r = \sqrt{R}, \quad (13)$$

395 where  $P_x, P_u, P_\theta$  are process covariance matrices for states, input and parameters.  $Q, V, W$  are  
 396 their corresponding process noise covariances, respectively and  $R$  is the observation noise  
 397 covariance. The square-root representations of these matrices are calculated (13), where the  
 398 “chol” operator represents a Cholesky factorization for efficient matrix square-rooting and  
 399 “diag” forms block diagonal matrix.

400 • *Time update step*

401 We evaluate the cubature points ( $i = 1, 2, \dots, m = 2N$ ):

$$\mathcal{X}_{i,t-1|t-1}^a = S_{t-1|t-1}^a \xi_i + \hat{x}_{t-1|t-1}^a, \quad (14)$$

402 where the set of sigma points  $\xi$  is pre-calculated at the beginning of algorithm (Eq. 7). Next, we  
 403 propagate the cubature points through the nonlinear dynamic system of process equations and  
 404 add noise components:

$$\mathcal{X}_{i,t|t-1}^{x,u,\theta} = \mathbf{F} \left( \mathcal{X}_{i,t-1|t-1}^{a(x)}, \mathcal{X}_{i,t-1|t-1}^{a(u)}, \mathcal{X}_{i,t-1|t-1}^{a(\theta)} \right) + \mathcal{X}_{i,t-1|t-1}^{a(q,v,w)}, \quad (15)$$

405 where  $\mathbf{F}$  comprises  $[\mathbf{f}(x_{t-1}, \theta_{t-1}, u_{t-1}), u_{t-1}, \theta_{t-1}]^T$  as expressed in process equation (9). The  
 406 superscripts distinguish among the components of cubature points, which correspond to the  
 407 states  $x$ , input  $u$ , parameters  $\theta$  and their corresponding noise variables ( $q, v, w$ ) that are all  
 408 included in the augmented matrix  $\mathcal{X}^a$ . Note that the size of new matrix  $\mathcal{X}_{i,t|t-1}^{x,u,\theta}$  is only  $(n_x +$   
 409  $n_u + n_\theta) \times m$ .

410 We then compute the predicted mean  $\hat{x}_{t|t-1}$  and estimate the square-root factor of predicted  
 411 error covariance  $S_{t|t-1}$  by using weighted and centered (by subtracting the prior mean  $\hat{x}_{t|t-1}$ )  
 412 matrix  $X_{t|t-1}$ :

$$\hat{x}_{t|t-1} = \frac{1}{m} \sum_{i=1}^m \mathcal{X}_{i,t|t-1}^{x,u,\theta}. \quad (16)$$

$$S_{t|t-1} = \text{qr}(X_{t|t-1}), \quad (17)$$

$$X_{t|t-1} = \frac{1}{\sqrt{m}} [\mathcal{X}_{1,t|t-1}^{x,u,\theta} - \hat{x}_{t|t-1}, \mathcal{X}_{2,t|t-1}^{x,u,\theta} - \hat{x}_{t|t-1}, \dots, \mathcal{X}_{m,t|t-1}^{x,u,\theta} - \hat{x}_{t|t-1}]. \quad (18)$$

413 The expression  $S = \text{qr}(X)$  denotes triangularization, in the sense of the QR decomposition<sup>1</sup>,  
 414 where resulting  $S$  is a lower triangular matrix.

415 • *Measurement update step*

416 During the measurement update step we propagate the cubature points through the measurement  
 417 equation and estimate the predicted measurement:

$$\mathcal{Y}_{i,t|t-1} = \mathbf{g}(\mathcal{X}_{i,t|t-1}^x, \mathcal{X}_{i,t|t-1}^u, \mathcal{X}_{i,t|t-1}^\theta) + \mathcal{X}_{i,t-1|t-1}^{a(r)}, \quad (19)$$

$$\hat{\mathcal{Y}}_{t|t-1} = \frac{1}{m} \sum_{i=1}^m \mathcal{Y}_{i,t|t-1}. \quad (20)$$

418 Subsequently, the square-root of the innovation covariance matrix  $S_{yy,t|t-1}$  is estimated by using  
 419 weighted and centered matrix  $Y_{t|t-1}$ :

$$S_{yy,t|t-1} = \text{qr}(Y_{t|t-1}), \quad (21)$$

$$Y_{t|t-1} = \frac{1}{\sqrt{m}} [\mathcal{Y}_{1,t|t-1} - \hat{\mathcal{Y}}_{t|t-1}, \mathcal{Y}_{2,t|t-1} - \hat{\mathcal{Y}}_{t|t-1}, \dots, \mathcal{Y}_{m,t|t-1} - \hat{\mathcal{Y}}_{t|t-1}]. \quad (22)$$

420 This is followed by estimation of the cross-covariance  $P_{xy,t|t-1}$  matrix and Kalman gain  $K_t$ :

$$P_{xy,t|t-1} = X_{t|t-1} Y_{t|t-1}^T, \quad (23)$$

$$K_t = (P_{xy,t|t-1} / S_{yy,t|t-1}^T) / S_{yy,t|t-1}. \quad (24)$$

421 The symbol / represents the matrix right divide operator; i.e. the operation  $A/B$ , applies the back  
 422 substitution algorithm for an upper triangular matrix  $B$  and the forward substitution algorithm for  
 423 lower triangular matrix  $A$ .

424 Finally, we estimate the updated state  $\hat{\mathcal{X}}_{t|t}$  and the square-root factor of the corresponding error  
 425 covariance:

$$\hat{\mathcal{X}}_{t|t} = \hat{\mathcal{X}}_{t|t-1} + K_t (\mathcal{Y}_t - \hat{\mathcal{Y}}_{t|t-1}), \quad (25)$$

$$S_{t|t} = \text{qr}([X_{t|t-1} - K_t Y_{t|t-1}]). \quad (26)$$

---

<sup>1</sup> The QR decomposition is a factorization of matrix  $X^T$  into an orthogonal matrix  $Q$  and upper triangular matrix  $R$  such that  $X^T = QR$ , and  $XX^T = R^T Q^T QR = R^T R = SS^T$ , where the resulting square-root (lower triangular) matrix is  $S = R^T$ .

426 The difference  $y_t - \hat{y}_{t|t-1}$  in Eq. (25) is called the innovation or the residual. It basically reflects  
 427 the difference between the actual measurement and predicted measurement (prediction error).  
 428 Further, this innovation is weighted by Kalman gain, which minimizes the posterior error  
 429 covariance.

430 In order to improve convergence rates and tracking performance, during parameter  
 431 estimation, a Robbins-Monro stochastic approximation scheme for estimating the innovations  
 432 (Ljung and Söderström, 1983; Robbins and Monro, 1951) is employed. In our case, this involves  
 433 approximation of square-root matrix of parameter noise covariance  $S_{w_t}$  by:

$$S_{w_t} = \sqrt{(1 - \lambda_w)S_{w_{t-1}}^2 + \lambda_w \tilde{K}_t (y_t - \hat{y}_{t|t-1})(y_t - \hat{y}_{t|t-1})^T \tilde{K}_t^T}, \quad (27)$$

434 where  $\tilde{K}_t$  is the partition of Kalman gain matrix corresponding to the parameter variables, and  
 435  $\lambda_w \in (0,1]$  is scaling parameter usually chosen to be a small number (e.g.  $10^{-3}$ ). Moreover, we  
 436 constrain  $S_{w_t}$  to be diagonal, which implies an independence assumption on the parameters. Van  
 437 der Merwe (2004) showed that the Robbins-Monro method provides the fastest rate of  
 438 convergence and lowest final MMSE values. Additionally, we inject process noise artificially by  
 439 annealing the square-root covariance of process noise with  $S_{q_t} = \text{diag}((1/\sqrt{\lambda_q} - 1)S_{t-1|t-1}^x)$ ,  
 440 using  $\lambda_q = 0.9995$ ,  $\lambda_q \in (0,1]$  (Arasaratnam and Haykin, 2008).

#### 441 *Backward smoothing pass*

442 The following procedure is a backward pass, which can be used for computing the smoothed  
 443 estimates of time step  $t$  from estimates of time step  $t + 1$ . In other words, a separate backward  
 444 pass is used for computing suitable corrections to the forward filtering results to obtain the  
 445 smoothing solution  $p(x_t, y_{1:T}) = \mathcal{N}(\hat{x}_{t|T}^s | \hat{x}_{t|T}^s, P_{t|T}^s)$ . Because the smoothing and filtering  
 446 estimates of the last time step  $T$  are the same, we make  $\hat{x}_{T|T}^s = \hat{x}_{T|T}$ ,  $S_{T|T}^s = S_{T|T}$ . This means the  
 447 recursion can be used for computing the smoothing estimates of all time steps by starting from  
 448 the last step  $t = T$  and proceeding backward to the initial step  $t = 0$ . To accomplish this, all  
 449 estimates of  $\hat{x}_{0:T}$  and  $S_{0:T}$  from the forward pass have to be stored and are then called at the  
 450 beginning of each time step of backward pass (28,29).

- 451 • *Square-root cubature RTS smoother*

452 Each time step of the smoother is initialized by forming an augmented state vector  $\hat{\mathbf{x}}_{t|t}^a$  and  
 453 square-root covariance  $S_{t|t}^a$ , using estimates from the SCKF forward pass,  $\hat{\mathbf{x}}_{t|T}$ ,  $S_{t|T}$ , and square-  
 454 roots covariance matrices of the noise components:

$$\hat{\mathbf{x}}_{t|t}^a = [\hat{\mathbf{x}}_{t|T}^T, 0, 0, 0, 0]^T, \quad (28)$$

$$S_{t|t}^a = \text{diag}\{S_{t|T}, S_{q,T}, S_v, S_{w,T}, S_r\}. \quad (29)$$

455 We then evaluate and propagate cubature points through nonlinear dynamic system (SDEs are  
 456 integrated in forward fashion):

$$\mathcal{X}_{i,t|t}^a = S_{t|t}^a \xi_i + \hat{\mathbf{x}}_{t|t}^a, \quad (30)$$

$$\mathcal{X}_{i,t+1|t}^{x,u,\theta} = \mathbf{F}(\mathcal{X}_{i,t|t}^{a(x)}, \mathcal{X}_{i,t|t}^{a(u)}, \mathcal{X}_{i,t|t}^{a(\theta)}) + \mathcal{X}_{i,t|t}^{a(q,v,w)}. \quad (31)$$

457 We compute the predicted mean and corresponding square-root error covariance matrix:

$$\hat{\mathbf{x}}_{t+1|t} = \frac{1}{m} \sum_{i=1}^m \mathcal{X}_{i,t+1|t}^{x,u,\theta}, \quad (32)$$

$$S_{t+1|t} = \text{qr}(X_{t+1|t}), \quad (33)$$

$$X_{t+1|t} = \frac{1}{\sqrt{m}} [\mathcal{X}_{1,t+1|t}^{x,u,\theta} - \hat{\mathbf{x}}_{t+1|t}, \mathcal{X}_{2,t+1|t}^{x,u,\theta} - \hat{\mathbf{x}}_{t+1|t}, \dots, \mathcal{X}_{m,t+1|t}^{x,u,\theta} - \hat{\mathbf{x}}_{t+1|t}]. \quad (34)$$

458 Next, we compute the predicted cross-covariance matrix, where the weighted and centered  
 459 matrix  $X'_{t|t}$  is obtained by using the partition  $(x, u, \theta)$  of augmented cubature point matrix  $\mathcal{X}_{i,t|t}^a$   
 460 and the estimated mean  $\hat{\mathbf{x}}_{t|t}^a$  before it propagates through nonlinear dynamic system (i.e. the  
 461 estimate from forward pass):

$$P_{x'x,t+1|t} = X'_{t|t} X_{t+1|t}^T, \quad (35)$$

$$X'_{t|t} = \frac{1}{\sqrt{m}} [\mathcal{X}_{1,t|t}^{a(x,u,\theta)} - \hat{\mathbf{x}}_{t|t}^{a(x,u,\theta)}, \mathcal{X}_{2,t|t}^{a(x,u,\theta)} - \hat{\mathbf{x}}_{t|t}^{a(x,u,\theta)}, \dots, \mathcal{X}_{m,t|t}^{a(x,u,\theta)} - \hat{\mathbf{x}}_{t|t}^{a(x,u,\theta)}]. \quad (36)$$

462 Finally, we estimate the smoother gain  $A_t$ , the smoothed mean  $\hat{\mathbf{x}}_{t|T}^s$  and the square-root  
 463 covariance  $S_{t|T}^s$ :

$$A_t = (P_{x'x,t+1|t} / S_{t+1|t}^T) / S_{t+1|t}, \quad (37)$$

$$\hat{\mathbf{x}}_{t|T}^s = \hat{\mathbf{x}}_{t|t}^{a(x,u,\theta)} + A_t (\hat{\mathbf{x}}_{t+1|T}^s - \hat{\mathbf{x}}_{t+1|t}), \quad (38)$$

$$S_{t|T}^S = \text{qr}([X_{t|t}' - A_t X_{t+1|t}, A_t S_{t+1|T}^S]). \quad (39)$$

464 Note that resulting error covariance  $S_{t|T}^S$  will be smaller than  $S_{t|t}$  from the forward run, as the  
 465 uncertainty over the state prediction is smaller when conditioned on all observations, than when  
 466 only conditioned on past observations.

467 This concludes our description of the estimation procedure, which can be summarized in the  
 468 following steps:

469 1) Evaluate the forward pass of the SCKF, where the continuous dynamic system of process  
 470 equations is discretized by an LL-scheme for all cubature points. Note that both time update  
 471 and measurement update steps are evaluated with an integration step  $\Delta t$ , and we linearly  
 472 interpolate between available observation values. In this case, we weight all noise  
 473 covariances by  $\sqrt{\Delta t}$ . In each time step of the filter evaluation we obtain predicted  
 474  $\{\hat{x}_{t|t-1}, \hat{u}_{t|t-1}, \hat{\theta}_{t|t-1}\}$  and filtered  $\{\hat{x}_{t|t}, \hat{u}_{t|t}, \hat{\theta}_{t|t}\}$  estimates of the states, parameters and the  
 475 inputs. These predicted estimates are used to estimate prediction errors  $e_t = y_t - \hat{y}_t$ , which  
 476 allows us to calculate the log-likelihood of the model given the data as:

$$\log p(y_{1:T}|\theta) = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \sum_{t=1}^T \left[ \log |S_{yy,t|t-1} S_{yy,t|t-1}^T| + \frac{e_t e_t^T}{S_{yy,t|t-1} S_{yy,t|t-1}^T} \right]. \quad (40)$$

477 2) Evaluate the backward pass of the SCKS to obtain smoothed estimates of the states  $\hat{x}_{t|T}^S$ , the  
 478 input  $\hat{u}_{t|T}^S$ , and the parameters  $\hat{\theta}_{t|T}^S$ . Again, this operation involves discretization of the  
 479 process equations by the LL-scheme for all cubature points.

480 3) Iterate until the stopping condition is met. We evaluate log-likelihood (40) at each iteration  
 481 and terminate the optimization when the increase of the (negative) log-likelihood is less than  
 482 a tolerance value of e.g.  $10^{-3}$ .

483 Before we turn to the simulations, we provide with a brief description of DEM, which is used for  
 484 comparative evaluations.

485 *Dynamic expectation maximization*

486 DEM is based on variational Bayes, which is a generic approach to model inversion (Friston  
 487 et al., 2008). Briefly, it approximates the conditional density  $p(\vartheta|y, m)$  on some model  
 488 parameters,  $\vartheta = \{x, u, \theta, \eta\}$ , given a model  $m$ , and data  $y$ , and it also provides lower-bound on  
 489 the evidence  $p(y|m)$  of the model itself. In addition, DEM assumes a continuous dynamic  
 490 system formulated in generalized coordinates of motion, where some parameters change with  
 491 time, i.e. hidden states  $x$  and input  $u$ , and rest of the parameters are time-invariant. The state-  
 492 space model has the form:

$$\begin{aligned} \tilde{y} &= \tilde{\mathbf{g}}(x, u, \theta) + \tilde{r} \\ D\tilde{x} &= \tilde{\mathbf{f}}(x, u, \theta) + \tilde{q}, \end{aligned} \quad (41)$$

493 where

$$\tilde{\mathbf{g}} = \begin{bmatrix} g = \mathbf{g}(x, u, \theta) \\ g' = g_x x' + g_u u' \\ g'' = g_x x'' + g_u u'' \\ \vdots \end{bmatrix}, \quad \tilde{\mathbf{f}} = \begin{bmatrix} f = \mathbf{f}(x, u, \theta) \\ f' = f_x x' + f_u u' \\ f'' = f_x x'' + f_u u'' \\ \vdots \end{bmatrix}. \quad (42)$$

494 Here,  $\tilde{\mathbf{g}}$  and  $\tilde{\mathbf{f}}$  are the predicted response and motion of the hidden states, respectively.  $D$  is  
 495 derivative operator whose first leading diagonal contains identity matrices, and which links  
 496 successive temporal derivatives  $(x', x'', \dots; u', u'', \dots)$ . These temporal derivatives are directly  
 497 related to the embedding orders<sup>2</sup> that one can specify separately for input ( $d$ ) and for states ( $n$ ) *a*  
 498 *priori*. We will use embedding orders  $d = 3$  and  $n = 6$ .

499 DEM is formulated for the inversion of hierarchical dynamic causal models with (empirical)  
 500 Gaussian prior densities on the unknown parameters of generative model  $m$ . These parameters  
 501 are  $\{\theta, \eta\}$ , where  $\theta$  represents set of model parameters and  $\eta = \{\alpha, \beta, \sigma\}$  are hyperparameters,  
 502 which specify the amplitude of random fluctuations in the generative process. These  
 503 hyperparameters correspond to (log) precisions (inverse variances) on the state noise ( $\alpha$ ), the  
 504 input noise ( $\beta$ ), and the measurement noise ( $\sigma$ ), respectively. In contrast to standard Bayesian  
 505 filters, DEM also allows for temporal correlations among innovations, which is parameterized by  
 506 additional hyperparameter  $\gamma$  called temporal precision.

---

<sup>2</sup> The term “embedding order” is used in analogy with lags in autoregressive modeling.

507 DEM comprises three steps that optimize states, parameters and hyperparameters  
508 receptively: The first is the  $D$ -step, which evaluates Eq. (41), for the posterior mean, using the  
509 LL-scheme for integration of SDEs. Crucially, DEM (and its generalizations) does not use a  
510 recursive Bayesian scheme but tries to optimize the posterior moments of hidden states (and  
511 inputs) through an generalized (“instantaneous”) gradient ascent on (free-energy bound on) the  
512 marginal likelihood. This generalized ascent rests on using the generalized motion (time  
513 derivatives to high order) of variables as part of the model generating or predicting discrete data.  
514 This means that DEM is a formally simpler (although numerically more demanding) than  
515 recursive schemes and only requires a single pass though the time-series to estimate the states.

516 DEM comprises additional  $E$  (expectation) and  $M$  (maximization) steps that optimize the  
517 conditional density on parameters and hyperparameters (precisions) after the  $D$  (deconvolution)  
518 step. Iteration of these steps proceeds until convergence. For an exhaustive description of DEM,  
519 see (Friston et al., 2008). A key difference between DEM (variational and generalized filtering)  
520 and SCKS is that the states and parameters are optimized with respect to (a free-energy bound  
521 on) the log-evidence or marginal likelihood, having integrated out dependency on the  
522 parameters. In contrast, SCKS optimizes the parameters with respect to the log-likelihood in  
523 Equation (40), to provide maximum likelihood estimates of the parameters, as opposed to  
524 maximum *a posteriori* (MAP) estimators. This reflects the fact that DEM uses shrinkage priors  
525 on the parameters and hyperparameters, whereas SCKS does not. SCKS places priors on the  
526 parameter noise that encodes our prior belief that they do not change (substantially) over time.  
527 This is effectively a constraint on the volatility of the parameters (not their values *per se*), which  
528 allows the parameters to ‘drift’ slowly to their maximum likelihood value. This difference  
529 becomes important when evaluating one scheme in relation to the other, because we would  
530 expect some shrinkage in the DEM estimates to the prior mean, which we would not expect in  
531 the SCKS estimates (see next section).

532 DEM rests on a mean-field assumption used in variational Bayes; in other words, it assumes  
533 that the states, parameters and hyperparameters are conditionally independent. This assumption  
534 can be relaxed by absorbing the parameters and hyperparameters into the states as in SCKS. The  
535 resulting scheme is called generalized filtering (Friston et al., 2010). Although generalized  
536 filtering is formally more similar to SCKS than DEM (and is generally more accurate), we have

537 chosen to use DEM in our comparative evaluations because DEM has been validated against  
538 EKF and particle filtering (whereas generalized filtering has not). Furthermore, generalized  
539 filtering uses prior constraints on both the parameters and how fast they can change. In contrast,  
540 SCKS and DEM only use one set of constraints on the change and value of the parameters,  
541 respectively. However, we hope to perform this comparative evaluation in a subsequent paper;  
542 where we will consider Bayesian formulations of cubature smoothing in greater detail and relate  
543 its constraints on changes in parameters to the priors used in generalized filtering.

544 Finally, for simplicity, we assume that the schemes have access to all the noise (precision)  
545 hyperparameters, meaning that they are not estimated. In fact, for SCKS we assume only the  
546 precision of measurement noise to be known and update the assumed values of the  
547 hyperparameters for fluctuations in hidden states and input during the inversion (see Eq. 27). We  
548 can do this because we have an explicit representation of the errors on the hidden states and  
549 input.

## 550 **Inversion of dynamic models by SCKF and SCKS**

551 In this section, we establish the validity and accuracy of the SCKF and SCKS scheme in  
552 relation to DEM. For this purpose, we analyze several nonlinear and linear continuous stochastic  
553 systems that were previously used for validating of DEM, where its better performance was  
554 demonstrated in relation to the EKF and particle filtering. In particular, we consider the well  
555 known Lorenz attractor, a model of a double well potential, a linear convolution model and,  
556 finally, we devote special attention to the inversion of a hemodynamic model. Even though some  
557 of these models might seem irrelevant for hemodynamic and neuronal modeling, they are  
558 popular for testing the effectiveness of inversion schemes and also (maybe surprisingly) exhibit  
559 behaviors that can be seen in models used in neuroimaging.

560 To assess the performance of the various schemes, we perform Monte Carlo simulations,  
561 separately for each of these models; where the performance metric for the statistical efficiency of  
562 the estimators was the squared error loss function (SEL). For example, we define the SEL for  
563 states as:

$$SEL(x) = \sum_{t=1}^T (x_t - \hat{x}_t)^2. \quad (43)$$

564 Similarly, we evaluate SEL for the input and parameters (when appropriate). Since the SEL is  
 565 sensitive to outliers; i.e. when summing over a set of  $(x_t - \hat{x}_t)^2$ , the final sum tends to be biased  
 566 by a few large values. We consider this a convenient property when comparing the accuracy of  
 567 our cubature schemes and DEM. Furthermore, this measure of accuracy accommodates the  
 568 different constraints on the parameters in DEM (shrinkage priors on the parameters) and SCKS  
 569 (shrinkage priors on changes in the parameters). We report the SEL values in natural logarithmic  
 570 space; i.e.  $\log(\text{SEL})$ .

571 Note that all data based on the above models were simulated through the generation function  
 572 in the DEM toolbox (*spm\_DEM\_generate.m*) that is available as part of SPM8  
 573 (<http://www.fil.ion.ucl.ac.uk/spm/>).

574 **Table 1.** State and observation equations for dynamic systems.

	$\mathbf{f}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta})$	$\mathbf{g}(\mathbf{x}, \boldsymbol{\theta})$
Lorenz attractor	$\begin{bmatrix} \theta_1 x_2 - \theta_1 x_1 \\ \theta_3 x_1 - 2x_1 x_3 - x_2 \\ 2x_1 x_2 + \theta_2 x_3 \end{bmatrix} \frac{1}{32}$	$x_1 + x_2 + x_3$
Double-well	$\frac{2x}{1+x^2} - \frac{x}{16} + \frac{u}{4}$	$\frac{1}{16}x^2$
Convolution model	$\theta_2 x + \theta_3 u$	$\theta_1 x$
Hemodynamic model	$\begin{bmatrix} \epsilon u - \kappa(h_1 - 1) - \chi(h_2 - 1)/h_1 \\ (h_1 - 1)/h_2 \\ \tau(h_2 - F(h_3))/h_3 \\ \tau(h_2 E(h_2) - F(h_3)h_4/h_3)/h_4 \end{bmatrix}$	$V_0[k_1(1 - x_4) + k_2(1 - x_4/x_3) + k_3(1 - x_3)]$

575

576 **Table 2.** Parameters of the generative model for the simulated dynamic systems.

		<b>Lorenz</b>	<b>Double-well</b>	<b>Convolution</b>	<b>Hemodynamic</b>
Observation-noise precision	Simulated	$\sigma = 1$	$\sigma = e^2$	$\sigma = e^8$	$\sigma = e^6$
	Prior pdf	$\sim \mathcal{N}(0,1)$	$\sim \mathcal{N}(0, e^{-2})$	$\sim \mathcal{N}(0, e^{-8})$	$\sim \mathcal{N}(0, e^{-6})$
State-noise precision	Simulated	$\alpha = e^{16}$	$\alpha = e^{16}$	$\alpha = e^{12}$	$\alpha = e^8$
Input-noise precision	Simulated	-	$\beta = \frac{1}{8}$	$\beta = e^{16}$	$\beta = e^8$

	Prior pdf	-	$\sim \mathcal{N}(0,1)$	$\sim \mathcal{N}(0,0.1)$	$\sim \mathcal{N}(0,0.1)$
Parameter-noise precision	Prior pdf <sup>3</sup>	$\sim \mathcal{N}(0,0.1)$	-	$\sim \mathcal{N}(0, 10^{-4})$	Tab. 3
Initial conditions	Simulated	$x_0 = [0.9, 0.8, 30]^T$	$x_0 = 1$	$x_0 = [0, 0]^T$	$x_0 = [0, 0, 0]^T$

577

578 *Lorenz attractor*

579 The model of the Lorenz attractor exhibits deterministic chaos, where the path of the hidden  
580 states diverges exponentially on a butterfly-shaped strange attractor in a three dimensional state-  
581 space. There are no inputs in this system; the dynamics are autonomous, being generated by  
582 nonlinear interactions among the states and their motion. The path begins by spiraling onto one  
583 wing and then jumps to the other and back in chaotic way. We consider the output to be the  
584 simple sum of all three states at any time point, with innovations of unit precision  $\sigma = 1$  and  
585  $\gamma = 8$ . We further specified a small amount of the state noise ( $\alpha = e^{16}$ ). We generated 120 time  
586 samples using this model, with initial state conditions  $x_0 = [0.9, 0.8, 30]^T$ , parameters  $\theta =$   
587  $[18, -4, 46.92]^T$  and an LL-integration step  $\Delta t = 1$ .

588 This sort of chaotic system shows sensitivity to initial conditions; which, in the case of  
589 unknown initial conditions, is a challenge for any inversion scheme. Therefore, we first compare  
590 SCKF, SCKS and DEM when the initial conditions  $x_0$  differ from the true starting values, with  
591 known model parameters. This simulation was repeated five times with random initializations  
592 and different innovations. Since we do not estimate any parameters, only a single iteration of the  
593 optimization process is required. We summarized the resulting estimates in terms of the first two  
594 hidden states and plotted their trajectories against each other in their corresponding state-space  
595 (Fig. 1A). It can be seen that all three inversion schemes converge quickly to the true trajectories.  
596 DEM provides the least accurate estimate (but still exhibits high performance when compared to  
597 EKF and particle filters (Friston, 2008a; Friston et al., 2008)). The SCKF was able to track the  
598 true trajectories more closely. This accuracy is even more improved by SCKS, where the initial  
599 residuals are significantly smaller, hence providing the fastest convergence.

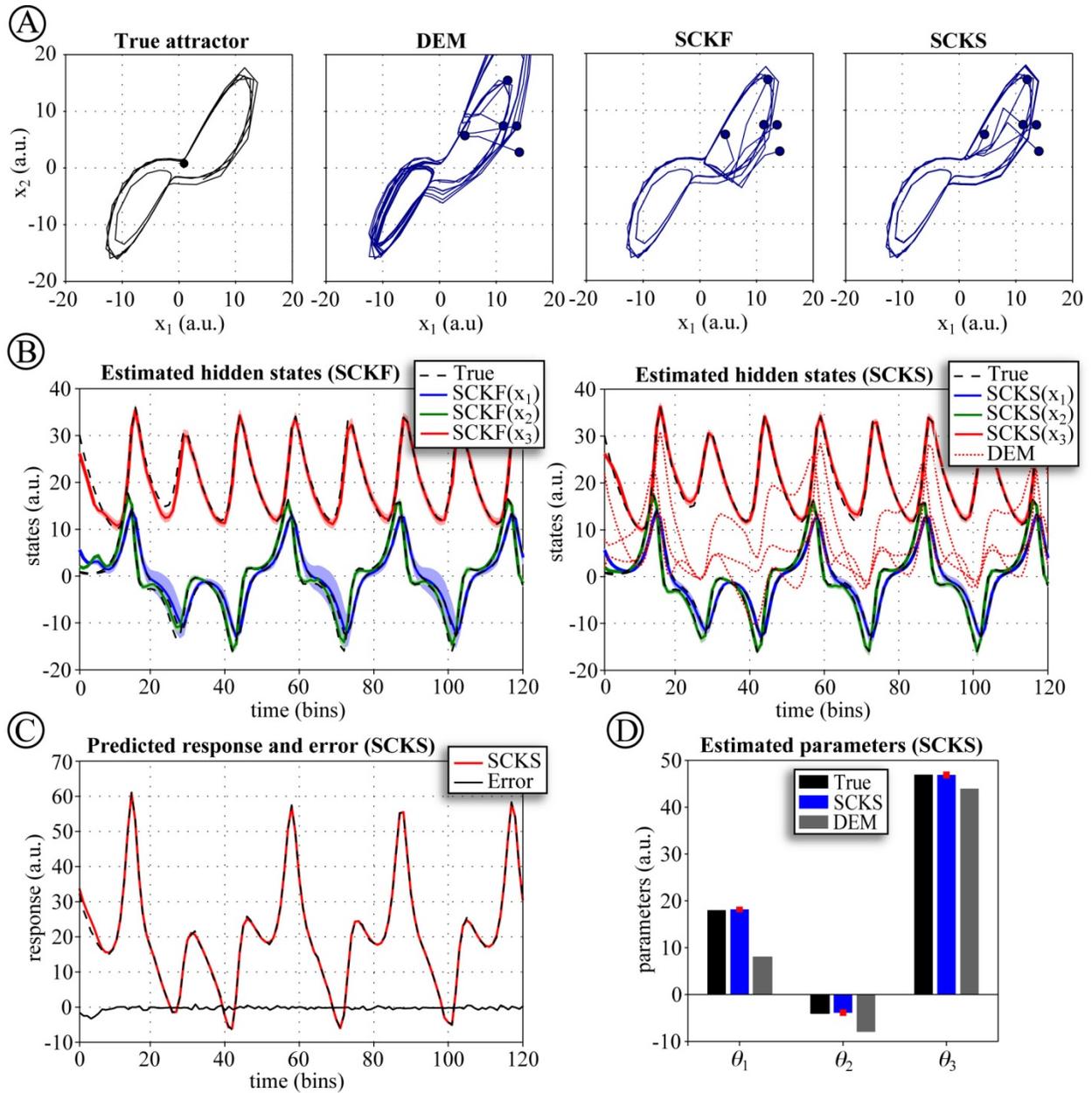
---

<sup>3</sup> Prior precision on parameter noise is used for initialization and during CKF step the parameter noise variance is estimated by Robbins-Monro stochastic approximation (27) with scaling parameter  $\lambda_w = 10^{-2}$  for the Lorenz attractor and  $\lambda_w = 10^{-3}$  for the convolution and hemodynamic models.

600 Next, we turned to testing the inversion schemes when both initial conditions and model  
601 parameters are unknown. We used initial state conditions  $x_0 = [2, 8, 22]^T$  and parameters  
602  $\theta_0 = [10, -8, 43]^T$ , where their true values were the same as above. We further assumed an  
603 initial prior precision on parameter noise  $p(\theta) = \mathcal{N}(0, 0.1)$ , and allowed the algorithm to iterate  
604 until the convergence. The SCKF and SCKS converged in 6 iteration steps, providing very  
605 accurate estimates of both states and parameters (Fig. 1B). This was not the case for DEM,  
606 which did not converge, exceeding the maximum allowed number of iteration, 50.

607 The reason for DEM's failure is that the updates to the parameters are not properly  
608 regularized in relation to their highly nonlinear impact on the trajectories of hidden states. In  
609 other words, DEM makes poor updates, which are insensitive to the highly nonlinear form of this  
610 model. Critically, SCKF and SCKS outperformed DEM because it uses an online parameter  
611 update scheme and were able to accommodate nonlinearities much more gracefully, through its  
612 cubature-point sampling. Heuristically, cubature filtering (smoothing) can be thought of as  
613 accommodating nonlinearities by relaxing the strong assumptions about the form of the  
614 likelihood functions used in optimizing estimates. DEM assume this form is Gaussian and  
615 therefore estimates its local curvature with second derivatives. A Gaussian form will be exact for  
616 linear models but not non-linear models. Conversely, cubature filtering samples this function  
617 over greater distances in state or parameter space and relies less on linear approximations

618



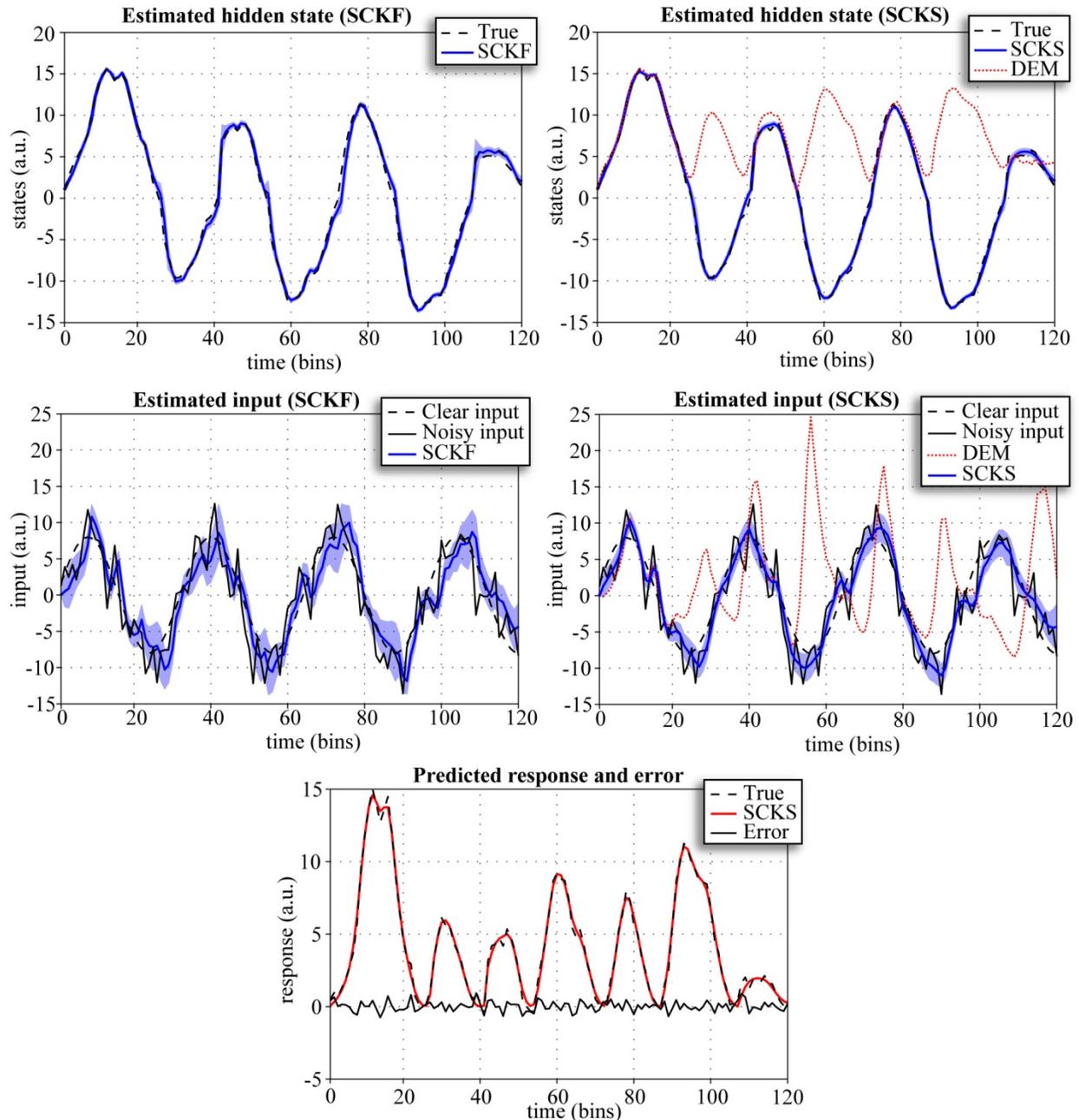
619

620 **Figure 1.** (A) The Lorenz attractor simulations were repeated five times, using different starting conditions (dots)  
 621 and different random innovations. The hidden states of this model were estimated using DEM, SCKF and SCKS.  
 622 Here, we summarize the resulting trajectories in terms of the first two hidden states, plotted against each other in  
 623 their corresponding state-space. The true trajectories are shown on the upper left. (B) The inversion of Lorenz  
 624 system by SCKF, SCKS and DEM. The true trajectories are shown as dashed lines, DEM estimates with dotted  
 625 lines, and SCKF and SCKS estimates with solid lines including the 90% posterior confidence intervals (shaded  
 626 areas). (C) Given the close similarity between the responses predicted by DEM and SCKS, we show only the result  
 627 for SCKS. (D) The parameters estimates are summarized in lower left in terms of their expectation and 90%  
 628 confidence intervals (red lines). Here we can see that DEM is unable to estimate the model parameters.

629 • *MC simulations:* To verify this result, we conducted a series of 100 Monte Carlo  
630 simulations under three different estimation scenarios. In the 1<sup>st</sup> scenario, we considered  
631 unknown initial conditions of hidden states but known model parameters. The initial conditions  
632 were sampled randomly from uniform distribution  $x_0 \sim \mathcal{U}(0,20)$ , and the true values were the  
633 same as in all previous cases. In the 2<sup>nd</sup> scenario, the initial states were known but the model  
634 parameters unknown, being sampled from the normal distribution around the true values  
635  $\theta_0 \sim \mathcal{N}(\theta_{true}, 10)$ . Finally, the 3<sup>rd</sup> scenario was combination of the first two; with both initial  
636 conditions and parameters unknown. In this case, the states were always initialized with  $x_0 =$   
637  $[2, 8, 22]^T$  and parameters sampled from the normal distribution. Results, in terms of average  
638  $\log(\text{SEL})$ , comparing the performance of SCKS and DEM are shown in Fig. 4.

### 639 *Double-well*

640 The double-well model represents a dissipative system with bimodal variability. What  
641 makes this system particularly difficult to invert for many schemes is the quadratic form of the  
642 observation function, which renders inference on the hidden states and their causes ambiguous.  
643 The hidden state is deployed symmetrically about zero in a double-well potential, which makes  
644 the inversion problem even more difficult. Transitions from one well to other can be then caused  
645 either by input or high amplitude fluctuations. We drove this system with slow sinusoidal input  
646  $u(t) = 8 \cdot \sin\left(\frac{1}{16}\pi t\right)$  and generated 120 time points response with noise precision  $\sigma = e^2$ , a  
647 small amount of state noise  $\alpha = e^{16}$ , and with a reasonable level of input noise  $\beta = 1/8$ . The  
648 temporal precision was  $\gamma = 2$  and LL-integration step again  $\Delta t = 1$ , with initial condition  
649  $x_0 = 1$ , and mildly informative (initial) prior on the input precision  $p(u) = \mathcal{N}(0,1)$ . We tried to  
650 invert this model using only observed responses by applying SCKF, SCKS and DEM. Fig. 2  
651 shows that DEM failed to estimate the true trajectory of the hidden state, in the sense that the  
652 state is always positive. This had an adverse effect on the estimated input and is largely because  
653 of the ambiguity induced by the observation function. Critically, the accuracy of the input  
654 estimate will be always lower than that of the state, because the input is expressed in  
655 measurement space vicariously through the hidden states. Nevertheless, SCKF and SCKS were  
656 able to identify this model correctly, furnishing accurate estimates for both the state and the  
657 input, even though this model represents a non-Gaussian (bimodal) problem (Fig. 2).



658  
 659 **Figure 2.** Inversion of the double-well model, comparing estimates of the hidden state and input from SCKF, SCKS  
 660 and DEM. This figure uses the same format as Fig. 1B,C. Again, the true trajectories are depicted with dashed lines  
 661 and the shaded area represents 90% posterior confidence intervals. Given the close similarity between the responses  
 662 predicted by DEM and SCKS, we show only the result for SCKS.

663 • *MC simulations:* To evaluate the stability of SCKS estimates in this context, we repeated  
 664 the simulations 100 times, using different innovations. It can be seen from the results in Fig. 4  
 665 that the SCKS estimates of the state and input are about twice as close to the true trajectories  
 666 than the DEM estimates. Nevertheless, the SCKS was only able to track the true trajectories of

667 the state and input completely (as shown in Fig 3.) in about 70% of all simulations. In remaining  
668 30% SCKS provided results where some half-periods of hidden state trajectories had the wrong  
669 sign; i.e. flipped around zero. At the present time, we have no real insight into why DEM fails  
670 consistently to cross from positive to negative conditional estimates, while the SCKS scheme  
671 appears to be able to do this. One might presume this is a reflection of cubature filtering's ability  
672 to handle the nonlinearities manifest at zero crossings. The reason this is a difficult problem is  
673 that the true posterior density over the hidden state is bimodal (with peaks at positive and  
674 negative values of hidden state). However, the inversion schemes assume the posterior is a  
675 unimodal Gaussian density, which is clearly inappropriate. DEM was not able to recover the true  
676 trajectory of the input for any simulation, which suggests that the cubature-point sampling in  
677 SCKS was able to partly compensate for the divergence between the true (bimodal) and assumed  
678 unimodal posterior.

679

### 680 *Convolution model*

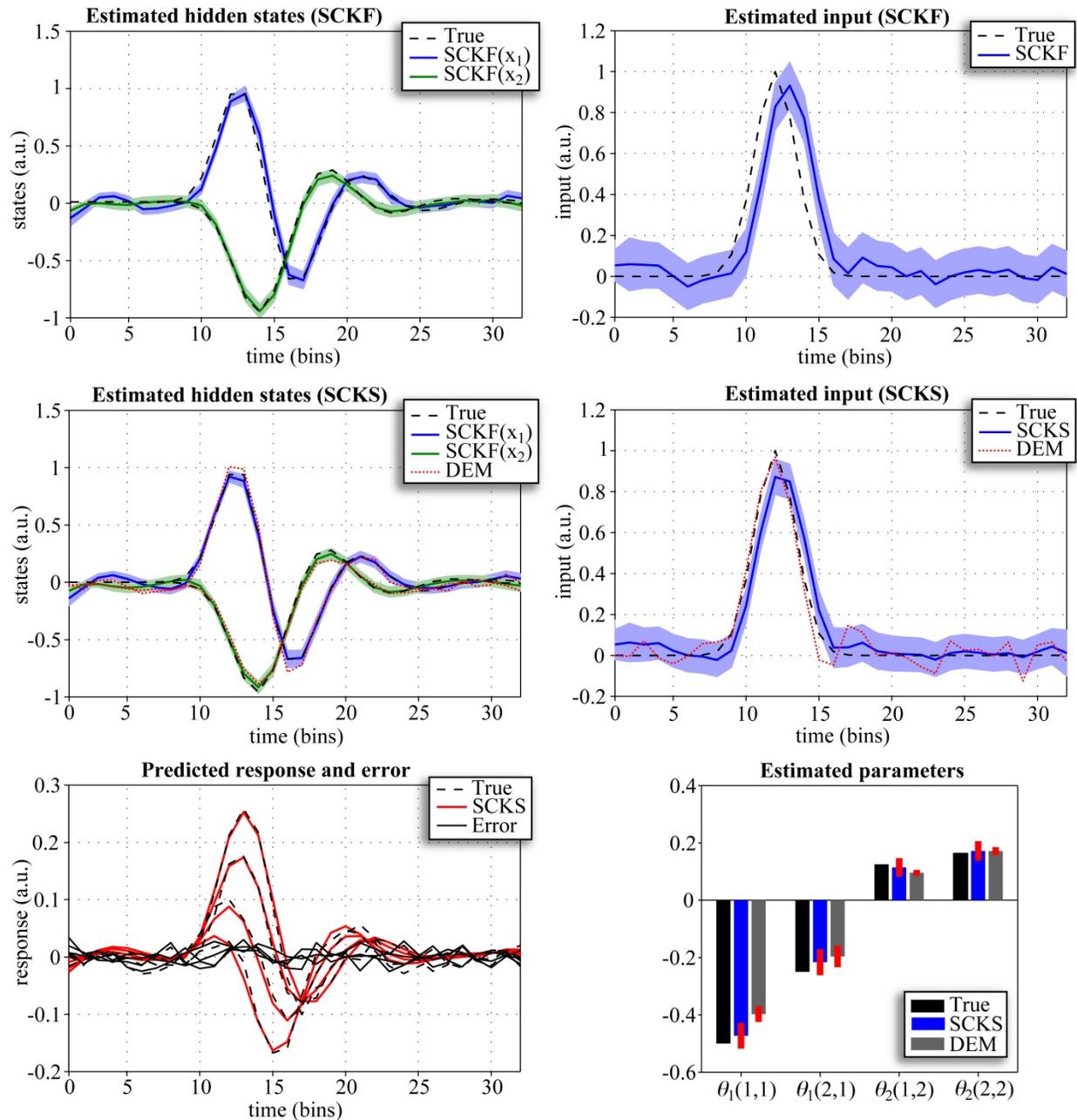
681 The linear convolution model represents another example that was used in (Friston, 2008a;  
682 Friston et al., 2008) to compare DEM, EKF, particle filtering and variational filtering. In this  
683 model (see Tab. 1), the input perturbs hidden states, which decay exponentially to produce an  
684 output that is a linear mixture of hidden states. Specifically, we used the input specified by  
685 Gaussian bump function of the form  $u(t) = \exp(\frac{1}{4}(t - 12)^2)$ , two hidden states and four output  
686 responses. This is a single input-multiple output system with the following parameters:

$$687 \theta_1 = \begin{bmatrix} 0.125 & 0.1633 \\ 0.125 & 0.0676 \\ 0.125 & -0.0676 \\ 0.125 & -0.1633 \end{bmatrix}, \theta_2 = \begin{bmatrix} -0.25 & 1.00 \\ -0.50 & -0.25 \end{bmatrix}, \theta_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

688 We generated data over 32 time points, using innovations sampled from Gaussian densities  
689 with precision  $\sigma = e^8$ , a small amount of state noise  $\alpha = e^{12}$  and minimal input noise  $\beta = e^{16}$ .  
690 The LL-integration step was  $\Delta t = 1$  and temporal precision  $\gamma = 4$ . During model inversion, the  
691 input and four model parameters are unknown and are subject to mildly informative prior  
692 precisions,  $p(u) = \mathcal{N}(0,0.1)$ , and  $p(\theta) = \mathcal{N}(0, 10^{-4})$ , respectively. Before initializing the  
693 inversion process, we set parameters  $\theta_1(1,1)$ ;  $\theta_1(2,1)$ ;  $\theta_2(1,2)$ ; and  $\theta_2(2,2)$  to zero. Fig. 3,  
694 shows that applying only a forward pass with SCKF does not recover the first hidden state and

695 especially the input correctly. The situation is improved with the smoothed estimates from  
696 SCKS, when both hidden states match the true trajectories. Nevertheless, the input estimate is  
697 still slightly delayed in relation to the true input. We have observed this delay repeatedly, when  
698 inverting this particular convolution model with SCKS. The input estimate provided by DEM is,  
699 in this case, correct, although there are more perturbations around the baseline compared to the  
700 input estimated by SCKS. The reason that DEM was able to track the input more accurately is  
701 that it has access to generalized motion. Effectively this means it sees the future data in a way  
702 that recursive update schemes (like SCKF) do not. This becomes important when dealing with  
703 systems based on high-order differential equations, where changes in a hidden state or input are  
704 expressed in terms of high-order temporal derivatives in data space (we will return to this issue  
705 later). Having said this, the SCKS identified the unknown parameters more accurately than  
706 DEM, resulting in better estimates of hidden states.

707

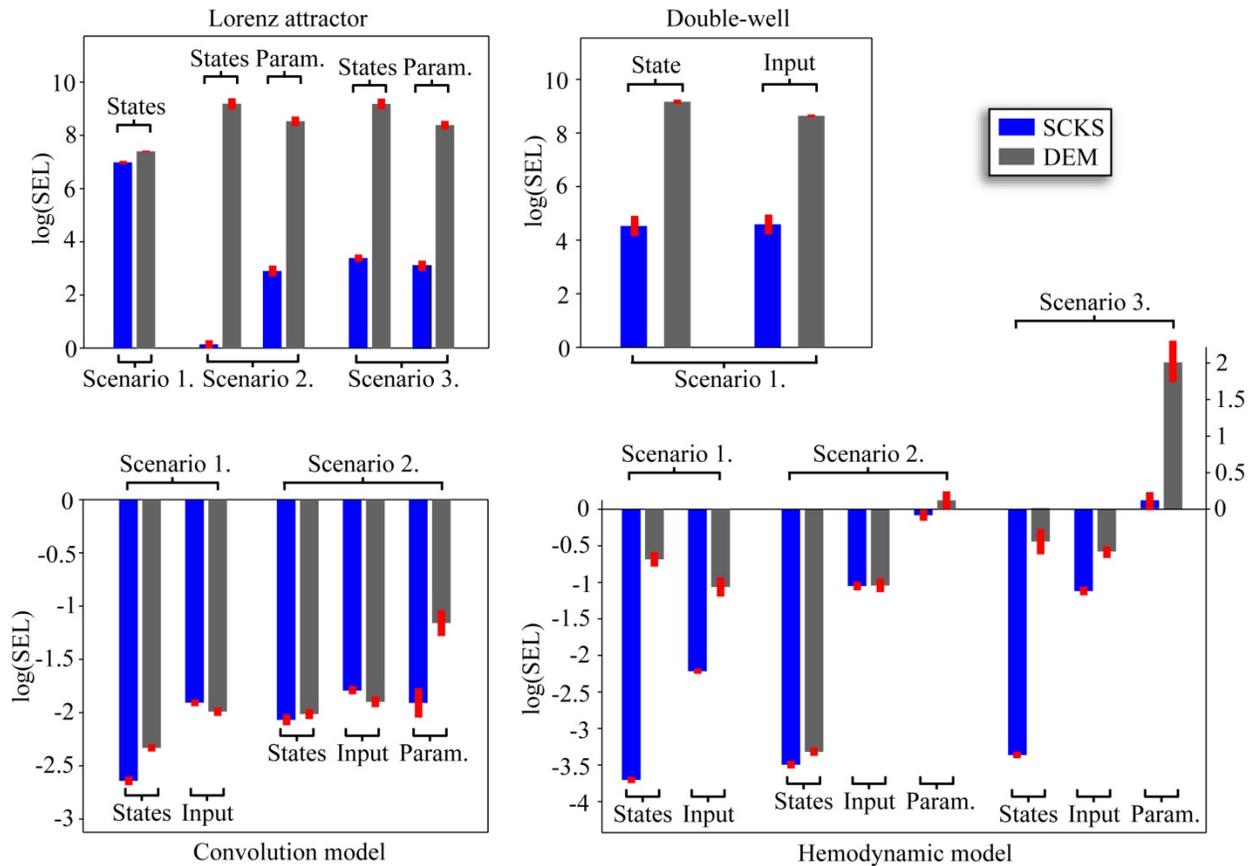


708

709 **Figure 3.** Results of inverting the linear convolution model using SCKF, SCKS and DEM; summarizing estimates  
 710 of hidden states, input, four model parameters and the response. This figure uses the same format as Fig. 1B,C,D.

711 • *MC simulations:* For Monte Carlo simulation we looked at two different scenarios. First,  
 712 we inverted the model when treating only the input as unknown, and repeated the simulations  
 713 100 times with different innovations. In the second scenario, which was also repeated 100 times  
 714 with different innovations, both input and the four model parameters were treated as unknown.  
 715 The values of these parameters were sampled from the normal distribution  $\theta_0 = \mathcal{N}(0,1)$ . Fig. 4,

716 shows that DEM provides slightly more accurate estimates of the input than SCKS. This is  
 717 mainly because of the delay issue above. However, SCKS again furnishes more accurate  
 718 estimates, with a higher precision on inverted states and markedly higher accuracy on the  
 719 identified model parameters.  
 720



721  
 722 **Figure 4.** The Monte Carlo evaluation of estimation accuracy using an average log(SEL) measure for all models  
 723 under different scenarios. The SEL measure is sensitive to outliers, which enables convenient comparison between  
 724 different algorithms tested on the same system. However, it cannot be used to compare performance among different  
 725 systems. A smaller log(SEL) value reflects a more accurate estimate. For quantitative intuition, a value of  
 726  $\log(\text{SEL}) = -2$  is equivalent to mean square error (MSE) of about  $2 \cdot 10^{-3}$  and a  $\log(\text{SEL}) = 7$  is a MSE of about  
 727  $7 \cdot 10^1$ .

728  
 729 *Hemodynamic model*

730 The hemodynamic model represents a nonlinear “convolution” model that was described  
 731 extensively in (Buxton et al., 1998; Friston et al., 2000). The basic kinetics can be summarized as  
 732 follows: Neural activity  $u$  causes an increase in vasodilatory signal  $h_1$  that is subject to auto-

733 regulatory feedback. Blood flow  $h_2$  responds in proportion to this signal and causes changes in  
734 blood volume  $h_3$  and deoxyhemoglobin content,  $h_4$ . These dynamics are modeled by a set of  
735 differential equations and the observed response is expressed as a nonlinear function of blood  
736 volume and deoxyhemoglobin content (see Tab. 1). In this model, the outflow is related to the  
737 blood volume  $F(h_3) = h_3^{1/\alpha}$  through Grubb's exponent  $\alpha$ . The relative oxygen extraction  
738  $E(h_2) = \frac{1}{\varphi} (1 - (1 - \varphi)^{1/h_2})$  is a function of flow, where  $\varphi$  is a resting oxygen extraction  
739 fraction. The description of model parameters, including the prior noise precisions is provided in  
740 Tab. 3.

741 **Table 3.** Hemodynamic model parameters.

Biophysical parameters of the state equations

	<b>Description</b>	<b>Value</b>	<b>Prior on noise variance</b>
$\kappa$	Rate of signal decay	$0.65 \text{ s}^{-1}$	$p(\theta_\kappa) = \mathcal{N}(0, 10^{-4})$
$\chi$	Rate of flow-dependent elimination	$0.38 \text{ s}^{-1}$	$p(\theta_\chi) = \mathcal{N}(0, 10^{-4})$
$\tau$	Hemodynamic transit time	0.98 s	$p(\theta_\tau) = \mathcal{N}(0, 10^{-4})$
$\alpha$	Grubb's exponent	0.34	$p(\theta_\alpha) = \mathcal{N}(0, 10^{-8})$
$\varphi$	Resting oxygen extraction fraction	0.32	$p(\theta_\varphi) = \mathcal{N}(0, 10^{-8})$
$\epsilon$	Neuronal efficiency	0.54	$p(\theta_\epsilon) = \mathcal{N}(0, 10^{-8})$

Fixed biophysical parameters of the observation equation

	<b>Description</b>	<b>Value</b>
$V_0$	Blood volume fraction	0.04
$k_1$	Intravascular coefficient	$7\varphi$
$k_2$	Concentration coefficient	2
$k_3$	Extravascular coefficient	$2\varphi-0.2$

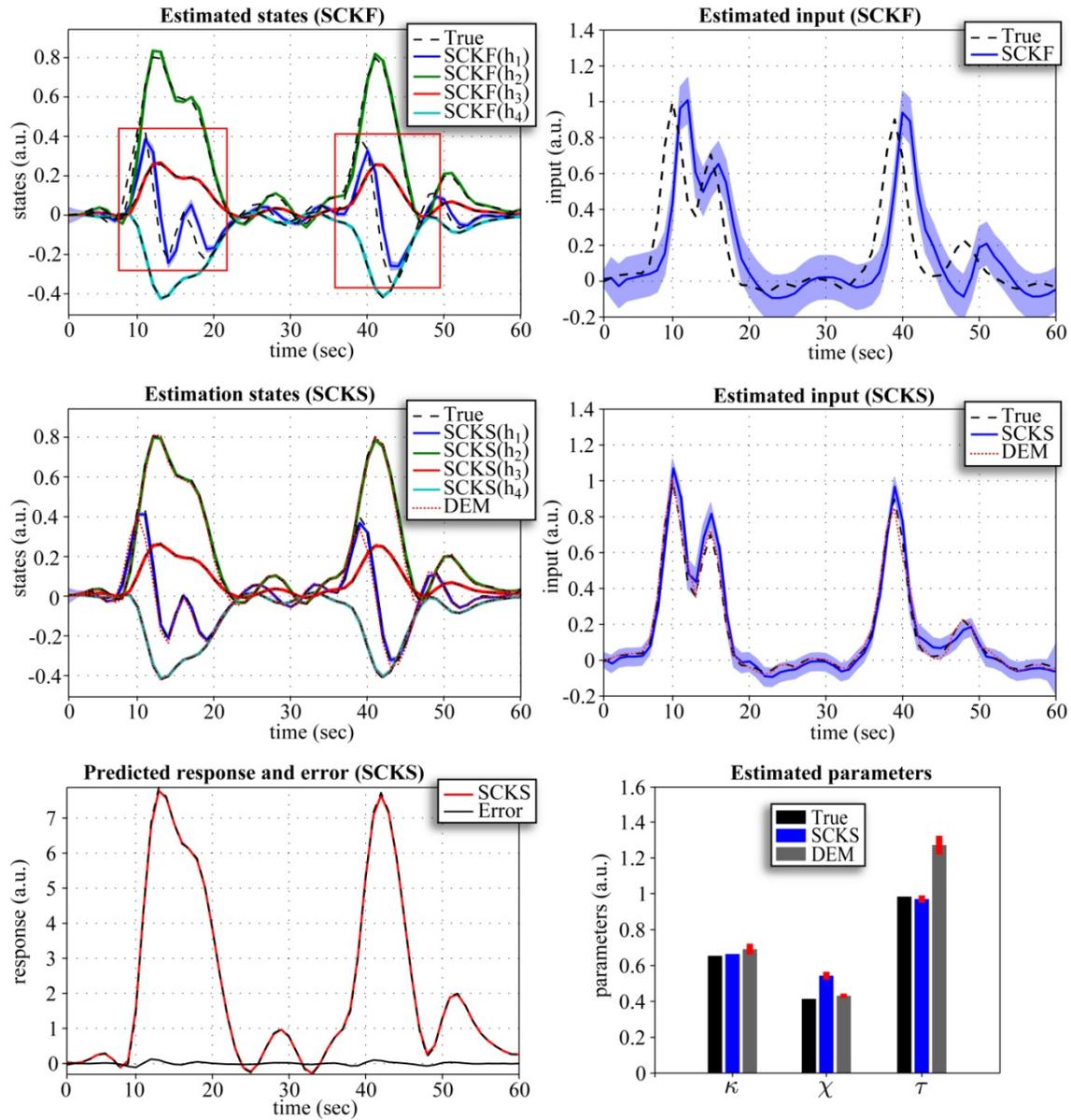
742

743 In order to ensure positive values of the hemodynamic states and improve numerical  
744 stability of the parameter estimation, the hidden states are transformed  $x_i = \log(h_i) \Leftrightarrow h_i =$   
745  $\exp(x_i)$ . However, before evaluating the observation equation, the log-hemodynamic states are  
746 exponentiated. The reader is referred to (Friston et al., 2008; Stephan et al., 2008) for a more  
747 detailed explanation.

748 Although there are many practical ways to use the hemodynamic model with fMRI data, we  
749 will focus here on its simplest instance; a single-input, single-output variant. We will try to  
750 estimate the hidden states and input through model inversion, and simultaneously identify model  
751 parameters from the observed response. For this purpose, we generated data over 60 time points  
752 using the hemodynamic model, with an input in the form of a Gaussian bump functions with  
753 different amplitudes centered at positions (10; 15; 39; and 48), and model parameters as reported  
754 in Tab. 2. The sampling interval or repeat time (TR) was equal to  $TR = 1$  sec. We added  
755 innovations to the output with a precision  $\sigma = e^6$ . This corresponds to a noise variance of about  
756 0.0025, i.e. in range of observation noise previously estimated in real fMRI data (Johnston et al.,  
757 2008; Riera et al., 2004), with a temporal precision  $\gamma = 1$ . The precision of state noise was  
758  $\alpha = e^8$  and precision of the input noise  $\beta = e^8$ . At the beginning of the model inversion, the true  
759 initial states were  $x_0 = [0,0,0,0]^T$ . Three of the six model parameters, specifically  $\theta = \{\kappa, \chi, \tau\}$ ,  
760 were initialized randomly, sampling from the normal distribution centered on the mean of the  
761 true values  $\theta_i = \mathcal{N}(\theta_i^{true}, 1/12)$ . The remaining parameters were based on their true values.  
762 The reasons for omitting other parameters from random initializations will be discussed later in  
763 the context of parameter identifiability. The prior precision of parameter noise are given in Tab.  
764 3, where we allowed a small noise variance ( $10^{-8}$ ) in the parameters that we considered to be  
765 known  $\{\alpha, \varphi, \epsilon\}$ ; i.e. these parameters can only experience very small changes during estimation.  
766 The parameter priors for DEM were as reported in (Friston et al., 2010) with the exception of  
767  $\{\alpha, \varphi\}$ , which we fixed to their true values.

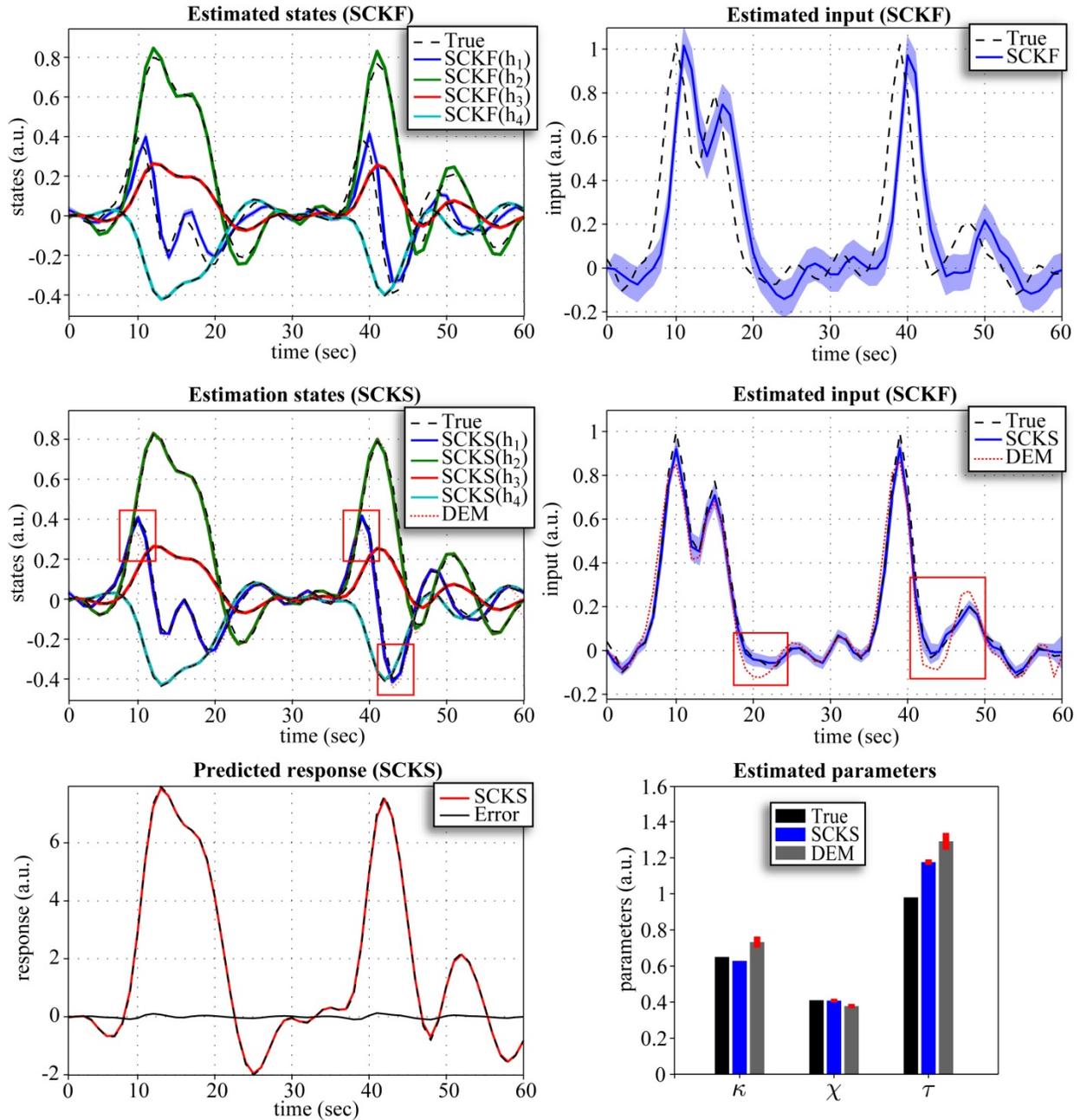
768 For model inversion we considered two scenarios that differed in the size of the integration  
769 step. First, we applied an LL-integration step of  $\Delta t = 0.5$ ; in the second scenario, we decreased  
770 the step to  $\Delta t = 0.2$ . Note that all noise precisions are scaled by  $\sqrt{\Delta t}$  before estimation begins.  
771 The same integration steps were also used for DEM, where we additionally increased the  
772 embedding orders ( $n = d = 8$ ) to avoid numerical instabilities. The results are depicted in Fig. 5  
773 and 6. It is noticeable that in both scenarios neither the hidden states nor input can be estimated  
774 correctly by SCKF. For  $\Delta t = 0.5$ , SCKS estimates the input less accurately than DEM, with  
775 inaccuracies in amplitude and in the decaying part of the Gaussian input function, compared to  
776 the true trajectory. This occurred even though the hidden states were tracked correctly. The  
777 situation is very different for  $\Delta t = 0.2$ : Here the results obtained by SCKS are very precise for

778 both the states and input. This means that a finer integration step had beneficial effects on both  
779 SCKF and SCKS estimators. In contrast, the DEM results did not improve. Here, including more  
780 integration steps between observation samples decreased the estimation accuracy for the input  
781 and the states. This means that DEM, which models high order motion, does not require the  
782 small integration steps necessary for SCKF and SCKS. Another interesting point can be made  
783 regarding parameter estimation. As we mentioned above, SCKS estimated the hidden states in  
784 both scenarios accurately, which might lead to the conclusion that the model parameters were  
785 also indentified correctly. However, although some parameters were indeed identified optimally  
786 (otherwise we would not obtain correct states) they were not equal to the true values. This is due  
787 to the fact that the effects of some parameters (on the output) are redundant, which means  
788 different sets of parameter values can provide veridical estimates of the states. For example, the  
789 effects of increasing the first parameter can be compensated by decreasing the second, to  
790 produce exactly the same output. This feature of the hemodynamic model has been discussed  
791 before in (Deneux and Faugeras, 2006) and is closely related to identifiably issues and  
792 conditional dependence among parameters estimates.



793

794 **Figure 5.** Results of the hemodynamic model inversion by SCKF, SCKS and DEM, with an integration step of  
 795  $\Delta t = 0.5$  and the first three model parameters were identified. This figure uses the same format as Fig 1B,C,D.



796

797 **Figure 6.** Results of the hemodynamic model inversion by SCKF, SCKS and DEM, with an integration step of  
 798  $\Delta t = 0.2$  and the first three model parameters were identified. This figure uses the same format as Fig 1B,C,D.

799 • *MC simulations:* We examined three different scenarios for the hemodynamic model  
 800 inversion. The simulations were inverted using an integration step  $\Delta t = 0.2$  for SCKF and SCKS  
 801 and  $\Delta t = 0.5$  for DEM. First, we focus on performance when the input is unknown, we have  
 802 access to the true (fixed) parameters and the initial states are unknown. These were sampled  
 803 randomly from the uniform distribution  $x_0 \sim \mathcal{U}(0,0.5)$ . In the second scenario, the input is again

804 unknown, and instead of unknown initial conditions we treated three model parameters  $\theta =$   
805  $\{\kappa, \chi, \tau\}$  as unknown. Finally in the last scenario, all three variables (i.e. the initial conditions,  
806 input, and three parameters) are unknown. All three simulations were repeated 100 times with  
807 different initializations of  $x_0, \theta_0$ , innovations, and state and input noise. From the MC simulation  
808 results, the following interesting behaviors were observed. Since the DEM estimates are  
809 calculated only in a forward manner, if the initial states are incorrect, it takes a finite amount of  
810 time before they converge to their true trajectories. This error persists over subsequent iterations  
811 of the scheme (E-steps) because they are initialized with the same incorrect state. This problem is  
812 finessed with SCKS: Although the error will be present in the SCKF estimates of the first  
813 iteration, it is efficiently corrected during the smoothing by SCKS, which brings the initial  
814 conditions closer to their true values. This enables an effective minimization of the initial error  
815 over iterations. This feature is very apparent from MC results in terms of  $\log(\text{SEL})$  for all three  
816 scenarios. When the true initial state conditions are known (2<sup>nd</sup> scenario), the accuracy of the  
817 input estimate is the same for SCKS and DEM, SCKS has only attained slightly better estimates  
818 of the states, hence also better parameter estimates. However, in the case of unknown initial  
819 conditions, SCKS is superior (see Fig 4).

820 • *Effect of model parameters on hemodynamic response and their estimation*

821 Although the biophysical properties of hemodynamic states and their parameters were  
822 described extensively in (Buxton et al., 1998; Friston et al., 2000), we will revisit the  
823 contribution of parameters to the final shape of hemodynamic response function (see Fig. 7A). In  
824 particular, our interest is in the parameters  $\theta = \{\kappa, \chi, \tau, \alpha, \varphi, \epsilon\}$ , which play a role in the  
825 hemodynamic state equations. We evaluated changes in hemodynamic responses over a wide  
826 range of parameters values (21 regularly spaced values for each parameter). In Fig. 7A, the red  
827 lines represent biologically plausible mean parameter values that were estimated empirically in  
828 (Friston et al., 2000), and which are considered to be the true values here (Tab. 3). The arrows  
829 show change in response when these parameters are increased. The first parameter is  $\kappa = 1/\tau_s$ ,  
830 where  $\tau_s$  is the time constant of signal decay. Increasing this parameter dampens the  
831 hemodynamic response to any input and suppresses its undershoot. The second parameter  
832  $\chi = 1/\tau_f$  is defined by the time constant of the auto-regulatory mechanism  $\tau_f$ . The effect of  
833 increasing parameter  $\chi$  (decreasing the feedback time constant  $\tau_f$ ) is to increase the frequency of

834 the response and lower its amplitude, with small change of the undershoot (see also the effect on  
835 the first hemodynamic state  $h_1$ ). The parameter  $\tau$  is the mean transit time at rest, which  
836 determines the dynamics of the signal. Increasing this parameter slows down the hemodynamic  
837 response, with respect to flow changes. It also slightly reduces response amplitude and more  
838 markedly suppresses the undershoot. The next parameter is the stiffness or Grub's exponent  $\alpha$ ,  
839 which is closely related to the flow-volume relationship. Increasing this parameter increases the  
840 degree of nonlinearity of the hemodynamic response, resulting in decreases of the amplitude and  
841 weaker suppression of undershoot. Another parameter of hemodynamic model is resting oxygen  
842 extraction fraction  $\varphi$ . Increasing this parameter can have quite profound effects on the shape of  
843 the hemodynamic response that bias it towards an early dip. This parameter has an interesting  
844 effect on the shape of the response: During the increase of  $\varphi$ , we first see an increase of the  
845 response peak amplitude together with deepening of undershoot, whereas after the value passes  
846  $\varphi = 0.51$ , the undershoot is suppressed. Response amplitude continues to grow until  $\varphi = 0.64$   
847 and falls rapidly after that. Additionally, the early dip starts to appear with  $\varphi = 0.68$  and higher  
848 values. The last parameter is the neuronal efficacy  $\epsilon$ , which simply modulates the hemodynamic  
849 response. Increasing this parameter scales the amplitude of the response.

850 In terms of system identification, it has been shown in (Deneux and Faugeras, 2006) that  
851 very little accuracy is lost when values of Grub's exponent and resting oxygen extraction fraction  
852  $\{\alpha, \varphi\}$  are fixed to some physiologically plausible values. This is in accordance with (Riera et al.,  
853 2004), where these parameters were also fixed. Grub's exponent is supposed to be stable during  
854 steady-state stimulation (Mandeville et al., 1999);  $\alpha = 0.38 \pm 0.1$  with almost negligible effects  
855 on the response within this range. The resting oxygen extraction fraction parameter is responsible  
856 for the early dip that is rarely observed in fMRI data. Its other effects can be approximated by  
857 combining parameters  $\{\kappa, \tau\}$ . In our case, where the input is unknown, the neuronal efficiency  
858 parameter  $\epsilon$  is fixed as well. This is necessary, because a change in this parameter is degenerate  
859 with respect to the amplitude of neuronal input.

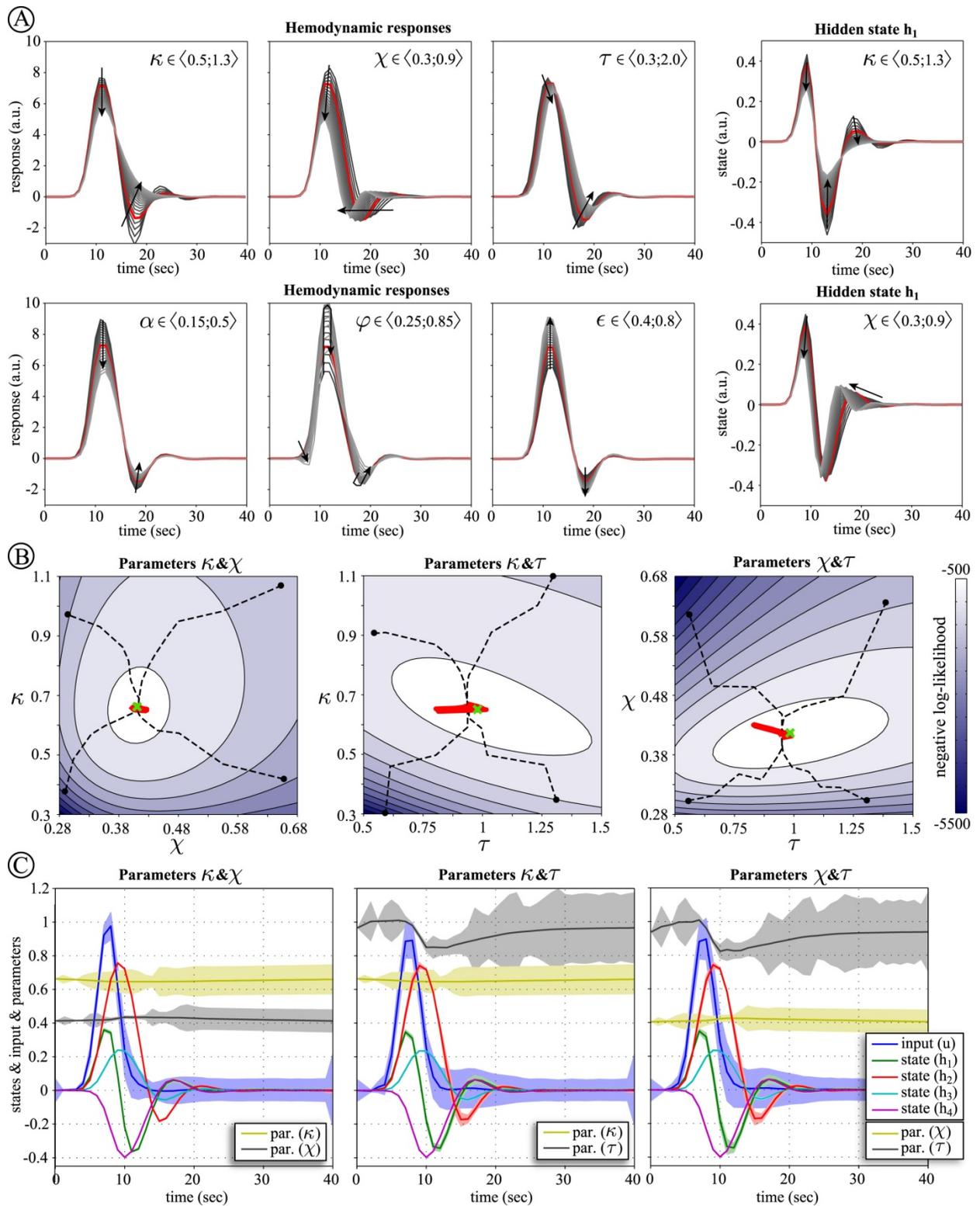
860 To pursue this issue of identifiability we examined the three remaining parameters  $\theta =$   
861  $\{\kappa, \chi, \tau\}$  in terms of the (negative) log-likelihood for pairs of these three parameters; as estimated  
862 by the SCKS scheme (Fig. 7B). The curvature (Hessian) of this log-likelihood function is, in  
863 fact, the conditional precision (inverse covariance) used in variational schemes like DEM and is

864 formally related to the Fisher Information matrix for the parameters in question. A slow  
865 curvature (shallow) basin means that we are conditionally uncertain about the precise value and  
866 that large changes in parameters will have relatively small effects on the observed response or  
867 output variables. The global optimum (true values) is marked by the green crosslet. To compute  
868 these log-likelihoods we ran SCKS for all combinations of parameters within their selected  
869 ranges, assuming the same noise precisions as in the hemodynamic simulations above (Tab. 3).  
870 Note that we did not perform any parameter estimation, but only evaluated log-likelihood for  
871 different parameter values, having optimized the states. Looking at the ensuing (color-coded)  
872 optimization manifolds, particularly at the white area bounded by the most inner contour, we can  
873 see how much these parameters can vary around the global optimum and still provide reasonably  
874 accurate predictions (of output, hidden states and input). This range is especially wide for the  
875 mean transient time  $\tau$ . One can see from the plot at the top of Fig. 7A that changing  $\tau =$   
876  $\langle 0.3; 2.0 \rangle$  over a wide range has little effect on the response. The region around the global  
877 maximum also discloses conditional dependencies and redundancy among the parameters. These  
878 dependencies make parameter estimation a generally more difficult task.

879         Nevertheless, we were curious if, at least under certain circumstances, the true parameter  
880 values could be estimated. Therefore, we allowed for faster dynamics on the parameters  $\{\kappa, \chi, \tau\}$   
881 by using higher noise variances ( $4 \cdot 10^{-4}$ ,  $2 \cdot 10^{-4}$ ,  $10^{-2}$ , respectively) and evaluated all three  
882 possible parameter combinations using SCKS. In other words, we optimized two parameters with  
883 the third fixed, over all combinations. These noise parameters were chosen after intensive  
884 testing, to establish the values that gave the best estimates. We repeated these inversions four  
885 times, with different initial parameter estimates selected within the manifolds shown in Fig. 7A.  
886 In Fig. 7B, we can see how the parameters moved on the optimization surface, where the black  
887 dashed line depicts the trajectory of the parameter estimates over successive iterations, starting  
888 from the initial conditions (black dot) and terminating around the global optimum (maximum).  
889 The red thick line represents the dynamic behavior of parameters over time during the last  
890 iteration. The last iteration estimate for all states, input and parameters is depicted in Fig. 7C.  
891 Here the dynamics of transit time ( $\tau$ ) is especially interesting; it drops with the arrival of the  
892 neuronal activation and is consequently restored during the resting period. This behavior is  
893 remarkably similar to that observed by Mandeville et al. (1999) in rat brains, where mean transit  
894 time falls during activation. Clearly, we are not suggesting that the transit time actually

895 decreased during activation in our simulations (it was constant during the generation of data).  
896 However, these results speak to the interesting application of SCKS to identify time-dependent  
897 changes in parameters. This could be important when applied to dynamic causal models of  
898 adaptation or learning studies that entail changes in effective connectivity between neuronal  
899 populations. The key message here is that if one can (experimentally) separate the time scale of  
900 true changes in parameters from the (fast) fluctuations inherent in recursive Bayesian filtering (or  
901 generalized filtering), it might be possible to estimate (slow) changes in parameters that are of  
902 great experimental interest.

903 In general, enforcing slow dynamics on the parameters (with a small noise variance) will  
904 ensure more accurate results for both states and input, provided the true parameters also change  
905 slowly. Moreover, we prefer to consider all parameters of hemodynamic state equations as  
906 unknown and limit their variations with high prior precisions. This allows us to treat all the  
907 unknown parameters uniformly; where certain (assumed) parameters can be fixed to their prior  
908 mean using an infinitely high prior precision.



909

910 **Figure 7.** (A) The top row depicts the effect of changing the hemodynamic model parameters on the response and  
 911 on the first hidden state. For each parameter, the range of values considered is reported, comprising 21 values.  
 912 (B) The middle row shows the optimization surfaces (manifolds) of negative log-likelihood obtained via SCKS for

913 combinations of the first three hemodynamic model parameters  $\{\kappa, \chi, \tau\}$ . The trajectories of convergence (dashed  
914 lines) for four different parameter initializations (dots) are superimposed. The true values (at the global optima) are  
915 depicted by the green crosshair and the dynamics of the parameters over the final iteration correspond to the thick  
916 red line. (C) The bottom row shows the estimates of hidden states and input for the corresponding pairs of  
917 parameters obtained during the last iteration, where we also show the trajectory of the parameters estimates over  
918 time .

919 • *Beyond the limits of fMRI signal*

920 One of the challenges in fMRI research is to increase a speed of brain volume sampling; i.e.  
921 to obtain data with a higher temporal resolution. Higher temporal resolution allows one to  
922 characterize changes in the brain more accurately, which is important in many aspects of fMRI.  
923 In this section, we will show that estimating unobserved (hidden) hemodynamic states and, more  
924 importantly, the underlying neuronal drives solely from observed data by blind deconvolution  
925 can significantly improve the temporal resolution and provide estimates of the underlying  
926 neuronal dynamics at a finer temporal scale. This may have useful applications in the formation  
927 of things like psychophysiological interactions (Gitelman et al., 2003).

928 In the hemodynamic model inversions above we did not use very realistic neuronal input,  
929 which was a Gaussian bump function and the data were generated with a temporal resolution of 1  
930 s. This was sufficient for our comparative evaluations; however in real data, the changes in  
931 underlying neuronal activation are much faster (possibly in the order of milliseconds) and may  
932 comprise a rapid succession of events. The hemodynamic changes induced by this neuronal  
933 activation manifest as a rather slow response, which peaks at about 4-6 s.

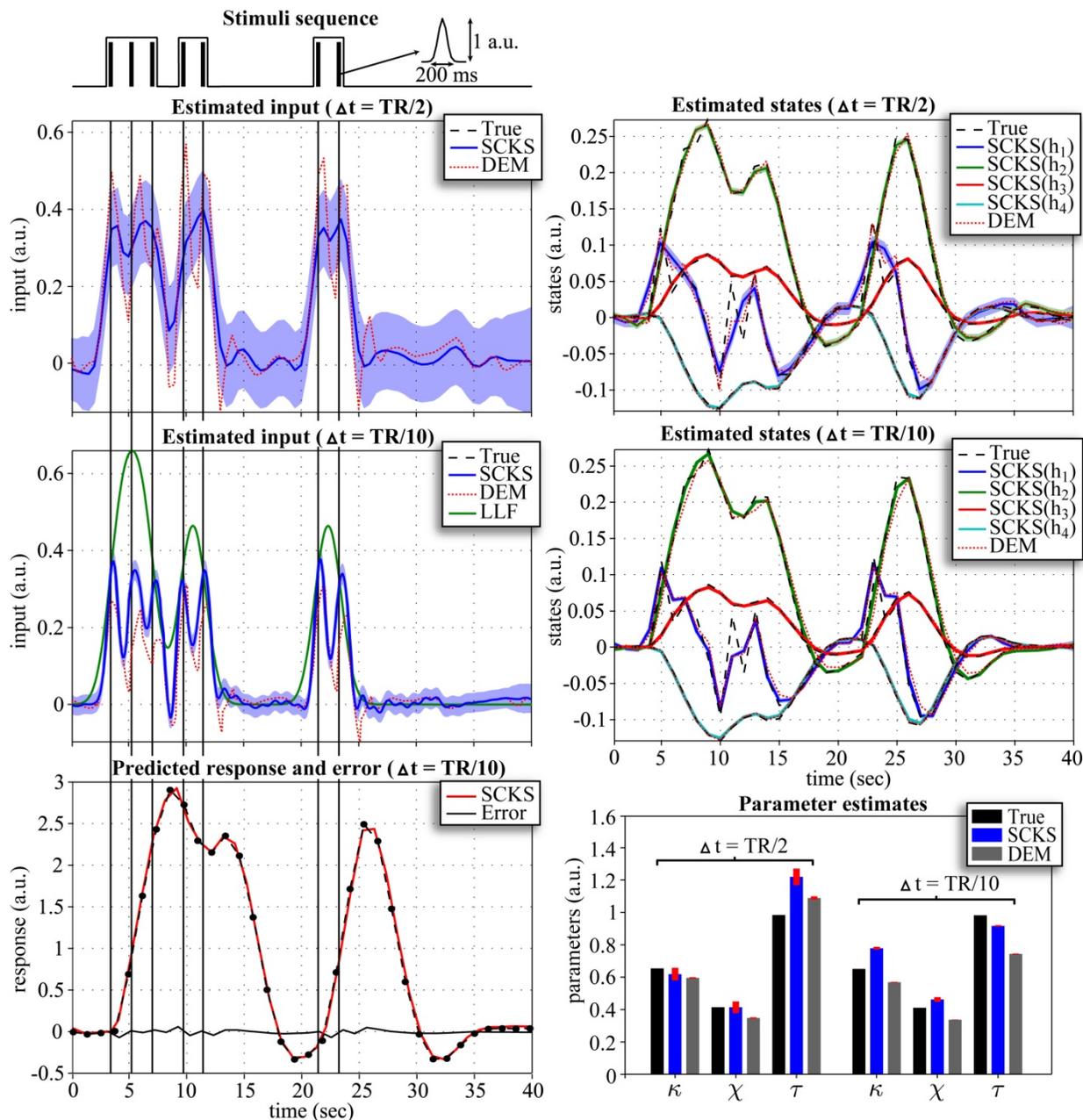
934 To make our simulations more realistic, we considered the following generation process,  
935 which is very similar to the simulation and real data used previously in Riera et al. (2004). First,  
936 we generated our data with a time step of 50 ms using the sequence of neuronal events depicted  
937 at the top of Fig. 8. These Gaussian-shaped neuronal events (inputs) had a FWTM (full-width at  
938 tenth of maximum) of less than 200 ms. Otherwise, the precisions on innovations, states noise,  
939 and input noise were identical to the hemodynamic simulations above. Next we down-sampled  
940 the synthetic response with a realistic TR = 1.2 s, obtaining data of 34 time points from the  
941 original 800. For estimation, we used the same priors on the input  $p(u) = \mathcal{N}(0,0.1)$  and  
942 parameters as summarized in Tab. 3.

943 Our main motivation was the question: How much of the true underlying neuronal signal  
944 can we recover from this simulated sparse observation, when applying either SCKS or DEM? To  
945 answer this, two different scenarios were considered. The first used an integration step  $\Delta t =$   
946  $TR/2 = 0.6$  s which had provided quite favorable results above. The top row of Fig. 8 shows the  
947 estimated input and states provided by SCKS and DEM. It can be seen that the states are traced  
948 very nicely by both approaches. For the input estimates, SCKS captures the true detailed  
949 neuronal structure deficiently, although the main envelope is correct. For DEM, the input  
950 estimate is much closer to the true structure of the neuronal signal, distinguishing all seven  
951 events. However, one can not overlook sharp undershoots that appear after the inputs. The reason  
952 for these artifacts rests on the use of generalized coordinates of motion, where the optimization  
953 of high order temporal derivatives does not always produce the optimal low order derivatives (as  
954 shown in the Fig. 8).

955 In the second scenario, where we decreased the integration step to  $\Delta t = TR/10 = 0.12$  s,  
956 we see that the SCKS estimate of the input has improved markedly. For DEM the input estimate  
957 is actually slightly worse than in the previous case. Recalling the results from previous  
958 simulations (Fig. 5 and 6) it appears that the optimal integration step for DEM is  $\Delta t = TR/2$ , and  
959 decreasing this parameter does not improve estimation (as it does for SCKS). Conversely, an  
960 excessive decrease of  $\Delta t$  can downgrade accuracy (without an appropriate adjustment of the  
961 temporal precision).

962 Here we can also compare our results with the results obtained in (Riera et al., 2004), where  
963 the LL-innovation technique was used with a constrained nonlinear optimization algorithm  
964 (Matlab's *fmincon.m* function) to estimate the neuronal activation. In our simulations the  
965 neuronal input was parameterized by a set of RBFs, regularly spaced with an inter-distance  
966 interval equal to  $TR$ , where the amplitudes of RBFs together with the first three hemodynamic  
967 model parameters, including noise variances, were subject to estimation. The resulting estimate  
968 is depicted by the solid green line at the bottom of Fig. 8. It is obvious that this only captures the  
969 outer envelope of the neuronal activation. Although this approach represented the most advanced  
970 technique at the time of its introduction (2004), its use is limited to relatively short time-series  
971 that ensures the number of parameters to be estimated is tractable.

972 We conclude that inversion schemes like DEM and especially SCKS can efficiently  
 973 reconstruct the dynamics of neuronal signals from fMRI signal, affording a considerable  
 974 improvement in effective temporal resolution.



975  
 976 **Figure 8.** Inversion of the hemodynamic model for more realistic neuronal inputs (top left) and fMRI observations  
 977 sampled with a TR = 1.2 s (bottom left – dotted line). The input and hidden states estimates obtained by SCKS and  
 978 DEM are shown for an integration step  $\Delta t = TR/2$  (top row) and  $\Delta t = TR/10$  (middle row). The parameter  
 979 estimates are shown on the bottom right. The best estimate of the input that could be provided by the local  
 980 linearization filter is depicted on the middle left panel by the solid green line.

## 981 **Discussion**

982 We have proposed a nonlinear Kalman filtering based on an efficient square-root cubature  
983 Kalman filter (SCKF) and RTS smoother (SCKS) for the inversion of nonlinear stochastic  
984 dynamic causal models. We have illustrated its application by estimating neuronal activity by (so  
985 called) blind deconvolution from fMRI data. Using simulations of different stochastic dynamic  
986 systems, including validation via Monte Carlo simulations, we have demonstrated its estimation  
987 and identification capabilities. Additionally, we have compared its performance with an  
988 established (DEM) scheme, previously validated in relation to EKF and particle filtering (Friston  
989 et al., 2008).

990 In particular, using a nonlinear model based on the Lorenz attractor, we have shown that  
991 SCKF and SCKS outperform DEM when the initial conditions and model parameters are  
992 unknown. The double-well model turned out (as anticipated) to be difficult to invert. In this case,  
993 both SCKF and SCKS could invert both states and input correctly, i.e. to track their true  
994 trajectories in about 70% of the simulations (unlike DEM). Both the Lorenz attractor and double-  
995 well system are frequently used for testing the robustness of new nonlinear filtering methods and  
996 provide a suitable forum to conclude that SCKF and SCKS show a higher performance in  
997 nonlinear and non-Gaussian setting than DEM. The third system we considered was a linear  
998 convolution model, where the performance of both inversion schemes was comparable. In contrast  
999 to the previous models, the SCKF alone was not sufficient for successful estimation of the states  
1000 and input. Although DEM provided a better estimate of the input, the SCKS was more precise in  
1001 tracking hidden states and inferring unknown model parameters.

1002 We then turned to the hemodynamic model proposed by Buxton et al. (1998) and completed  
1003 by Friston et al. (2000), which comprises nonlinear state and observation equations. The  
1004 complexity of this model, inherent in a series of nonlinear differential equations (i.e. higher order  
1005 ODEs) makes the inversion problem fairly difficult. If the input is unknown, it cannot be easily  
1006 solved by a forward pass of the SCKF or any other standard nonlinear recursive filter. It was  
1007 precisely this difficulty that motivated Friston et al. (2008) to develop DEM by formulating the  
1008 deconvolution problem in generalized coordinates of motion. The same problem motivated us to  
1009 derive a square-root formulation of the Rauch-Tung-Striebel smoother and solve the same  
1010 problem with a recursive scheme.

1011 Both DEM and SCKS (SCKF) use an efficient LL-scheme for the numerical integration of  
1012 non-autonomous multidimensional stochastic differential equations (Jimenez et al., 1999). Using  
1013 simulations, we have demonstrated that for a successful inversion of the hemodynamic model,  
1014 SCKS requires an integration step of at least  $\Delta t = TR/2$  for the accurate estimation of hidden  
1015 states, and preferably a smaller integration step for an accurate inference on the neuronal input.  
1016 Unlike SCKS, DEM provides the best estimates of the input when the integration step is  
1017  $\Delta t = TR/2$ . This is because it uses future and past observations to optimize a path or trajectory  
1018 of hidden states, in contrast to recursive schemes that update in a discrete fashion. Nevertheless,  
1019 with smaller integration steps, SCKS affords more precise estimates of the underlying neuronal  
1020 signal than DEM under any integration step. Additionally, in the case of more realistic  
1021 hemodynamic simulations we have shown that with the smaller integration step of about  
1022  $\Delta t = TR/10$  we were able to recover the true dynamics of neuronal activity that cannot be  
1023 observed (or estimated) at the temporal resolution of the measured signal. This takes us beyond  
1024 the limits of the temporal resolution of hemodynamics underlying the fMRI signal.

1025 An interesting aspect of inversion schemes is their computational cost. Efficient  
1026 implementations of SCKS with the integration step of  $\Delta t = TR/10$  (including parameter  
1027 estimation) are about 1.3 times faster than DEM (with an integration step of  $\Delta t = TR/2$  and a  
1028 temporal embedding  $n = 6$  and  $d = 3$ ). If the integration step is the same, then SCKS is about 5  
1029 times faster, which might have been anticipated, given that DEM is effectively dealing with six  
1030 times the number of (generalized) hidden states.

1031 We have also examined the properties of parameter identification of hemodynamic model  
1032 under the SCKS framework. Based on the previous experience (Deneux and Faugeras, 2006;  
1033 Riera et al., 2004), we constrained the hemodynamic model by allowing three parameters to  
1034 vary; i.e. rate of signal decay, rate of flow-dependent elimination, and mean transit time. The  
1035 remaining parameters were kept (nearly) constant, because they had only minor effects on the  
1036 hemodynamic response function.

1037 Our procedure for parameter identification uses a joint estimation scheme, where both  
1038 hidden states and parameters are concatenated into a single state vector and inferred  
1039 simultaneously in dynamic fashion. The SCKS is iterated until the parameters converge.  
1040 Moreover, the convergence is enhanced by a stochastic Robbins-Monro approximation of the

1041 parameter noise covariance matrix. This enabled very efficient parameter identification in all of  
1042 the stochastic models we considered, including the hemodynamic model. However, specifically  
1043 in the case of the hemodynamic model, we witnessed a particular phenomenon, which was also  
1044 reported by Deneux et al. (2006). Put simply, the effects of some parameters on the  
1045 hemodynamic response are degenerate, in that different combinations can still provide accurate  
1046 predictions of observed responses. In this context, we have shown in Fig. 7A that different sets  
1047 of parameters can produce a very similar hemodynamic response function. This degeneracy or  
1048 redundancy is a ubiquitous aspect of model inversion and is usually manifest as conditional  
1049 dependency among the parameter estimates. The problem of conditional dependencies is usually  
1050 finessed by optimizing the model in terms of its evidence. Model evidence ensures that the  
1051 conditional dependences are suppressed by minimizing complexity (which removes redundant  
1052 parameters). In our setting, we are estimating both states and parameters and have to contend  
1053 with possible conditional dependences between the states and parameters. In principle, this can  
1054 be resolved by comparing the evidence for different models and optimizing the parameterization  
1055 to provide the most parsimonious model. We will pursue this in a subsequent paper, in which we  
1056 examine the behavior of model evidence, as estimated under cubature smoothing. It should be  
1057 noted, that this work uses models that have already been optimized over the past few years, so  
1058 that they provide the right balance of accuracy and complexity, when trying to explain typical  
1059 fMRI data. However, we may have to revisit this issue when trying to estimate the hidden  
1060 neuronal states as well as parameters.

1061 There are further advantages of SCKS compared to DEM. Since DEM performs inference  
1062 on states and input in a forward manner only, it is sensitive to misspecification of initial  
1063 conditions. Critically, recent implementations of DEM (Friston et al., 2008) start each iteration  
1064 with the same initial values of the states and the input, resulting in significant error at the initial  
1065 phase of deconvolution. This is not the case for SCKS, which, by applying smoothing backward  
1066 step, minimizes the initial error and converges to the true initial value over iterations. Next, DEM  
1067 can produce sharp undershoots in the input estimate when the hidden states or their causes  
1068 change too quickly. The SCKS does not have this problem. However, the use of generalized  
1069 motion enables DEM to be applied online. Additionally, this framework also allows DEM to  
1070 model temporal dependencies in the innovations or fluctuations of hidden states, which might be  
1071 more plausible for biological systems. In Kalman filtering, these fluctuations are generally

1072 assumed to be Markovian. Having said this, it is possible to cast dynamical models in  
1073 generalized coordinates of motion as classical Markovian models, where the innovations are  
1074 successively colored before entering the state equation (see Eq. 3 in (Friston, 2008b)).

1075 Based on our MC simulations, we conclude that in general SCKS provided a more accurate  
1076 inversion of nonlinear dynamic models, including estimation of the states, input and parameters,  
1077 than DEM. Since DEM has been shown to outperform EKF and particle filtering, it makes the  
1078 SCKS the most efficient blind nonlinear deconvolution schemes for dynamic state-space models.

1079 Finally, all evaluations of the proposed approach, including the comparison with DEM, were  
1080 performed under the assumption that SCKS algorithm had access to the true precision parameter  
1081 on the measurement noise and DEM had access to precisions on all noise components. However,  
1082 for application to the real data we have to be able to estimate these precision parameters as well.  
1083 DEM is formulated as a hierarchical dynamic model, which allows for an elegant triple inference  
1084 on hidden states, input, parameters and hyperparameters. In the case of SCKS we have  
1085 introduced dynamic approximation techniques for the efficient estimation of the parameter state  
1086 noise covariance matrices. We also observed that the input noise variance can be considered  
1087 time-invariant, with a reasonable value (for the hemodynamic model) of about  $V = 0.1$ . This  
1088 value seemed to be consistent over different levels of noise and different input. The last  
1089 outstanding unknown quantity is the measurement noise covariance. We have found a robust  
1090 solution (Särkkä and Hartikainen, Under revision; Särkkä and Nummenmaa, 2009) that  
1091 combines the variational Bayesian method with the nonlinear Kalman filtering algorithm for the  
1092 joint estimation of states and time-varying measurement noise covariance in a nonlinear state-  
1093 space model. We have implemented this approach for our SCKS scheme with a minimal increase  
1094 in computational cost. Although this variational Bayesian extension was not utilized in our  
1095 proposal (for simplicity), it is now part of SCKS algorithm for future application to the real data.

1096 There are several application domains we hope to explore within our framework: Since  
1097 SCKS can recover the underlying time course of synaptic activation, we can model effective  
1098 connectivity at synaptic (neuronal) level. Because no knowledge about the input is necessary,  
1099 one can use this scheme to invert the dynamic causal models on the resting state data, or pursue  
1100 connectivity analyses in the brain regions that are dominated by endogenous activity  
1101 fluctuations, irrespective of task-related responses. We will also consider conventional

1102 approaches to causal inference that try to identify the direction of the information flow between  
1103 different brain regions (e.g. Granger causality, dynamic Bayesian networks, etc.). In this context,  
1104 one can compare the analysis of deconvolved hidden (neuronal) states with explicit model  
1105 comparison within the DCM framework. Another challenge would be to exploit the similarity  
1106 among neighboring voxels in relation to their time courses. There are thousands of voxels in any  
1107 volume of the human brain, and the judicious pooling of information from multiple voxels may  
1108 help to improve accuracy of our deconvolution schemes. Last but not least, we hope to test  
1109 variants of the hemodynamic model, starting with extension proposed by Buxton et al. (2004),  
1110 which accounts for non-steady-state relationships between CBF and CBV arising due to  
1111 viscoelastic effects. This is particularly interesting here, because we can, in principle,  
1112 characterize these inconstant relationships in terms of time-varying parameter estimates afforded  
1113 by our recursive schemes.

1114         The Matlab code for our methods (including estimation of measurement noise covariance),  
1115 which is compatible with the subroutines and variable structures used by the DEM in SPM8, is  
1116 available from the authors upon request.

## 1117 **Conclusion**

1118         In this paper, we have introduced a robust blind deconvolution technique based on the  
1119 nonlinear square-root cubature Kalman filter and Rauch-Tung-Striebel smoother, which allows  
1120 an inference on hidden states, input, and model parameters. This approach is very general and  
1121 can be applied to the inversion of any nonlinear continuous dynamic model that is formulated  
1122 with stochastic differential equations. This first description of the technique focused on the  
1123 estimation of neuronal synaptic activation by generalized deconvolution from observed fMRI  
1124 data. We were able to estimate the true underlying neuronal activity with a significantly  
1125 improved temporal resolution, compared to the observed fMRI signal. This speaks to new  
1126 possibilities for fMRI signal analysis; especially in effective connectivity and dynamic causal  
1127 modeling of unknown neuronal fluctuations (e.g. resting state data).

1128         We validated the inversion scheme using difficult nonlinear and linear stochastic dynamic  
1129 models and compared its performance with dynamic expectation maximization; one of the few

1130 methods that is capable of this sort of model inversion. Our approach afforded the same or better  
1131 estimates of states, input, and model parameters, with reduced computational cost.

## 1132 **Acknowledgments**

1133 This work was supported by the research frame no. MSM0021630513 and no.  
1134 MSM0021622404 and also sponsored by the research center DAR no. 1M0572, all funded by the  
1135 Ministry of Education of the Czech Republic. Additional funding was provided by NIH grant no.  
1136 R01EB000840 from the USA. KJF was funded by the Wellcome Trust. We would like to thank  
1137 to Jorge Riera for providing his implementation of LL-innovation algorithm.

## 1138 **References**

- 1139 Aguirre, G.K., Zarahn, E., D'esposito, M., 1998. The variability of human, BOLD hemodynamic  
1140 responses. *Neuroimage* 8, 360-369.
- 1141 Arasaratnam, I., Haykin, S., 2008. Nonlinear Bayesian Filters for Training Recurrent Neural Networks.  
1142 *MICAI 2008: Advances in Artificial Intelligence*, 12-33.
- 1143 Arasaratnam, I., Haykin, S., 2009. Cubature Kalman Filters. *IEEE Transactions on Automatic Control* 54,  
1144 1254-1269.
- 1145 Attwell, D., Buchan, A.M., Charkpak, S., Lauritzen, M., MacVicar, B.A., Newman, E.A., 2010. Glial and  
1146 neuronal control of brain blood flow. *Nature* 468, 232-243.
- 1147 Berns, G.S., Song, A.W., Mao, H., 1999. Continuous functional magnetic resonance imaging reveals  
1148 dynamic nonlinearities of "dose-response" curves for finger opposition. *Journal of Neuroscience* 19, 1-6.
- 1149 Birn, R.M., Saad, Z.S., Bandettini, P.A., 2001. Spatial heterogeneity of the nonlinear dynamics in the  
1150 fMRI BOLD response. *Neuroimage* 14, 817-826.
- 1151 Biscay, R., Jimenez, J.C., Riera, J.J., Valdes, P.A., 1996. Local linearization method for the numerical  
1152 solution of stochastic differential equations. *Annals of the Institute of Statistical Mathematics* 48, 631-  
1153 644.
- 1154 Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex  
1155 of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine* 34, 537-541.
- 1156 Bucy, R.S., Senne, K.D., 1971. Digital synthesis of non-linear filters. *Automatica* 7, 287-298.
- 1157 Buxton, R.B., Uludag, K., Dubowitz, D.J., Liu, T.T., 2004. Modeling the hemodynamic response to brain  
1158 activation. *Neuroimage* 23, S220-S233.
- 1159 Buxton, R.B., Wong, E.C., Frank, L.R., 1998. Dynamics of blood flow and oxygenation changes during  
1160 brain activation: the balloon model. *Magnetic Resonance in Medicine* 39, 855-864.
- 1161 David, O., 2009. fMRI connectivity, meaning and empiricism Comments on: Roebroeck et al. The  
1162 identification of interacting networks in the brain using fMRI: Model selection, causality and  
1163 deconvolution. *NeuroImage*, In Press.
- 1164 David, O., Guillemain, I., Sallet, S., Reyt, S., Deransart, C., Segebarth, C., Depaulis, A., 2008.  
1165 Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS Biology* 6,  
1166 2683-2697.
- 1167 Deneux, T., Faugeras, O., 2006. Using nonlinear models in fMRI data analysis: model selection and  
1168 activation detection. *NeuroImage* 32, 1669-1689.

1169 Doucet, A., De Freitas, N., Gordon, N., 2001. Sequential Monte Carlo methods in practice. Springer  
1170 Verlag.

1171 Fernandez-Prades, C., Vila-Valls, J., 2010. Bayesian Nonlinear Filtering Using Quadrature and Cubature  
1172 Rules Applied to Sensor Data Fusion for Positioning. In proceeding of IEEE International Conference on  
1173 Communications, pp. 1-5.

1174 Friston, K., 2008a. Variational filtering. *Neuroimage* 41, 747-766.

1175 Friston, K., 2009. Dynamic casual modeling and Granger causality Comments on: The identification of  
1176 interacting networks in the brain using fMRI: Model selection, causality and deconvolution. *Neuroimage*,  
1177 In Press.

1178 Friston, K.J., 2002. Bayesian estimation of dynamical systems: an application to fMRI. *NeuroImage* 16,  
1179 513-530.

1180 Friston, K.J., 2008b. Hierarchical models in the brain. *PLoS Computational Biology* 4, e1000211.

1181 Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. *Neuroimage* 19, 1273-1302.

1182 Friston, K.J., Mechelli, A., Turner, R., Price, C.J., 2000. Nonlinear responses in fMRI: the Balloon model,  
1183 Volterra kernels, and other hemodynamics. *NeuroImage* 12, 466-477.

1184 Friston, K.J., Stephan, K.E., Daunizeau, J., 2010. Generalised Filtering. *Mathematical Problems in*  
1185 *Engineering* 2010.

1186 Friston, K.J., Trujillo-Barreto, N., Daunizeau, J., 2008. DEM: a variational treatment of dynamic systems.  
1187 *Neuroimage* 41, 849-885.

1188 Gitelman, D.R., Penny, W.D., Ashburner, J., Friston, K.J., 2003. Modeling regional and  
1189 psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage* 19,  
1190 200-207.

1191 Handwerker, D., Ollinger, J., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across  
1192 subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21, 1639-1651.

1193 Haykin, S.S., 2001. Kalman filtering and neural networks. Wiley Online Library.

1194 Hinton, G.E., van Camp, D., 1993. Keeping the neural networks simple by minimizing the description  
1195 length of the weights. In proceedings of COLT-93. ACM, pp. 5-13.

1196 Hu, Z., Zhao, X., Liu, H., Shi, P., 2009. Nonlinear analysis of the BOLD signal. *EURASIP Journal on*  
1197 *Advances in Signal Processing* 2009, 1-13.

1198 Iadecola, C., 2002. CC commentary: intrinsic signals and functional brain mapping: caution, blood  
1199 vessels at work. *Cerebral Cortex* 12, 223-224.

1200 Ito, K., Xiong, K., 2002. Gaussian filters for nonlinear filtering problems. *Automatic Control, IEEE*  
1201 *Transactions on* 45, 910-927.

1202 Jaakkola, T.S., 2000. Tutorial on variational approximation methods. *Advanced mean field methods:*  
1203 *theory and practice*, 129-159.

1204 Jacobsen, D., Hansen, L., Madsen, K., 2008. Bayesian model comparison in nonlinear BOLD fMRI  
1205 hemodynamics. *Neural computation* 20, 738-755.

1206 Ji, Z., Brown, M., 2009. Joint State and Parameter Estimation For Biochemical Dynamic Pathways With  
1207 Iterative Extended Kalman Filter: Comparison With Dual State and Parameter Estimation. *Open*  
1208 *Automation and Control Systems Journal* 2, 69-77.

1209 Jimenez, J.C., 2002. A simple algebraic expression to evaluate the local linearization schemes for  
1210 stochastic differential equations\* 1. *Applied Mathematics Letters* 15, 775-780.

1211 Jimenez, J.C., Ozaki, T., 2003. Local linearization filters for non-linear continuous-discrete state space  
1212 models with multiplicative noise. *International Journal of Control* 76, 1159-1170.

1213 Jimenez, J.C., Shoji, I., Ozaki, T., 1999. Simulation of stochastic differential equations through the local  
1214 linearization method. A comparative study. *Journal of Statistical Physics* 94, 587-602.

1215 Johnston, L.A., Duff, E., Mareels, I., Egan, G.F., 2008. Nonlinear estimation of the BOLD signal.  
1216 *NeuroImage* 40, 504-514.

1217 Julier, S., Uhlmann, J., Durrant-Whyte, H.F., 2002. A new method for the nonlinear transformation of  
1218 means and covariances in filters and estimators. *Automatic Control, IEEE Transactions on* 45, 477-482.

1219 Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *Journal of basic*  
1220 *Engineering* 82, 35-45.

1221 Kaminski, P., Bryson Jr, A., Schmidt, S., 1971. Discrete square root filtering: A survey of current  
1222 techniques. *Automatic Control, IEEE Transactions on* 16, 727-736.

1223 Kloeden, P.E., Platen, E., 1999. *Numerical Solution of Stochastic Differential Equations, Stochastic*  
1224 *Modeling and Applied Probability*, 3rd editon. Springer.

1225 Krüger, G., Glover, G.H., 2001. Physiological noise in oxygenation sensitive magnetic resonance  
1226 imaging. *Magnetic Resonance in Medicine* 46, 631-637.

1227 Lauritzen, M., 2001. Relationship of spikes, synaptic activity, and local changes of cerebral blood flow.  
1228 *Journal of Cerebral Blood Flow & Metabolism* 21, 1367-1383.

1229 Li, P., Yu, J., Wan, M., Huang, J., 2009. The augmented form of cubature Kalman filter and quadrature  
1230 Kalman filter for additive noise. *IEEE Youth Conference on Information, Computing and*  
1231 *Telecommunication, YC-ICT '09. IEEE*, pp. 295-298.

1232 Ljung, L., Söderström, T., 1983. *Theory and practice of recursive identification*. MIT Press, Cambridge,  
1233 MA.

1234 Logothetis, N.K., 2002. The neural basis of the blood–oxygen–level–dependent functional magnetic  
1235 resonance imaging signal. *Philosophical Transactions of the Royal Society of London. Series B:*  
1236 *Biological Sciences* 357, 1003.

1237 MacKay, D.J.C., 1995. Developments in probabilistic modelling with neural networks-ensemble learning.  
1238 In *proceedings of 3rd Annual Symposium on Neural Networks, Nijmegen, Netherlands*, pp. 191-198.

1239 Magistretti, P., Pellerin, L., 1999. Cellular Mechanisms of Brain Energy Metabolism and their Relevance  
1240 to Functional Brain Imaging. *Philos Trans R Soc Lond B Biol Sci* 354, 1155-1163.

1241 Mandeville, J.B., Marota, J.J.A., Ayata, C., Zaharchuk, G., Moskowitz, M.A., Rosen, B.R., Weisskoff,  
1242 R.M., 1999. Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. *Journal of*  
1243 *Cerebral Blood Flow & Metabolism* 19, 679-689.

1244 Mechelli, A., Price, C., Friston, K., 2001. Nonlinear coupling between evoked rCBF and BOLD signals: a  
1245 simulation study of hemodynamic responses. *NeuroImage* 14, 862-872.

1246 Miller, K.L., Luh, W.M., Liu, T.T., Martinez, A., Obata, T., Wong, E.C., Frank, L.R., Buxton, R.B.,  
1247 2001. Nonlinear temporal dynamics of the cerebral blood flow response. *Human Brain Mapping* 13, 1-12.

1248 Murray, L., Storkey, A., 2008. Continuous time particle filtering for fMRI. *Advances in Neural*  
1249 *Information Processing Systems* 20.

1250 Nelson, A.T., 2000. Nonlinear estimation and modeling of noisy time-series by dual Kalman filtering  
1251 methods. Ph.D thesis, Oregon Graduate Institute of Science and Technology.

1252 Norgaard, M., Poulsen, N.K., Ravn, O., 2000. New developments in state estimation for nonlinear  
1253 systems. *Automatica* 36, 1627-1638.

1254 Ozaki, T., 1992. A bridge between nonlinear time series models and nonlinear stochastic dynamical  
1255 systems: a local linearization approach. *Statistica Sinica* 2, 113–135.

1256 Riera, J., Watanabe, J., Kazuki, I., Naoki, M., Aubert, E., Ozaki, T., Kawashima, R., 2004. A state-space  
1257 model of the hemodynamic approach: nonlinear filtering of BOLD signals. *NeuroImage* 21, 547-567.

1258 Robbins, H., Monro, S., 1951. A stochastic approximation method. *The Annals of Mathematical Statistics*  
1259 22, 400-407.

1260 Roebroeck, A., Formisano, E., Goebel, R., 2009a. Reply to Friston and David:: After comments on: The  
1261 identification of interacting networks in the brain using fMRI: Model selection, causality and  
1262 deconvolution. *NeuroImage*.

1263 Roebroeck, A., Formisano, E., Goebel, R., 2009b. The identification of interacting networks in the brain  
1264 using fMRI: Model selection, causality and deconvolution. *Neuroimage*, In Press.

1265 Simandl, M., Dunik, J., 2006. Design of derivative-free smoothers and predictors. In *proceeding of the*  
1266 *14th IFAC Symposium on System Identification, SYSID06*, pp. 1240-1245.

1267 Sitz, A., Schwarz, U., Kurths, J., Voss, H.U., 2002. Estimation of parameters and unobserved components  
1268 for nonlinear systems from noisy time series. *Physical Review E* 66, 1-9.

1269 Sotero, R.C., Trujillo-Barreto, N.J., Jiménez, J.C., Carbonell, F., Rodríguez-Rojas, R., 2009.  
1270 Identification and comparison of stochastic metabolic/hemodynamic models (sMHM) for the generation  
1271 of the BOLD signal. *Journal of computational neuroscience* 26, 251-269.  
1272 Stephan, K.E., Kasper, L., Harrison, L.M., Daunizeau, J., den Ouden, H.E.M., Breakspear, M., Friston,  
1273 K.J., 2008. Nonlinear dynamic causal models for fMRI. *NeuroImage* 42, 649-662.  
1274 Särkkä, S., Hartikainen, J., Under revision. Extension of VB-AKF to Estimation of Full Covariance and  
1275 Non-Linear Systems. *IEEE Transactions on Automatic Control*.  
1276 Särkkä, S., Nummenmaa, A., 2009. Recursive noise adaptive Kalman filtering by variational Bayesian  
1277 approximations. *Automatic Control, IEEE Transactions on* 54, 596-600.  
1278 Valdes Sosa, P.A., Sanchez Bornot, J.M., Sotero, R.C., Iturria Medina, Y., Aleman Gomez, Y., Bosch  
1279 Bayard, J., Carbonell, F., Ozaki, T., 2009. Model driven EEG/fMRI fusion of brain oscillations. *Human*  
1280 *brain mapping* 30, 2701-2721.  
1281 Van der Merwe, R., 2004. Sigma-point Kalman filters for probabilistic inference in dynamic state-space  
1282 models. Ph.D.thesis, Oregon Graduate Institute of Science and Technology.  
1283 Wu, Y., Hu, D., Wu, M., Hu, X., 2005. Unscented Kalman filtering for additive noise case: augmented vs.  
1284 non-augmented. *IEEE Signal Processing Letters* 12, 357-359.

1285

1286

1287

1288