



HAL
open science

Hemodynamic Models: Investigation and Application to Brain Imaging Analysis

Thomas Deneux

► **To cite this version:**

Thomas Deneux. Hemodynamic Models: Investigation and Application to Brain Imaging Analysis. Human-Computer Interaction [cs.HC]. Ecole Polytechnique X, 2006. English. NNT: . tel-00457464

HAL Id: tel-00457464

<https://pastel.hal.science/tel-00457464>

Submitted on 17 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École Doctorale Polytechnique

Ph.D in Computer Science

Thomas Deneux

**Hemodynamic Models:
Investigation and Application to Brain
Imaging Analysis**

Advisor: Olivier FAUGERAS

Jury:

Mme. Line Garnero - reviewer
M. John Mayhew - reviewer
M. Habib Benali - reviewer
M. Nikos Paragios - reviewer
M. Gilles Dowek
M. Olivier Faugeras
M. Jean-Baptiste Poline - invited member
M. Patrick Chauvel - invited member

Acknowledgments

I would like to thank the many people who helped be during my PhD work, through discussions, collaborations, encouragements... Approximately in the order I met them:

Olivier Faugeras, who accompanied me throughout this thesis - I am particularly grateful to his energy and enthusiasm that were very helpful to lead all these projects and collaborations, and encourage me at difficult times,

Renaud Keriven, Miklos Santha and Jacques Stern for their help in the first choices at the time of my Master,

Jean-Baptiste Poline, Line Garnero, Sylvain Baillet, Anne-Lise Paradis who helped me to learn something about brain imaging,

Lars-Kai Hansen gave me a great confidence by inviting me in Copenhagen, and Daniel Jacobsen was a nice companion during this first collaboration,

Lars Hanson, Torben Lund, and everybody at the Hvidovre Hospital in Copenhagen, in particular for offering me an fMRI acquisition,

Many thanks to Ivo Vanzetta for sharing his monkey with me, and for the nice and fruitful collaboration we had,

The whole team of Guillaume Masson in Marseilles with whom I had many good times,

Jean-Luc Anton, Muriel Roth and Bruno Nazarian at the MRI center, hôpital La Timone in Marseille, for their welcome and help in fMRI acquisitions,

Maxime Descoteaux, for not complaining too much on being abandoned inside the MRI scanner, and the other fellows, Cécile Galléa, François Grimbert, Adrien Wohrer, Sylvain Vallaghe, who did not have that chance,

Bertrand Thirion and Jean Daunizeau for useful discussions,

Christian Bénar and Jean Gotman, to whom I apologize for not having been able to pursue on our projects,

the whole Odyssée team, and in particular the other fellows in Paris, for their company, discussions, fun, and everything: Guillaume Charpiat, Geoffray Adde, Gheorghe Postelnicu, Pierre Maurel, Irène Fasiello, Romain Brette, Patrick Labatut, Jonathan Touboul; thanks for the helping hands during the last rushing hours,

Maxime Deurbergue, Bénédicte Pavot, Marion Fabre, and hundred times Allison Geier for their corrections of the English in the PhD manuscript,

All the friends and family who supported and encouraged me, or even gave me a

helping hand,

The reviewers who accepted the rude job of reading the thesis: Line Garnero, Habib Benali, with whom I had nice discussions, Nikos Paragios, who accepted even though asked at the very last minute, and a special thank to John Mayhew and Ying Zheng who helped me much at the end of this work,

And Gilles Dowek, Jean-Baptiste Poline and Patrick Chauvel who accepted to be part of the defence committee.

Synopsis (English)

Abstract

An accurate analysis of functional MRI measurements requires a precise understanding of the physiological processes involved in this measure. This PhD work shows both investigations in hemodynamic models and algorithms to use such models in the analysis of brain imaging measurements.

A particular concern with functional MRI is the temporal modelling of the responses to neural activity. Today, the most standard analysis methods use the General Linear Model framework, which supposes a linear relationship between brain activity and the BOLD response. We show how it is possible alternatively to use nonlinear models in data analysis. Our estimation of parameters by energy minimization is the equivalent to linear regression, and our adaptation of the Fisher statistical test enables activation detection, hypothesis testing, and eventually comparison between different models.

We then have extended our methods to the analysis of multiple modalities data, and in particular, proposed a method to estimate the cortical activity underlying simultaneous fMRI and EEG measurements. We were able to achieve accurate estimation on synthetic data.

Additionally to these methodological researches, we have investigated the model equations with an Optical Imaging experiment. We have focused on the dynamic of the blood flow, which is at a crossroad between electrical, metabolism and oxygenation processes. We have identified specific questioning facts about this dynamic, such as nonlinearity with respect to electrical synaptic activity, and delays with respect to the blood volume response. Furthermore, we have conceived a new method for estimating fast erythrocyte motions in the blood from intrinsic optical imaging signal, which might provide a new useful measurement of this blood flow.

In the following synopsis, we summarize the main features of this PhD work, through highlighting the purposes, methods, conclusions and implications of each chapter (please note that the first part, which entails three chapter, is an introduction to the PhD). Also, we try to present an objective criticism of this work, through mentioning both its original contributions and its weaknesses. We hope that this summary will help the reader to rapidly navigate through the thesis, while understanding the relations between its different components.

PART II: USING NONLINEAR MODELS IN fMRI

Chapter 4: Model Identification

Objectives

- To estimate the parameters of physiological models from fMRI signals.
- To evaluate the estimation precision.
- To evaluate the ability of specific physiological models to fit real datasets.

Methods

- Physiological models are expressed as dynamical equation systems.
- An intuitive and a more mathematical approach to study the stability around equilibrium of a dynamical system are proposed.
- Our estimation algorithm relies on a Maximum Likelihood (least square) estimation.
- A classical sensitivity analysis quantifies the system identifiability through investigating its sensitivity to variations of the parameters.
- Synthetic data is used to validate the methods, and compare the Maximum Likelihood estimation to a Kalman Likelihood algorithm.
- An fMRI experiment is done to compare a linear model with eight variations of the models proposed by Buxton et al.[33, 32] and Friston et al.[59]: a damped oscillator modelling of the flow response [59] is compared to a convolution with a gamma-variate function [32], the effects of compliance and neural habituation models are investigated [32], and we tried to fix the values of some parameters. Parameter estimation is achieved both on “high-quality” average responses to the stimulation, and on individual-voxel raw time courses.

Results

- The Kalman Likelihood algorithm did not produce better estimations than the Maximum Likelihood one on synthetic data; we thus used the latter in the real data analysis since it is much faster.
- The physiological models can account for the measured BOLD responses: the Signal to Noise Ratios of their fit to the data are all better than the one of the linear model, and they vary according to which particular model is used.
- All these models are poorly identifiable, i.e. the parameter values can vary significantly while producing very small changes in the BOLD output. All the more since we estimated the output scaling factor, which interacts with the input scaling factor? This seems to be a particular concern, since it prevents us from being able to find the exact values of some parameters, because of their interactions with others.

Original contributions

- Our analytical computation of the derivative of the system output with respect to parameters allows a faster and more precise likelihood maximization algorithm.
- We underline in this chapter problems about the identifiability of parameters: the estimation cannot guarantee the values of the estimated parameters to be exact (anyway, who would use fMRI signal to estimate the oxygen extraction ratio at rest or other physiological parameter?) The next chapter shows how the most important thing we need is to be able to characterize the system dynamic rather than finding the exact parameter values.

Weaknesses

- There is currently no model for the BOLD measure at 3 Tesla. Since our experiment was done with a 3T scanner, we used the model proposed for a 1.5T scanner in the literature [32], $y = V_0(a_1(q - 1) - a_2(v - 1))$, but we considered its parameters unknowns (we actually supposed $a_2/(a_1 + a_2) = 1/10$).
- The algorithm did not use priors on the parameters, which is probably the only way to increase the identifiability of parameters.

Conclusion and implications

- Accurate fits can be obtained between BOLD measurements and physiological model predictions, even though the exact parameter values of these models cannot be extracted. It opens the way to using nonlinear models more widely in the analysis of fMRI datasets, which is investigated in the next chapter.
- Using priors is probably the only way to provide exact estimations of the parameter values from fMRI signals.
- The different models we used showed different fits to the measured data. The statistical comparison between models will be investigated in the next chapter.

Chapter 5: Nonlinear Hypothesis Testing and Model Selection

Objectives

- To develop statistical methods to analyze fMRI data in a physiological models framework instead of a linear convolution one.
- To evaluate them for detection activation purposes.
- To use them also for the selection between different physiological models.

Methods

- Because our models are nonlinear, and in order to use the same Fisher test as in the SPM software, we first estimate the model parameters and then linearize the equation with respect to parameters.
- We developed a non-hierarchical test to enable the comparison between two non-hierarchical models (i.e. where none is a particular case of the other), which was not possible with the original F-test. This non-hierarchical test relies on the building of a third more general model such that the two models to compare are particular cases of the former.
- These tests were applied to the same data as in the previous chapter.

Results

- There was no difference between the activation detection obtained with non-linear models or with linear models.
- All model enhancements were judged statistically significant by the tests when applied on the high-quality average responses or even on the time courses of the most activated voxels: the hypothesis of a volume compliance effect, and of neural habituation were validated.
- Friston's damped oscillator modelling of the flow response was preferred to Buxton's convolution with a gamma-variate function, though the difference was not great.
- The linear model and the simpler physiological models were selected when the tests were applied to noisier voxel time courses, because they entail less degrees of freedom.

Original contributions

- This work shows how nonlinear models can be used in fMRI studies.
- We recommend to use statistical tests based on Signal to Noise Ratios, such as the Fisher test we have developed, rather than on the a posteriori distribution of one specific parameter as was proposed by Friston et al.[58]. The non-identifiability of the parameters indeed implies that their a posteriori distribution are very flat.
- The hierarchical and non-hierarchical tests allow model selection.

Weaknesses

- Because of the linearization, the p-values computed using our statistical tests may not be exact, since the models are strongly nonlinear with respect to their physiological parameters. More accurate methods should scan the whole space of parameters (Metropolis-Hastings Markov-Chain Monte Carlo [127]).
- We supposed a white measure noise instead of estimating noise autocorrelations which could change the p-values also.

Conclusion and implications

- This chapter is a first step toward the use of nonlinear models in the analysis of fMRI data. We have shown the possibility to adapt statistical tests in the nonlinear framework, which is an encouragement to pursue efforts for the modelling of the BOLD response.
- We addressed the identifiability issue raised in the previous chapter: the most important thing with parameter estimation is that it allows an accurate characterization of the system dynamic, even though the parameter estimates are not exact.
- Several enhancements of the first model proposed by Buxton et al. were validated by the statistical tests. This raises some questions in terms of physiological models: Do in vivo measurements confirm the compliance effect? The neural habituation effect? Could the nonlinearities observed in fMRI in the short time range be caused by nonlinearities in the flow response rather than by neural habituation processes? Aren't both Friston's and Buxton's modelling of the flow response over-simplistic? The next chapter addresses some of these questions.
- The use of linear models for some applications with noisy data, such as localization, is justified.

PART III: INVESTIGATION ON THE CEREBRAL BLOOD FLOW IN OPTICAL IMAGING

Chapter 6: Dynamics and Nonlinearities of the Flow Response

Objectives

- To investigate whether the nonlinearities we observed in fMRI could not be attributed to nonlinear effects in the relation between neural activity and the flow response, rather than to neural habituation, using more informative Optical Imaging measurements.
- To investigate the temporal relation between the flow and volume responses.

Methods

- We use concurrent recordings of blood volume (with intrinsic optical signals measured at one isosbestic wavelength) and blood flow (with a laser-Doppler probe), and additional recordings of neural electrical activity using Voltage-Sensitive Dye (VSD).
- We copied the experimental design used in the fMRI experiment, excepted we explored responses to shorter stimulations.
- The measurements are described qualitatively first, and then used to fit new models including some nonlinear saturation effects.

Results

- We found a linear relation between the stimulus length and the amplitude of the electrical neural activity as measured with the VSD technique, a nonlinear relation between the amplitudes of electrical activity and flow activity, and a linear relation between flow and volume amplitudes.
- The flow response appears to be delayed with respect to the blood volume response, which is in contradiction with models supposing of flow-driven volume response. Moreover, we observed initial undershoots in the flow responses.
- The addition of some nonlinear saturation effect to Buxton's or Friston's linear models of the flow response enabled us to better predict the amplitudes of the flow, but not the delays.

Original contributions

- We have shown that existing linear models for the flow response cannot account for the responses we measured, which suggests it is highly important today to study the blood flow dynamic.

Weaknesses

- Only one animal was used for this experiment.
- The nonlinear model we propose remains too simple.

Conclusion and implications

- Two main reasons can explain the nonlinearities between VSD and flow measurements. Either the flow response depends on an additional neural activity not detected by the VSD (e.g. spiking), which is nonlinear with respect to the stimulation length. Or there are nonlinearities in the mechanisms involved in the flow response.
- The delay between volume and flow responses, and the initial undershoot observed in the flow responses, might have an explanation already proposed by Lauritzen et al.[115]: the increase in the blood volume in the arterioles is linked to a decrease of the outflow rather than an increase of the inflow.
- These questions highlight the necessity for modelling each compartment separately. It would be very profitable then to have a spatial measure of the flow, which would allow distinguishing the different compartments, rather than a measure at a unique point. The next chapter will propose a method that intends to produce such a spatial measure.

Chapter 7: Bi-dimensional Flowmetry with Intrinsic Recordings

Objectives

- To estimate the velocity of the blood inside the vessels and the variations in time of this velocity, through tracking the motion of Red Blood Cells (RBCs).

Methods

- The method uses image processing techniques to estimate the motion produced by Red Blood Cells (RBCs) along the blood vessels in volume images (intrinsic signals at 570nm).
- We first co-register the frames of the image sequences to correct for brain movements, and apply temporal and spatial filterings to enhance them.
- The data is extracted from manually-segmented vessels as 2D time-space arrays
- We use the *structure tensor* technique to estimate the directions of trajectory lines inside these arrays. The structure tensor is a common tool in image processing usually used for detecting directions.
- We tested the ability of the method to detect the blood pulsation in a first experiment with no particular stimulation, and where laser-Doppler flow and intrinsic signal volume were acquired simultaneously.
- We tested the ability of the method to detect sensory-induced responses in a second experiment with visual stimulations and volume recordings only.

Results

- The technique detected accurately the velocity changes induced by the heart pulsation.
- We also obtained responses evoked by a visual stimulation. The responses among a same category of vessels had similarities: arteries showed an early blood velocity response; arterioles and venules showed a later response, often preceded by an initial undershoot; veins showed very little response.

Original contributions

- The motion of RBCs have already been observed in vasculature recordings of the brain or the eye [73], and some studies aimed at estimating an average velocity of the RBCs. We used image processing techniques to investigate the variations over time of this velocity.
- The technique we used and our implementation had to be optimized to take into account the high dimensionality of the data. Indeed, the co-registration and velocity estimation procedures have to be repeated over all frames, all trial repetitions, and all vessels (velocity estimation only), and the data have high temporal and spatial resolutions are needed (sampling at 200Hz and pixels are 10 μ m size).

Weaknesses

- The structure tensor estimates an average gradient direction instead of a maximum gradient direction, which can lead to errors when several directions are present in the 2D time-space images (movement artefact, aliasing effects...). We are currently developing Gabor filters to overcome this problem.
- The method does not take into account the whole width of the vessels, but only a one-dimensional cut along them.
- The experiment was done with only one monkey, which is not enough to validate the obtained time courses.

Conclusion and implications

- This method is promising for estimating the blood velocity - and hence blood flow - responses in different vascular compartments (not the smaller ones however).
- Further developments in the image processing techniques are necessary for more precise estimates of the RBCs' velocity.
- The first stimulation-induced responses we obtained raise several interesting questions. In particular, they show the necessity to consider the different vascular compartments in hemodynamic models for the BOLD response rather than the venous compartment only.

PART IV: USING NONLINEAR MODELS IN EEG-fMRI FUSION

Chapter 8: EEG-fMRI Fusion using Kalman Filtering

Objectives

- To develop an algorithm for the “fusion” of simultaneous EEG and fMRI measurements, to estimate the spatio-temporal neural activity on the cortical surface which underlies these measurements.

Methods

- The framework includes physiological models for the EEG and fMRI measures, as well as an auto-regressive model for the neural activity. These models are gathered in a unique stochastic differential equations system.
- The Kalman filter and smoother techniques take into account the spatial and temporal information from both modalities, and integrate them in a symmetric manner.
- The method is demonstrated on synthetic data, where the input is the sum of a neural activity generated according to the auto-regressive model, and an impulse with much larger amplitude. This checks the algorithm’s ability to estimate a neural activity that does not fit the model.
- Though this data is simulated, it relies on a realistic head model obtained from an anatomical MRI image, which allows to investigate how the method can handle large amounts of data.

Results

- The Kalman filter and smoother can be applied to either the EEG measurements alone, the fMRI measurements alone, or all together. We compared the estimation results in the three cases. The EEG alone result shows an accurate temporal localization of the activity peak, but a diffuse spatial localization on the cortex. Symmetrically, the estimation using the fMRI alone shows a diffuse temporal localization of the peak, but a focused spatial localization. In the EEG+fMRI estimation, the spatial localization is exact, whereas the temporal localization looks like a compromise between EEG alone and fMRI alone solutions.
- The activity peak, which was not generated according to the auto-regressive model for neural activity, was not estimated completely accurately, which indicates that the estimation highly depends on the model supposed for neural activity.
- The fMRI and EEG brought very complementary information in the estimation of the activity outside the peak: the fMRI gives information on the low-frequency content of the activity time courses, and the EEG brings additional information on their high-frequency content.

Original contributions

- Choice of the Kalman filter as an efficient way to solve a huge inverse problem.
- We developed a cortex downsampling algorithm which keeps track of an interpolation between original and decimated mesh (it is based on the edge contraction technique, which originates from the algorithmic geometry community).
- Important code development in C++ was necessary to optimize the time and memory costs of the algorithm.

Weaknesses

- The method was only tested on synthetic data.
- The relationship between electrical and metabolic activities was assumed to be only a linear gain, which is not realistic (power relations would be preferable for example). It is probable that the original Kalman filter could have difficulties to handle a strongly nonlinear relationship, and that enhancements (Sigma-Point Kalman Filter, Particle Filter... [165]) should be considered also.
- We did not design a proper EM algorithm to estimate the physiological parameters and noise variances of the models. Rather, we used the same parameters for the estimation as for the generation of data. Note that the estimation took hours even though.

Conclusion and implications

- This chapter demonstrates that fMRI and EEG, despite of their very different temporal and spatial resolutions, can theoretically be associated to estimate an underlying cortical activity, and be very complimentary for such an estimation.
- It also demonstrates that the high dimensionality of the corresponding inverse problem can be addressed thanks to adequate estimation choice and implementation.
- The road toward real data application will probably require improvements in the models of the neural activity, and the electrical / metabolic relationships. Other filtering techniques than the Kalman filter may be considered too.

General Conclusion

The PhD work addressed two kinds of questions about hemodynamic models:

Why use biologically plausible models in brain imaging analysis?

- In chapters 4 and 5 we proved the superiority of hemodynamic models on linear convolution models for predicting the signals measured in fMRI.
- Analysis methods generally used in the General Linear Model (GLM) framework can be adapted to the nonlinear dynamical models framework: linear regression becomes model identification (chapter 4), and statistical tests become model selection (chapter 5).
- Dynamic models should be used to address specific questions on temporal aspects of the cortical activity (e.g. neural habituation), but linear models are still preferable to address simple cognitive questions (e.g. activation detection).
- There is another advantage of using biologically plausible models for the integration of data issued from different imaging modalities (such as EEG-fMRI fusion as in chapter 8). They enable the estimation of the underlying cortical activity, again in the dynamical model framework, which appears to be both rigorous and simple.

Which additional modelling are required?

- In chapter 6, we proved that existing models of the blood flow are over-simplistic: they should take into account non-linear phenomena such as flow saturation, and include complex interactions between flow and blood volume, instead of a simple inductive interaction.
- It is necessary in particular to consider each vascular compartment separately: arteries, capillaries, veins. This observation is confirmed by the technique we developed in chapter 7 to estimate the blood flow in individual vessels.
- It ought to be noted that this new technique is an original result of a fruitful interaction between neuroscience and computer vision.
- In addition to this mechanical aspect of hemodynamics, it is necessary to better understand on which component of the neurones electrical activity the blood flow response depends.

It is probable that future progresses in both modelling and analysis techniques will result in a more and more acute interpretation of fMRI data and of other data relying on hemodynamic processes.

Synopsis (Français)

Résumé

L'enjeu de la présente thèse est de proposer de nouvelles méthodes d'analyse des données d'imagerie cérébrale acquises en Imagerie par Résonance Magnétique fonctionnelle (IRMf). Elle s'est concentrée en particulier sur la compréhension des signaux temporels mesurés en IRMf et leur lien avec l'activité cérébrale. En effet, les variations du signal que l'on observe en IRMf sont dues à des changements de l'afflux du sang dans le cerveau et de l'oxygénation de ce sang. Ces changements sont liés à l'activité des neurones, et l'on nomme ce phénomène la *réponse hémodynamique*. Cette réponse hémodynamique fait l'objet d'un important effort de modélisation, de manière à mieux pouvoir interpréter les données d'IRMf. Et cette thèse contient des travaux liés à la fois à la modélisation pour elle-même, avec l'étude de certains détails des modèles hémodynamiques, et à la fois à l'utilisation de ces modèles pour l'analyse des données, avec en particulier l'analyse des données IRMf et la fusion entre des données d'IRMf et d'Electroencéphalographie (EEG).

Ainsi, la première partie de la thèse est consacrée à l'utilisation de modèles hémodynamiques en IRMf. En effet, aujourd'hui, les méthodes standard d'analyse de données d'IRM fonctionnelle utilisent le Modèle Général Linéaire (GLM), qui suppose une relation linéaire entre l'activité des neurones, la réponse hémodynamique et les mesures IRMf. Nous montrons qu'il est aussi possible d'utiliser des modèles plus plausibles du point de vue biologique, et éventuellement non linéaires pour analyser les données. A la place de la régression linéaire utilisée habituellement, nous proposons une identification de modèle basées sur une minimisation d'énergie, et nous proposons d'adapter les tests de Fisher utilisés habituellement dans le cadre du GLM pour pouvoir réaliser dans le nouveau cadre la détection d'activations, le test d'hypothèses cognitives, ainsi que des comparaisons entre différents modèles.

La seconde partie quant à elle est expérimentale : nous avons étudié les équations de différents modèles hémodynamiques grâce à des expériences d'Imagerie Optique chez le singe éveillé, dans le cadre d'une collaboration avec Ivo Vanzetta dans l'équipe "Dynamique de la perception visuelle et de l'action" au CNRS Marseille. Nous nous sommes intéressés en particulier à la dynamique du flux sanguin, qui est de première importance car elle fait le lien entre les activités électriques et métaboliques et les changements du volume et de l'oxygénation du sang. Nous avons mis en évidence des aspects de la réponse hémodynamique qui ne sont pas prévus par les modèles actuels, tels qu'une non linéarité de cette réponse du flux par rapport à l'intensité de la réponse électrique. Par

ailleurs, dans le cadre de la même collaboration, nous avons conçu une méthode pour estimer la vitesse des globules rouges dans les vaisseaux sanguins filmés en Imagerie Optique, qui constitue une nouvelle technique de mesure de ce flux sanguin.

Enfin, dans la troisième partie, nous avons étendu les méthodes présentées dans la première partie à l'analyse de données de modalités multiple, et en particulier, proposons une méthode pour estimer l'activité cérébrale à partir d'enregistrements simultanés en IRMf et en EEG. Cette méthode est validée sur des données synthétiques.

Le présent synopsis résume les points importants de ces travaux : les objectifs, les méthodes, les conclusions et conséquences pour chaque chapitre (en-dehors des trois premiers chapitres qui sont consacrés à introduire le travail). Nous avons également tenté d'en présenter une critique objective, en mentionnant à la fois ce qui constitue des contributions originales et les faiblesses restantes. Nous espérons que ce résumé permettra au lecteur de se repérer rapidement dans cette thèse, et de bien comprendre les relations entre ses différentes composantes.

PARTIE II: UTILISATION DE MODELES NON-LINEAIRES EN IRMf

Chapitre 4: Identification de Modèles

Objectifs

- Estimer les paramètres d'un modèle physiologique (de la réponse hémodynamique) à partir de mesures en IRM fonctionnelle.
- Evaluer la précision d'une telle estimation.
- Evaluer l'adéquation entre des modèles hémodynamiques et les réponses réelles observées.

Méthodes

- Le formalisme des systèmes dynamiques est utilisé pour décrire des modèles physiologiques.
- Deux approches sont proposées pour étudier la stabilité d'un tel système dynamique autour de l'équilibre (l'une intuitive et la seconde plus mathématiques).
- Notre algorithme d'estimation des paramètres repose sur un critère de maximum de vraisemblance.
- Une analyse de sensibilité permet de quantifier l'identifiabilité du système, en étudiant sa sensibilité aux variations des paramètres.
- Cette méthode est validée sur des données synthétiques, et est comparée à une seconde reposant sur un critère de vraisemblance de Kalman.
- Une expérience est réalisée en IRM fonctionnelle pour comparer un modèle linéaire et huit variations des modèles proposés par Buxton et al.[33, 32] et Friston et al.[59] : nous comparons deux modélisations différentes du flux sanguin (l'une de la forme d'un oscillateur amorti [59] et l'autre d'une convolution avec une fonction gamma [32], nous étudions les effets de viscosité du volume sanguin et d'habituation neuronale [32], et nous essayons de fixer les valeurs de certains paramètres. Ces estimations de paramètres sont réalisées à la fois sur des données de haute qualité (réponses moyennées pour toutes les stimulations et tous les sujets), et sur les décours temporels individuels.

Résultats

- L'algorithme reposant sur le critère de vraisemblance de non linéaires ne produit pas de meilleurs résultats que celui reposant sur un simple maximum de vraisemblance; nous avons donc utilisé le second pour l'analyse de données réelles en raison de sa rapidité.
- Les modèles physiologiques se révèlent en adéquation avec les réponses mesurées : ils ont tous un meilleur rapport signal sur bruit par rapport aux données réelles que le modèle linéaire, ce rapport lui-même pouvant varier d'un modèle à l'autre.

- Tous les modèles sont peu identifiables, c'est à dire que les valeurs de leurs paramètres peuvent varier beaucoup sans que cela n'induisse de variations importantes en sortie du modèle. Cela est du en partie au fait que nous estimons à la fois un facteur d'échelle en entrée et en sortie. Cela empêche de déterminer les valeurs exactes de certains paramètres, à cause de leurs interactions avec les autres paramètres.

Contributions originales

- Notre méthode analytique pour le calcul de la dérivée de la sortie du système par rapport aux paramètres augmente la rapidité et la précision de l'algorithme de maximisation de la vraisemblance
- Nous mettons en évidence dans ce chapitre des problèmes liés à l'identifiabilité des paramètres : on ne peut pas garantir d'estimation exacte de la valeur des paramètres (ce qui n'est pas si étonnant pour certains paramètres : comment pourrait-on espérer que l'IRMf seule puisse permettre d'estimer par exemple la valeur de la fraction d'extraction d'oxygène au repos ?). Le chapitre suivant montre comment cette non identifiabilité n'est pas incompatible avec une caractérisation correcte de la dynamique du système.

Faiblesses

- Notre expérience a été réalisée sur un scanner 3 Tesla, mais il n'existe actuellement que des modèles de la mesure BOLD à 1.5 Tesla. Nous avons donc utilisé un modèle à 1.5 Tesla [32], $y = V_0(a_1(q-1) - a_2(v-1))$, en supposant ses paramètres inconnus (modulo la relation $a_2/(a_1 + a_2) = 1/10$).
- Nous n'avons pas utilisé d'a priori pour l'estimation des paramètres, ce qui serait probablement l'unique moyen d'augmenter leur identifiabilité.

Conclusion et conséquences

- Les modèles physiologiques permettent de prédire correctement les mesures BOLD, mais dans le même temps il n'est pas possible d'estimer les valeurs exactes des paramètres de ces modèles. Cela ouvre la voie à l'utilisation de modèles non linéaires pour l'analyse de données d'IRMf, ce qui est l'objet du chapitre suivant.
- L'utilisation d'a priori sur les paramètres est probablement le seul moyen d'estimer plus exactement leurs valeurs à partir des signaux IRMf.
- Les différents modèles utilisés montrent des différences dans la qualité de la prédiction. La comparaison statistique entre différents modèles est également étudiée dans le chapitre suivant.

Chapitre 5: Test d'Hypothèses et Sélection de Modèle dans le Cadre non Linéaire

Objectifs

- Développer des méthodes statistiques pour analyser des données IRMf en utilisant des modèles physiologiques plutôt que des modèles de convolution linéaire.
- Evaluer ces méthodes dans le cadre de la détection d'activation.
- Utiliser ces méthodes également pour la sélection entre différents modèles physiologiques.

Méthodes

- Pour pouvoir utiliser les mêmes tests de Fisher que dans le logiciel SPM dans le nouveau contexte d'un modèle non linéaire, nous estimons d'abord les paramètres du modèle, puis linéarisons les équations par rapport aux paramètres autour de l'estimée.
- Comme ce test de Fisher ne permet pas de comparer des modèles non hiérarchiques (c'est à dire telle que le premier est un cas particulier du second), nous proposons un test non hiérarchique basé sur la construction d'un troisième modèle plus général dont les deux premiers sont des cas particuliers.
- Ces tests sont appliqués aux données du chapitre précédent.

Résultats

- Il n'y a pas de différence entre les résultats de détection d'activation en utilisant des modèles linéaires ou non linéaires
- Toutes les complexifications des modèles que nous avons étudiés - à savoir l'hypothèse d'un effet de viscosité du volume sanguin, et celle d'une habitude neuronale - ont été validées par les tests statistiques, appliqués aux réponses moyennées ou même aux décours temporels des voxels les plus activés.
- La modélisation du flux proposée par Friston (oscillateur amorti) a été préférée à celle proposée par Buxton (convolution avec une fonction gamma), quoique la différence fût faible.
- Le modèle linéaire et les modèles physiologiques les plus simples en revanche sont préférés lorsque les tests sont appliqués aux décours temporels de voxels bruités ou peu activés, car ces modèles ont moins de degrés de liberté.

Contributions originales

- Ce travail montre que l'on peut utiliser des modèles non linéaires dans l'analyse d'expériences en IRM fonctionnelles.

- Nous recommandons d'utiliser des tests statistiques basés sur un critère de rapport signal sur bruit, tel que le test de Fisher, plutôt que des tests basés sur une distribution a posteriori marginale d'un paramètre spécifique, comme le proposent Friston et al.[58]. En effet, la non identifiabilité d'un tel paramètre implique que sa distribution a posteriori est très plate.
- Les tests hiérarchique et non hiérarchique que nous proposons permettent de choisir le modèle le plus statistiquement significatif.

Faiblesses

- A cause de l'étape de linéarisation des modèles, les p-values calculées pour réaliser les tests statistiques peuvent ne pas être exactes, car les modèles sont fortement non linéaires par rapport aux paramètres physiologiques. Des méthodes plus robustes devraient être envisagées comme l'utilisation d'un critère d'information d'Akaike (AIC), ou même l'échantillonnage de tout l'espace des paramètres (Metropolis-Hastings Markov-Chain Monte Carlo [127]).
- Nous supposons un bruit blanc au lieu d'estimer des autocorrélations qui pourraient également modifier la valeur de la p-value.

Conclusion et conséquences

- Ce chapitre est une première étape vers l'utilisation de modèles non linéaires dans l'analyse de données d'IRM fonctionnelle. Nous avons montré la possibilité d'adapter des tests statistiques à ce contexte non linéaires, ce qui constitue un encouragement à poursuivre les efforts de modélisation de la réponse BOLD.
- Nous avons pu contourner le problème d'identifiabilité des modèles signalé dans le chapitre précédent : le plus important au sujet de l'estimation des paramètres est qu'elle permette une caractérisation exacte du système dynamique, tandis que la valeur exacte des paramètres importe moins.
- Plusieurs ajouts par rapport au modèle initial de Buxton et al. ont été validés par les tests statistiques. Cependant, ils soulèvent encore des questions d'ordre physiologiques : Est-ce que les mesures in vivo confirment l'hypothèse de viscosité ? Celle de l'habituation neuronale ? Les non linéarités observées en IRMf dans les échelles de temps courtes ne pourraient-elles pas être causées par une non linéarité de la réponse du flux plutôt qu'un phénomène d'habituation neuronale ? Est-ce que donc les deux modèles du flux proposés par Friston et Buxton ne seraient pas trop simples ? Ces deux questions sont l'objet du chapitre suivant.
- L'utilisation de modèles linéaire quant à elle est justifiée pour certaines applications sur des données fort bruitées, par exemple pour la localisation d'activité.

PARTIE III: ETUDE DU FLUX SANGUIN EN IMAGERIE OPTIQUE

Chapitre 6: Dynamique et Non-linéarités de la Réponse du Flux Sanguin

Objectifs

- Etudier si les non linéarités observées en IRM fonctionnelle ne peuvent pas être attribuées à des effets non linéaires dans la réponse du flux sanguin à l'activité neuronale plutôt qu'à une habitude neuronale, grâce à une expérience en Imagerie Optique.
- Etudier les relations temporelles entre les réponses du flux et du volume sanguin.

Méthodes

- Nous utilisons des enregistrements simultanés du volume sanguin (mesuré avec des signaux optiques intrinsèques à une longueur d'onde isosbétique) et du flux sanguin (mesuré avec une sonde non hiérarchique), et des enregistrements séparés de l'activité électrique (mesurée grâce à un gel sensible au voltage - technique VSD).
- Le protocole expérimental est similaire à celui des expériences précédentes en IRMf, excepté qu'il se concentre sur de courtes durées de stimulation visuelle.
- Les mesures sont d'abord décrites de manière qualitative, puis nous proposons un nouveau modèle non linéaire du flux pour tenter de capturer les effets non linéaires de saturation.

Résultats

- Nous avons observé une relation linéaire entre la durée de la stimulation visuelle et l'amplitude de l'activité électrique des neurones mesurée avec la technique VSD, une relation non linéaire entre l'amplitude de cette activité électrique et de la réponse du flux, et une relation linéaire entre les amplitudes du flux et du volume.
- La réponse du flux apparaît en retard par rapport à la réponse du volume, ce qui est en contradiction avec les modèles qui supposent une réponse du volume déterminée par celle du flux. De plus, nous observons une petite réponse négative initiale du flux.
- La modélisation d'un phénomène de saturation permet de mieux prédire l'amplitude du flux, mais pas les délais.

Contributions originales

- Nous avons montré que les modèles linéaires actuellement proposés pour la réponse du flux sanguin ne peuvent pas rendre compte de certains phénomènes, ce qui suggère d'étudier prioritairement cette réponse du flux.

Faiblesses

- Cette expérience n'a été réalisée que sur un seul animal.
- Le modèle de saturation que nous proposons est probablement encore trop simple.

Conclusion et conséquences

- Principalement, deux raisons permettent d'expliquer les non linéarité observées entre les réponses électriques (VSD) et hémodynamiques (flux et volume). Soit la réponse hémodynamique dépend d'autres aspects de l'activité neuronale non détectés par la technique VSD (par exemple, l'émission de spikes), qui eux-mêmes seraient sujet à une habituation neuronale. Soit la réponse du flux sanguin à l'activité neuronale présente des effets non linéaires de saturation.
- Le délais entre les réponses du volume et du flux sanguin, et la réponse initiale négative du flux, pourraient avoir une explication déjà proposée par hiérarchique et al.[115] : le volume sanguin augmente d'abord dans les capillaires et les artéioles non pas sous l'effet d'une augmentation du flux entrant, mais sous celui d'une diminution du flux sortant.
- L'ensemble de ces questions souligne l'importance de modéliser chaque compartiment vasculaire séparément. Il serait très intéressant pour cela d'avoir une mesure spatiale du flux à la place d'une mesure en un seul point, ce qui permettrait de distinguer ces différents compartiments. Le chapitre suivant propose une méthode qui produit une telle mesure spatiale.

Chapitre 7: Mesure Bidimensionnelle du Flux en Enregistrements Intrinsèques

Objectifs

- Estimer la vitesse du sang à l'intérieur des vaisseaux et les variations de cette vitesse dans le temps, grâce à la détection du mouvement des globules rouges (RBCs).

Méthodes

- Le principe est d'utiliser des techniques de traitement d'image classiques pour estimer le mouvement des globules rouges le long des vaisseaux dans des images de volume mesurées avec des signaux optiques intrinsèques à 570nm.
- Les séquences d'images sont tout d'abord recalées pour corriger les mouvements de l'ensemble du cerveau, puis des filtrages temporels et spatiaux permettent de mieux faire ressortir le mouvement le long des vaisseaux.
- Les données de vaisseaux individuels segmentés manuellement sont extraites dans des matrices 2D temps-espace.
- Nous utilisons la technique du *tenseur de structure* pour détecter les directions des lignes de trajectoires à l'intérieur de ces matrices. Le tenseur de structure est un outil classique pour la détection de directions en traitement d'image.
- La méthode a été utilisée avec une première expérience sans stimulation, pour détecter les pulsations cardiaques, avec en contrôle une mesure simultanée du flux grâce à une sonde non hiérarchique
- Elle a ensuite été testée dans une seconde expérience utilisant des stimulations visuelles.

Résultats

- La technique a permis de détecter correctement des changements de la vitesse du sang induits par la pulsation cardiaque.
- Elle a également permis de détecter correctement des réponses évoquées par une stimulation visuelle. Les réponses observées sont similaires à l'intérieur d'une même catégorie de vaisseaux : on observe une réponse rapide dans les artères, une réponse plus tardive dans les artérioles et les vénules (voire une courte réponse négative initialement), et peu de réponse dans les veines.

Contributions originales

- Le mouvement des globules rouges avait déjà été observé en imagerie optique [73], et certaines études avaient déjà estimé une vitesse moyenne des globules rouges. Nous avons montré qu'il est possible d'estimer les variations dans le temps de cette vitesse grâce à des techniques de traitement d'image.

- Le choix que nous avons fait du tenseur de structure et son implémentation ont été optimisés pour tenir compte de la très grande dimensionnalité des données. En effet, les procédures de recalage et d'estimation de la vitesse doivent être répétées pour toutes les images de tous les essais, et pour tous les vaisseaux, alors que les données elles-mêmes ont une haute résolution temporelle (200Hz) et spatiale ($10\mu\text{m}$).

Faiblesses

- Le tenseur de structure estime une direction moyenne du gradient dans les images 2D temps-espace, et non la direction dans laquelle le gradient est maximal, ce qui peut conduire à des erreurs lorsque plusieurs directions sont présentes (artefacts de mouvements, effets d'aliasing, ...). D'autres techniques pourraient être envisagées pour palier ce problème (filtre de Gabor).
- La méthode ne prend pas en compte toute l'épaisseur des vaisseaux, mais seulement une coupe unidimensionnelle le long d'eux. La encore des algorithmes plus complexes peuvent être envisagés.
- L'expérience n'a été réalisée que sur un animal, ce qui n'est pas suffisant pour valider les décours temporels obtenus.

Conclusion et conséquences

- Cette méthode est prometteuse pour l'estimation de la vitesse du sang - et de là du flux sanguin - et ce dans différents compartiments vasculaire (mais pas les capillaires cependant qui sont trop petits).
- De nouveaux développements des techniques de traitement d'image appliquées à ce problème pourront probablement améliorer les réponses obtenues.
- Les premières réponses évoquées que nous avons obtenues soulèvent d'intéressantes questions physiologiques, et prouvent à nouveau l'importance de considérer les différents compartiments vasculaire dans la modélisation de la réponse hémodynamique plutôt que le seul compartiment veineux.

PARTIE IV: UTILISATION DE MODELES NON LINEAIRES POUR LA FUSION EEG-IRMF

Chapitre 8: Fusion EEG-IRMF avec le filtrage de Kalman

Objectifs

- Développer un algorithme pour la “fusion” d’enregistrements simultanés en EEG et IRMF, de manière à estimer l’activité spatio-temporelle des neurones sur la surface du cortex sous-jacente à ces mesures.

Méthodes

- Nous utilisons des modèles physiologiques des mesures EEG et IRMF, ainsi qu’un modèle autorégressif simple de l’activité neuronale. Ces modèles sont regroupés sous le formalisme d’un unique système d’équations différentielles stochastiques.
- Les techniques du filtre et du lisseur de Kalman permettent de tenir en compte les informations à la fois temporelles et spatiales des deux modalités, et de les intégrer d’une manière symétrique.
- La méthode est validée sur des données synthétique, où l’entrée du modèle (l’activité neuronale) a été choisie comme la somme d’une activité générée selon le modèle autorégressif, plus une activité impulsionnelle de forte amplitude. Ainsi, nous pouvons observer la capacité de l’algorithme à estimer une activité neuronale qui ne respecte pas le modèle autorégressif a priori.
- Ces données sont synthétiques, mais reposent néanmoins sur un modèle de tête réaliste obtenu à partir d’une image IRM anatomique, ce qui permet de tester la capacité de la méthode à gérer des données de dimensionnalité réaliste.

Résultats

- L’algorithme (filtre et lisseur de Kalman) peut être appliqué soit aux mesures EEG seules, soit aux mesures IRMF seules, soit aux deux ensemble. Nous avons comparé les résultats obtenus dans les trois cas. Les résultats en EEG seule sont précis en terme de localisation temporelle du pic d’activité, mais sa localisation est diffuse sur le cortex. De manière symétrique, l’estimation basée sur l’IRMF seule est focale sur le plan spatial, mais diffuse sur le plan temporel. Pour l’estimation en EEG+IRMF, la localisation spatiale est exacte, tandis que la localisation temporelle montre un compromis entre les solutions EEG seule et IRMF seule.
- L’IRMF et l’EEG apportent des informations complémentaires également pour l’estimation de l’activité autorégressive (en-dehors du pic d’activité) : l’IRMF donne une information sur le contenu basse fréquence des décours temporels de l’activité, tandis que l’EEG apporte une information supplémentaire sur leur contenu haute fréquence.

Contributions originales

- Le choix du filtre de Kalman comme un moyen efficace de résoudre un problème inverse de dimensionnalité très grande.
- Nous avons développé parallèlement un algorithme de décimation de cortex qui conserve les paramètres de l'interpolation entre les maillages original et décimé, basé sur la technique de contraction d'arête, connue dans la communauté de géométrie algorithmique.
- Un important développement en C++ a été nécessaire pour optimiser les coût en temps et en mémoire de l'algorithme.

Faiblesses

- La méthode a été testée sur des données synthétiques seulement.
- La relation entre les activités électrique et métabolique est supposée linéaire, ce qui n'est pas réaliste (une loi de puissance par exemple serait déjà préférable). En fait, le filtre de Kalman pourrait avoir des difficultés dans le cas d'une relation très non linéaires, et il serait préférable alors d'utiliser des techniques de filtrage plus élaborées (filtre particulaire [165], ...).
- Nous n'avons pas utilisé d'algorithme EM pour estimer également les paramètres physiologiques et les variances des bruits dans le modèle. Au contraire, nous avons utilisé pour l'estimation les mêmes paramètres que pour la génération des données. Cela est du au problème du temps d'estimation qui déjà dans ce contexte simple s'élevait à plusieurs heures.

Conclusion et conséquences

- Ce chapitre montre que l'IRM fonctionnelle et l'EEG, malgré les importantes différences de leurs résolutions temporelles et spatiales, peuvent théoriquement être associées pour l'estimation de l'activité corticale sous-jacente, et être alors très complémentaires.
- Il montre également que la haute dimensionnalité du problème inverse correspondant peut être gérée grâce à de bons choix d'estimation et d'implémentation.
- L'application à des données réelles demandera probablement au préalable d'améliorer les modèles de l'activité électrique et de la relations entre les activités électrique et métabolique. D'autres techniques de filtrage que le simple filtre de Kalman devraient être envisagées également.

Conclusion générale

Ce travail de thèse a abordé deux grandes questions autour des modèles hémodynamiques

Quel est l'intérêt d'utiliser des modèles hémodynamiques en imagerie cérébrale ?

- Les chapitres 4 et 5 prouvent la supériorité des modèles dynamiques sur les modèles de convolution linéaire pour prédire les réponses observées en IRMf.
- Les méthodes d'analyse utilisées dans le cadre du Modèle Linéaire Général (GLM) peuvent être adaptées aux modèles dynamiques non linéaires: la régression linéaire devient de l'identification de modèle (chapitre 4) et les tests statistiques deviennent de la sélection de modèle (chapitre 5).
- Les modèles dynamiques devraient donc être utilisés pour répondre à des questions précises sur des aspects temporels de l'activité corticale (par exemple, concernant une adaptation neuronale), mais les modèles linéaires restent préférable pour répondre à des questions simples (par exemple, pour détecter des activations).
- Il y a un autre avantage à utiliser des modèles biologiquement plausibles, pour l'intégration de plusieurs modalités (la fusion EEG-IRMf entre autres dans le chapitre 8). Ils permettent en effet d'estimer l'activité neuronale dans un cadre rigoureux et finalement simple, celui encore des systèmes dynamiques.

Quels futurs efforts de modélisations ?

- Dans le chapitre 6, on a montré que les modèles actuels pour le flux sanguin étaient trop simples : il devraient d'une part tenir compte de phénomènes non linéaires de saturation du flux, et prévoir des interactions complexes (au lieu de simplement inductive) entre le flux et le volume sanguin.
- Il est nécessaire en particulier de considérer séparément les différents compartiments sanguins : artères, capillaires, veines. Cette constatation est confirmée par la technique mise en oeuvre dans le chapitre 7 pour estimer le flux sanguin dans chaque vaisseau individuel.
- On note d'ailleurs que cette nouvelle technique est le résultat original d'une interaction fructueuse entre les neurosciences et la vision par ordinateur.
- En plus de cet aspect mécanique de la modélisation hémodynamique, il est nécessaire de mieux comprendre de quelles composantes de l'activité électrique dépend exactement le flux sanguin.

On peut espérer que les prochains progrès concernant à la fois la modélisation et sa prise en compte dans l'analyse du signal permettront une utilisation de plus en plus fine des techniques d'imagerie comme l'IRMf basées sur des phénomènes hémodynamiques.

Contents

I	Introduction	3
1	Brain Imaging	5
1.1	Physiological and Physical Basis of Measure Modalities	5
1.1.1	The Neuron	6
1.1.2	Macroscopic Electrophysiology	6
1.1.3	Metabolism	7
1.2	EEG Methods	9
1.2.1	Physics of EEG and MEG	9
1.2.2	Forward Problem	10
1.2.3	Inverse Problem	11
1.2.4	Real Data Processing	13
1.3	fMRI Methods	14
1.3.1	Physics of the Magnetic Resonance Imaging	14
1.3.2	Pre-processing	19
1.3.3	Hypothesis-Driven Analysis Methods	21
1.4	Optical Imaging Methods	26
1.4.1	Physics of the Optical Imaging	28
1.4.2	Acquisition Methods	29
1.4.3	Data Analysis	31
2	Physiological Models of the hemodynamic response	33
2.1	Physiological Models	33
2.1.1	Neural Basis of the BOLD Signal	33
2.1.2	Metabolism	35
2.1.3	Flow Response	38
2.1.4	Oxygen Extraction from the Capillaries	39
2.1.5	Visco-Elastic Properties of the Vessels and the Venous Balloon .	41
2.1.6	The BOLD Signal	43
2.2	Empirical Observations and Predictions of the BOLD Response	44
3	State of the Art of EEG/MEG-fMRI Fusion	49
3.1	Understanding the problem	49
3.1.1	Physiological Relation between Electrical and Hemodynamic Responses	50
3.1.2	EEG/MEG-fMRI fusion: what should be estimated ?	50
3.1.3	Methodology	51
3.2	EEG/MEG and fMRI Share a Spatial Information	51
3.2.1	Convergence Evidence	53

3.2.2	fMRI-Guided EEG/MEG	53
3.2.3	Symmetric Spatial Integration	54
3.3	EEG/MEG and fMRI Share a Temporal Information	55
3.3.1	EEG-Guided fMRI	55
3.3.2	Symmetric Temporal Integration	56
3.4	EEG/MEG and fMRI Share a Spatio-Temporal Information	56
3.4.1	Trial by Trial Variations in Response Amplitude	56
3.4.2	Concurrent Estimation of the Spatio-Temporal Neural Activity	57
II Using Nonlinear Models in fMRI		59
4	Model Identification	61
4.1	Model Formulation	61
4.1.1	Stochastic Differential Equations (SDE) Terminology	61
4.1.2	A First Model	62
4.1.3	System Stability	63
4.1.4	A Second Model	64
4.1.5	A Note on Integration	66
4.2	Parameter Estimation	66
4.2.1	The Least Square Approach	67
4.2.2	Derivative of the System Output with respect to Parameters	68
4.2.3	Handling Confound Effects	68
4.2.4	Simulations	69
4.2.5	Including Evolutive Noise: Kalman Filtering	71
4.3	Quantifying Estimation Precision	74
4.3.1	A Sensitivity Analysis	75
4.3.2	A Posteriori Probabilities on Parameters in a Linearized Framework	77
4.4	Experimental Results	79
4.4.1	Experimental Design	79
4.4.2	Data Analysis	81
4.4.3	Qualitative Description of the Estimated Responses	82
4.4.4	Fitting Models to Mean Responses	84
4.4.5	Sensitivity Analysis of the Mean Responses	86
4.4.6	Estimation on Voxel Raw Time Course	88
4.5	Conclusion	90
5	Nonlinear Hypothesis Testing and Model Selection	91
5.1	Statistical Test on Nonlinear Models	91
5.1.1	Hierarchical Model Comparison	92
5.1.2	Application to Activation Detection	93
5.1.3	Application to Other Cognitive Questions	94
5.1.4	Non-Hierarchical Model Comparison	94
5.2	Experimental Results	95
5.2.1	Model Selection on Average Responses	95
5.2.2	Activation Maps	96
5.2.3	Model Comparison on Voxel Raw Time Courses	96
5.3	Discussion	99

5.3.1	Choice of Hemodynamic Model	99
5.3.2	Using Dynamical Models in fMRI	101
5.3.3	Assumptions Used	101
5.4	Conclusion	101
III	Investigation on the Cerebral Blood Flow in Optical Imaging	103
6	Dynamics and Nonlinearities of the Flow Response	105
6.1	Introduction	105
6.2	Methods	106
6.2.1	Experimental Paradigm	106
6.2.2	Optical Imaging Setup	107
6.2.3	Signal Processing	108
6.3	Results	109
6.3.1	Responses to the Stimulations	109
6.3.2	Non-linearities	110
6.3.3	Flow and Volume Dynamics	116
6.4	Discussion	116
6.4.1	Non-Linearity of the Hemodynamic Response	116
6.4.2	Blood Flow Dynamics	117
6.5	Conclusion	118
7	Bi-dimensional Flowmetry with Intrinsic Recordings	119
7.1	Introduction	119
7.2	Method	121
7.2.1	Data Pre-processing	121
7.2.2	Detecting Hemoglobin Trajectories with the Structure Sensor	125
7.2.3	Conducted Experiments	127
7.3	Results	127
7.3.1	Extraction of RBC-Motion Information	127
7.3.2	Sensory-Evoked Changes in RBCs' Speed	128
7.4	Discussion	129
7.4.1	RBCs-Motion Estimation	129
7.4.2	Application for Physiological modelling	130
7.4.3	Conclusion	130
IV	Using Nonlinear Models in EEG-fMRI Fusion	133
8	EEG-fMRI Fusion using Kalman Filtering	135
8.1	Methods	135
8.1.1	Why Choose the SDE Formulation ?	135
8.1.2	Physiological Models	137
8.1.3	Kalman Filter and Kalman Smoother	138
8.1.4	Implementation	141
8.2	Results on Synthetic Data	141
8.2.1	Cortex Mesh Downsampling	141
8.2.2	Artificial Data	142

8.2.3	Results	143
8.3	Discussion	147
8.3.1	Combining EEG and fMRI Measures with Filtering Techniques .	147
8.3.2	Open Questions	147
8.4	Conclusion	148

List of Symbols

Dynamic Variables

u	Stimulation pattern
N	Neural activity
j	Electrical activity on the cortical surface
CMRO_2	Cerebral metabolism of oxygen
m	Normalized CMRO_2
\bar{C}^c	Average oxygen concentration in the capillaries
E	Oxygen extraction ration in the capillaries
g	Ratio between tissue and capillary plasma oxygen concentration
CBF, F	Cerebral blood flow (general or venous compartment)
f	Normalized CBF
$\text{CBV}, V, \text{Hbt}$	Cerebral blood volume fraction (general or venous compartment)
v	Normalized CBV
Q, Hbr	Deoxyhemoglobin content (general or venous compartment)
q	Normalized deoxyhemoglobin content
S	fMRI BOLD signal
y	Fractional BOLD signal change
z	EEG measure

Model Parameters

Neural response and habituation

ϵ	Neural efficiency
τ_I	Inhibitory time constant
κ_N	Inhibitory gain factor

Neurovascular coupling (Friston et al.[59])

κ_s	Flow decay time constant
κ_f	Flow feedback time constant

Neurovascular coupling (Buxton et al.[32])

τ_f	Width of CBF impulse response
f_1	Normalized CBF response to sustained neural activation
τ_m	Width of CMRO_2 impulse response
m_1	Normalized CMRO_2 response to sustained neural activation
n	Steady-state flow-metabolism relation ($f - 1 = n(m - 1)$)
δ_t	Delay of CBF response relative to CMRO_2 response
χ	$\chi = \epsilon(f_1 - 1)$

Capillary oxygen extraction

C^a	Arteriolar blood oxygen concentration
E_0	Baseline oxygen extraction
g_0	Baseline ratio of tissue and arterial plasma oxygen concentration
ϕ	Capillary transit time constant
\bar{C}_0^a	Baseline oxygen concentration in the capillaries
ρ	Tissue transport time constant

Venous dynamic

α	Grubbs' parameter (steady-state flow-volume relation: $v = f^\alpha$)
τ	Venous transit time
$\tau_{visc}^+/\tau_{visc}^-$	Hysteresis venous compliance constants

Other physiological parameters

V_0	Baseline blood volume fraction
-------	--------------------------------

BOLD measure (at 1.5T: Buxton et al.[32])

a_1	Weight for deoxyhemoglobin change
a_2	Weight for volume change
b	$b = V_0(a_1 + a_2)$

Model functions

h	BOLD impulse function
h_f	CBF impulse function (Buxton et al.[32])
h_m	CMRO ₂ impulse function (Buxton et al.[32])
G	Lead-field matrix for the EEG forward problem

Stochastic Dynamical Equations

x, X	Hidden-states of the system
y, Y	Output of the system
ξ	Evolution noise (variance is usually Q)
η	Measure noise (variance is usually Σ or R)
e	Cumulated error (including evolution and measure noise, variance is usually Σ)

Part I

Introduction

Chapter 1

Brain Imaging

In this chapter, we propose a brief overview of different brain investigation techniques, as well as their physiological support. Then, we shall describe in more detail the three modalities we were more interested in, Electro-Encephalography, functional Magnetic Resonance Imaging and Optical Imaging. We will also explain how the measures provided by these modalities are analyzed.

1.1 Physiological and Physical Basis of Measure Modalities

Neuroscientists sometimes like to define the ultimate goal of their research as “deciphering the neural code”. It is tempting indeed to consider the brain as a machine whose mechanisms could be dissected, analyzed and at last explained. However, notwithstanding the fact that the brain complexity cannot be compared to that of any of our computers, a first major obstacle to its comprehension is the difficulty to observe the way it manipulates “information”. Existing techniques for the measure of the cortical activity rely upon very modern and diverse sciences and technologies. However, each of them can only access a part of the desired information. As a consequence it is necessary to optimize the precision of the analysis based on each technique, and to know how to benefit from their diversity.

It is usual to classify these brain measurement techniques according to the following criteria:

- the **spatial resolution** (please note that in addition to the minimum spatial resolution, some techniques are also limited by the maximum spatial extent they can investigate at the same time)
- the **temporal resolution**
- the **invasiveness** (how aggressive is the technique for the subject?)

Of course, no technique can simultaneously offer the best spatial and temporal resolutions (and be non-invasive at the same time). Each one has its specific advantages. Figure 1.1 shows a summary of these criteria for the most usual techniques. We shall present now the necessary physiological and physical notions needed to describe these techniques, beginning with the best spatial resolution.

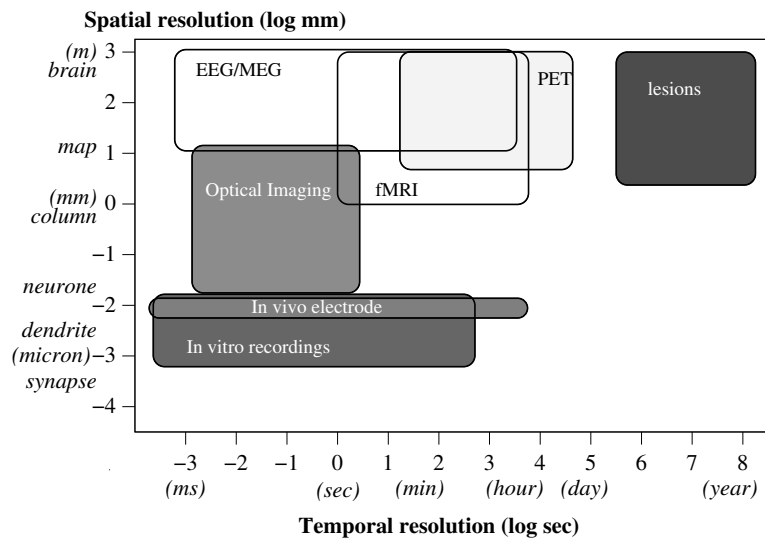


Figure 1.1: Temporal and spatial resolution of the most usual brain investigation techniques. The colour expresses how much invasive they are (from white: non invasive, to dark grey: highly invasive).

1.1.1 The Neuron

Neurons are the basic units of the cortical network. Their number is about 10^{11} in the brain. A single neuron integrates information collected from other neurons by its dendrites through synaptic connections, and then sends itself excitatory or inhibitory signals via strong discharges (*action potentials*, or *spikes*), that propagate through its axon and are transmitted to other neurons dendrites. More precisely, synaptic excitation consists of ionic channels opening, which induce changes in the intracellular potential (variations are negative or positive according to the excitatory or inhibitory nature of the synaptic connection). The unbalance between intra- and extra-cellular potentials is then called *membrane polarization*. When this polarization reaches a certain threshold (around -40mV), additional ionic channels are opened and cause the spike emission, which propagates through the axon afterward. The mechanisms that govern membrane depolarization are subject to a high modelling effort [82, 105].

Neurons activity can be monitored *in vivo* by introducing electrodes in animals cortex (usually rats, cats or monkeys, anesthetized or awake), which measure variations of the extra-cellular potential. Spikes emitted by the surrounding neurons are clearly visible in the recorded *electrophysiological signals*, and it is possible to move the electrode close enough to a single cell to record its activity. Spikes offer a measure of neuron activity.

Besides, the low-frequency content of electrodes measurement (*low field potential*, LFP) is admitted to represent a sum of the intracellular potential variations of every neurons in a small region. The intracellular potential of a single neuron can hardly be recorded *in vivo*; this can be done however with cells cultures.

1.1.2 Macroscopic Electrophysiology

Ionic currents described above (displacement of charged particles) are the source of the *Electro- and Magneto-encephalography* (EEG and MEG) measures on the scalp. Spike

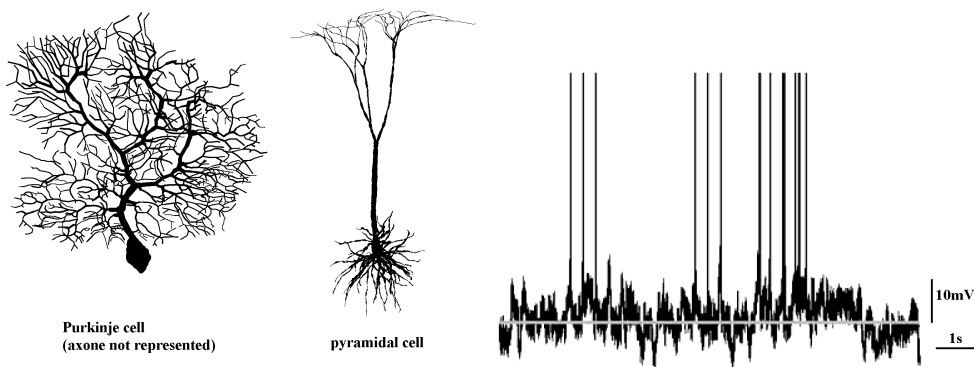


Figure 1.2: Examples of the variability of neurones shapes (left, adapted from [138]), and of the intra-cellular potential time course (right, adapted from [175]).

propagations are not detected, because they actually generate two currents in opposite directions. These ionic currents are called *primary*, and they generate *secondary* currents throughout the whole head, which close current loops. The potential differences between two EEG scalp electrodes are mainly caused by these secondary currents. On the contrary, the contribution of secondary currents to the magnetic field is small compared to that of primary currents, so MEG is usually assumed to measure intracellular primary currents.

Actually, ionic currents in single neurons cannot be detected on the scalp, unless they are synchronized and add themselves to each other in a constructive way. This is actually the case in the cerebral cortex, where assemblies of 10^4 neurons with similar functions are organized in macro-columns. This is not the case however for the central grey kernels, which have a star organization.

Today, EEG measure is done through the use of a latex helmet on which electrodes are set. Electrical contact with the scalp is ensured by the use of a conductor gel. The number of electrodes is about 20 in the common displays, but can go up to 64, 128 or even 256. Since EEG only measures potential differences between electrodes, the choice of a reference electrode is of crucial importance. Sometimes, additional electrodes are placed below the ears to serve as reference.

The MEG measure requires more costly installations: magnetic sensors are displayed in a bowl above the head. Because of the very low level of the magnetic field created by the brain, supra-conductors metals are required, and are to be maintained at cold temperatures with liquid helium.

EEG and MEG have the great advantage of being totally non-invasive, and as such can be used on humans. They also benefit from a high temporal resolution, since the measured signals can be sampled at arbitrary rates. 1KHz sampling is enough to get the fastest features like evoked responses or oscillations. On the spatial side however, cortical sources localization is a serious concern, and its precision does not go below a few centimetres.

1.1.3 Metabolism

In the last 20 years, brain imaging has known a rapid development with the use of new imaging techniques. These techniques do not measure directly neural activity, but secondary effects that take place in the metabolic chain to supply the brain in energy.

Brain energy consumption represents 20 percents of the whole body consumption. It is believed that the most energy consuming effects are the activation of the sodium pump to create the cell polarization and the repacking of neurotransmitters after synaptic communications. This energy is supplied by the consumption of glucose and oxygen.

Thus, during the activation of a region, glucose and oxygen consumption increase. At the same time, complex signalling mechanisms lead to an increased blood flow in the vasculature that drains the region. This flow response actually exceeds the metabolic demand, so that the blood in the brain vasculature is finally more oxygenated during activations than during rest periods. All the effects related to the blood (increased vessels volume, variation of blood oxygenation) are referred as the *hemodynamic response*. The temporal aspect of these effects is an important concern, since they are not immediate, can last for up to 30 seconds, and exhibit characteristic shapes.

Several imaging techniques attempt to measure variations of these quantities. Initially, the *Tomography by Emission of Positrons* (TEP) was used to measure variations of the blood volume. It had the disadvantage to require the absorption of radioactive substances that played the role of markers.

Its usage has been replaced progressively by the *Magnetic Resonance Imaging* (MRI), which produces three-dimensional images of the head, while being completely non-invasive. The physical principles of MRI are exposed section 1.3.1. Basically, the MRI signal originates in the hydrogen atoms, mostly present in water molecules. However, under particular settings for the MRI acquisition sequence ($T2^*$ ponderation), it is possible to obtain a signal modulated by the blood deoxygenation. This measure, called *functional MRI* (fMRI), offers a high spatial resolution (under the usual magnetic field strength of 1.5 or 3 Tesla, it is possible to obtain one millimetre side voxels). Besides, the temporal aspect is weak, since it requires 2 seconds or more to acquire the images of a whole brain (at most, one can go down to hundred milliseconds when acquiring a few planar slices only).

MRI scanners can also be used for *Magnetic Resonance Spectroscopy* (MRS). This procedure skims through a large spectrum of resonance frequencies, and thus recovers signals for other atoms than hydrogen. It allows measures of other metabolic quantities such as glucose, lactate... that are more closely related to neural activity than hemodynamic variables, and moreover could inform on the nature of the activity. Unfortunately, MRS acquisition times are long (several minutes at least), which makes its temporal resolution much worse than that of fMRI.

These non-invasive imaging modalities aside, there are also invasive techniques that allow recording the dynamics of various metabolic and hemodynamic quantities on animals. Among them, *Optical Imaging* consists in monitoring the brain surface with an optical camera. The brain should be illuminated with light at the appropriate wavelength; then it is possible to measure blood volume and blood oxygenation dynamics in the first layers of the cortex. It is also possible to dispose on the brain surface a *voltage-sensitive dye* (VSD), which allows recordings of the brain's electrical activity: when illuminated at a specific wavelength, this dye has fluorescence properties that depend on the electrical potential of the neurons. The recorded fluorescence signals are believed to be connected to the LFP measures. Lastly, the use of a *Laser-Doppler probe*, which measures particles dispersion velocities, allows one to record blood flow dynamics.

1.2 EEG Methods

Since EEG collects a signal by the mean of electrodes on the scalp, the localization in the brain of measured activity is a major issue. Solving this localization problem, also called *inverse problem* supposes knowing how electrical sources in the brain generate currents that propagate throughout the head and lead to the electrical potentials being measured on the scalp in EEG (or magnetic field being measured in MEG). This is called the *forward problem*, which is determined by the quasi-static approximation of Maxwell equations. Note that on account of the quasi-static approximation it is not necessary to consider time evolutions in these two problems: measures at a given time t only depend on the brain's electrical activity at the same time t . An overview and unified analysis of different forward and inverse problems can be found in (Mosher et al.[142]).

1.2.1 Physics of EEG and MEG

We consider the head to be a volume conductor with a non-uniform anisotropic conductivity tensor σ . Three equations systems have to be considered. First, the general Maxwell equations in the medium:

$$\begin{cases} \epsilon \nabla \mathbf{E} & = \rho \\ \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} & = 0 \\ \nabla \mathbf{B} & = 0 \\ \mu \nabla \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} & = \mathbf{J}. \end{cases} \quad (1.1)$$

E and B are the electric and magnetic fields, ρ and J are the charge and current densities, and ϵ and μ the medium permittivity and permeability constants ($\mu = \mu_0$, the permeability of free space).

Secondly the charges conservation equation:

$$\nabla \mathbf{J} + \frac{\partial \rho}{\partial t} = 0. \quad (1.2)$$

Thirdly, the determination of propagation currents (notated $\mathbf{J}^{(\mathbf{v})}$) in a conductor medium:

$$\mathbf{J}^{(\mathbf{v})} = \sigma \mathbf{E}. \quad (1.3)$$

Since the primary currents in the cortex vary slowly compared to propagation speeds in the Maxwell equations, the time derivative terms in (1.1) and (1.2) can be ignored, which constitutes the quasi-static approximation. In particular, the second Maxwell equation states that the electrical field curling is zero; then it can be expressed as the gradient of an *electrical potential*:

$$\mathbf{E} = -\nabla V.$$

It is convenient to split the currents in the brain \mathbf{J} into two terms: the primary currents ($\mathbf{J}^{(\mathbf{p})}$) are those of the sources located on the cortex, whereas the secondary currents ($\mathbf{J}^{(\mathbf{v})}$) are those resulting of the propagation throughout the head (1.3):

$$\mathbf{J} = \mathbf{J}^{(\mathbf{p})} + \mathbf{J}^{(\mathbf{v})} = \mathbf{J}^{(\mathbf{p})} + \sigma \mathbf{E} = \mathbf{J}^{(\mathbf{p})} - \sigma \nabla V. \quad (1.4)$$

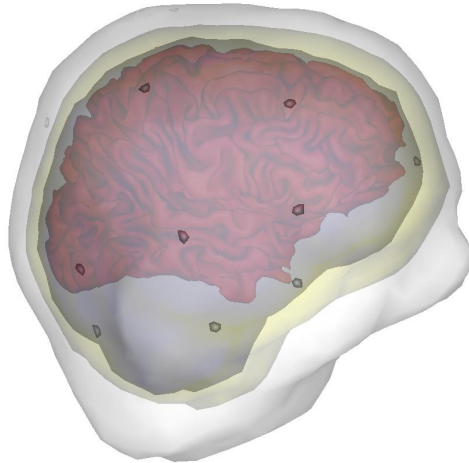


Figure 1.3: *Realistic head model*: the skin, the two surfaces of the skull and the cortex surfaces were segmented in the anatomic MRI. The spots on the skin surface indicate the positions of the EEG electrodes. The cortical electrical sources are supposed to be positioned on the brain surface and oriented orthogonally to it.

We can then re-write that the current divergence is zero (equation (1.2), where the time derivative term has been dropped according to the quasi-static assumption):

$$\nabla \mathbf{J}^{(p)} - \nabla(\sigma \nabla V) = 0. \quad (1.5)$$

This last equation relates the electrical potential (and in particular its values at the EEG electrodes) to the currents generated on the cortex.

We also give a brief insight into the MEG (Magneto-Encephalography) equations. In a similar way as for electrical field, the fourth Maxwell equation (“magnetic field divergence is zero”) allows \mathbf{B} to be written as the curl of a potential: $\mathbf{B} = \nabla \times \mathbf{A}$. Then the third Maxwell equation can be solved analytically [78], for \mathbf{A} and then for \mathbf{B} , leading to the following integral expression of magnetic field at position \mathbf{x} :

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \int \mathbf{J}(\mathbf{x}') \times \frac{\mathbf{x} - \mathbf{x}'}{\|\mathbf{x} - \mathbf{x}'\|^3} dx' \quad (1.6)$$

Once equation (1.5) has been solved for the electrical potential V , it is straightforward to obtain the magnetic field at the MEG sensor locations, by replacing the current term \mathbf{J} by its decomposition between primary and secondary currents in (1.4). Actually, the major contribution comes from the primary currents, and it would be the only one if conductivities were homogeneous.

1.2.2 Forward Problem

The forward problem consists in predicting the electric potentials on the scalp measured in EEG, and the magnetic field outside the head measured in MEG, given sources activity on the cortex. In other words, it consists in solving equations (1.5) for the potential V given the primary currents J and the conductivities σ , and then (1.6) for the magnetic field \mathbf{B} . As there is no analytical solution to the first equation, it needs to be solved numerically.

Two main methods can be distinguished: *Finite Elements Method* (FEM) consists in dividing the whole head volume in small tetrahedra with a given conductivity; equation (1.5) is applied to each tetrahedron, which produces a linear system of equations that can be solved numerically [142, 52]. This system is sparse because each tetrahedron interacts only with its neighbours.

Boundary Elements Method (BEM), in turns, takes advantage of the conductivities homogeneity of the brain compartments: the head is divided into nested regions with constant conductivities: brain, skull, scalp (figure 1.3). These regions can be segmented in anatomical MRI images. Then, only the boundaries between these compartments need to be tessellated.

Indeed, in the BEM framework, one can derive from (1.5) equations involving the boundaries potentials only. Let us denote by $\Omega_0, \Omega_1, \dots$ the different compartments (sources being all located in Ω_0), $\sigma_0, \sigma_1, \dots$ their conductivities, and S_{01}, S_{12}, \dots their boundaries. Then the potential on surface $S_{i(i+1)}$ verifies [64]:

$$(\sigma_i + \sigma_{i+1})V(\mathbf{x}) = 2\sigma_0 V_0(\mathbf{x}) + \frac{1}{2\pi} \sum_j (\sigma_{j+1} - \sigma_j) \int_{S_{j(j+1)}} V(\mathbf{x}') dS(\mathbf{x}'), \quad (1.7)$$

where $V_0(\mathbf{x})$ is the potential created by the primary currents:

$$V_0(\mathbf{x}) = \frac{1}{4\pi\sigma_0} \int_{\Omega_0} \mathbf{J}(\mathbf{p}) \frac{\mathbf{x} - \mathbf{x}'}{\|\mathbf{x} - \mathbf{x}'\|^3} dx'.$$

Equations on all boundaries form a linear system. Once it is solved [25, 1], it gives the potential values on each boundary. The magnetic field then can also be determined by a boundaries derivation of (1.6):

$$\mathbf{B}(\mathbf{x}) = \mathbf{B}_0(\mathbf{x}) + \frac{\mu_0}{4\pi} \sum_i (\sigma_{i+1} - \sigma_i) \int_{S_{i(i+1)}} V(\mathbf{x}') \frac{\mathbf{x} - \mathbf{x}'}{\|\mathbf{x} - \mathbf{x}'\|^3} dS(\mathbf{x}'), \quad (1.8)$$

where $\mathbf{B}_0(\mathbf{x})$ is the magnetic field created by the primary currents:

$$\mathbf{B}_0(\mathbf{x}) = \frac{\mu_0}{4\pi} \int_{\Omega_0} \mathbf{J}(\mathbf{p}) \times \frac{\mathbf{x} - \mathbf{x}'}{\|\mathbf{x} - \mathbf{x}'\|^3} dx'.$$

1.2.3 Inverse Problem

The inverse problem consists, conversely, in estimating the electrical activity of the sources given EEG and/or MEG measures (note that it can thus achieve, in a natural manner, “fusion” between EEG and MEG data). However, there are different sorts of inverse problems, depending on the nature of the sources we are looking for (see Hämäläinen et al., 1993 [78] and Baillet et al., 2001 [14] for review, and Grova et al.[74] for a comparison of methods). Basically, we can distinguish between *dipole models*, *beamforming* and *distributed dipoles models*.

We saw in section 1.1 that the electrical activity in the brain that can be detected in EEG/MEG takes place in the grey matter, and requires assemblies of 10^4 neurons to have a synchronized activity. In the dipole models, the sum of synchronized activities of such a neuron assembly in a small region is approximated by a dipole of current. Inverse problem consists then in localizing a dipole inside the brain, or a small set of dipoles, that best explains the measured data.

In distributed models, electric activity in the brain is modelled with a large set of small current dipoles. These dipoles are located in the grey matter, and approximate the local cortical column organization. Since the grey matter has nearly a constant small width (2,5mm), it can be approximated by a surface, and hence the dipoles are all positioned on this surface. Moreover, on account of the vertical cortical columns organization, the currents should always be orthogonal to that surface.

The beamforming methods are slightly different, since they scan each source independently and compute their activation, in order to optimize a certain criterion. In contrast to the distributed models, they do not try to fit all the observed data at once. Concerning these methods, we will let the reader refer to reviews in [167, 166].

It is possible to use a common formalization for dipole and distributed models that takes into account the fact that the forward model is linear with respect to source intensities and nonlinear with respect to source localizations and orientations. Let's denote here by $j = (j_i)_{i \leq n}$ the intensities of the sources, $l = (l_i)_{i \leq n}$ their localizations and $o = (o_i)_{i \leq n}$ their orientations. We also denote by $z(l, o, j)$ the EEG/MEG measure predicted by the forward model, which depends on the source characteristics, and z the actual measure. The forward model allows constructing a matrix G , called the *lead field* matrix, which depends (non-linearly) on source localizations and orientations, such that the predicted measure is:

$$z(l, o, j) = G(l, o)j \quad (1.9)$$

The inverse problem for dipole models consists in estimating all the characteristics \hat{l} , \hat{o} and \hat{j} of a small number of dipoles, by minimizing the prediction error:

$$(\hat{l}, \hat{o}, \hat{j}) = \text{Argmin} \|z - G(l, o)j\|.$$

Since the lead field matrix $G(l, o)$ depends nonlinearly on the positions and orientations of the sources, minimization necessitates to use iterative methods.

The inverse problem for distributed models consists in estimating the intensities of a large number of dipoles whose locations and orientations are fixed. The matrix G , which depends on these locations and orientations, is already determined. Thus we have:

$$\hat{j} = \text{Argmin}_j \|z - Gj\|. \quad (1.10)$$

But this inverse problem is under-determined: the number of unknowns in the source vector j is much larger than the number of measures in z . There are infinity of solutions such that the prediction error is zero. It is necessary then to introduce an additional constraint on the sources, by using a regularization term in the minimized energy:

$$\hat{j} = \text{Argmin}_j \|z - Gj\|^2 + R(j). \quad (1.11)$$

Please note that it is possible to use a specific a priori covariance C on the noise at the electrodes, which defines another norm $(z - Gj)^T C (z - Gj)$ instead of $\|z - Gj\|^2$. This appears naturally when using a Bayesian framework [13].

$$\hat{j} = \text{Argmin}_j (z - Gj)^T C^{-1} (z - Gj) + R(j).$$

When the regularization term is bilinear ($R(j) = j^T \Sigma^{-1} j$), it is called a *Tikhonov regularization*, and the estimation can be obtained directly:

$$\hat{j} = (G^T C^{-1} G + \Sigma^{-1})^{-1} G^T C^{-1} z.$$

The regularization term can have several different expressions:

- The simpler regularization term is the norm of j , scaled with an appropriate constant λ . It may have several drawbacks: for example, an activity deep inside a cortical sulcus will tend to be explained by more superficial and smaller norm sources.
- Using a bilinear regularization term with $\Sigma = (\text{diag}(G^T G))^{-1}$ normalizes the columns of the matrix G to account for deep sources by penalizing sources too close to the sensors [91].
- The regularization can include other *anatomical constraints*, in particular, impose a continuity of the activity over the cortex. Adde et al.[2] worked also on anisotropic smoothness energies, that assumed a spatial continuity, though permitting discontinuities between adjacent regions. Since the regularization was not bilinear, energy minimization required iterative methods.
- *Functional constraints* can be used also, by the means of an *a priori* ponderation of sources variances based on other imaging modalities like fMRI [11].

1.2.4 Real Data Processing

The signal measured in EEG actually contains a large amount of noise. This noise has many possible origins:

- There are a lot of *physiological noises*, i.e. potential fluctuations that are not due to neural activity (respiration, muscles contraction, ...); eye blinking in particular generates currents that are hundred times stronger than those related to neural activity.
- Additional *measure noises* can be induced, by electrode movements for example.
- Lastly, during a cognitive experiment, the signal related to the experimental paradigm is drowned amidst those of all other *ongoing cerebral activities*.

To recover a signal of interest from experimental data, the electrode time courses are often averaged over all repetitions of the experimental stimulus, this averaging being *triggered* by the times of stimulus presentation. The averaged electrode time courses obtained by this method are called *evoked potentials*. This allows keeping the signal component related to the stimulation and repeated identically at each presentation, while other independent signals tend to be cancelled by summation.

The evoked potentials usually show one or several specific features, such as a response peak after a constant delay, or oscillations. Then, one can apply the inverse problem at each time point, or alternatively, select only a few remarkable time instants according to these features. At every selected time, the inverse problem will estimate the cortical electric activity that accounts for electrode potentials at that time.

The evoked potentials technique drawback is that it only keeps activities induced by the stimulation that are repeated identically at each trial. For example, it will tend

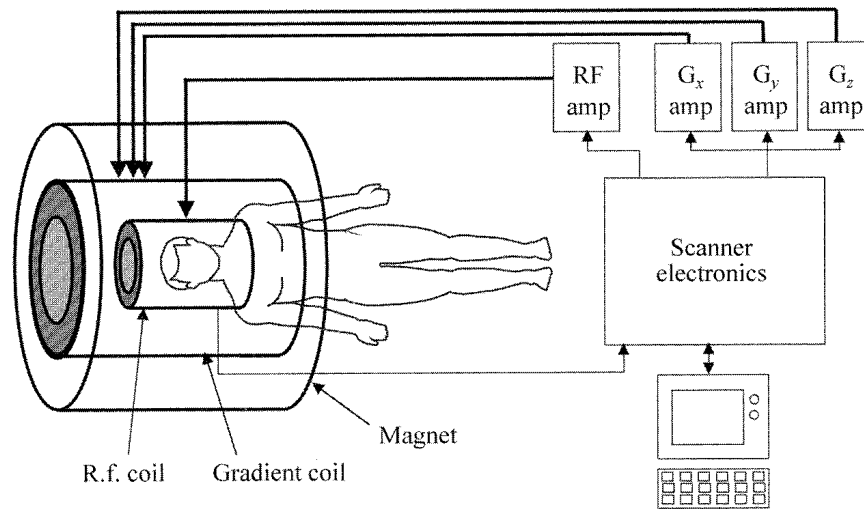


Figure 1.4: *Diagram of an MRI scanner*: the magnet produces the static magnetic field B_0 ; the gradient coil produces linear spatial variations of this static field, in order to achieve the slice selection, as well as the encoding in the Fourier space of the measured resonance signals; the RF coil produces the transient excitatory radio-frequency pulse and detects the resonance signals. Figure issued from Jezzard and Clare, 2001 [93].

to suppress activities whose latencies with respect to the stimulus vary between trials. Other methods like time-frequency analysis overcome these problems.

Before averaging the data, some pre-processing can be operated to it also: frequencies filtering can remove undesirable effects like low-frequency drifts, or artefacts due to electrical installation at the sector frequency (50Hz in France). Also, a selection can be done manually in the data, to discard trials containing large artefacts (like those where eye blinking is observed).

1.3 fMRI Methods

1.3.1 Physics of the Magnetic Resonance Imaging

Nuclear Magnetic Resonance (NMR) is the basic physical phenomenon used for Magnetic Resonance Imaging. Though this is a quantic phenomenon, it can be explained by using classical analogies. Nucleus particles (protons and neutrons) have a quantic property called *spin* that can be considered as a magnetic dipole moment (i.e. it behaves like a tiny magnet). Basically, a MRI scanner creates a strong constant magnetic field, with which the particles spins align. Image acquisition consists first in exciting the spins by a transient perturbation of the magnetic field, called the *Radio Frequency* (RF) pulse, that induces a rotative movement of the spins; then one measures the radio frequency signal produced by the rotating spins. The contrasts obtained in the resulting image originate in the differences between return to equilibrium time constants of the various tissues. A review of these phenomena and of the specific techniques for acquiring MRI images can be found in [93].

ting the spins by a transient perturbation of the magnetic field, called the *Radio Frequency* (RF) pulse, that induces a rotative movement of the spins; then one measures the radio frequency signal produced by the rotating spins. The

contrasts obtained in the resulting image originate in the differences between return to equilibrium time constants of the various tissues. A review of these phenomena and of the specific techniques for acquiring MRI images can be found in [93].

Nuclear Magnetic Resonance Principle

Let us describe this more precisely. In the absence of magnetic field, the spins of the particles do not have a preferred direction, and the average magnetic moment M in a sample of molecules is zero. But in the presence of a magnetic field B_0 , the possible energy states of the nucleus spin are quantized, i.e. only a small number of states are allowed. For the hydrogen atom, which is the one on which most MRI techniques focus upon, two states are possible, and correspond to the alignment of the spin magnetic moment parallel or anti-parallel to the external magnetic field. As the anti-parallel state has a higher energy level than the parallel state, there will be a small preference for the parallel state, so that there will be a resulting magnetic moment aligned with B_0 . Note however that this preference is very small: as an example, at 1.5 Tesla, and at physiological temperatures, only 10 spins in every 1,000,000 contribute to the net magnetic moment of the sample. Scanner currently used on humans create a 1.5 Tesla or 3 Tesla static magnetic field. However scanners with higher fields exist, which are usually used on animals (5T, 7T, and up to 12T).

The second important phenomenon is magnetic resonance. When the external magnetic field is perturbed, the spin magnetic moments will precess around their average direction at a frequency ν_0 determined by the field intensity and the *gyromagnetic ratio* of the atom, according to the *Larmor* law:

$$\nu_0 = \gamma|B_0|.$$

The *Radio Frequency pulse* (RF) technique consists in applying in addition to B_0 a small oscillating field B_1 that is transverse (i.e. orthogonal) to B_0 : if the frequency of these oscillations is equal to the Larmor frequency above, then the nucleus spins in the sample will have a coherent behaviour, resulting in a net magnetization vector rotating at the Larmor frequency. The transverse component of this vector is rotating in the (x, y) plane, and thus induces a radio signal, while the longitudinal component (in the same direction z as B_0) is stationary. The angle of inclination of the vector away from the z axis is proportional to the length of the RF pulse: a short RF pulse tips the magnetization by a few degrees and only a feeble radio signal is observed. A longer 90° pulse brings the magnetization totally in the transverse plane, and the signal observed is maximal. On the contrary, a 180° pulse, of twice the length, inverts the magnetization on the z axis and produces no component in the transverse plane.

The Free Induction Decay

After the Radio Frequency excitation is switched off, the net magnetization returns to its equilibrium state, i.e. aligns with B_0 , and the radio signal observed resulting from the transverse component decays to zero. This return to equilibrium is exponential, and is characterized by several time constants called relaxation times that are at the origin of MRI contrasts:

- The T1 relaxation time characterizes the recovery of the longitudinal magnetization due to the static magnetic field B_0 . It may vary across tissues, because it

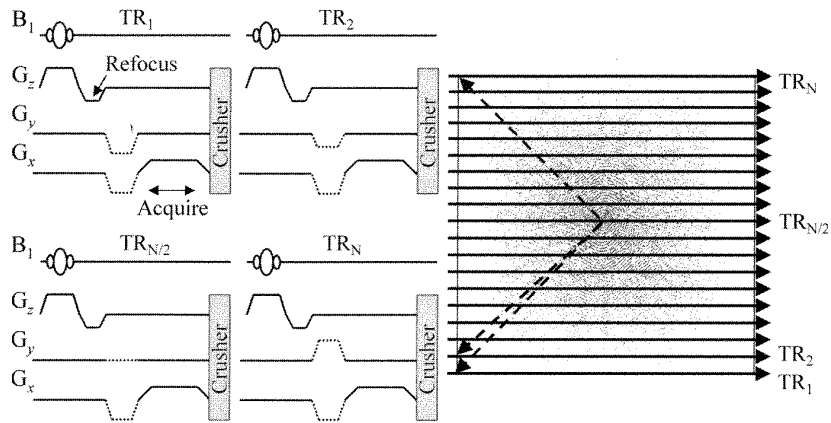


Figure 1.5: Pulse sequence diagram (left) and k-space representation (right) of the FLASH acquisition sequence. Figure issued from Jezzard and Clare, 2001 [93].

depends on thermal agitations and their frequencies (whether they are close to the Larmor frequency or not).

- The T_2 and T_2^* relaxation times characterize the decay of the transverse magnetization. They are actually shorter than the T_1 relaxation because they are not only due to the re-inclination of spin moments toward the z axis, but also to the dephasing of the rotating components in the transverse (x, y) plane, due to random fluctuations of the Larmor frequencies inside the sample. These fluctuations have two main causes: local inhomogeneities of the static magnetic field and fluctuations at a molecular level. It is actually possible to cancel the dephasing caused by magnetic field local inhomogeneities by using the *spin echo* technique. For that reason, there are two different relaxation times: T_2 is the one when using the spin echo, and thus characterizes the signal decay due to fluctuation at the molecular level only, whereas T_2^* is shorter and reflects mostly the effect of the magnetic field's local inhomogeneities. The T_2 process is sensitive to molecular motions, but depends less on their frequencies than T_1 . The T_2^* process is sensitive to features that perturb the magnetic field, like boundaries between different tissues, or agents with particular paramagnetic properties. This is the basis of the BOLD contrast used in fMRI.

Slice Selection

We now see how it is possible to measure the resonance of nuclei so as to produce three-dimensional images whose contrast will reflect the relaxation properties of the different tissues, and how to adjust the ponderation between these different relaxations (T_1, T_2, T_2^*). The key idea for that is to make the Larmor resonance frequency of atoms vary spatially by modifying the static field B_0 with the application of an additional magnetic field, called magnetic gradient, oriented in the same direction and whose intensity varies along one spatial dimension. The same coil can be used for the transmission of the gradient and the reception of the radio signal.

It is possible first to induce resonance in one slice only of the measured tissues: this is called the *slice selection* technique. For example, by applying a magnetic gradient G_z along the z direction, the static magnetic field becomes:

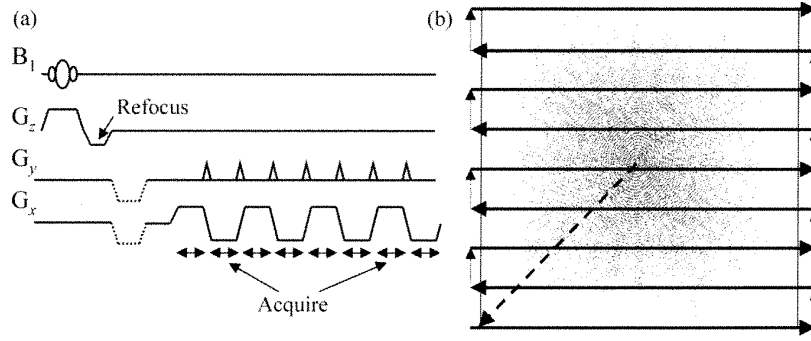


Figure 1.6: Pulse sequence diagram (left) and k-space representation (right) of the EPI acquisition sequence. Figure issued from Jezzard and Clare, 2001 [93].

$$B(z) = B_0 + G_z z,$$

and the Larmor frequencies of atoms depend of their positions:

$$\nu(z) = \gamma B_0 + \gamma G_z z.$$

Then, the Radio Frequency pulse is chosen to contain only a specific range of frequencies, to excite the atoms with specific z coordinates. Usually a sinus cardinal excitation pulse is chosen, because its Fourier transform is a car box: the frequencies it entails are uniformly distributed in the chosen range. An excitation in the frequency range $f \pm df$ will induce resonance in the slice

$$z = \frac{f - \gamma B_0}{\gamma G_z} \pm \frac{df}{\gamma G_z}.$$

The Concepts of Spatial Encoding and k-Space

It is also possible to apply magnetic gradients during the relaxation, along the other directions (x and y in our example). Then the frequencies of the magnetization vector precession vary spatially, and it becomes possible to encode the signal issued from every pixels of the slice in the Fourier space. Let us show this more precisely. We denote by $\rho(x, y)$ the contribution to signal of the point of the slice in position (x, y) . In the absence of magnetic gradient, the radio signal received by the coil is:

$$S(t) = \iint \rho(x, y) e^{2\pi i \gamma B_0 t} dx dy.$$

When adding to B_0 two gradients in the x and y directions that can vary during time $G_x(t)$ and $G_y(t)$, the static magnetic field becomes:

$$B(x, y, t) = B_0 + G_x(t)x + G_y(t)y,$$

the precession frequency at time position (x, y) and at time t :

$$\nu(x, y, t) = \gamma B_0 + \gamma G_x(t)x + \gamma G_y(t)y,$$

and the signal received by the coil:

$$S(t) = \iint \rho(x, y) e^{2\pi i(\gamma B_0 t + \gamma \int_0^t G_x(u) du + \gamma \int_0^t G_y(u) du)} dx dy.$$

For convenience, we use the notation:

$$k_x(t) = \gamma \int_0^t G_x(u) du$$

$$k_y(t) = \gamma \int_0^t G_y(u) du.$$

It appears then that the measured signals are Fourier transforms of the intensity ρ , modulated by the ν_0 frequency:

$$S(t) = \iint \rho(x, y) e^{2\pi i(k_x(t)x + k_y(t)y)} dx dy e^{2\pi i\nu_0 t}.$$

Slice Acquisition

A slice acquisition consists then in demodulating the signal by the ν_0 frequency, and measuring signals for values of k_x and k_y spanning the so-called *k-space*, i.e. the space of all frequencies of the 2D Fourier transform of the ρ intensity. Then the image ρ is obtained by the inverse Fourier transform of measures in this k-space.

Figures 1.5 and 1.6 show two standard scanning sequences: the gradient echo sequences FLASH and EPI. During a FLASH sequence, one line in the k-space is acquired after each RF pulse:

- The sequence begins with the slice selection described above, where a gradient is applied in the z direction during the pulse.
- There is also a slice refocus negative gradient, called *gradient echo*, which ensures that the spins at the top and at the bottom of the slice have the same phase (since the initial gradient induced dephasing).
- Immediately following this slice selection, the k-space coordinate of the signal for the slice is $k_x = k_y = 0$.
- Then the selection of the line in the k-space is achieved by applying a gradient in the y direction, which is called *phase encoding*.
- At the same time a negative gradient in the x direction is applied to move toward the origin of the k_x coordinates.
- During the acquisition, a gradient is applied in the x direction so that a whole line is acquired for all k_x coordinates: this is called the *frequency encoding*.

During an EPI (Echo Planar Imaging) sequence, the whole k-space is sampled, which means one entire slice is acquired, after each RF pulse. It is of particular interest for fast fMRI acquisitions (see below). To acquire a whole volume using EPI sequences, two main choices are possible concerning the slice acquisition order: *sequential* (in the natural increasing or decreasing order) or *interleaved* (first the odd numbers and then the even numbers). The interleaved method intends to avoid artefacts due to imprecise

slice selection: when a slice is selected with the gradient method, it is possible that some atoms in the neighbouring slices are excited also, because of local field distortions. Then it leads to artefacts if one of these slices is acquired just after. However the interleaved method also has drawbacks, in particular for the motion correction pre-processing in fMRI described below.

fMRI Signal

In functional MRI, the signal of interest relies in the T2* contrast. Indeed, the oxyhemoglobin is diamagnetic and has little influence on the local magnetic field, whereas the deoxyhemoglobin (i.e. hemoglobin that delivered its oxygen molecule to the brain) is paramagnetic, and distorts the magnetic field in its neighbourhood, thus shortening the T2* relaxation time and decreasing the MR signal. There is in permanence a substantial amount of deoxyhemoglobin in the venous vessels, yielding an attenuation of the MR signal. When the brain is activated, as discovered by the pioneering works of Ogawa and his colleagues [148], the local blood flow increases more than the oxygen metabolism, which results in a more oxygenated blood. This reduction of deoxyhemoglobin concentration leads to a signal increase (a few percents at 1.5 Tesla, and 5~10% at 3 Tesla).

Typically, EPI acquisition sequences are used for fMRI, because they have the advantage of being fast: one slice can be acquired in 30 to 100ms, and a whole brain in 2 to 4s. The drawback is that these image cannot have a high spatial resolution: usual fMRI slices have 64 by 64 pixels.

1.3.2 Pre-processing

fMRI data consists in a set of functional 3D images acquired with a given repetition time (RT). Before analyzing these images a number of pre-processing is required. The two most drastic effects that need to be corrected are the motion of the head inside the scanner during acquisition and the different acquisition times of the slices in a same volume. Unfortunately, it is very difficult to correctly take into account these two effects simultaneously.

The registration of head movements consists of two steps:

- *Motion estimation*: this motion is often supposed rigid (which is actually an approximation since intrinsic artefacts of the EPI sequence introduce non-rigid distortion in the images). Giving an estimate of it consists then in the estimation of six parameters (three translation parameters and three rotation parameters). This is usually achieved by minimizing the square difference of successive images with a reference image. A more robust method was proposed in [56], who reduced the impact of regions with large intensity differences by the use of a non-quadratic and slow-increasing energy function.
- *Motion correction*: it consists in re-interpolating the images by using the estimated motion parameters. However, Freire et al. [55] recommended not to correct motions if they are smaller than voxel sizes. Indeed motion correction may introduce undesired changes in the data.

Each slice in an fMRI volume is acquired at a different time. However it is convenient for the analysis to suppose that each volume corresponds to a measure at a given instant.

For that reason, data pre-processing often includes a temporal re-interpolation of each voxel time courses, also called *slice timing*. The interpolation proposed by the SPM software consists in applying a dephasing to the signals by passing by the Fourier space. As it does not bring new information in the data but on the contrary may delete some, an alternative simply consists in including the slice acquisition delays in the data analysis.

There is an unavoidable problem that arises when the subject head moved inside the scanner during acquisition: What should be applied first ? Should it be slice timing ? But in that case the temporal interpolation is realized with measures belonging to different locations in the brain. Should it be motion correction ? But in that case, the spatial interpolation will mix voxels from different slices and the subsequent slice timing will prove improper, more especially if slices are acquired in interleaved order. The best choice is probably the first when slices were acquired in interleaved order, and the second when they were acquired in sequential order.

Additional pre-processing may be desired, depending on the subsequent analysis:

- *Normalization*: it is often necessary to register the data issued from single subjects between each other and onto some reference coordinates, in order to perform *group analysis*, i.e. ask cognitive questions over a population, investigate intra-subject variability, or eventually compare results with standard knowledge on the brain functions based on anatomy. This normalization consists of a non-rigid transformation based on the correlation between the anatomical image of the subjects and a template image. However, the accuracy of such registration in terms of anatomical, and *a fortiori* of functional correspondence can be contestable.
- *Spatial smoothing*: it is sometimes interesting to smooth the data, to increase the Signal to Noise Ratio of the voxels time courses, on condition that the patterns of activity encompass a cluster of voxels. It relies on the fact that the summation cancels the uncorrelated noises in different voxels, while it consolidates activity patterns. However it will reduce the spatial resolution of fMRI and tend to make activity patterns with a small spatial extent disappear. Also, it proves useful when smoothing the data to include anatomical information rather than using an isotropic filter. Andrade et al.[7] proposed to use an anisotropic smoothing parallel to the brain surface, to avoid mixing signals from different tissues. Other anatomically-informed method have been proposed, like parcellation [53] and anatomical basis functions [100].
- *Removing global effects* consists in subtracting its average value to each image. It is intended to remove some nuisance effects occurring in the whole image. But studies showed that it may also remove activation patterns if the latter have a large spatial extent [46].
- *Temporal filtering*: whereas low-pass filtering can be employed to improve Signal to Noise Ratio, but reduce the temporal resolution of the signals, high-pass filtering is used to remove undesired slow-varying trends from the signals. These nuisance effects can have several origins: scanner varying parameters like the warming of the gradient coils during an experiment, biological rhythms like respiration [92], metabolic processes like the decline of oxygen reserve in a region [9]. It can be performed by removing low-frequency compounds in a basis of functions, by using the Fourier transform, or by applying recursive filters. Since some part of the

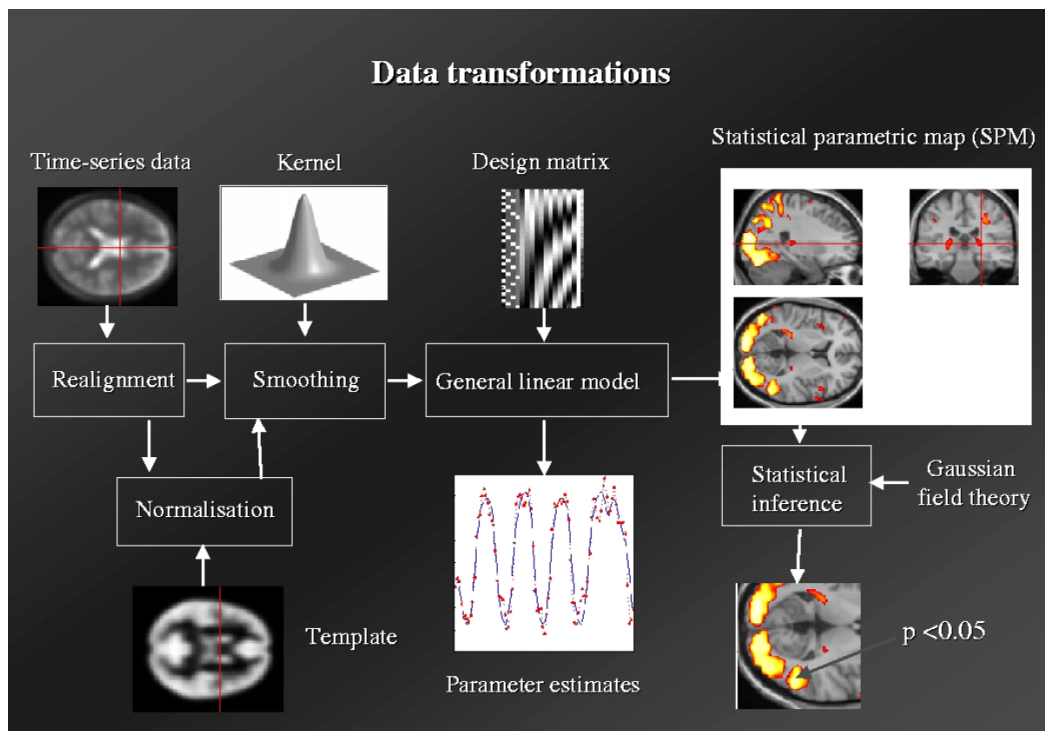


Figure 1.7: Typical data analysis performed with the SPM software. Figure and caption issued from the PhD thesis of Bertrand Thirion [162].

activation may have been affected by this filtering, it often proves important to take it into account during the analysis.

1.3.3 Hypothesis-Driven Analysis Methods

In this section we describe the standard analysis in the General Linear Model (GLM) framework, as it is implemented in the Statistical Parametric Mapping (SPM) software. A complete description can be found in [181].

Let us recall that an fMRI acquisition consists of a four-dimensional data set with images acquired with a given repetition time TR. Actually, we are particularly interested in the variations of the MR signal across time, which reflect - partly at least - the variations of the blood oxygenation, hence the name *Blood Oxygenation Level Dependent* (BOLD) signal. On the contrary, the absolute values of the measures and their variation inside the images are less interesting, since they also reflect other causes of the local field distortions responsible for the T2* contrast such as tissues properties, which bear no relation with the brain activity. It is important to keep in mind which effects are responsible for these temporal fluctuation of the signals:

- The BOLD signal variations reflect the sum of all brain activities that induce changes in the metabolism of a region as well as, through mechanism that are still misunderstood, in the blood flow that irrigates the region; please note that some changes in the brain activity may also induce zero or negative hemodynamic changes.

- There are several other physiological causes for hemodynamic fluctuations, such as heart beat and respiration.
- Artefacts may appear because of the subject motion, if it has not been corrected properly during the pre-processing step: the measured time courses in a given voxel then comes from different brain regions, and the signal can change drastically because of the different tissue properties.
- The scanner measure also introduces artefact fluctuations: blank measure noises and slow-varying drifts due to the heating of the gradients.

This list is not exhaustive, but it reminds us that the neural activity we want to study, and especially the one that is in relation with the experiment, only represents a small part of the measured signals. Under this condition, there are two main approaches to the analysis of fMRI data:

- The *exploratory methods* intend - as the name indicates - to explore the different signals present in the data and decompose them in order to isolate the effects of different physiological or measure nuisances, and time courses possibly caused by a neural activity; generally, they are *multivariate*, which means they are applied to every voxels together, to take advantage on the repetition of same signals across them; they may be useful for a preliminary analysis of experimental data, to sharpen the intuitions of experimenters on the nature of the cerebral activity.
- On the other hand, the *hypothesis-driven methods* are dedicated to answer specific questions on the cortical activation; generally, they consist in validating or rejecting a model of the response to the experimental conditions; these models can be often assessed voxel-wise and the question becomes a localization question: where does the stimulation elicit an activation? which cortical area is implicated in a particular cognitive task ? such methods are then called *univariate*.

We shall focus here on hypothesis-driven methods, which are the basis of most fMRI studies. We shall briefly present their different steps: first, one builds a model of the fMRI signals based on the experimental conditions, the hypothesis to test, and a modelling of the hemodynamic processes; secondly, one fits this model to the measured signals voxel-wise; and thirdly, one assesses with statistical tests the significance of the different parts of the model on account of the accuracy of the fit and the quantity of noise in the dataset.

These steps are greatly simplified when one considers that the BOLD effect consecutive to neural activity can be obtained with a linear convolution, as we shall explain below. This constitutes the General Linear Model (GLM) framework. As a consequence, the effects of different activities are supposed to add themselves up with no particular interactions.

Experimental Design

Since fMRI analysis relies on the temporal variations of the signals to infer on the underlying activity rather than on absolute values, it is important that the experiment include control conditions along with the experimental conditions, the only difference between the two conditions being the presence or not of the particular cognitive task that

is studied. Historically, experiment consisted of alternating periods (or *epochs*) where the subject endured different conditions, for example experiencing visual stimulations, performing a specific motor action, or - which is the simplest control condition - just resting inside the scanner.

Other types of experimental designs have appeared since. As an alternative to epochs, instantaneous activities can also be used, which are then called *events* (e.g. very fast stimulation or action, or transition between two conditions). It is also possible to make a parameter vary in a stimulation, like for example the intensity of a contrast in a visual stimulation; such experiments are called *parametric*. The activity does not even need to rely on external stimulations. This is the case with acquisitions on the fMRI correlates of EEG rhythms [68], or with studies on epileptic patients [69]. Then the fMRI signals will be predicted through the use of the measurements from other brain acquisition modalities, the EEG in particular.

In all cases, what we call the experimental design is the set of condition sequences. These sequences are functions of time; they will be box-car functions in the case of epoch conditions, sets of Dirac impulses in the case of event conditions, or real-valued functions in the case of parametric paradigms or other modalities measurement.

Model Construction

The principle of hypothesis-driven methods is to predict the shape the BOLD signal should have in response to each experimental condition. Then, thanks to the linear approximation, the signal in each voxel should be a pondered sum of the responses to the different conditions, plus noise.

Let us denote by u_1, u_2, \dots the different condition sequences. It is supposed that the response to the condition u_i should be proportional to the convolution product $h * u_i$, where h is an empirical Hemodynamic Response Function (HRF), which accounts for the different metabolic, hemodynamic and physical effects involved in the BOLD response. This HRF h is chosen by experience, its estimation being the subject of numerous studies. It is often modelled as the difference of two gamma functions [112] parameterized to fit the general empirical shape, with a first positive response that reaches its maximum at about four seconds, followed by a smaller undershoot that lasts for about twenty seconds (see chapter 2). It is possible however to add some flexibility in the model by allowing this response function to be itself a pondered sum of basis functions h_1, h_2, \dots : the response to condition u_i is then $(\sum_k \alpha_{ij} h_j) * u_i$.

As a result, and according to the assumption that the BOLD signal responds linearly to neural activity, the signal y in a given voxel is modelled as the sum of the responses to the different conditions, plus a constant uninteresting value, plus noise:

$$y(t) = \sum_{i,j} \alpha_{ij} (h_j * u_i)(t) + \alpha_0 + \eta(t).$$

η is then the sum of all the other signal components, including physiological and physical noises. It is generally assumed to be Gaussian, and may be coloured. Since among these nuisance effects there are often low-frequency drifts, it is convenient to add to the model a set of low-frequency basis functions that will capture them. We denote by c_1, \dots, c_m these basis functions (usually polynomial or sinus basis functions are used), and $\alpha'_1, \dots, \alpha'_m$ their ponderations in the signal, which leads to

$$y(t) = \sum_{i,j} \alpha_{ij} (h_j * u_i)(t) + \sum_k \alpha'_k c_k(t) + \eta(t).$$

Let us note that if there is a constant component among the c_k basis functions, the constant α_0 is not needed any more.

This last model can be displayed in a matrix form: we denote by X the matrix whose column are all the $h_j * u_i$ and all the c_k time courses, and β the vector of all the α_{ik} and all α'_k . We get

$$\boxed{y = X\beta + \eta.} \quad (1.12)$$

Linear Regression

The linear regression consists then in estimating the model parameters entailed in β given the measured signal y . We suppose here that η is a white Gaussian noise process.

$$\eta \sim \mathcal{N}(0, \sigma^2 I)$$

If it is not the case (i.e. $\eta \sim \mathcal{N}(0, \Sigma)$ with a non-trivial variance matrix Σ), then it is possible to whiten the signal y by multiplying the left and right parts of the equation by the “inverse root square” of Σ (i.e. a matrix $\Sigma^{-1/2}$ such that $\Sigma^{-1/2}\Sigma^{-1/2} = \Sigma^{-1}$):

$$\Sigma^{-1/2}y = (\Sigma^{-1/2}X)\beta + \Sigma^{-1/2}\eta.$$

We get an equivalent equation, where the new noise $\Sigma^{-1/2}\eta$ is white.

Under the Gaussian and whiteness assumption for η , the Maximum Likelihood estimator of β is given by the pseudo-inverse matrix multiplication

$$\boxed{\hat{\beta} = (X^T X)^{-1} X^T y.}$$

It is a non-biased estimator of β (i.e. its mean is β), and its variance is

$$E(\hat{\beta}\hat{\beta}^T) = (X^T X)^{-1} X^T \sigma^2 X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}.$$

Please note that we will also need to estimate the noise variance for deriving the statistical tests below. A non-biased estimator is

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})^T (y - X\hat{\beta})}{n - p},$$

where n is the number of sampling points of the measure y , p is the rank of the matrix X , and $\nu = n - p$ is the number of degrees of freedom associated with the linear regression. Please note that $y - X\hat{\beta}$ is the orthogonal projection of the noise η in a subspace of \mathbb{R}^n whose dimension is ν . Hence $\hat{\sigma}^2$, which is its square norm normalize by ν , follows a χ^2 law with ν degrees of freedom.

Hypothesis Testing

The General Linear Model framework naturally provides two useful statistical tests that can be used to assess hypothesis on the Model. They have different roles. The Fisher test on the one hand serves to test a subspace of the model, i.e. test whether some part of the model is significantly present in the measures. In particular, testing the whole model is equivalent to asking whether there is any activity in the considered voxel that is related to the experimental paradigm. The Student test on the other hand is used to test a *contrast*, i.e. a linear combination of the parameters, and to test its positivity. For example, one can use a contrast on one parameter only, to test whether it is significantly positive; or it can be used to test whether one parameter is significantly larger than another.

Let us begin with the Student test. It is defined by a contrast γ , the question asked being: is $\gamma^T \beta$ significantly positive? The tests consist in supposing a *null hypothesis* \mathcal{H}_0 , which specifies that $\gamma^T \beta$ is actually zero, and compute the probability under this hypothesis of the value $\gamma^T \hat{\beta}$ obtained after parameter estimation. More accurately, the Student test examines the probability of this product normalized by its estimated variance:

$$T = \frac{\gamma^T \hat{\beta}}{\sqrt{\gamma^T \hat{\sigma}^2 (X^T X)^{-1} \gamma}} \sim \mathcal{T}(\nu).$$

Under hypothesis \mathcal{H}_0 , the *statistic* T follows a Student law with ν degrees of freedom. The test hence consists in calculating the *p-value* of the obtained value for T , i.e. the probability that a random variable with a distribution $\mathcal{T}(\nu)$ exceeds this value. If this probability is below a certain threshold, the hypothesis \mathcal{H}_0 is invalidated, and the test is positive.

We now present the Fisher test. Basically, it is used to test a null hypothesis \mathcal{H}_0 , which specifies that only a subpart of the model is sufficient to describe the measured time courses y :

$$y = X_1 \beta_1 + \eta_1,$$

where X_1 is a matrix with $q < p$ columns and whose image is a sub-space of the image of X . The signal can then be decomposed in three orthogonal terms: the part explained by the two models, the part explained by the original model but not by the new simpler one, and the part that remains unexplained by the two models:

$$y = X_1 \hat{\beta}_1 + (X \hat{\beta} - X_1 \hat{\beta}_1) + (y - X \hat{\beta}).$$

Under the hypothesis \mathcal{H}_0 , the terms $(X \hat{\beta} - X_1 \hat{\beta}_1)$ and $(y - X \hat{\beta})$ are two independent Gaussian distributions, and the ratio of their square norms, corrected by their dimensions follows a Fisher law with q and $n - p$ degrees of freedom:

$$F = \frac{n - p}{q} \frac{(X \hat{\beta} - X_1 \hat{\beta}_1)^T (X \hat{\beta} - X_1 \hat{\beta}_1)}{(y - X \hat{\beta})^T (y - X \hat{\beta})} \sim \mathcal{F}(q, n - p).$$

To test whether the hypothesis \mathcal{H}_0 is acceptable, it is necessary to calculate the *p-value* of the F statistic under the $\mathcal{F}(q, n - p)$ distribution, i.e. the probability that a variable that follows this distribution exceed the obtained value of F . If it is less than a

threshold probability, the hypothesis is rejected. Please note that different expressions exist for this F-test.

In the SPM software, the F-test is defined by a set of linear combinations of the model regressors. For example, let us consider q such linear combinations $\gamma_1, \gamma_2, \dots, \gamma_q$. We denote by γ the $p \times q$ matrix formed by these vectors. Then the square matrix $\Gamma = \gamma(\gamma^T \gamma)^{-1} \gamma^T$ defines an orthogonal projector, such that $X(I - \Gamma)$ is the projection of X on a sub-model that ignores the additional modelling in $X\gamma$. $X(I - \Gamma)$ defines the null hypothesis \mathcal{H}_0 : it is a square matrix whose dimension is p but whose rank is q , and which plays the same role as X_1 above.

In the two statistical tests we have described, activation is determined voxel-wise, and the threshold value of the test has been chosen so that the probability of false positive is small (for example < 0.05). However, when considering all the voxels as a whole, the same statistical threshold would lead to a much higher probability of false positives. Indeed, the probability, under the hypothesis there is actually no activation in the whole cortex, that at least one voxel is declared activated will be much higher. The statistical thresholds need then to be adjusted (lowered). However, this adjustment should take into consideration the spatial smoothness in the data. Worsley et al.[183, 182] calculated the accurate adjustment through the use of the *random field theory* [3].

modelling of Nonlinear Interactions

Since nonlinear effects prove to happen in fMRI [24, 41, 20, 67, 139, 173], especially when experimental conditions are presented closely, it is necessary to take them into account in fMRI analysis. As an alternative to using non-linear physiological models (that we will present later), some authors proposed to build regressors in the GLM framework that account for nonlinear interactions.

Friston et al.[60] proposed to keep only bilinear interaction terms, according to a Volterra expansion of the BOLD response truncated to the second order. The method they propose to estimate these bilinear interactions is similar to that for estimating the linear HRF response: the second order Volterra kernel is supposed to be possibly decomposed in a small number of basis functions, which are convolved bi-linearly with the condition sequences, in exactly the same way as the HRF can be expressed over a small number of basis functions that are convolved with the condition sequence. The resulting time courses serve as additional regressors for the fMRI analysis.

Since these Volterra kernels do not have much physiological meaning, Wager et al.[176] proposed to model nonlinear effects in SPM studies by adding a small set of specific regressors that have a less general form than Volterra kernels, but are designed according to usually observed nonlinearities in data. They estimated which nonlinear effects were generally present in the signals, and the resulting regressors they obtained included amplitude attenuation or delay effects.

One of our major work has been to show that it is possible to use nonlinear models in fMRI analysis rather than additional linear regressors which are supposed to account for the possible nonlinear effects.

1.4 Optical Imaging Methods

Optical imaging entails new imaging techniques that allow us to visualize the functioning brain at unprecedented resolutions. Specifically, there are two techniques; the first is

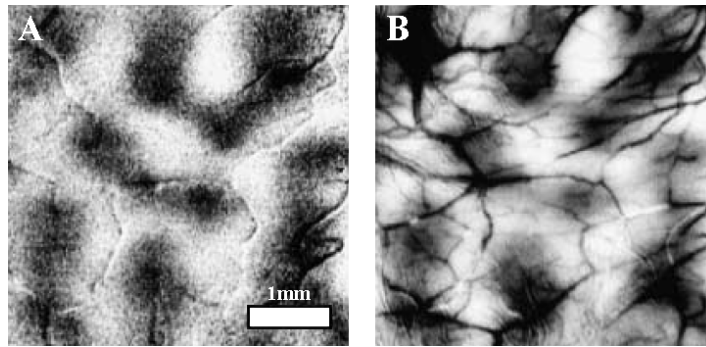


Figure 1.8: A comparison of ocular dominance maps derived from oximetry and blood volume changes: the contrast between the signal changes when either the contralateral or the ipsilateral eye is open makes the ocular dominance cortical columns become apparent. Left: contrast obtained with oxymetric changes (recorded at 605nm); right: with blood volume changes (570nm). Figure modified from Vanzetta et al., 2001 [172].

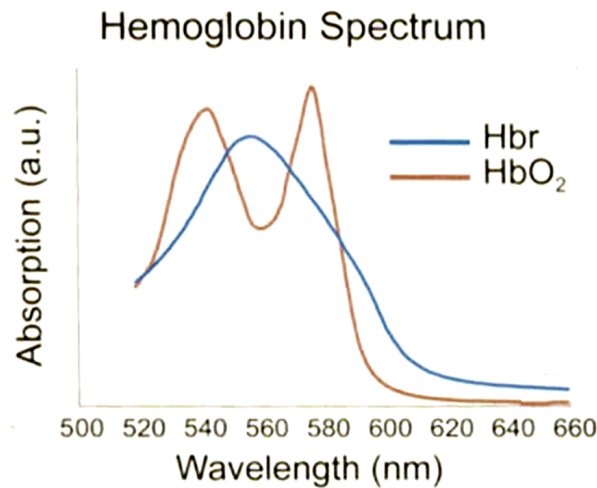


Figure 1.9: Absorption spectrums of oxygenated (HbO₂) and deoxygenated (Hbr) hemoglobin.

based on *intrinsic signals*, and the second is based on *voltage-sensitive dyes*. These two techniques consist in video monitoring the brain surface, and produce observations at the scale of assemblies of neurons. Indeed, the spatial resolution mainly depends on the optical properties of the magnifying lens and of the video camera, and can be superior to 50 microns. The recording in a single pixel is thus composed of activity contributions from 250-500 neurons. In particular, optical imaging allows the visualization of the functional organization of the cortical columns in a given cortical region. Understanding the functional processing performed by a given ensemble of neurons is a necessary step before being able to discover the principles underlying the neural code. General reviews on the optical imaging methods can be found in [70, 72]. Figure 1.8 show an example result with optical imaging, which enables the visualization of the cortical columns organization of the V1 cortex.

1.4.1 Physics of the Optical Imaging

The activity-dependent intrinsic signals originate in the light absorption properties of the various substances in the brain tissues. In particular, the fluctuations of hemodynamic quantities such as hemoglobin induce changes in the reflected light, which was discovered by the pioneering experiments of Kelin [96] and Millikan [140] (see [38] for a detailed review on these light absorption and scattering effects). These intrinsic signals can be detected by optical monitoring [71].

These light absorption changes depend on the wavelength at which the brain surface is illuminated, according to the respective absorption spectra of oxygenated and deoxygenated hemoglobin (figure 1.9). The basic intrinsic imaging then consists in selecting which specific wavelength should be used according to the variations of hemodynamic quantity measures. In order to record the total amount of hemoglobin (blood volume), one has to choose a so-called *isosbestic* wavelength, i.e. at which oxy- and deoxyhemoglobin have the same absorption ratio (green light, $\lambda=570\text{nm}$). Conversely, to record oxygenation changes, it is necessary to use the wavelength at the peak of the difference spectra between the two quantities (orange light, $\lambda=600\text{nm}$).

A power-law relates the intensity of reflected light to the concentration of the *chromophore* (i.e. the substance which absorbs the light at the used wavelength), known as the Beer-Lambert law. The detected intensity I depends of the illumination intensity I_0 , the concentration of chromophore c and its absorption coefficient γ , and the length of the path travelled by the photons through the tissues L .

$$I = I_0 e^{-\gamma L c}. \quad (1.13)$$

As a result, the relative variations of the chromophore concentration can be determined only up to an unknown constant:

$$\frac{\Delta I}{I} = -\gamma L \Delta c.$$

On the other hand, it is possible to illuminate the brain surface with white light, and then use *optical imaging spectroscopy* [130] to decompose the reflected light according to frequencies, through the use of a spectrograph positioned between the cortex and the camera (the camera then records “spatiospectral” 2D images, in which one dimension stands for the wavelengths, and only one dimension is left for space, which implies it is only possible to monitor a one-dimensional slice of the image at the same time). The analysis of spectroscopy data consists in estimating the absorption spectrum at each time point, and then fit it to the theoretical absorption spectra of oxy- and deoxyhemoglobin, in order to estimate both quantities. This spectrum estimation is actually quite difficult and requires specific calculations, since it must take into account the non-linear variations with wavelength of the optical path length [134, 137, 106]. However, provided this estimation is achieved accurately, then it is possible to bypass the problem of the unknown scaling constant mentioned above.

Intrinsic signals, since they are related to slow hemodynamic processes, do not enable the exploration of cortical dynamics and the time visualization of neuronal activity. However, it is possible to use fast extrinsic probes: the cortex preparation can be stained with a suitable voltage-sensitive dye (VSD) [39, 161]. The dye molecules bind to the external surface of neurons and act as molecular transducers that transforming changes in membrane potential into optical signals. Optical imaging with voltage-sensitive dyes

permits the visualization of cortical activity with a time resolution below the millisecond and a spatial resolution of 50-100 microns [89, 83]. It is important to note that this VSD technique measures a sum of the intra-cellular potentials rather than their spiking activity. Hence, it provides an information on neuronal processing, including sub-threshold activities and focused in particular on the synaptic potentials since the dye preferentially binds to the dendritic terminations of neurons. As a consequence, the VSD signals should be closer to the electrophysiological measurements of low-field potential (LFP) than to those of multi-unit activity (MUA). Nevertheless, VSD should have a better spatial resolution than LFP, since the dye molecules are directly bound to the cell membranes, whereas the LFP is an extra-cellular measure and thus represents an average of cell potentials over a larger region. A study by Jin et al.[94] has shown a good concordance between the two measurements, highlighting however differences in the waveforms during different activities, which suggests that they are sensitive to different aspects of the synchronization across the population.

Lastly, since the preparation of animals for optical imaging experiments is very similar to the conventional preparation for *in vivo* electrophysiological experiments, it is possible to acquire to insert an electrode and acquire electrical measures simultaneously with optical imaging recordings. It is also possible to insert a laser-Doppler probe, which allows estimating microcirculatory blood flow in a small tissue sample, through a measure of molecules dispersion based on the Doppler effect [49]. Such a probe does not penetrate inside the tissue, but it is approached to the cortical surface at a distance of about one millimetre.

1.4.2 Acquisition Methods

Animal Preparation

Optical imaging can be used on small animals such as rats, cats or monkeys. The animal needs to endure a surgical operation and a trepanation of the skull: a small hole is drilled in the skull. The dura mater, which is a natural protective membrane between the cortex and the skull, is replaced by a transparent silicone sheet to guarantee optical access while properly protecting the brain against infections and mechanical traumas. A chamber is sealed onto the bone using dental cement (figure 1.10). This chamber is closed with a transparent Plexiglas cover and filled with a transparent gel (agar). If it is filled perfectly, i.e. without any air bubbles or cerebrospinal fluid droplets, this arrangement provides an ideal optical interface and, at the same time, stabilizes the brain. Opening and re-closing the chamber allows experimenters to perform routine cleaning and eventual treatment of the implant. For a detailed description of all surgical and maintenance procedures see [8]. The European and NIH guidelines stipulate the way animals have to be taken care of, and imposes limitations on treatments which might injure them.

It is possible to work with anesthetized or awake animals. It is preferable in general to use awake animals since their behaviour is closer to real conditions. However, anesthetized animals permit easier experimentation procedures since they are unconscious and thus do not move. In particular, it not possible with awake animals to acquire signals for long times (more than 7~8 seconds) during visual stimulation where they are asked to fixate a target point. Let us note that, to maintain the animal motivation to keep the fixation, it is usual to reward him after each successful trial, with fruit juice for example.

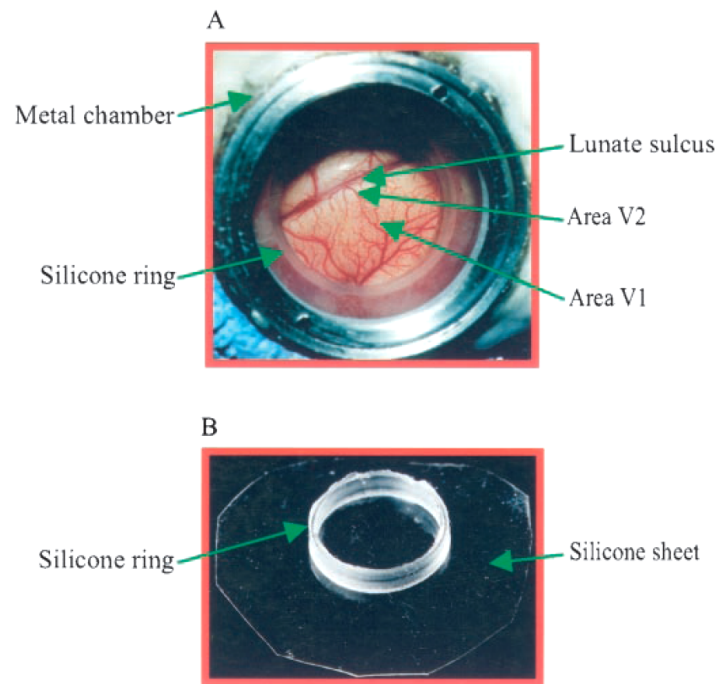


Figure 1.10: Cranial window and artificial dura mater (figure and caption issued from Shtoyerman et al., 2000 [158]). (A) The exposed cortex as seen through the transparent silicon membrane. The lunate sulcus is in the top left part of the cranial window. V1 and V2 are thus available for optical imaging. This picture was taken 5 months after insertion of the artificial dura over the exposed cortex. (B) Enlargement of the artificial silicon dura in the cranial window shown in (A). The silicon ring in the center stabilizes the artificial dura and prevents the real dura from growing on the imaging area.

Experimental Setup

The figure 1.11 shows the experimental setup generally used for optical imaging experiments. The animal head is held rigidly in a stereotaxic frame in the case of the anesthetized animal or a head holder for the awake animal. It is recommended to use a vibration isolation table to attenuate vibrational noises. A computer controls the stimulation (usually, visual presentations are projected on a screen), the cortex illumination and the video camera acquisitions.

To attenuate the signal noises due to physiological rhythms (heart beat and respiration especially), it is useful to synchronize the stimulations and acquisitions with them. Heart-beat synchronization is achieved by making each new trial begin at the time of a heart pulsation, detected with any pulse-monitoring device.

Optical Recording and Illumination

The optical setup consists of a macroscope, which can generally be obtained by using two lenses separated with the appropriate distance, and a video camera. Digital CCD cameras have the advantage over traditional video cameras to have a better signal-to-noise ratio (close to 1000:1) at low light levels. This is quite important since the signal variations can be very small compared to baseline values. And the light level is a major limiting factor, and when one needs high resolutions (for VSD acquisitions for instance), a compromise

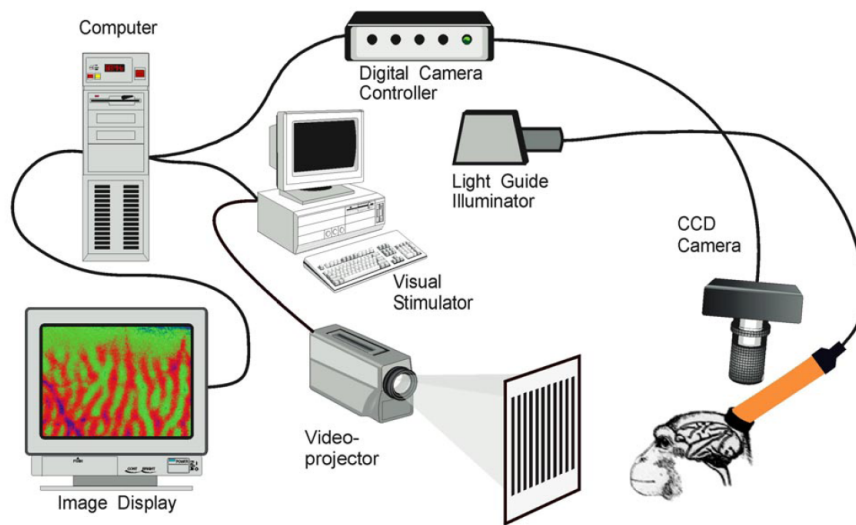


Figure 1.11: The experimental setup. Figure adapted from Ts'o et al., 1990 [164].

needs to be found between frame acquisition rate, number of pixels, magnification and signal-to-noise ratio.

Only the cortex should be illuminated, which can be achieved by the use of light guides. The light must be stable across time (fluctuation should be less than 1:1000), which requires a high-quality regulated power supply. The wavelength selection is achieved by the use of interference filters. When experimenting with high light intensities, it is desirable to control the illumination so that the cortex is illuminated only during data acquisitions. This can be achieved through using electromechanical shutter.

The details specific to the voltage-sensitive dye technique (preparation and data acquisition) can be found in Slovin and Grinvald, 2002 [160].

1.4.3 Data Analysis

A dataset from optical imaging measurements consist in a set of tri-dimensional images (space \times space \times time, and as many images as trials). Adequate averaging is required to isolate the signals which are related to the experimental conditions.

Trial Averaging

The trials corresponding to a specific condition are averaged, to keep the part of recorded signals that is evoked by this condition, while the other components should tend to cancel by summations.

However, the obtained time courses might entail features that are not specific to the stimulation, but rather to the acquisition procedure (e.g. heart beat response when the acquisition is synchronized with the heart pulsation, or response to the presentation of the fixation point) For that reason, it is necessary to use and estimate the response to a blank condition, i.e. to the acquisition procedure alone, with no particular stimulation. The spatio-temporal response specific to a given stimulation is obtained by dividing the average response to that condition by the average response to the blank condition (division is used instead of subtraction on account of the Beer-Lambert power law

(1.13) which relates the measured signals to actual hemodynamic changes). Since the acquisition conditions can change across experimental sessions, it is necessary to realize this correction with the blank condition independently for each session.

It is also possible to divide the responses to two different conditions in order to stress the contrast between the responses to these two conditions (in the same way as contrasts are formed in fMRI analysis). As an example, the figure 1.8 displays the results of an ocular dominance experiment [172]. Here, the two conditions consists of visual stimulations presented to only one eye. The responses for each eye have been averaged over trials as explained above, and then the average response to the contralateral stimulation (ipsilateral eye was shut) has been divided by the average response to the ipsilateral stimulation (contralateral eye was shut). The displayed maps show a temporal averaging of the responses between 1 and 9s, and make ocular dominance cortical columns appear.

Estimation Standard Deviation

It is often necessary to quantify the error in the estimated response, i.e. the noise which has not been suppressed by trial averaging (see for example the figure 6.2 in chapter 6). This error is estimated through the variance of individual responses. Because of the normalization with the blank condition, it is necessary to bin the trials in packets containing both stimulated and blank conditions, and to compute an estimated response for each bin using the averaging and normalization by the blank condition explained above. Let us notate y_1, \dots, y_k these estimations in each bin, and \bar{y} their mean. The variance $\hat{\sigma}$ of the y_i can be estimated as:

$$\hat{\sigma}^2(t) = \text{var}(f_i(t)) = \frac{1}{k-1} \sum_{i=1}^k (\bar{y}(t) - f_i(t))^2.$$

And the standard error e on the final estimation \bar{y} is:

$$\begin{aligned} e^2(t) = \text{var}(\bar{y}(t)) &= \frac{1}{k} \sum_{i=1}^k \text{var}(f_i(t)) \\ e(t) &= \frac{1}{\sqrt{k}} \hat{\sigma}(t). \end{aligned}$$

Chapter 2

Physiological Models of the hemodynamic response

In this chapter, we will detail the different mechanisms involved in the BOLD response to neural activity. They include changes in the metabolism, the cerebral blood flow, volume and oxygenation, and the magnetic resonance properties of the vessels and tissues.

Additionally, we will show with simulations how these models can account for the different features of the empirically observed hemodynamic response.

2.1 Physiological Models

The figure 2.1 summarizes the main phenomena responsible for the BOLD response, which we shall describe here.

Basically, the neuron glucose and oxygen consumptions increase during energy-demanding tasks. This consumption is referred as Cerebral Metabolism Rate of glucose and oxygen (CMRGL and CMRO2). The two molecules are extracted from the blood during its transit inside the capillaries.

In order to meet this need for additional glucose and oxygen supply, the cerebral blood flow (CBF) increases within a small delay, in a much larger proportion than CMRGL or CMRO2 (the mechanisms underlying this flow response appear to be much complex, and are not directly linked to the lacks of glucose or oxygen). Under this flow pressure, the venous vessels dilate, according to their visco-elastic properties: the “Balloon Model” describes how the cerebral blood volume (CBV) is related to the CBF.

The main consequence for the cerebral deoxyhemoglobin content (HbR for “reduced” hemoglobin) dynamic is that the CBF effect predominates over those of CMRO2 and CBV: blood becomes more oxygenated (HbR decreases). At last, this results in an increased MRI signal, since the $T2^*$ decay due to the venous HbR content is attenuated.

We shall give more details on these processes, and in particular present various modelling attempts in the form of dynamical models. Reviews of the BOLD signal and its underlying neurovascular regulations can be found in [81, 124, 144].

2.1.1 Neural Basis of the BOLD Signal

In the human brain, when a particular area is recruited to carry out a task, a group of neurons in this area fires action potentials. However, there may be groups of neu-

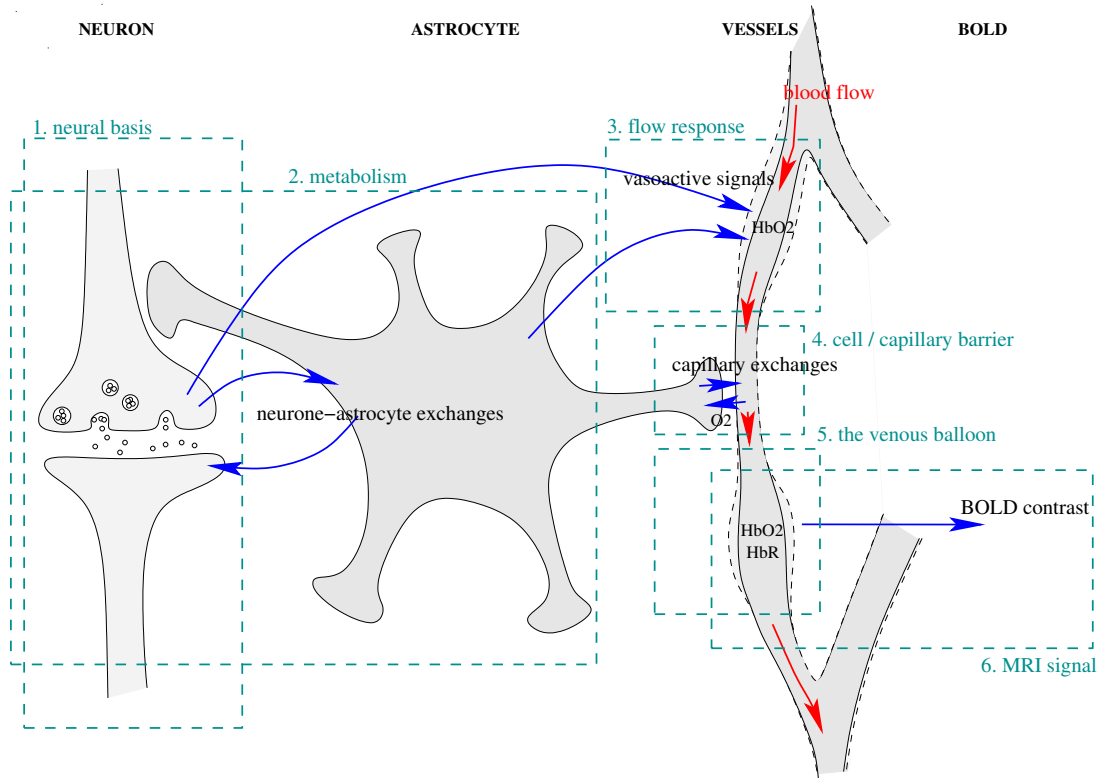


Figure 2.1: The physiological processes underlying the BOLD response essentially involve the neurons, the brain vasculature, and the astrocytes (small cells playing an important role in the different exchanges of molecules and ions between the two former cells). The following questions are critical for modelling these processes. *1. Neural basis of the hemodynamic response:* Which aspect of the neural activity (spikes, synaptic activity...) is the most directly linked to the hemodynamic changes? *2. Metabolism:* Which are the most energy-demanding tasks and how is the energy produced? *3. Arterial vessel dilatation and cerebral blood flow response:* Which kind of signals do control the resistance of the arterial vessels and how closely are they related to the metabolic demand? *4. Oxygen extraction from the capillaries:* To what extent does the extraction of oxygen from the hemoglobin molecules depend on the blood flow, the metabolic demand, and possible tissue reserves? *5. Visco-elastic properties of the vessels:* What is the dynamic relation between blood flow and blood volume in the vessels in general, and what is the implication for the venous deoxyhemoglobin content (HbR)? *6. Measure in fMRI:* How does the BOLD signal depend on blood volume and deoxyhemoglobin, and can it be reasonably assumed that the major part of the signal comes from the veins?

rons which do not fire action potentials, but nevertheless utilize oxygen. For instance, neurons at sub-threshold levels of activation, or neurons with simultaneous excitation and inhibition inputs. As a consequence, it is valuable to ask which changes in neural activity do influence the BOLD signal.

It is possible to acquire *electrophysiological measures* by the means of electrodes inserted in the animal brain (they may be used on humans too, under particular circumstances such as epilepsy diagnosis). The measured electrical signals can be separated into two parts according to their frequency content. The high-frequency (300-1500Hz) *multi-unit activity* (MUA) captures the spiking activity of neurons in a small neighbourhood of the electrode tip [118]. The low-frequency (40-130Hz) *low-field potential* (LFP), on the other hand, is believed to reflect the superposition of synchronized dendritic currents, averaged over a larger volume of tissue [141].

Several studies have shown significant correlations between the amplitudes of the BOLD responses and both the MUA and LFP activities in a same region [102]. Some of these studies achieved fMRI measures on humans and electrophysiological measures on monkeys [79, 80, 151]. Others also used EEG [147, 27], or optical imaging and Laser-Doppler flow recordings [135, 6].

Logothetis et al.[122] used simultaneous electrophysiological and fMRI acquisitions on macaque monkeys to compare how well LFP and MUA could predict the fMRI measurements. Whereas in some recording sites, LFP and MUA were similar and both predicted the fMRI response well, in some others, the two differed, and LFP signals were found to better correlate fMRI than MUA. In particular, in those sites, MUA showed transient responses which returned to baseline within 2-4s though the stimulus stayed on for a longer time; at the same time the LFP and fMRI signals were both sustained throughout the stimulus presentations, which lasted for up to 20 seconds. Hence, Logothetis et al.concluded that BOLD fMRI signals, “reflect the input and intracortical processing of a given area rather than its spiking output.”

Whereas a good concordance was generally found between LFP and hemodynamic responses, some nonlinear relations between the two of them have been identified as well. In some cases, a threshold effect was highlighted [90, 48, 157]. Mathiesen et al.[135, 115] even found a sigmoid dependence of the CBF measurements with the LFP measurements when varying the frequency of electric stimulations.

In addition, the BOLD response may reflect the activity of a network in a region, which in some cases is more than just the sum of the activities of each individual unit. For instance, the fMRI signal may reflect changes in neuronal synchrony without a concomitant rise in mean firing rate [57]. Also, Kilner et al.[101] proposed a heuristic model of the dependence of fMRI on the oscillations of a network. They concluded that an increased BOLD signal should correspond to an acceleration of the oscillations, and thus to a shift in the measured oscillations toward the high frequencies.

2.1.2 Metabolism

In the empirical observations we have mentioned, it appeared that the BOLD response depends more on intracortical processing than on neuron firing rates, even though in many situations, both are correlated.

In accordance with these observations, the studies on brain metabolism showed that the most energy-consuming tasks are related to the synaptic activity, rather than to the spike generation (Creutzfeld [40] estimated the latter to represent a maximum of

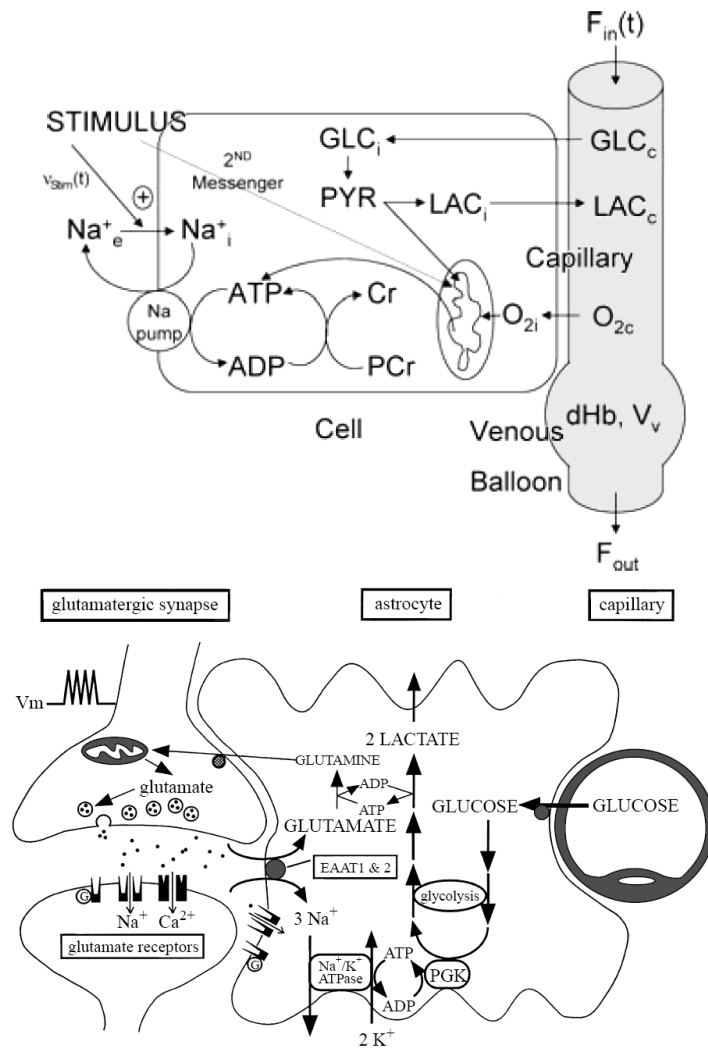


Figure 2.2: *Brain metabolism*: The upper part of the figure, issued from Aubert and Costalat, 2002 [9], illustrates the main processes of the metabolism, including glucose aerobic and anaerobic consumptions (respiration and glycolysis), ATP production and the Na⁺ pump. The bottom part of the figure, issued from Magistretti and Pellerin, 1999 [128], depicts in detail the glycolysis and the on the cycle of the glutamate neurotransmitter.

3% of the cortical energy consumption). This synaptic activity can be summarized as (1) when action potentials arrive in the axonic termination of the source neuron, the latter releases neurotransmitters into the synaptic cleft and (2) these neurotransmitters activate receptors on the target neuron and induce changes in ionic gradients. To maintain a continuous neural activity, energy-consuming processes restore the ionic gradients and repack the neurotransmitter molecules into the source neuron [144].

These processes involve both the neurons and the *astrocytes*. Astrocytes are stellate cells that have ramifications around the neurone synapses and around the capillary vessels. There are about ten astrocytes for one neuron, and they play an important role in the various chemical reactions as well as molecular and ionic exchanges with neurons and capillaries. In particular, they possess receptors to glutamate, one of the most current neurotransmitters, and participate in its reuptake toward the source neuron.

A detailed review of the cell metabolism can be found in [128]. In summary, the energy is supplied by glucose consumption, which is achieved by both aerobic and anaerobic processes. The aerobic process is called *glycolysis*, and is thought to take place in the astrocytes, while the anaerobic process, also called *respiration*, occurs in the mitochondria of the neuron. The glucose is transformed into lactate, and serves to produce ATP molecules, which will themselves bring energy to various processes - including in particular the Na^+ , K^+ pump which restores ionic gradients, and the reuptake system of glutamate and other neurotransmitters. The blood circulating in the brain provides the glucose and oxygen molecules needed by these processes, and drains the produced lactate. These exchanges take place around the capillary vessels. Figure 2.2 shows two diagrams issued from [9] and [128] which illustrate these various mechanisms.

As a result, the metabolism of a brain region can be characterized by the consumption of glucose and oxygen, usually denoted by CMRGL and CMRO2. As far as the BOLD response is concerned, we are particularly interested in the dynamic of oxygen. Upon neural activation, one would expect CMRO2 to increase, and hence also the deoxyhemoglobin in the capillaries and the veins, which should result in a decrease of the BOLD signal. The contrary happens, due to an increase in the cerebral blood flow (CBF) that overcompensates for the decrease in oxygen, delivering an oversupply of oxygenated blood [54]. On the other hand, the glucose supply does appear to match the consumption. See [66] for a review on the changes in oxygen metabolism under increased neural activity.

The different mechanisms involved in metabolism have been put in equation by Aubert and Costalat [9]. In particular, they simulated the BOLD responses obtained with various neural activity patterns and various hypothesis on the CMRO2 changes. However, this modelling was very complex and implied at least 15 evolving variables. For some applications, it is desirable to model the changes related to neural activation more simply, all the more since the only metabolism quantity we are interested in for models of the BOLD response is CMRO2. For instance, Davis et al.[45] proposed to model the CMRO2 response with a step function, while Buxton et al.[32] proposed a convolution of the neural activity with a gamma-variate function:

$$\begin{cases} \text{CMRO2} &= 1 + (m_1 - 1) h_m * u \\ h_m(t) &= \frac{1}{6\tau_m} \left(\frac{t}{\tau_m}\right)^3 e^{-\frac{t}{\tau_m}}, \end{cases} \quad (2.1)$$

where two parameters are used: m_1 controls the scaling of the response (it is equal to the plateau value of the CMRO2 response to long stimulations), and τ_m controls its

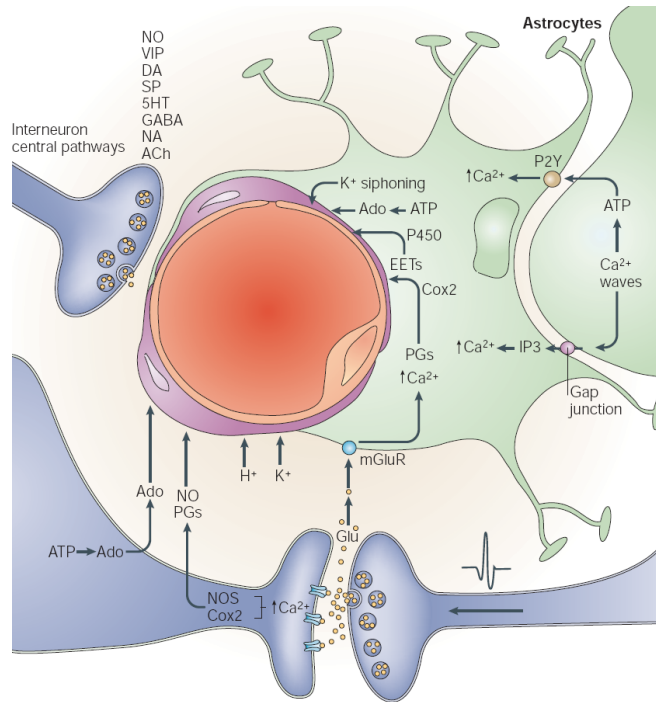


Figure 2.3: Control on the blood flow. Figure issued from the review on neurovascular regulation by Iadecola [88].

speed.

2.1.3 Flow Response

As previously mentioned, the CBF response to neural activity does not match the need for additional oxygen supply, but exceeds it. Moreover, it appears that the control of blood flow is not directly linked to the metabolism; rather, it is originated, at least partly, from the synaptic communications. Basically, the flow is regulated by the contractions and extensions of smooth muscle cells which surround the arteriole vessels. The tension of these muscles itself can be modulated by a wide range of vasodilator substances. The review by Iadecola [88] highlights the fact that activity-induced flow responses require complex signalling mechanisms that involve not only neurons but also astrocytes and vascular cells.

As a matter of fact, among the vasoactive agents, there are substances which are also involved in the generation of the extracellular ionic currents (such as K^+ and H^+), neurotransmitters (such as acetylcholine, GABA, catecholamines and neuropeptides). Other neurotransmitters, such as glutamate, are not vasoactive, but stimulate the production of powerful vasodilators, including nitric oxide (NO), through calcium-mediated enzymatic activation. The increase in ATP metabolism that is associated with neural activation also leads to production of the potent vasodilator adenosine. Moreover, it was recently found that, though synaptic signalling and not energy deficit is the main cause of the flow response, there exist other factors which are linked to the intracellular energy metabolism. Indeed, the lactate/pyruvate ratio, which is an indicator of energy needs, regulates CBF through signalling that involves NO and protein kinase C. Numerous references can be found in [88] (references 10-43).

In addition, observations on the spatial extent of the CBF increase highlighted the concurrent effects of several mechanisms. The existence of a fast and very localized flow responses observed in some cases cannot be explained uniquely by the diffusion of vasoactive metabolites in the tissues, and requires a specific action of astrocytes releasing vasoactive factors close to the vessels. At the same time, it appears that the vasodilation in the area of activity is propagated upstream by ‘intramural’ vascular signalling, in order to increase the blood supply in the whole region irrigated by upstream pial arteries. Otherwise the local CBF increase would result in a CBF drop in the neighbouring areas.

Because these mechanisms are highly complex, and since a large part of them is still unknown, there exists no precise dynamic model of the flow response to neural activity. Actually, when it is used fMRI models, the flow dynamic is often modelled quite roughly with a trapezoidal shape: for instance Aubert and Costalat [9], who were mainly interested in slow metabolic dynamics, used in their simulations a flow response, which increased linearly during 5 seconds after the stimulation begin, then reached a plateau and decreased linearly after the stimulation end.

More elaborated models describe the flow response as a linear function of neural activity. Friston et al.[59] proposed a damped oscillator model, by introducing the notion of a “flow inducing signal” s , which is actually the flow derivative:

$$\begin{cases} \dot{f} &= s \\ \dot{s} &= \epsilon u - \kappa_s s - \kappa_f (f - 1). \end{cases} \quad (2.2)$$

In this model, u denotes the neural activity, supposed to be zero at rest state, f denotes the flow value relative to (i.e. divided by) its value at rest state, and three parameters are used: neural efficacy ϵ , decay time constant κ_s and autoregulatory feedback time constant κ_f . This system is able to produce plausible time courses, with an initial increase and a slower return to baseline.

Buxton et al.[32] proposed a convolution of the neural activity with a gamma-variate function, similar to (2.1):

$$\begin{cases} f &= 1 + (f_1 - 1) h_f * u \\ h_f(t) &= \frac{1}{6\tau_f} \left(\frac{t}{\tau_f}\right)^3 e^{-\frac{t}{\tau_f}}. \end{cases} \quad (2.3)$$

Like above, f_1 controls the scaling of the response and τ_f controls its speed.

2.1.4 Oxygen Extraction from the Capillaries

As mentioned above, the exchanges between the brain and the blood occur around the capillary vessels. We are particularly interested in the extraction of oxygen from hemoglobin molecules, since it turns the latter into deoxyhemoglobin molecules whose paramagnetic properties are responsible for the BOLD signal. Inside the blood, most of the oxygen molecules are bound to hemoglobin molecules; however a small amount of remaining O_2 molecules are free inside the blood plasma. These molecule can pass through the pores of the capillary wall and then diffuse into the brain tissues.

For hemodynamic models, the quantity which is important to determine concerning this transit in the capillaries, is the ratio of oxygen extracted from the blood, denoted by E . At steady state, it is related to the metabolism of oxygen, i.e. its consumption

by the brain cells, according to the equation [148]

$$\frac{E \cdot \text{CBF}}{\text{CMRO}_2} = \frac{E_0 \cdot \text{CBF}_0}{\text{CMRO}_{2_0}}, \quad (2.4)$$

where the zero subscript denotes values at rest. This equation has been widely used to estimate CMRO₂ given measures of CBF and inferences on E in optical imaging [136, 95] - actually E cannot be calculated directly, but the ratio between oxygenated and deoxygenated blood measurements can give information on it.

Buxton and Frank [31] proposed a model of the dynamic of this oxygen delivery to the brain. The particularity of their model is that they supposed the large imbalance of flow and oxygen metabolism to reflect a tight coupling in the presence of a limitation of O₂ availability. This means that all oxygen that diffuses from the capillaries is effectively metabolized, and there is no back-flux from the tissues to the blood. In addition to this first hypothesis (a), they used the following assumptions: (b) the volume of the capillaries remains constant when the flow increases (there is no *capillary recruitment*), (c) the ratio between the pool of dissolved O₂ in the plasma and O₂ bound to hemoglobin is constant (this means that exchanges between the two pools are very rapid), (d) the diffusivity of oxygen across the capillary wall is constant: each O₂ molecule in plasma has a probability per unit time k of being extracted (in particular, it implies that the oxygen concentration in the tissues is zero).

As a consequence of these assumptions, the extraction of oxygen E at time t only depends on the flow at the same time. They could obtain its value by writing an equation of the diffusion of oxygen at a given time and location inside the capillary, and integrating this equation along the path of an amount of blood through the capillary:

$$E = 1 - (1 - E_0)^{\text{CBF}_0/\text{CBF}}.$$

Actually, this tight coupling between flow and oxygen metabolism contradicts many observations of an *initial dip* in fMRI or optical imaging (see next section), which is thought to be due to an early increase of CMRO₂, before the flow response. Please note that recently, Buxton and his colleagues [32] admitted that oxygen extraction does not necessarily need to be coupled with the flow, and they proposed a more complete model where they used the steady-state relation (2.4) to determine oxygen extraction and allowed CMRO₂ to have a time course independent on CBF (equations (2.1) and (2.3)).

Zheng et al.[185] proposed a more detailed model, the “Oxygen tissue transport” (OTT) model, where several assumptions of the Buxton and Frank model were relaxed. In particular, they supposed that the concentration of O₂ in the tissues is non-zero, so that the diffusivity from the capillaries depends on that concentration, and the changes of this concentration depend both on the oxygen diffusion and on the metabolic demand, which is independent on the flow. As a consequence, their model entails two more dynamic variables: the ratio of tissue oxygen concentration and the plasma oxygen concentration at the arterial end g , and the average oxygen concentration in the capillaries \bar{C}^c . Moreover, they considered that the mean transit time inside the capillaries was non-negligible, and thus took into account the fact that a small amount of blood that leaves the capillaries at time t had entered them at a time $t - \tau$ where CBF, g and \bar{C}^c were possibly different. As a result, they obtained the system:

$$\begin{cases} \frac{\phi}{f} \dot{E} &= -E + (1-g) \left(1 - \left(1 - \frac{E_0}{1-g_0} \right)^{1/f} \right) \\ \frac{\phi}{f} \dot{\bar{C}}^c &= -\bar{C}^c - \frac{C^a E}{\log \left(1 - \frac{E}{1-g} \right)} + C^a g \\ \frac{\rho}{E_0} \dot{g} &= \left(\frac{\bar{C}^c - g C^a}{\bar{C}_0^c - g_0 C^a} - 1 \right) - M, \end{cases}$$

where M is the metabolic demand (CMRO2), and the parameters are: g_0 , \bar{C}_{B0} , E_0 , are the values at rest of the varying quantities, ϕ is the inverse of transit time at rest, C^a is the concentration of oxygen at the arterial end, and ρ is a constant taking into account the volumes of capillary and tissue, transit time and transport lags.

2.1.5 Visco-Elastic Properties of the Vessels and the Venous Balloon

We saw how neural activation induces changes in the cerebral blood flow and the extraction of oxygen from the capillaries, according to mechanisms which are still under investigation. The last variable we need to characterize in order to fully account for the BOLD signals is the density of blood in a fixed volume, usually referred as cerebral blood volume (CBV) or *perfusion*. Blood is a viscous fluid travelling inside vessel which themselves have specific visco-elastic properties. As a result, the flow and volume dynamics depend on each other, according to more or less complex relations.

CBV Dynamics

At steady-state, numerous studies were in agreement with an early work of Grubb et al.[75], who found a power-law relation between flow and volume:

$$V = F^\alpha, \quad (2.5)$$

where the *Grubb's parameter* α is generally estimated between 0.25 and 0.4.

However, during transient phases, this steady-state relation makes no sense, since when the volume is varying, the flow cannot be the same throughout. Indeed, the variation of the blood volume V inside a given vascular compartment can be expressed as the difference between entering and exiting flows:

$$\dot{V} = F_{in} - F_{out}.$$

When using relative quantities (i.e. normalized by the values at rest), denoted by lower-case letters, it gives:

$$\dot{v} = \frac{1}{\tau} (f - f_{out}), \quad (2.6)$$

where $\tau = CBV_0/CBF_0$ is the mean transit time in the compartment.

There exist mainly two reference models: the *Balloon model* proposed by Buxton et al.[33], and the *Windkessel model* proposed by Mandeville et al.[131]. The Balloon Model uses equation (2.6) to determine the volume dynamic, which is controlled by the definition of the f_{out} term. The Windkessel model proposes an analogy between blood pressures and flows in the vessels on the one hand, and potential differences and current intensities in an electrical circuit on the other hand.

We focus here on the Balloon model. Buxton and colleagues supposed the flow leaving the arteries and entering the capillaries given, and assumed that the volume of capillaries remained unchanged upon activation, which implies that the flow entering the veins is the same as that entering the capillaries. On the other hand, the volume of the veins increases under the pressure of a larger inflow, hence the name “balloon”. They ran simulations with various definitions for the f_{out} function. Whereas Friston et al.[59] used the steady state expression in their dynamical model $f_{out} = v^{1/\alpha}$, where α is the Grubb’s parameter, Buxton et al.[30] proposed to add a resistance of the vessel to volume changes, with the introduction of a *viscosity* term τ_{visc} :

$$f_{out} = v^{1/\alpha} + \frac{1}{\tau_{visc}}v,$$

Moreover, as this resistance to fast variation can differ whether volume is increasing or decreasing, τ_{visc} was allowed to take two different values, leading to the final hysteresis scheme

$$\dot{v} = \begin{cases} \frac{1}{\tau} (f - v^{\frac{1}{\alpha}} - \frac{1}{\tau_{visc}^+} \dot{v}) & \text{if } \dot{v} \geq 0 \\ \frac{1}{\tau} (f - v^{\frac{1}{\alpha}} - \frac{1}{\tau_{visc}^-} \dot{v}) & \text{if } \dot{v} \leq 0. \end{cases} \quad (2.7)$$

This viscosity parameter allowed the model to account for *compliance* effects, i.e. a delay of the changes of volume with respect to those of flow, which is generally admitted to be the cause of the poststimulus undershoot observed in fMRI.

Alternatively, Mayhew and his colleagues proposed a modified Balloon/Windkessel model which used an additional state variable to model this delayed compliance [107]. They also proposed, in their “three-compartment” model, to model a balloon effect in the arteries (but with no compliance), which allowed them to better tune the dependence of optical imaging volume and oxygenation measurements on the arterial, capillary and venous compartments [184]. However, though CBV is presumably delayed with respect to CBF in the veins, it might be different in other compartments. As a matter of fact, Malonek et al. found in a study on anesthetized cats that CBF changes lagged after CBV changes by one to two seconds [129]. They suggested an early capillary recruitment which would be obtained by a momentary reduction of the outflow from the capillaries.

Venous Deoxyhemoglobin Content

In order to predict the BOLD signal measured in a cortical region, it is necessary to determine the deoxyhemoglobin content HbR of the veins. Like the vein volume, its variations are determined as the difference between entering and exiting quantities. The entering deoxyhemoglobin is the product between the entering flow F_{in} and the deoxyhemoglobin concentration at the end of the capillaries EC^a . Conversely, the exiting deoxyhemoglobin is the product between the outflow F_{out} and the deoxyhemoglobin concentration in the veins (it is assumed that the blood inside the veins is mixed sufficiently fast to be homogeneous with a deoxyhemoglobin concentration equal to HbR/V):

$$HbR = EC^a F_{in} - \frac{HbR}{V} F_{out}.$$

Using relative values (we denote by q the deoxyhemoglobin content relative to rest $q = \text{HbR}/\text{HbR}_0$), it yields:

$$\dot{q} = \frac{1}{\tau} \left(\frac{E}{E_0} f - \frac{q}{v} f_{out} \right), \quad (2.8)$$

2.1.6 The BOLD Signal

The BOLD signal is due to the paramagnetic property of deoxyhemoglobin, which shortens the T_2^* relaxation time of the MR signals because of the increased inhomogeneity of the magnetic field it induces. This effect was first described by Ogawa and his colleagues [148]. Then several studies developed precise models of this phenomenon, and proposed estimations or calibration methods for the parameters of the models [23, 44]. More recently, Buxton et al. proposed a simple expression, which we shall briefly present here.

Basically, the MR signal can be modelled with an exponential

$$S = S_{\max} e^{-TE R_2^*},$$

where S_{\max} is the effective spin density, and $R_2^* = 1/T_2^*$ is the transverse relaxation rate constant. Actually, only a small part of R_2^* should be attributed to the deoxyhemoglobin. However, this part remains sufficient to produce observable signal variations. We denote by $R_2^*(0)$ the value of R_2^* if no deoxyhemoglobin was present, R the additional relaxation produced by deoxyhemoglobin, and R_0 this additional relaxation at rest state. We have:

$$\frac{S}{S_0} = e^{-TE(R-R_0)} \approx 1 - TE(R - R_0).$$

The relaxation produced by deoxyhemoglobin should be proportional to the magnitude of the field distortion, and hence to the deoxyhemoglobin content of the voxel. However, because of diffusion effects, which affect the smaller vessels, two different power-law relationships are more accurate to model the dependencies of the relaxation on the blood volume and on the deoxyhemoglobin concentration in the blood. Additionally, since these diffusion attenuate the effects of deoxyhemoglobin in small vessels, the signals due to capillaries are usually ignored, and only the veins are considered. The final expression for the BOLD signal we will use is the linear approximation as it appears in [32]:

$$\frac{S}{S_0} = 1 + V_0(a_1(1 - q) - a_2(1 - v)),$$

where v and q are the relative venous volume and deoxyhemoglobin changes, V_0 is the ratio of blood per unit of volume at rest, and a_1 and a_2 are two constants which depend on the oxygen extraction at rest E_0 and on some scanner and acquisition properties. The values estimated by Obata et al. [146] for a magnetic field of 1.5 Tesla with $TE = 40\text{ms}$ and $E_0 = 0.4$ are $a_1 = 3.4$ and $a_2 = 1.0$. It appears that, as expected, the major part of the BOLD signal is due to the venous deoxyhemoglobin content; however, volume also plays a non-negligible role, since for two similar deoxyhemoglobin contents, the one which is enclosed in the smallest volume (i.e. the most concentrated one) is the one which attenuates the signal the most. This is due to diffusion effects (these effects become less effective at higher magnetic fields).

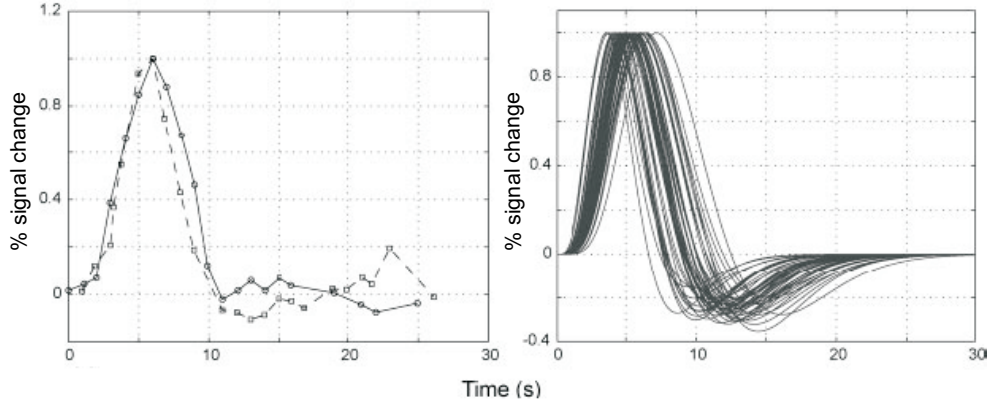


Figure 2.4: Time course of the BOLD response. Left: measurements from experiments in the motor cortex (Liu et al., 2000 [119]) (circles) and in the visual motor cortex (Glover et al., 1999 [67]) (squares), in response to a movement or a visual stimulation of 2s. Right: Theoretical hemodynamic response functions modelled by two gamma-variate functions. The distributions of curves were created by randomly perturbing the parameters by 10%. Figure adapted from the review by Logothetis, 2004 [124].

There is currently no model for the BOLD measure at 3T. In our work, we choose to use the equation above as an approximation for the measure of the BOLD signal at 3T, and we impose a relationship between a_1 and a_2 , to avoid too many unknown parameters. On account on the fact that diffusion effects, and hence volume effects are smaller when the intensity of the magnetic field increases, we suppose $a_2/(a_1 + a_2) = 1/10$ (instead of $a_2/(a_1 + a_2) = 1/4.4$ at 1.5T), and introduce the new parameter $b = V_0(a_1 + a_2)$, so that the new measure equation can write:

$$\frac{\Delta S}{S_0} = b(.9(1 - q) - .1(1 - v)).$$

2.2 Empirical Observations and Predictions of the BOLD Response

The Hemodynamic Response Function (HRF) is the theoretical BOLD response to a short neural activity (figure 2.4 on the left). The linearity assumption used in most fMRI analysis states that the BOLD response to any neural activity pattern can be obtained by convolving this pattern with the HRF (this linearity hypothesis is discussed in the next section). Thus, if the linearity assumption did hold, it would be sufficient to know the HRF shape to interpret, and even de-convolve [67], the measured time courses.

The two main characteristic features of the HRF are a strong positive response which starts 1~2s after the stimulation and reaches its maximum at 4~5s after the stimulation, and a subsequent drop below baseline which can last for twenty seconds. Lange and Zeger [112] proposed to use the difference of two Gamma distributions to respectively model the positive response and negative undershoot of this HRF (figure 2.4 on the right). It involves five parameters (n_1, t_1, n_2, t_2, a):

$$h(t) = \left(\frac{t}{n_1 t_1}\right)^{n_1} e^{-\frac{t}{t_1}} - a \left(\frac{t}{n_2 t_2}\right)^{n_2} e^{-\frac{t}{t_2}}.$$

Now we shall discuss several characteristic features of this response, and show simulation results from the models previously described, either issued from the literature or from our own simulations. We will pay a particular attention to the departures from the nonlinearity assumptions which have been reported in several studies.

Positive Response and Undershoot

To better understand the Balloon Model, we ran some simple simulations and tried to isolate specific features of the BOLD response, displayed in figure 2.5.

In the part A of the figure, the flow was simulated using the damped oscillator proposed by Friston and colleagues (equation (2.2)), the oxygen metabolism was obtained by the derivations by Buxton and Frank, 1997 $CMRO2 = f(1 - E_0)^{1/f}/E_0$, the cerebral blood volume was obtained with (2.7) but with no compliance effect, and the deoxyhemoglobin content was obtained with (2.8). In the part B of the figure, the flow and oxygen metabolism were simulated using the gamma-variate convolutions proposed by Buxton and colleagues (equations (2.3) and (2.1)), but CBF and CMRO2 were still tightly coupled since the same functions were used (only the amplitude differed, by the ratio $CBF/CMRO2 = 2.5$). Actually, neither of the two simulations appears to be satisfying: in A, we can observe the characteristic BOLD undershoot, but this undershoot is already present in the flow time course, which is not the typical shape observed for CBF. In B, on the contrary, no undershoot is present at all.

All subsequent stimulations use the second flow modelling. In the parts C and D we investigated the role of the compliance parameter τ_{visc} in equation (2.7). We separated the effect of τ_{visc}^+ , which denotes a resistance against volume increase (in C: little effect is observed on the BOLD signal), and the effect of τ_{visc}^- , which denotes a resistance against volume decrease. The latter had a strong influence on the BOLD and induced a prolonged undershoot, because of a higher deoxyhemoglobin content than at rest.

The Initial Dip

The question of the initial dip is particularly important. In optical imaging it has been possible to observe reliably an increase in deoxygenated blood in the first seconds of neural activity [169, 158, 171], probably caused by an early increase in CMRO2, before the CBF response had the time to bring new oxygenated blood. Moreover, it appeared in these studies that this early activity-dependent increase in deoxyhemoglobin concentration was much more colocalized with the neural activity than the delayed increase in oxyhemoglobin concentration, such suggesting that imaging based on the negative dip response in fMRI studies could significantly increase their spatial resolution. We show in our simulation (parts E and F of the figure) how changes in the delay of the CMRO2 response with respect to that of the flow response, but also in its amplitude, can produce such an initial dip. Please note however that there is not an agreement yet on the reliability on this initial dip in the BOLD responses [29].

Nonlinearities of the BOLD Response

Most fMRI data analysis rely on the fact that the BOLD signal are supposed to respond linearly to neural activity. In fact, numerous fMRI studies have considered the question of the validity range of this linearity hypothesis. They agreed on the fact that the assumption holds for stimulation durations or interstimulus intervals (ISI) larger than a

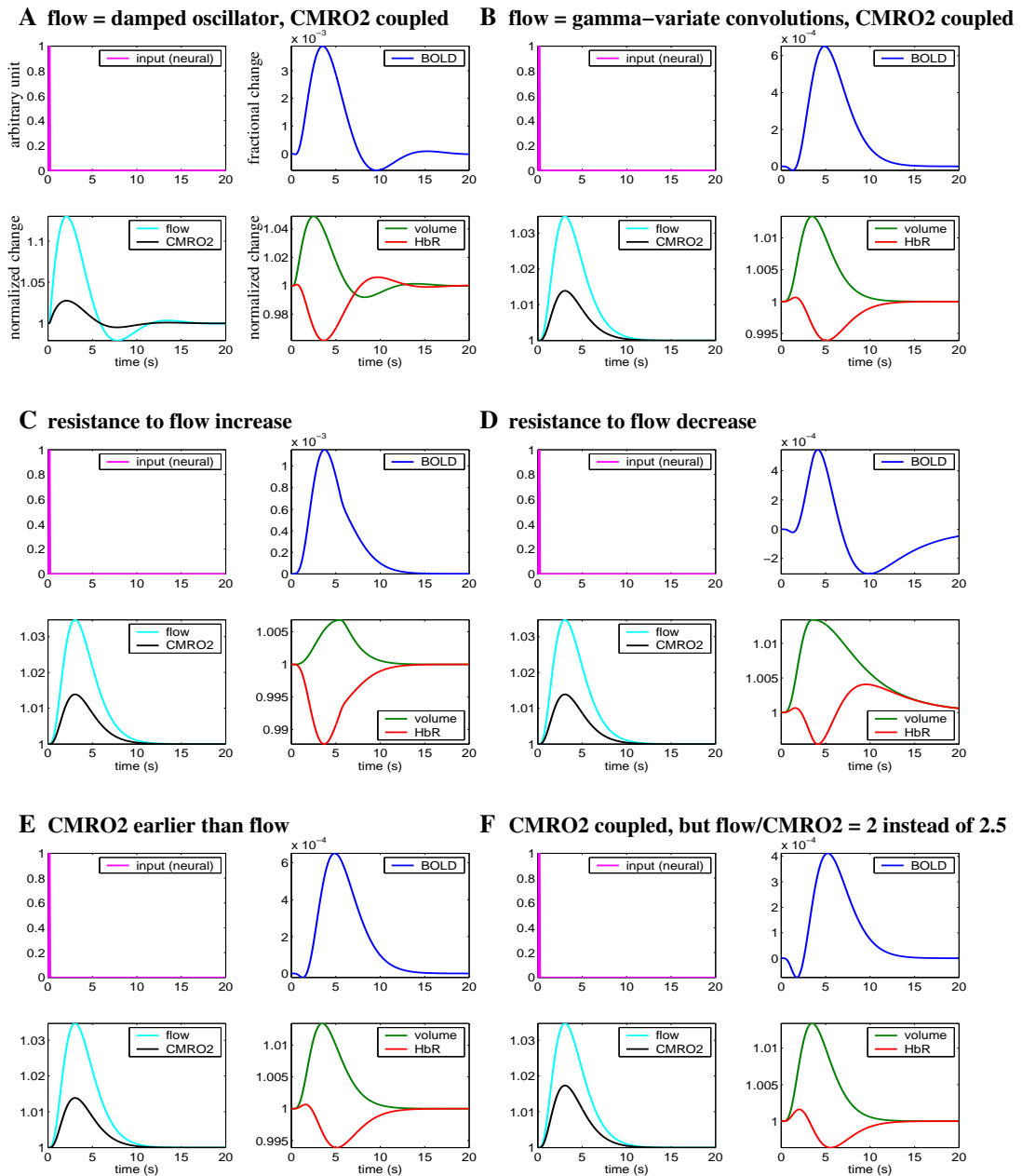


Figure 2.5: Simulations of hemodynamic and BOLD time courses using the Balloon Model (see text).

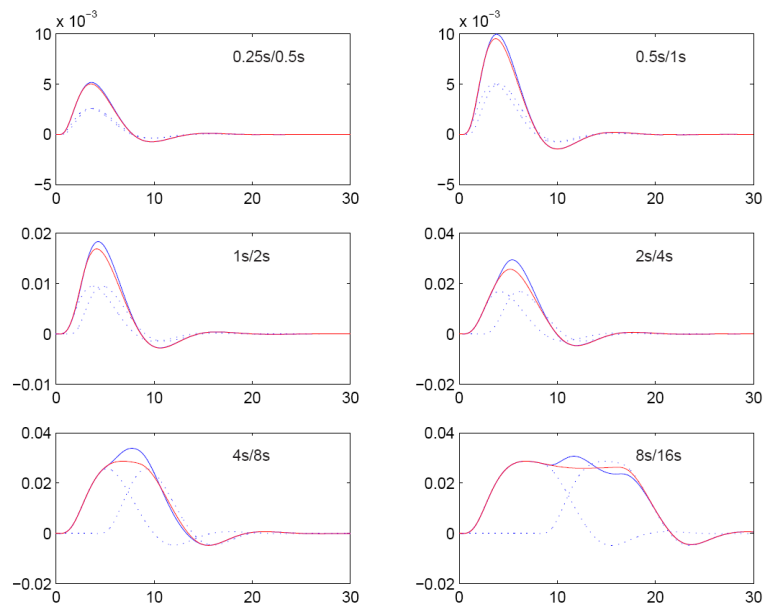


Figure 2.6: By comparing responses to a given stimulation length with the linear prediction from responses to stimulations twice shorter, one can assess the nonlinearities of the model dynamic. Here, the maximum nonlinear effect occurs for the 4s/8s comparison.

threshold. The value of this threshold varies among studies : 2-3 seconds [24, 41] to 4-6 seconds [20, 67, 139, 173]. Studies involving other measurement modalities established that some nonlinearities in the BOLD were not present at the neural level, and hence were due to hemodynamic effects: blood flow measurement in humans via Arterial Spin Labelling [139, 146], or electrical activity in animals [90].

Therefore, it is highly valuable to ask physiological questions about these nonlinearities: How can they be explained? At which level do they occur: neural, signalling, vascular? We showed on a simulations (figure 2.6) that the Balloon model can produce some nonlinearities also, by comparing responses to specific stimulation lengths with their predictions from the responses to shorter stimulations. These questions shall be developed in the chapters 4, 5 6.

Chapter 3

State of the Art of EEG/MEG-fMRI Fusion

Our work on hemodynamic models was intended in particular to enable new analysis methods of data issued from different brain imaging modalities. We were particularly interested in the fusion of EEG, MEG and fMRI data, since they are today the most powerful, non-invasive techniques which can be used on human. We propose here a review of existing methods for the integration of EEG/MEG and fMRI, and try to give a better understanding of the subject.

3.1 Understanding the problem

Brain imaging modalities span a very large panel of spatial and temporal resolutions, from microns to centimetres, and from milliseconds to years. However, no technique alone gathers all these resolutions together; only a subset of them. It is of great interest then to know how to integrate data issued from multiple modalities in order to benefit higher spatial and temporal resolutions.

On the one hand, EEG and MEG have a high temporal resolution, down to the millisecond, but the localization is much more problematic, since measures of electric potentials and magnetic field are done outside the head, and signals recorded on each sensor potentially depend on activities everywhere on the cortex. The inverse problem intends to localize the activities on the cortex, but this estimation remains uncertain. On the other hand, fMRI offers a much better spatial resolution, since the BOLD signals are already measured voxel-wise. At the same time, its temporal resolution is only several seconds, due to the smoothing by the hemodynamic process, and to the sparse sampling (100ms to acquire a single slice, and up to 3s for an entire brain). By the way, let us note that the fusion between EEG and MEG data has a natural formulation in the inverse problem, since they are driven by the same currents on the cortical surface. For that reason, from now on, the notation EEG/MEG will represent indifferently both or only one of the two modalities.

Hence, fusion between EEG/MEG and fMRI is potentially very valuable. However, it is not a straightforward thing. The two modalities measure very different consequences of the neural activity indeed, and the gap between their resolutions (temporal resolution particularly) makes designing fusion methods difficult. As a result, to develop fusion methods, it is necessary to ask the three following questions:

- How are the measures related physiologically ?
- Which information do we want to extract from multimodal data ?
- And only thereafter: How can this information be extracted ?

3.1.1 Physiological Relation between Electrical and Hemodynamic Responses

EEG/MEG is a measure of synchronized electrical activities, while the BOLD effect in fMRI is related to the metabolism (energy consumption) and delayed by the hemodynamic processes. Some activities are likely to be detected by only one of the two modalities. For example:

- An energy consuming activity which does not imply the synchronization of local neuronal assemblies, or which occurs deep inside a cortical sulcus, will probably only elicit a BOLD response.
- Inversely, some changes in the electrical dynamics of the brain could occur at a constant metabolism level, and then only be seen in EEG.

Next comes the question of how electrical activity is related to metabolic changes: is it for example the current intensity, or the power in a given frequency band, which best correlates the energy consumption [133, 101]? And then, how is the BOLD dynamic related to this energy consumption (see the previous chapter)?

3.1.2 EEG/MEG-fMRI fusion: what should be estimated ?

In neuropsychiatric experiments, the problem of estimating neuronal activity always comes with a specific cognitive question. For instance, in EEG/MEG, the data processing will not be the same whether one is interested in evoked potentials, i.e. activities time-locked to the experimental paradigm, or in responses in the frequency domain, or in an ongoing activity which is independent on any stimulation. Nor when one is interested in the temporal aspects only (latencies, frequencies...), or in sources localization (need to solve the inverse problem). Also, in typical parametric fMRI analysis, for instance, the purpose is often to find regions performing a specific cognitive task by the means of statistical tests, rather than by estimating neuronal activity in every voxel.

In a similar way, the various methods proposed for EEG/MEG-fMRI fusion do not estimate the same aspects of the brain activity. Basically, we observed two main criteria to classify different fusion approaches.

First, it is critical to know whether the method estimates an activity linked to a repeated external stimulation or not. In that case, electrical and hemodynamic measures do not need to be acquired simultaneously, since the activity patterns are supposed to be reproduced identically at every stimulus presentation (of course, this is on condition that the environment - inside the scanner or not - does not influence the subject behaviour). On the contrary, if the random trial-by-trial variations of the responses are to be taken into account, or if there is no deterministic repeated stimulation, simultaneous acquisition is necessary, so that both modalities can measure the same activity fluctuations. In that case, the only possible electrical measure is EEG, since concurrent MEG and fMRI acquisition is not really feasible !

Secondly, approaches can be distinguished according to whether they suppose that the different modalities share spatial, temporal or spatio-temporal information. Indeed, some methods will suppose that both EEG/MEG and fMRI give information on source locations, but that only the electrical measures give a pertinent information on their time courses. Inversely, others will consider that EEG signals at the electrodes can have - in a sense to be specified - a correlation with BOLD time courses, but will use fMRI only for sources localization (no inverse problem). Note that in these two cases, even though the shared information is only spatial or temporal, the estimation result will combine the two aspects. Please also note that there is a logical link between the two criteria presented in this subsection: doing simultaneous EEG-fMRI acquisition implicitly implies that they share some temporal information.

3.1.3 Methodology

Several authors have described different levels of integration between multiple-modality data. Horwitz and Poeppel [86] proposed to distinguish between three stages. The first one, called *convergence validation* only consists in comparing analysis results obtained on each modality data alone (activation maps essentially), and in discussing their similarities. Despite their simplicity, it is necessary to conduct such observations in order to validate the hypothesis concerning the link between electrical and hemodynamic measures of cerebral activity.

Next, so-called *asymmetric integration* methods consist in using the result of a first analysis with one of the modalities, and then utilizing this result as a constraint for the analysis of the second one. Historically, the first fusion attempts consisted in constraining the inverse problem in EEG/MEG by using the activation map found in fMRI (Friston and colleagues call this *integration through constraint* [101]). More recently, the inverse was explored also, where the EEG signals recorded during simultaneous acquisitions were used to build fMRI regressors (Friston and colleagues call this *integration through prediction*).

These approaches do not suppose explicitly an unknown cortical activity underpinning the diverse measures. On the contrary, *symmetric integration* methods do require a common forward model that links underlying neuronal dynamics to measured hemodynamic and electrical responses; this dynamic is estimated by a concurrent integration of both measures.

Table 3.1.3 classifies the different existing fusion approaches (convergence validations excluded), according to all the considerations above.

We shall now present these methods, classified according to the shared information criterion. This does not mean that they will appear in an order of increasing complexity or modelling effort. Other reviews on methods for the fusion between EEG/MEG and fMRI can be found in [63, 76]. The review by Nunez and Silberstein [145] focuses on the physiological basis of the relation between the electrophysiological (EEG/MEG) and the hemodynamic (fMRI) measures. The review by Salek-Haddadi et al.[155] focuses on methods which have applications in epilepsy.

3.2 EEG/MEG and fMRI Share a Spatial Information

As we mentioned above, assuming that EEG and fMRI only share spatial information (no temporal mapping between the two measures) implies that there is no need for si-

Method category	Authors	Integration sense	Simultaneous acquisition required ?	Shared information	How is the BOLD predicted ?	What is estimated ?
fMRI-guided EEG/MEG	Babiloni et al.[11], Ahlfors et al.[5], ...	fMRI \rightarrow EEG	no (event-related paradigm)	spatial pattern	stim * HRF	activation map + evoked electrical responses
symmetric spatial integration	Daunizeau et al.[43]	EEG \leftrightarrow fMRI	no (event-related paradigm)	spatial pattern	stim * HRF	activation map + evoked electrical and hemodynamic responses
EEG-guided fMRI	Goldman et al.[68] (α rhythm), Goman et al.[69] (epilepsy), Riera et al.[152], ...	EEG \rightarrow fMRI	yes (ongoing activity paradigm)	temporal patterns	power in frequency domain * HRF (α rhythm) or δ * HRF (epilepsy)	activation map corresponding to EEG temporal patterns
symmetric temporal integration	Martinez et al.[133]	EEG \leftrightarrow fMRI	yes (ongoing activity paradigm)	temporal patterns	power in frequency domain * HRF	temporal patterns + EEG and fMRI activation maps (can be different)
trial-by-trial amplitude variation	Lahaye et al.[111]	EEG \leftrightarrow fMRI	yes (event-related paradigm variations presentations)	spatial pattern + trial amplitude variations	stim . trial amplitudes) * HRF	activation map + evoked electrical and hemodynamic responses + amplitudes
spatio-temporal modelling	Trujillo et al.[163], Halchenko et al.[77], Deneux and Faugeras [47]	EEG \leftrightarrow fMRI	yes or no (both types of paradigms possible)	spatio-temporal sources activity	sources time courses * HRF or physiological nonlinear models	spatio-temporal sources activity

Table 3.1: Classification of EEG-fMRI approaches according to the criteria (1) shared spatial and/or temporal information, and (2) asymmetric/symmetric method.

multaneous data acquisition. Thus, all the methods in this section deal with estimations of an activity related to a specific repeated stimulation.

3.2.1 Convergence Evidence

Numerous studies showed similarities between electrical and hemodynamic analysis in the localization of regions which participate in a specific cognitive task. An extensive list of such experiments that used EEG/MEG and fMRI conjointly is given in the review by Halchenko et al.[76]. They involve a large variety of perceptual and cognitive functions: visual perception [159, 168], motor activation [4, 116, 156], somatosensory mapping [156], auditory [85] or visual oddball tasks [109], target detection [143], face perception [84], sleep [87], language [174]. Generally, the electrophysiological and fMRI measurements do agree on the location of the activity [4]. However, some studies reported a significant (1-5cm) displacements between the locations obtained from fMRI analysis or EEG/MEG inverse problem [16, 117]. Furthermore, such displacements could be very consistent across the experiments of different researchers (for instance motor activations [108, 104, 156]).

Results showing such displacements are not easy to interpret. They indicate that fMRI and EEG do find an activity in a same (extended) region. But it is not clear whether the difference in the precise localization is due to a bias in the estimation or whether there really are two foci of activity implicated in the same task. In any case, fusion algorithm should be flexible enough to allow estimations of activities seen by the EEG or the fMRI only.

3.2.2 fMRI-Guided EEG/MEG

The first attempts in integrating fMRI and EEG data were made by constraining the ill-posed inverse problem in EEG/MEG with activation maps obtained in a former fMRI analysis. This has a simple formulation using a Bayesian framework for the EEG/MEG inverse problem, and by tuning the a priori variance of the sources according to the fMRI analysis results (see the first chapter, section 1.2.3). The EEG/MEG measure m is a linear function of current intensities on the cortex surface, s , as described by the forward problem, plus additional measure noise n :

$$z = Gj + \eta.$$

If the measure noise η is supposed Gaussian with variance Σ , and if the a priori distribution of the sources is Gaussian also, with variance C , i.e.

$$\begin{aligned} p(\eta) &= \mathcal{N}(0, \Sigma) \\ p(j) &= \mathcal{N}(0, C), \end{aligned}$$

then the a posteriori distribution of the sources given the measure is

$$p(j|z) = \mathcal{N}((G^T \Sigma^{-1} G + C^{-1})^{-1} G^T \Sigma^{-1} z, (G^T \Sigma^{-1} G + C^{-1})^{-1}).$$

A priori on the activation of sources issued from the fMRI study can be set by tuning the diagonal terms C_{ii} of their variance matrix C . The inverse problem will locate primary currents responsible for the measures, preferentially in the sources with high variance. To allow some activities to be seen by the electric measures only, all sources should have a minimum variance, even in regions not activated in fMRI [62].

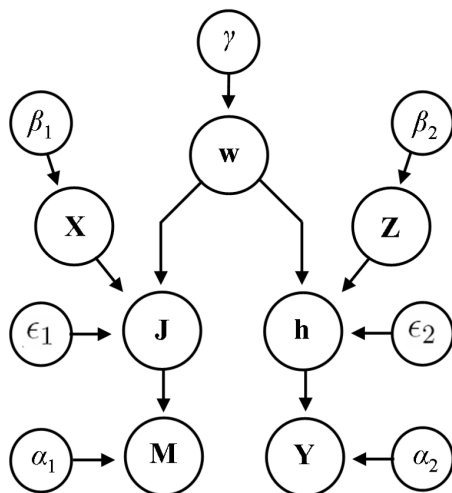


Figure 3.1: Forward model for EEG and fMRI in [43]. **(1)** A vector w represents the spatial profile of sources intensities, common to EEG and fMRI measures. **(2)** Electrical and hemodynamic temporal patterns of the response to the stimulation, X and Z , are supposed independent from each other. The brain surface is divided in regions (about twenty) according to the Broca classification: X and Z can vary across regions, but are supposed constant in a same region - this is the way to express that there can be several activity loci, but the activity in a given locus is synchronized. **(3)** By multiplying these electrical and hemodynamic temporal patterns X and Z by the voxels activity intensities w , one gets specific electrical and hemodynamic time courses in every voxels, notated S and H . **(4)** EEG signals M are obtained by applying the forward problem to S . **(5)** fMRI signals Y are obtained by convolving the stimulation sequence by H .

As an example, Babiloni et al.[11, 10] set

$$C_{ii} = 1 + \lambda\alpha_i,$$

where α_i is the statistically significant percentage signal increase of the fMRI signal during the task compared to the rest state, and λ is a hyper-parameter that controls the fMRI constraint.

Several variants have been proposed [42, 13, 120]. Ahlfors and Simpson proposed a “subspace regularization” technique [5], in which the cortical sources are chosen from all possible solutions describing the EEG/MEG measures as the one which minimizes a distance to a subspace defined by the fMRI data. Actually, all these techniques can be reformulated in order to match the Bayesian derivation above.

3.2.3 Symmetric Spatial Integration

Asymmetric methods rely on two steps: first, a weight is computed for every cortex element according to their fMRI activation, then these weights are updated during the EEG/MEG analysis to result in a map of cortical activity informed by the two modalities.

A symmetric integration of the modalities would rather consist in estimating these weights from the two measures together. Daunizeau et al.[43] proposed such a method. They first constructed a forward model where the electrical and hemodynamic temporal patterns were independent, but their intensities depended on a common spatial profile

w. Figure 3.1 depicts this model. These temporal and spatial patterns are estimated using Bayesian inference, according to an iterative algorithm. The advantages of using a Bayesian framework is that the estimation algorithm can be derived simply from the model formulation. Hyper-parameters, like noise variances, can be estimated as well. Finally, it enables statistical test derivations to assess the activation of individual sources.

3.3 EEG/MEG and fMRI Share a Temporal Information

Techniques integrating electrical and fMRI temporal aspects have emerged with the advent of simultaneous EEG and fMRI acquisitions. Though one could imagine applications on evoked responses, the existing methods presented here are rather interested in an ongoing activity during stimulation-free experiments. There is of course no MEG recordings in these cases.

3.3.1 EEG-Guided fMRI

The principle of integration through prediction is to use EEG measures to build regressors for the fMRI analysis. Namely, one or several time courses are extracted from the EEG measures at the electrodes, or from a subset of electrodes. The signal(s) is(are) convolved with a canonical hemodynamic function, and serve(s) as (a) regressor(s) for the General Linear Model. Standard fMRI analysis then finds regions whose BOLD signals are correlated - positively or negatively - with the EEG prediction(s).

For the moment, this has been applied successfully under two main experimental conditions. The first one consisted in leaving a subject with his eyes closed inside the scanner, and then using the power of EEG signal in a specific frequency band (particularly, for α oscillations) to correlate fMRI measures [37, 68, 113, 114]. Negative correlations were found between alpha oscillations and the BOLD signal in parietal and frontal areas, while positive correlations were found, but were contradictory between studies. Please note that these result do not necessarily locate the generators of oscillations, but rather regions whose activity is related to them. Note also that Kilner and Friston [101] have predicted heuristically that increased activity leads to a frequency shift of signals spectral signatures towards higher frequencies, which is in accordance with these negative correlations between the BOLD signals and the low-frequency α oscillations.

The second experimental condition is concerned with the diagnosis of epileptic subjects [178, 117, 21, 69, 177], who are asked to remain sedentary inside the scanner. These people experience abnormal electrical activity in the regions of the brain affected by epilepsy, manifested by frequent discharges or spikes. Thus, acquiring simultaneously EEG and BOLD signals is highly valuable. Indeed, EEG time courses serve to detect the times at which these discharges occur. The signal extracted from EEG is then a succession of instants (diracs), which is convolved with the HRF and serves as an fMRI regressor. This method offers new opportunities to locate epileptogenic regions (more detail can be found in chapter 8).

3.3.2 Symmetric Temporal Integration

The methods exposed above were asymmetric, in the sense that time courses were imposed by the EEG, and the fMRI was then used to locate the sources based on these signals. Martinez-Montes et al.[133] proposed a method to extract temporal features common to simultaneously acquired EEG and fMRI signals in a symmetric manner. fMRI data was considered a 2D array (voxels by time), whereas EEG data was considered a 3D array (electrodes by frequencies by time). The EEG data, expressed in the frequency domain is first convolved with a standard HRF, to make the signals comparable. Then the signals were decomposed as sums of *atoms*: an fMRI atom y_k is the product of a single spatial pattern u_{sk} by a temporal pattern v_{tk} : it is actually a rank 1 matrix,

$$y_k = u_{sk}v_{tk}^T.$$

(the authors actually did not use the transposition sign). In a similar way, an EEG atom z_k is the product of spatial (electrodes), frequencies and temporal patterns a_{dk} , b_{wk} and c_{tk} . It is a 3D array, hence in mathematical terms it would be defined with a tensor product.

$$z_k = a_{dk}b_{wk}c_{tk}$$

The authors proposed an algorithm which realized the atom decompositions by imposing the temporal patterns in EEG and fMRI decompositions to be maximally correlated (the vectors v_{tk} and c_{tk} must have maximum covariance).

They estimated these atoms using a tri-linear Partial Least Square algorithm. As a result, they were able to extract time courses present both in the EEG frequency powers and in the fMRI signals. They produced the maps corresponding to estimated temporal patterns by doing a GLM study in fMRI, and solving the inverse problem in EEG, and compared the two results. It is important to note that the algorithm did not use any constraint on similarity of these maps.

3.4 EEG/MEG and fMRI Share a Spatio-Temporal Information

Now, we want to take into account the fact that both temporal and spatial information given by electrical and hemodynamic measures can be integrated concurrently.

3.4.1 Trial by Trial Variations in Response Amplitude

Lahaye et al. presented an original method to integrate simultaneously-acquired EEG and fMRI in a stimulation-controlled experiment [111]. They proposed to take into account the fact that, during an experimental study, the neural responses to successive repetitions of the same condition might vary in amplitude, due to attention modulations for example. Thus EEG and fMRI measurements were supposed to share spatial information on source locations, as well as a simple temporal information, which was the amplitude variations of successive responses.

The method is close to the asymmetric integration-through-constraint methods presented above, i.e. the inverse problem in EEG is tuned by source covariance priors issued

from fMRI analysis (more precisely they tune the electrode measures covariance instead of sources covariance, but it is actually equivalent). The difference is that these covariances are not imposed straight ahead; an iterative algorithm discerns for every source whether the fMRI constraint should really be applied to it, depending on whether trial-by-trial amplitude variations for this source are correlated in EEG and fMRI or not. Although the formulation is quite complex, this method is a good way to take advantage on the simultaneous acquisition of EEG and fMRI.

3.4.2 Concurrent Estimation of the Spatio-Temporal Neural Activity

A more general integration of all the information contained in EEG/MEG and fMRI was proposed by Trujillo et al.[163]. They gathered forward models for all modalities in a single Bayesian framework and derived algorithms to estimate the spatio-temporal unknown, i.e. neural activity time courses of every source on the cortical surface. Their method can be applied indifferently on average evoked responses obtained in (non-simultaneous) EEG, MEG and fMRI experiments, or on an ongoing activity recorded by simultaneous EEG-fMRI measurements.

In their forward model, the relation between electrical activity on the cortex and that measured in EEG and MEG was given by the usual electrophysiological forward model, while that between metabolic activity and the BOLD measure was given by an HRF convolution. But they admitted ignorance about the link between these electrical and metabolic activities and an upstream common neural activity and supposed a simple linear link. The results on simulated and experimental data from a somatosensory MEG/fMRI experiment confirmed the applicability of Bayesian formalism to the multimodal analysis, even under the set of simplifying assumptions mentioned above. In particular, they had to face extensively large matrices.

Halchenko et al.[77] proposed a linear programming algorithm to solve the EEG/MEG-fMRI inverse problem. It enabled them in particular to use a nonlinear quadratic relation between electrical and hemodynamic activities. As a matter of fact, traditional least-square minimization techniques would not perform accurately on account of such a heavily nonlinear dependence. Their approximations allowed them to re-write the system in the form of a linear minimization problem under constraints. They applied their methods successfully to a reduced synthetic dataset, and are currently on the way to apply it to real-experiment data.

Kiebel and Friston [98, 99] recently proposed to apply the Statistical Parametric Mapping (SPM) techniques to EEG data. Hence they built a common framework for the EEG and fMRI measures that they believe will enable them to achieve fusion between the two modalities in a near future. Among the choices about this framework they had to deal with, they preferred to achieve parameter estimation on a voxel-by-voxel basis (mass-univariate framework) rather than on all voxels together (multivariate framework). This choice was motivated by the drop of the number of variance parameters that needed to be estimated then. However, it implies to first solve the EEG inverse problem and thereafter analyze each voxel independently. Consequently, the covariance information between the distributions of sources (which states in particular that several sources configurations can account for the same EEG measure) is ignored. All the same, though it loses this covariance information, the method appears to take into account spatial and temporal information from the EEG and fMRI measures.

We shall propose in chapter 8 a method for fusion between EEG and fMRI that

belongs to this category. The forward model for fMRI however will rely on a physiology-related dynamical system, and the algorithm we use, based on Kalman filtering, will allow the reduction of data dimensionality.

Part II

Using Nonlinear Models in fMRI

Chapter 4

Model Identification

In this chapter, we will show how to estimate the physiological parameters of a hemodynamic model using brain imaging measures (we will use fMRI in this chapter, and optical imaging in chapter 6). There are several complexity levels in such an estimation, depending on the assumptions on the noise in the data, and on various priors on the parameter values. Besides, whatever the method used, parameter estimation remains limited on behalf of the noise in the data and intrinsic identifiability properties of the dynamical systems.

To begin with, we will present the general mathematical framework for hemodynamic models: *stochastic differential equations systems*. It will be applied to two different versions of the Balloon Model (we will actually see that models expressed with convolutions do not strictly fit this framework), and show how to address some dynamical properties of such systems like their stability around equilibrium.

Secondly, we will present the Maximum Likelihood estimation method we have developed, accompanied with examples on simulation data, as well as other possible estimation frameworks.

Lastly, we will address the more difficult question of the estimation accuracy: is it possible to completely characterize the a posteriori distribution of parameters ? Some approximations will be presented, and in particular we propose an analysis that quantifies the sensitivity of model output with respect to each parameter.

All these methods have been applied to a flashing checkerboard experiment that will be the subject of the fourth part.

4.1 Model Formulation

4.1.1 Stochastic Differential Equations (SDE) Terminology

The physiological models we described in chapter 2 can be viewed as input-state-output systems [59]:

$$\begin{cases} \dot{x}(t) &= F(x(t), u(t), \theta) + \xi(t) \\ y(t) &= G(x(t), \theta) + \eta(t), \end{cases} \quad (4.1)$$

where

- The input u is the external stimulation driving neural activity.

- The state x denotes the set of all hidden variables associated with the hemodynamic process, such as metabolism, blood flow, venous deoxyhemoglobin content; in this chapter we will consider that there is an independent process at each spatial location (in practice each voxel), which means that x denotes the hidden variables at one particular location.
- The output y is the set of measures obtained by brain imaging techniques; in this chapter y will be the fMRI signal at the same location.
- F is the evolution function, which accounts for all physiological processes driving the system.
- ξ is an evolution noise, which accounts for random perturbations of these physiological processes; in the strict SDE formalism, it is assumed white and Gaussian.
- G is the measure function, explaining how the measure y depends on the hidden variables; in this chapter, it will describe the dependence of the MR signal on hemodynamic quantities.
- η is a measure noise, which accounts for physiological and physical effects randomly altering this measure; it is also assumed white and Gaussian.
- θ is a set of physiological and physical parameters on which evolution and measure functions depend.

We are particularly interested in the parameters of this system. Indeed, whereas F and G describe general dynamical properties of the physiological processes, it is the set of parameters θ that tunes this dynamic. These parameters can account for biological and physical properties of the body and measurement equipment, which may vary across experimentations, subjects or even cortical regions; they can also quantify the cognitive response of the brain to stimulation (e.g. intensity, latency of the activation). The problem of estimating these parameters from the measured data is a *system identification* problem, and it will be the topic of this chapter.

4.1.2 A First Model

The Balloon Model proposed by Buxton and al. [33] and completed by [59] (flow dynamic) describes the dynamics of a “flow inducing signal” s , the blood flow f , the blood venous volume v , the veins deoxyhemoglobin content q (these values are normalized and thus equal 1 at rest), and the BOLD signal y :

$$\begin{cases} \dot{s} &= \epsilon u - \kappa_s s - \kappa_f (f - 1) \\ \dot{f} &= s \\ \dot{v} &= \frac{1}{\tau} (f - v^{1/\alpha}) \\ \dot{q} &= \frac{1}{\tau} (f \frac{1 - (1 - E_0)^{1/f}}{E_0} - v^{1/\alpha - 1} q) \\ y &= V_0 (a_1 (1 - q) - a_2 (1 - v)). \end{cases} \quad (4.2)$$

Here, the set of hidden variables x is a four-elements vector

$$x = \begin{pmatrix} s \\ f \\ v \\ q \end{pmatrix}.$$

Please note that we did not make noises appear in the equation. Also note that the two first equations stand for a unique second order equation for the flow, but it was necessary to add the flow derivative $s = \dot{f}$ to the state variables to match the general formulation in (4.1):

$$\ddot{f} = \epsilon u - \kappa_s \dot{f} - \kappa_f (f - 1).$$

Let us describe the set of parameters θ . The neural efficiency parameter ϵ quantifies the activation since it determines to what extent the system does respond to stimulation. Then there are physiological parameters: the flow decay κ_s , the flow time constant κ_f , the venous transit time τ , Grubb's parameter α , the oxygen extraction at rest E_0 and the blood volume fraction at rest V_0 , which may vary across brain regions and across subjects; and scanner-dependent parameters: a_1 and a_2 (as we mentioned in the introduction, we will either use the known values for a_1 and a_2 at 1.5T, or suppose $a_2/(a_1 + a_2) = 1/10$ at 3T, which implies we need to estimate the product $b = V_0(a_1 + a_2)$).

4.1.3 System Stability

We are interested here in assessing the *stability* of a dynamical system: the hemodynamic main effect should be roughly a smoothing of the input; for that reason, there should not be any special dynamics like bifurcations, limit cycles... rather, when the input becomes constant, the system should converge toward a unique stable equilibrium point.

As an example, we shall examine here each equation of the Balloon Model above.

The flow dynamic equation is a pure linear damped oscillator. It can then be computed exactly by a convolution

$$f(t) = 1 + k * u(t).$$

If we assume $\Delta = \kappa_s^2 - 4\kappa_f < 0$, we have:

$$k(t) = \epsilon e^{-\frac{\kappa_s}{2}t} \cos\left(\frac{\sqrt{4\kappa_f - \kappa_s^2}}{2}t\right)$$

(if we had had $\Delta > 0$, k would have been of a different form, with exponentials only).

Since it is a linear convolution, the flow does not have any special dynamic (when the input is constant, the flow converges necessarily to the equilibrium point $1 + \epsilon u_0/\kappa_f$).

Volume only depends on flow. If we note $v(f) = f^\alpha$, \dot{v} in (4.2) has the same sign as $v(f) - v$. The equation looks like an exponential decay to steady state, though it is nonlinear. If the input is constant, the flow and the volume necessarily converge to their equilibrium points $(1 + \epsilon u_0/\kappa_f)$ and $(1 + \epsilon u_0/\kappa_f)^\alpha$. In a similar way, if we note $q(v, f) = f \frac{1 - (1 - E_0)^{1/f}}{E_0} v^{1-1/\alpha}$, \dot{q} is the same sign as $q(v, f) - q$. If the input is constant the deoxyhemoglobin content eventually converges to an equilibrium point.

As a conclusion, the Balloon Model (4.2) is stable since it will always converge toward an equilibrium when its input becomes constant.

From a more mathematical point of view, the system stability can be assessed through examining the eigenvalues of the Jacobian of the evolution function F at the equilibrium point x_0 . Again, we show this in the example.

When the system is driven by a constant input u_0 , then there is a unique x_0 such that $F(x_0, u_0, \theta) = 0$:

$$x_0 = \begin{pmatrix} 0 \\ 1 + \frac{\epsilon u_0}{\kappa_f} \\ \left(1 + \frac{\epsilon u_0}{\kappa_f}\right)^\alpha \\ \frac{1 - (1 - E_0)^{1/(1 + \frac{\epsilon u_0}{\kappa_f})}}{E_0} \left(1 + \frac{\epsilon u_0}{\kappa_f}\right)^\alpha \end{pmatrix}.$$

This equilibrium point is called *stable* when the system evolution converges back toward it after small perturbations. This is the case if and only if the Jacobian of the evolution function with respect to x has real negative or complex with negative-real-part eigenvalues. In our example, this Jacobian is:

$$\frac{\partial F}{\partial x} = \begin{pmatrix} -\kappa_s & -\kappa_f & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\tau} & -\frac{x_3^{1/\alpha-1}}{\alpha\tau} & 0 \\ 0 & \frac{\partial F_4}{\partial x_2} & \frac{1}{\tau} \left(1 - \frac{1}{\alpha}\right) x_3^{1/\alpha-2} x_4 & -\frac{x_3^{1/\alpha-1}}{\tau} \end{pmatrix},$$

with

$$\frac{\partial F_4}{\partial x_2} = \frac{1}{\tau} \left(\frac{1 - (1 - E_0)^{1/x_2}}{E_0} + \frac{\log(1 - E_0)(1 - E_0)^{1/x_2}}{E_0 x_2} \right),$$

and its eigenvalues evaluated at x_0 are

$$\left\{ -\frac{\kappa_s + \sqrt{\kappa_s^2 - 4\kappa_f}}{2}, -\frac{\kappa_s - \sqrt{\kappa_s^2 - 4\kappa_f}}{2}, -\frac{(1 + \frac{\epsilon u_0}{\kappa_f})^{1-\alpha}}{\alpha\tau}, -\frac{(1 + \frac{\epsilon u_0}{\kappa_f})^{1-\alpha}}{\tau} \right\}$$

(they can be obtained as follows: note that the matrix $\frac{\partial F}{\partial x}$ is trigonal by blocks with block sizes equal to 2, 1, and 1; the four eigenvalues are respectively the 2 eigenvalues of the first 2x2 block, and the third and fourth diagonal terms). Since the physiological parameters are always positive, either these eigenvalues are real and negative or they have negative real parts: the system is always stable around the equilibrium.

4.1.4 A Second Model

We will apply our system identification procedures not only to the Balloon Model above, but also to several enhancements of it, as they appear in [32]. Here we show how to bring these enhancements together in a single system formulation and how to simplify it.

Three more variables are considered: the metabolism (CMRO2) m becomes an independent variable instead of the flow-locked expression $f \frac{1 - (1 - E_0)^{1/f}}{E_0}$; the neural activity N is the output of a simple habituation model (with a neural inhibition I) instead of the stimulus-locked expression ϵu . Flow and metabolism are not described by an evolution equation any more, but as convolutions (denoted by $*$) of neural activity with gamma-variate functions:

$$\begin{cases} N &= \max(\epsilon u - I, -N_0) \\ \dot{I} &= \frac{1}{\tau_I}(\kappa_n N - I) \\ f &= 1 + (f_1 - 1) h_f(t - \delta_t) * N \\ m &= 1 + (m_1 - 1) h_m(t) * N \\ \dot{v} &= \frac{1}{\tau} (f - (v^{1/\alpha} + \tau_{visc} \dot{v})) \\ \dot{q} &= \frac{1}{\tau} (m - \frac{q}{v} (v^{1/\alpha} + \tau_{visc} \dot{v})) \\ y &= V_0(a_1(1 - q) - a_2(1 - v)), \end{cases} \quad (4.3)$$

with

$$\begin{cases} h_f(t) &= \frac{1}{6\tau_f} \left(\frac{t}{\tau_f}\right)^3 e^{-\frac{t}{\tau_f}} \\ h_m(t) &= \frac{1}{6\tau_m} \left(\frac{t}{\tau_m}\right)^3 e^{-\frac{t}{\tau_m}}. \end{cases}$$

Additional parameters are the steady-state neural activity N_0 , the inhibitory time constant τ_I , the inhibitory gain factor κ_n , the normalized CBF and CMRO2 responses to sustained activity f_1 and m_1 , the delay δ_t between CMRO2 and CBF responses, the widths τ_f and τ_m of the CBF and CMRO2 impulse responses, and the volume viscoelastic time constants τ_{visc}^+ and τ_{visc}^- (a hysteresis rule is authorized for the volume dynamics: the viscosity parameter τ_{visc} can take 2 different values whether $\frac{\partial v}{\partial t} > 0$ ($\tau_{visc} = \tau_{visc}^+$) or $\frac{\partial v}{\partial t} < 0$ ($\tau_{visc} = \tau_{visc}^-$)).

We notice that the new dynamical system includes temporal convolutions that cannot be expressed in terms of differential equations. We will need then to keep a hybrid form with both differential equations and convolutions, which means that the model does not strictly satisfy the SDE formalism. We shall mention later how to handle these convolutions for parameter estimation. For now we shall propose some simplifications of these equations.

First, the neural habituation model can be analytically solved in

$$N = \max(\epsilon h_N * u, -N_0),$$

with

$$h_N(t) = D - \frac{\kappa_n}{\tau_I} e^{-\frac{\kappa_n + 1}{\tau_I} t}$$

and D denotes the Dirac impulse at $t = 0$.

Secondly, to eliminate the indetermination that exists between the parameters ϵ , f_1 and m_1 , we will use two new parameters instead: $\xi = \epsilon(f_1 - 1)$ and $n = (f_1 - 1)/(m_1 - 1)$, which is the steady-state flow-metabolism ratio.

Thirdly, let us note that the volume evolution equation can be transformed in

$$\dot{v} = \frac{1}{\tau + \tau_{visc}} (f - v^{\frac{1}{\alpha}}) = \begin{cases} \frac{1}{\tau + \tau_{visc}^+} (f - v^{\frac{1}{\alpha}}) & \text{if } f^\alpha > v \\ \frac{1}{\tau + \tau_{visc}^-} (f - v^{\frac{1}{\alpha}}) & \text{if } f^\alpha < v. \end{cases}$$

Hence, the new Balloon Model formulation (4.3) is equivalent to

$$\begin{cases} \dot{v} &= \frac{1}{\tau + \tau_{visc}}(1 + \xi h_f * \max(h_N * u, -N_0) - v^{1/\alpha}) \\ \dot{q} &= \frac{1}{\tau}(1 + \frac{\xi}{n} h_m * \max(h_N * u, -N_0) - \frac{q}{v}(v^{1/\alpha} + \tau_{visc}\dot{v})) \\ y &= V_0(a_1(1 - q) - a_2(1 - v)). \end{cases} \quad (4.4)$$

4.1.5 A Note on Integration

In order to solve the differential equations, we used a discretization step of 5ms in our implementations of models. We did not need to use a more computationally costly Runge-Kutta method because physiological quantities, and hence the evolution function, are changing slowly.

Please also note that we always suppose that the variables are at rest state at the beginning of integration, which implied for our experiment that the fMRI acquisition starts after a sufficiently long resting period in the scanner.

4.2 Parameter Estimation

We describe here the more general parameter estimation framework. However, the procedure we chose assumes some simplifications of this framework, which we will justify. We will also describe other existing methods at different moments in this chapter.

Let us be reminded of the stochastic differential equations system:

$$\begin{cases} \dot{x}(t) &= F(x(t), u(t), \theta) + \xi(t) \\ y(t) &= G(x(t), \theta) + \eta(t), \end{cases}$$

The system noises ξ and η are white Gaussian noises:

$$\xi(t) \sim \mathcal{N}(0, Q) \quad \text{and} \quad \eta(t) \sim \mathcal{N}(0, R).$$

θ denotes the parameters of the system, and variances Q and R are hyper-parameters. We denote $\tilde{\theta} = \{\theta, Q, R\}$.

The most complete parameter estimation procedure would consist in determining the a posteriori probability distribution of parameters given the measure $p(\theta|y)$. It requires to have a priori distributions of parameters and hyper-parameters values, $p(\tilde{\theta})$, but if it is not the case, it is also possible to use degenerate non-informative a priori distributions. Then the a posteriori distribution is obtained by the Bayes rule:

$$p(\tilde{\theta}|y) = \frac{p(y|\tilde{\theta})p(\tilde{\theta})}{\int_{\tilde{\theta}'} p(y|\tilde{\theta}')p(\tilde{\theta}')}$$

Nevertheless, this calculation is highly difficult and cannot be obtained exactly, owing to the nonlinearity of the system. In particular, a complete estimation procedure would require the following steps:

- To take into account the SDE formulation in the computation of conditional probability $p(y|\tilde{\theta})$ - and in particular the presence of evolutive noise, it is necessary to use the Kalman filter. The extended Kalman filter, to be more accurate, since the F and G functions are nonlinear with respect to the state x . Anyway these functions will be linearized locally by the extended Kalman filter, so that all probabilities remain Gaussian.

- Since the system equations are not linear with respect to parameters either, the a posteriori distribution of θ cannot be obtained by the Gaussian calculus. The only known way to approach it is to span the whole space of possible parameter values, by the use of adequate Monte Carlo Markov Chains (MCMC).

These methods have a high computational cost, whereas we would rather need fast estimation procedures that could be applied in every voxels in a masked brain. It motivated us to use several simplifications that appeared to be reasonable.

4.2.1 The Least Square Approach

The major assumption we made was to ignore the evolution noise ξ . As a matter of fact, if the nonlinear effects of the model with respect to input and hidden states are small enough, it is possible to approximate this evolution noise by an additional coloured measure noise. Indeed, let us use a linear approximation, where functions F and G are replaced by matrices multiplications:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) + \xi(t) \\ y(t) &= C x(t) + \eta(t). \end{cases}$$

This system solves in:

$$\begin{aligned} x(t) &= \int_0^t e^{A(t-s)} B u(s) + \xi(s) ds \\ y(t) &= C \int_0^t e^{A(t-s)} B u(s) + \xi(s) ds + \eta(t) \\ &= [C \int_0^t e^{A(t-s)} B u(s) ds] + [C \int_0^t e^{A(t-s)} \xi(s) ds + \eta(t)]. \end{aligned} \quad (4.5)$$

The system output can thus be decomposed into two terms, the first one being the output of a system with no noise; the second one being a coloured Gaussian noise.

Although the models we use are nonlinear, we shall assume that their output can still be expressed as the sum of a deterministic term and a coloured Gaussian noise. This deterministic term is the result of integrating the differential system and is nonlinear with respect to input and parameters.

$$y = f(u, \theta) + e, \quad e \sim \mathcal{N}(0, \Sigma). \quad (4.6)$$

Besides, we do not try to compute the whole a posteriori distribution of parameters, but only to find the Maximum Likelihood estimate, i.e. θ that maximizes the conditional probability of the measured data:

$$\begin{aligned} \hat{\theta} &= \operatorname{argmax}_{\theta} p(y|\theta) \\ &= \operatorname{argmin}_{\theta} -\log p(y|\theta) \\ &= \operatorname{argmin}_{\theta} \frac{1}{2}(f(u, \theta) - y)^T \Sigma^{-1}(f(u, \theta) - y). \end{aligned}$$

At this point, it is necessary to have information on the noise temporal variance Σ . For example, it is possible to assume :

$$\Sigma = \sigma^2 \Sigma_0,$$

where Σ_0 determines the autocorrelation, possibly estimated from the data, and σ^2 is an unknown parameter. Under such assumption, we need to minimize the energy

$$\mathcal{E}(\theta) = (f(u, \theta) - y)^T \Sigma_0^{-1} (f(u, \theta) - y). \quad (4.7)$$

To minimize this quadratic energy $\mathcal{E}(\theta)$, we use a Levenberg Marquardt algorithm [149, 132], implemented in the Matlab function 'lsqcurvefit'. The algorithm needs at each iteration step the Jacobian $\frac{\partial f}{\partial \theta}$ of the system output with respect to the parameters. We will show now how to compute this derivation exactly by solving a new differential system issued from the system definition.

4.2.2 Derivative of the System Output with respect to Parameters

$f(u, \theta)$ denotes the output (i.e. the set of all $y(u, \theta, t)$) of the deterministic system

$$\begin{cases} \dot{x}(u, \theta, t) &= F(x(u, \theta, t), u(t), \theta) \\ y(u, \theta, t) &= G(x(u, \theta, t), \theta). \end{cases} \quad (4.8)$$

Since $x(t, u, \theta)$ and $y(t, u, \theta)$ are defined by a differential system, it is also possible to compute $\frac{\partial x}{\partial \theta}(t, u, \theta)$ and $\frac{\partial y}{\partial \theta}(t, u, \theta)$ with a new differential system .

Differentiating both sides of (4.8) with respect to θ , we get the new system

$$\begin{cases} \frac{\partial \dot{x}}{\partial \theta} &= \frac{\partial F}{\partial x}(x, u(t), \theta) \frac{\partial x}{\partial \theta} + \frac{\partial F}{\partial \theta}(x, u(t), \theta) \\ \frac{\partial y}{\partial \theta} &= \frac{\partial G}{\partial x}(x, \theta) \frac{\partial x}{\partial \theta} + \frac{\partial G}{\partial \theta}(x, \theta). \end{cases}$$

This system can be integrated numerically, using the initial conditions $\frac{\partial x}{\partial \theta}(t=0) = 0$ (at time $t=0$, the state variables are at rest and do not depend upon θ), which leads to an exact evaluation of $\frac{\partial f}{\partial \theta} = (\frac{\partial y}{\partial \theta}(t, u, \theta))_{0 \leq t \leq T}$.

We saw above, on an example model formulations (4.4), that the evolution function F can entail temporal convolutions of the input u . It is actually not a problem, since the derivative of such terms with respect to parameters is easily given by:

$$\frac{\partial(H * u)}{\partial \theta} = \frac{\partial H}{\partial \theta} * u + H * \frac{\partial u}{\partial \theta} = \frac{\partial H}{\partial \theta} * u.$$

It is more computationally efficient and rigorous to use this method than computing this derivative by finite differences. Indeed, the new differential system state, $\frac{\partial x}{\partial \theta}$, is a n_x (number of hidden states) by n_p (number of parameters) matrix, and it is faster to integrate the new system (4.2.2) once than integrating at least n_p times the initial system equation (4.1) with an n_x vector state.

4.2.3 Handling Confound Effects

It is often useful when estimating parameters to ignore a certain set of time course components in real datasets, such as low frequencies.

Let us note C the matrix whose columns are the undesirable components. Then $p_C = I - C(C^T C)^{-1} C^T$ is the projector orthogonal to these confounds. Ignoring them consists in fitting $p_C f(u, \theta)$ to $p_C y$ instead of fitting $f(u, \theta)$ to y .

The new energy to minimize with respect to θ writes

Exact parameter values used for simulation						
	ϵ	κ_s	κ_f	τ	α	E_0
	.1	.65	.4	1	.4	.4
Estimation on exact model output						
initial guess	.2	1	1	2	.6	.6
2 iterations	.41	1.24	.97	1.31	.56	.49
8 iterations	.32	1	.69	1.10	.6	.49
50 iterations	.11	.65	.4	.92	.47	.42
150 iterations	.1	.65	.4	.99	.41	.4
Estimation on noisy model output						
initial guess	.2	1	1	2	.6	.6
2 iterations	.4	1.32	.98	1.25	.55	.48
8 iterations	.36	1.15	.72	.99	.62	.47
50 iterations	.23	.6	.32	.41	.65	.52
150 iterations	.27	.59	.31	.35	.64	.56

Table 4.1: Parameter values corresponding to estimations in figure 4.1.

$$\mathcal{E}(\theta) = (p_C(f(u, \theta) - y))^T \Sigma_0^{-1} (p_C(f(u, \theta) - y)),$$

and the gradient of $p_C f(u, \theta)$ against parameters for using Levenberg-Marquardt algorithm is $p_C \frac{\partial f}{\partial \theta}(u, \theta)$, with $\frac{\partial f}{\partial \theta}(u, \theta)$ computed as explained above.

4.2.4 Simulations

We show some example estimations on simulated data. We used a stimulation time course with many different durations and inter-stimulus intervals, sampled from suitable exponential distributions, in order to elicit various response patterns. The system input was thus the box-car function displayed in the upper left-hand corner of figure 4.1, and we used the model (4.2) to generate the synthetic BOLD response. We generated two time courses: one is the deterministic output of the system without noise; the other is the output of the same system, where we included some evolution and measure noises (measure noise standard deviation was 40% of that of the model output), and we used the parameter values found in [59]. Parameters ϵ , κ_s , κ_f , τ , α and E_0 were estimated, while measure parameters V_0 , a_1 and a_2 were fixed to the values indicated in [33].

The figure 4.1 shows several iterations of the estimation, starting with initial parameter values distant from the solution. Estimated values are displayed in table 4.1. A first estimation was run to fit the deterministic simulation (part B of the figure): then the true parameter values could be found exactly. A second estimation was run to fit the noisy simulation (part C of the figure). In this example, it is notable that the estimated parameters were quite different from the true ones, whereas the fitted response is quite similar to the deterministic simulation (both time courses are compared in part D of figure 4.1: the standard deviation of the difference is 9% of that of either signal), which means that different parameter sets can produce similar dynamics. Note however that some parameters however are estimated correctly: for example, we shall prove later in this chapter that parameters κ_f and κ_s estimates are more accurate than others in general.

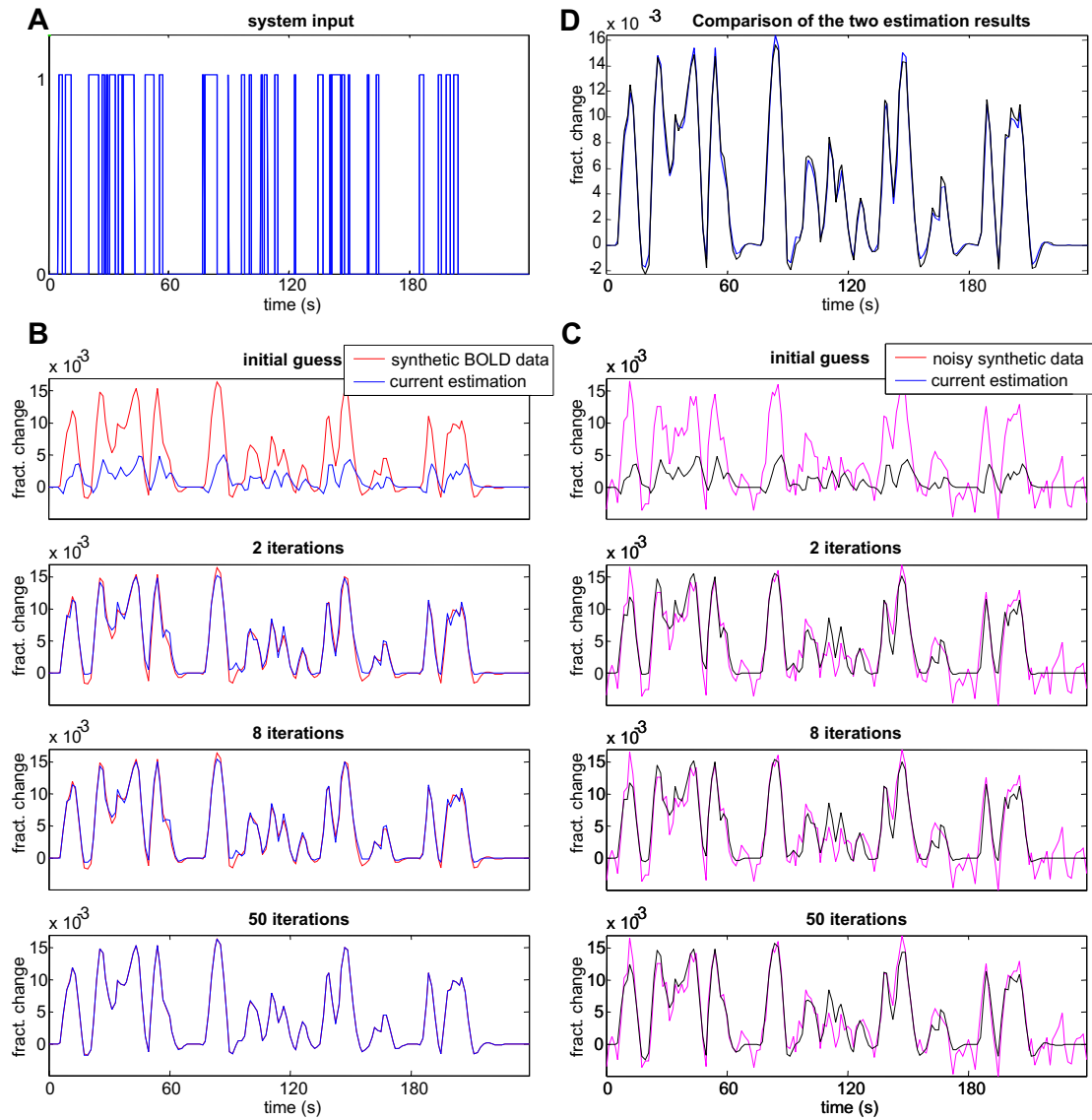


Figure 4.1: Parameter estimation on simulated data. (A) input used for the stimulation. (B) several steps of the estimation procedure on the noiseless model output. (C) same steps for the estimation on a model output with evolution and measure noises. (D) comparison between the last fitting curves of the two estimations: they are very similar though the parameter estimate are much different.

4.2.5 Including Evolutive Noise: Kalman Filtering

Here we give a brief insight into how it is possible to estimate the hidden state time courses, and use this information to compute the likelihood of the measured data. This methodology was used in [153] in order to estimate the model parameters.

Extended Kalman Filter: Estimating Hidden-State Time Courses

The *Kalman filter* is used to estimate the hidden states of a SDE system given its input and output. Originally it is intended to be applied to linear systems, and then allows to determine explicitly the (Gaussian) distribution of the hidden states at every time t given all the measures of outputs until t . For nonlinear models, the *extended Kalman filter* consists in approximating the system equations to their Taylor expansion at the first order at each iteration step. We will describe these techniques in more details' in chapter 8, and we also recommend the lecture notes by Max Welling [180]. Here, we shall summarize the main results and show an example with the same simulation as above.

For convenience' sake, we shall write the system in a discretized form:

$$\begin{cases} x_{k+1} &= x_k + F(x_k, u_k, \theta)dt + \xi_k \\ y_k &= G(x_k, \theta) + \eta_k, \end{cases}$$

with

$$\xi_k \sim \mathcal{N}(0, Q) \quad \text{and} \quad \eta_k \sim \mathcal{N}(0, R).$$

The Kalman filter is an iterative algorithm that estimates in turn the probability distributions of x_k given measures before instant k , and that of x_k given measures until instant k . The two repeated procedures are called *evolution update* and *measure update*. All these probabilities are always Gaussian, thanks to the linearization procedures:

$$\begin{aligned} p(x_k | y_1, \dots, y_{k-1}, \tilde{\theta}) &\sim \mathcal{N}(\hat{x}_k^{k-1}, P_k^{k-1}) \\ p(x_k | y_1, \dots, y_k, \tilde{\theta}) &\sim \mathcal{N}(\hat{x}_k^k, P_k^k) \end{aligned}$$

To compute the distribution of hidden-states given all measures (past and future), $p(x_k | y, \tilde{\theta})$, it is necessary to apply a second backward iterative algorithm called the *Kalman smoother*.

$$p(x_k | y_1, \dots, y_n, \tilde{\theta}) \sim \mathcal{N}(\hat{x}_k^n, P_k^n)$$

Kalman Likelihood of the Measure

We want to derive the likelihood of the measure given the parameters and hyper-parameters, $p(y | \tilde{\theta})$, without ignoring evolutive noise. It is possible to derive it by making the hidden-state distributions estimated by the Kalman filter appear in its development [180]. It consists in splitting the likelihood into the conditional likelihoods of every individual measure:

$$\begin{aligned}
p(y|\theta) &= p(y_1|\tilde{\theta}) p(y_2|y_1, \tilde{\theta}) p(y_3|y_1, y_2, \tilde{\theta}) \dots p(y_n|y_1 \dots y_{n-1}, \tilde{\theta}) \\
&= \prod_{k=1}^n p(y_k|y_1 \dots y_{k-1}, \tilde{\theta}) \\
&= \prod_{k=1}^n \int p(y_k, x_k|y_1 \dots y_{k-1}, \tilde{\theta}) dx_k \\
&= \prod_{k=1}^n \int p(y_k|x_k, y_1 \dots y_{k-1}, \tilde{\theta}) p(x_k|y_1 \dots y_{k-1}, \tilde{\theta}) dx_k \\
&= \prod_{k=1}^n \int p(y_k|x_k, \tilde{\theta}) p(x_k|y_1 \dots y_{k-1}, \tilde{\theta}) dx_k.
\end{aligned}$$

Since the Kalman filter allows to determine the Gaussian probability

$$p(x_k|y_1 \dots y_{k-1}, \tilde{\theta}) \sim \mathcal{N}(\hat{x}_k^{k-1}, P_k^{k-1}),$$

and since we have

$$p(y_k|x_k, \tilde{\theta}) \sim \mathcal{N}(G(x_k), R),$$

it is possible to show, after approximating G at the first order around \hat{x}_k^{k-1} :

$$\begin{aligned}
G(x_k) &= G(\hat{x}_k^{k-1}) + C(x_k - \hat{x}_k^{k-1}) \\
C &= \frac{\partial G}{\partial x}(\hat{x}_k^{k-1}),
\end{aligned}$$

that the terms inside the product have the Gaussian distributions

$$p(y_k|y_1 \dots y_{k-1}, \tilde{\theta}) \sim \mathcal{N}(G(\hat{x}_k^{k-1}), C R + P_k^{k-1} C^T).$$

This proves that the likelihood of the measure $p(y|\tilde{\theta})$ can be computed. It is then possible to estimate the model parameters and hyper-parameters by maximizing this likelihood.

We propose an intuitive explanation of the difference between the ‘‘deterministic’’ likelihood we used in (4.2.1) for our estimation procedure and this new Kalman likelihood. In the first case, the hidden-states are supposed to depend deterministically on the model input and model parameters, so that the likelihood only expresses how distant the predicted fMRI signal is from the actual measure. In the second case, hidden-states are allowed to be driven by random fluctuations, so that computing the likelihood consists first in estimating these fluctuations with the Kalman filter, and then the likelihood expresses how important these fluctuations are, and how distant the fMRI signal predicted by the estimated hidden-states is from the actual measure.

Algorithm

We summarize the estimation procedure using the Kalman likelihood

1. Start with a first guess of parameters and hyper-parameters.

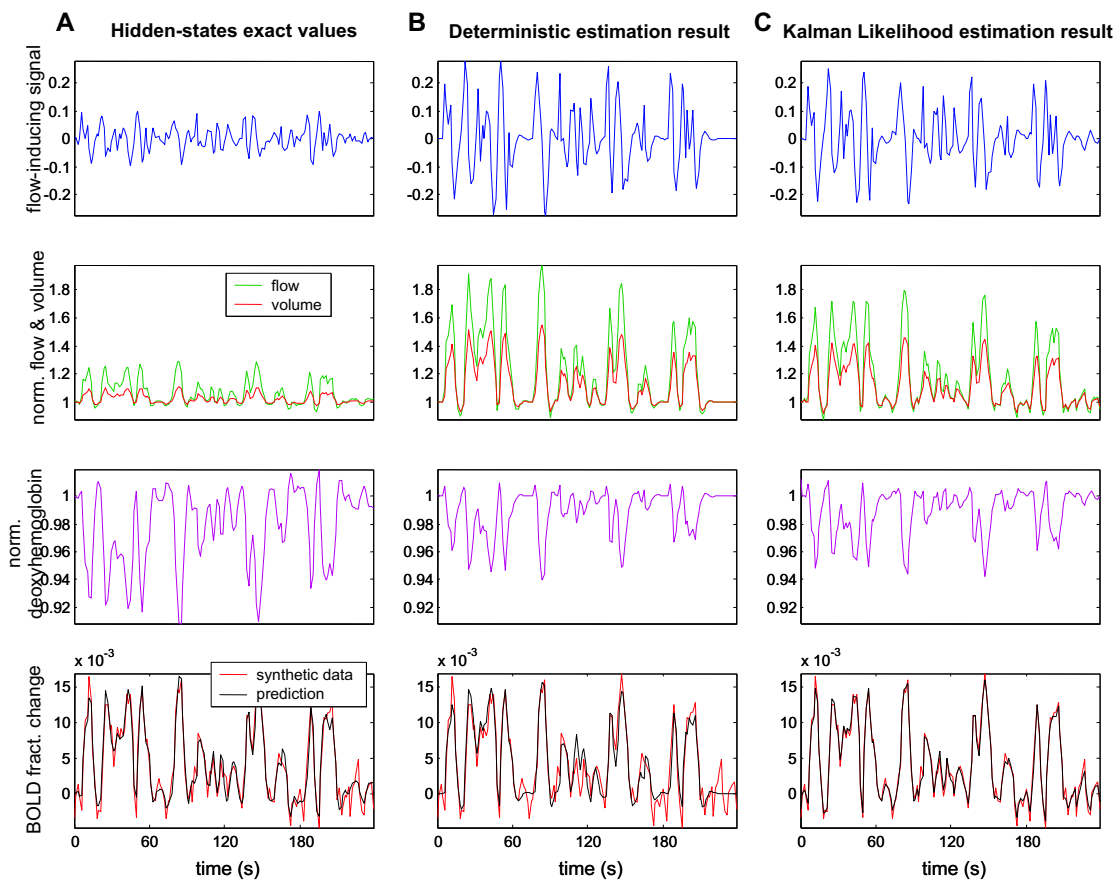


Figure 4.2: Comparison of the hidden-state time courses. (A) actual time courses from the simulation. (B) prediction using the deterministic parameter estimates. (C) Kalman filter estimation using the Kalman likelihood parameter estimates.

2. Estimate the hidden-state a posteriori distribution with the extended Kalman filter.
3. Compute the Kalman likelihood of the measured data.
4. Update the estimation: this is achieved using the simplex search method [110] implemented in the Matlab function `fminsearch`.
5. Steps 2-4 are repeated until the simplex diameter is less than a specified tolerance.

Simulations

Figure 4.2 shows an example estimation on the same simulation as in figure 4.1, and table 4.2 shows the estimated parameter values. In the column (A) of the figure, the three first displays show the actual hidden-states time courses from the simulation, and the last display compares the fMRI signal predicted by volume and deoxyhemoglobin time course (in black) with the actual time course obtained through adding measure noise (in red). In the column (B) are displayed the hidden-states obtained with the deterministic estimation in the last section - by applying the system equations to the input and the estimated parameters - and its bottom panel compares the fMRI signal predicted by

Exact parameter values used for simulation					
ϵ	κ_s	κ_f	τ	α	E_0
.1	.65	.4	1	.4	.4
Deterministic estimation on noisy model output					
.27	.59	.31	.35	.64	.56
Kalman likelihood estimation on noisy model output					
.25	.63	.34	.46	.65	.53

Table 4.2: Comparison of parameter estimates obtained by the deterministic and Kalman likelihood methods (see figure 4.2).

these hidden-states with the noisy simulation result. The column (C) shows hidden-states estimated by the Kalman filter after Kalman likelihood estimation of parameters and hyper-parameters has been achieved and again compares the estimated versus actual fMRI courses.

It first appears in this figure that both estimation procedures did not estimate accurately the hidden states: the response of the flow to the stimulation, characterized by the neural efficiency parameter ϵ is too strong, even though it leads to variations of deoxyhemoglobin and BOLD with a correct amplitude, thanks to the tuning of other parameters. Besides, some details of the actual hidden-state time courses were found by the Kalman method and not by the simpler method. For example, at the end of the stimulation period, the deterministic estimation returns to baseline because there is no more input (first display in column (B)), but actually there was still some random ongoing activity (same display in (A)) that the Kalman method could partly recover (in (C)). Some more subtle details can also be found.

On the other hand, the Kalman prediction of the BOLD response overfits the noisy stimulation: it is closer to it (last display in (C)) than the prediction from actual hidden-states (in (A)). And the parameter estimation accuracy is not better than that of the first estimation.

As a result, we did not find a striking superiority of the Kalman likelihood methods, as far as parameters are concerned. On the contrary, these methods imply a large number of degrees of freedom since the whole hidden-states time courses are estimated along with model parameters and hyper-parameters, which makes the appreciation of results more difficult. Moreover they are less computationally efficient, since they require to run a Kalman filter at each iteration step, and since there is no easy way to compute the energy gradient with respect to parameters as we did in section 4.2.2. These are the reasons why we chose to use our Maximum Likelihood estimation procedure that ignores evolution noises.

4.3 Quantifying Estimation Precision

We saw in our simulations that parameter estimation encounters a serious obstacle in terms of accuracy of the obtained values. It seems that quite different parameter sets can produce very similar system output. This asks the question of the *system identifiability*: do we have enough information once we know the system input u and output y to determine the parameter values? Is there a unique solution θ to the equation $y = f(u, \theta)$?

More precisely, the question is rather to quantify the identifiability or non-identifiability: how much can we change θ while keeping $f(u, \theta)$ close enough to y ? Indeed, there might be only one solution to the equation $y = f(u, \theta)$, since we saw, when the estimation was performed of a noiseless simulation, that it was able to find the exact values of parameters. But in the same time, it appears that the effects on the output of some of the parameters do interfere with those of others, making the system little identifiable.

The extreme case with a non-unique solution would be for example if the scale factor on the input (neural efficacy ϵ) and that on the output (V_0) were estimated from data with an input low enough to make the linear approximation of the model hold. Indeed, increasing the first could be compensated by decreasing the second by the same factor to produce exactly the same output. It would not be possible to estimate these 2 parameters, but only their product.

We propose here an analysis that inspects the sensitivity of the model to variations of parameters. This analysis, since it is based on a linearization of the system around the Maximum Likelihood estimate, leads naturally to a framework for deriving a posteriori probabilities of the parameters, according to the Bayesian principles. However we must keep in mind that this linearization with respect to parameters is a coarse simplification.

4.3.1 A Sensitivity Analysis

We want to investigate to what extent the system output is sensitive to changes in one given parameter. Let us note θ_i this parameter, and θ_2 the rest of parameters ($\theta = \{\theta_i, \theta_2\}$). Let us also notate $J = \frac{\partial f}{\partial \theta}(u, \theta)$ the derivative of system output (obtained according to the calculation in section 4.2.2, J_i its i^{th} column and J_2 the matrix consisting of the remaining columns. For a small parameter change $d\theta$ we have

$$f(u, \theta + d\theta) = f(u, \theta) + Jd\theta = f(u, \theta) + J_i d\theta_i + J_2 d\theta_2.$$

For a small change $d\theta_i$ of θ_i , f varies by $J_i d\theta_i$; however, if J_i is not orthogonal to the other Jacobian components J_2 , part of this variation can be compensated by a change in the other parameters: $d\theta_2 = -J_2^+ J_i d\theta_i$, where $J_2^+ = (J_2^T J_2)^{-1} J_2^T$ denotes the pseudo-inverse of J_2 . We then have:

$$\min_{d\theta_2} \|f(u, \theta + d\theta) - f(u, \theta)\| = \|(I - J_2 J_2^+) J_i d\theta_i\| = \pi_i |d\theta_i|,$$

with

$$\pi_i = \|(I - J_2 J_2^+) J_i\| = \sqrt{J_i^T (I - J_2 J_2^+) J_i}.$$

π_i denotes how much the system output is sensitive to variations of parameter θ_i . Hence, the bigger π_i , the more identifiable θ_i is. This also means that, for a given percentage x , if θ_i changes less than $\pi_i^{-1} x \|f(u, \theta)\|$, one can adjust the other parameters θ_2 to make the model output vary by less than $x\%$. Given an input u and an initial parameter set θ_0 , our sensitivity analysis consists in considering the sensitivity intervals $[\theta_{0i} - \pi_i^{-1} x \|f(u, \theta_0)\|, \theta_{0i} + \pi_i^{-1} x \|f(u, \theta_0)\|]$, with $x = 1$ to 5%. They are not confidence intervals for parameter estimation. Rather they indicate that the system output is very little sensitive to changes of θ_i inside these intervals. π_i can actually be expressed in a simpler manner from the Fisher information matrix $J^T J$, as we shall show below:

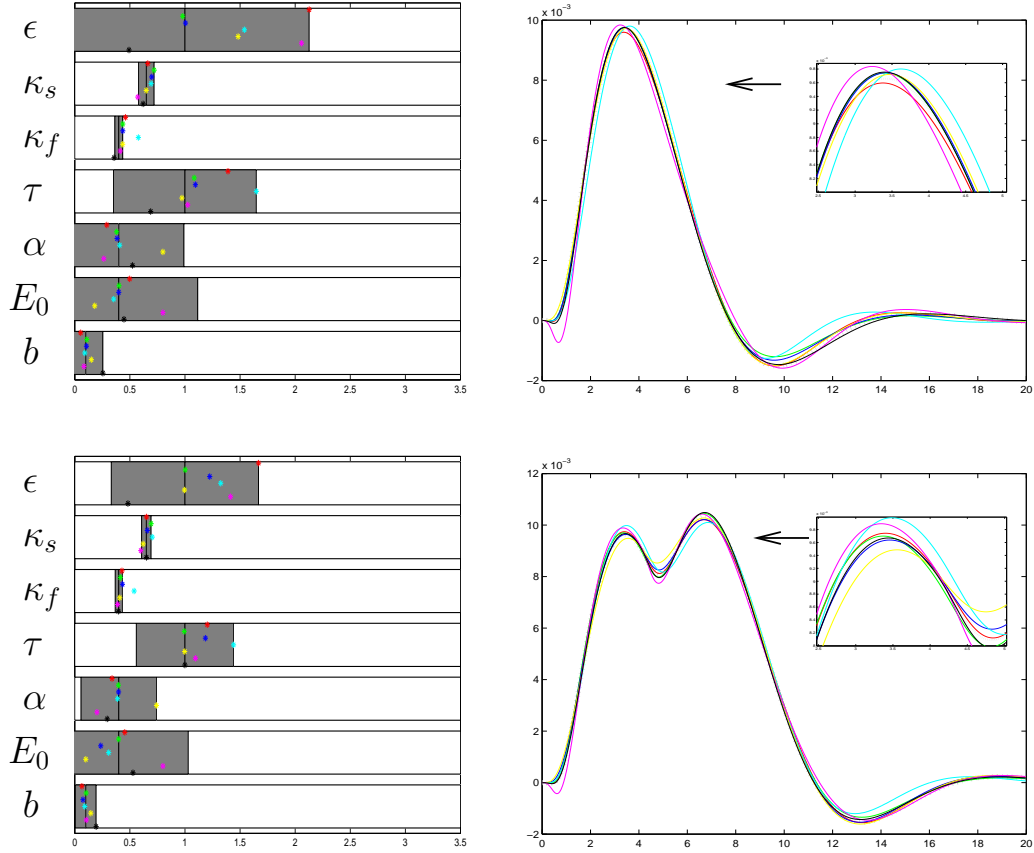


Figure 4.3: Sensitivity analysis for two different system inputs around a given θ_0 . Top: response to an impulse. Bottom: response to 2 consecutive impulses. Left: 2%-signal-change sensitivity intervals - colour stars show different parameter sets with one parameter constrained to be at the edge of its sensitivity interval (e.g., red corresponds to ϵ being fixed and the other parameters computed with $d\theta_2 = -J_2^+ J_i d\theta_i$). Right: output variations for these parameter sets compared to the output for the reference θ_0 (bold dashed line). Values of all parameters and percentages of signal changes are given in table 4.3.

$$\pi_i = \frac{1}{\sqrt{((J^T J)^{-1})_{ii}}}.$$

Figure 4.3 shows such a sensitivity analysis with $x = 2\%$ for two different inputs (a single impulse and two successive impulses). The sensitivity intervals are represented in the left column of the figure. Each of the seven parameters (encoded with different colours) was fixed to one of the two bounds of its sensitivity interval; then the values of the other six parameters were computed from $d\theta_2 = -J_2^+ J_i d\theta_i$ and resulting time course were plotted (with the same colour code). The figure clearly shows that very different parameter sets can result in very similar system outputs (table 4.3 shows the obtained parameter sets and the output variations). It also appears that the sensitivity depends on the input complexity: in the second case the parameters are more identifiable, because the effects of the different parameters can be more diverse and hence less correlated. For that reason, the experimental design we present later uses a large panel of ISI and stimulus duration to increase identifiability.

parameter	ϵ	κ_s	κ_f	τ	α	E_0	b	% output change
θ_0	1	.65	.4	1	.4	.4	.1	
ϵ fixed	2.13	.66	.46	1.39	.29	.5	.05	2.1%
κ_s fixed	.98	.72	.44	1.08	.38	.4	.11	3.3%
κ_f fixed	1	.7	.44	1.1	.38	.4	.11	2.3%
τ fixed	1.54	.69	.58	1.65	.41	.35	.09	9.4%
α fixed	1.48	.65	.44	0.97	.8	.18	.15	2.2%
E_0 fixed	2.06	.57	.41	1.03	.26	.8	.08	7.9%
$V_0(a_1 + a_2)$ fixed	0.49	.62	.36	.69	.52	.45	.26	2%

parameter	ϵ	κ_s	κ_f	τ	α	E_0	b	% output change
θ_0	1	.65	.4	1	.4	.4	.1	
ϵ fixed	1.67	.65	.43	1.2	.34	.45	.07	1.48%
κ_s fixed	1	.69	.41	0.99	.4	.4	.1	1.7%
κ_f fixed	1.22	.66	.43	1.19	.4	.24	.07	1.4%
τ fixed	1.32	.7	.54	1.44	.39	.31	.09	4.5%
α fixed	1	.62	.41	1	.74	.1	.15	2.9%
E_0 fixed	1.41	.6	.39	1.1	.2	.8	.1	3.8%
$V_0(a_1 + a_2)$ fixed	0.48	.65	.4	1	.29	.53	.19	1.3%

Table 4.3: Parameter values for the sensitivity analysis in figure 4.3: quite different parameter sets can lead to similar system outputs. The output variation is not exactly 2% when one parameter is fixed to the edge of its sensitivity interval, because these intervals were calculated using first order approximations with respect to parameters.

As a precaution, we must finally insist upon the fact that we only discussed identifiability at a local scale, i.e. we only considered one minimum of the energy and approximated locally the shape of possible model outputs with the tangent plane. However, since the shape can be more complicated, indetermination can be even worse than the one resulting from the discussion above.

4.3.2 A Posteriori Probabilities on Parameters in a Linearized Framework

To estimate the a posteriori distribution of the model parameters, we will use the same assumption as above, i.e. ignore the evolutive noise and only consider a coloured measure noise e . Of course, to derive an a posteriori distribution, first it is necessary to have an a priori information. However, we can use a flat a priori distribution, or even a non-informative degenerate uniform distribution. See [97] on the use of such methods in fMRI data analysis and [22] for more theoretical details.

$$y = f(u, \theta) + e, \quad \begin{array}{l} e \sim \mathcal{N}(0, \Sigma) \\ \theta \sim \mathcal{U}(\mathbb{R}). \end{array}$$

At this point, it is necessary to use the same linearization of the system output as above with respect to parameters:

$$f(u, \theta) = f(u, \hat{\theta}) + J_{\hat{\theta}}(\theta - \hat{\theta}). \quad (4.9)$$

Then we can calculate the a posteriori distribution of parameter θ using Bayesian

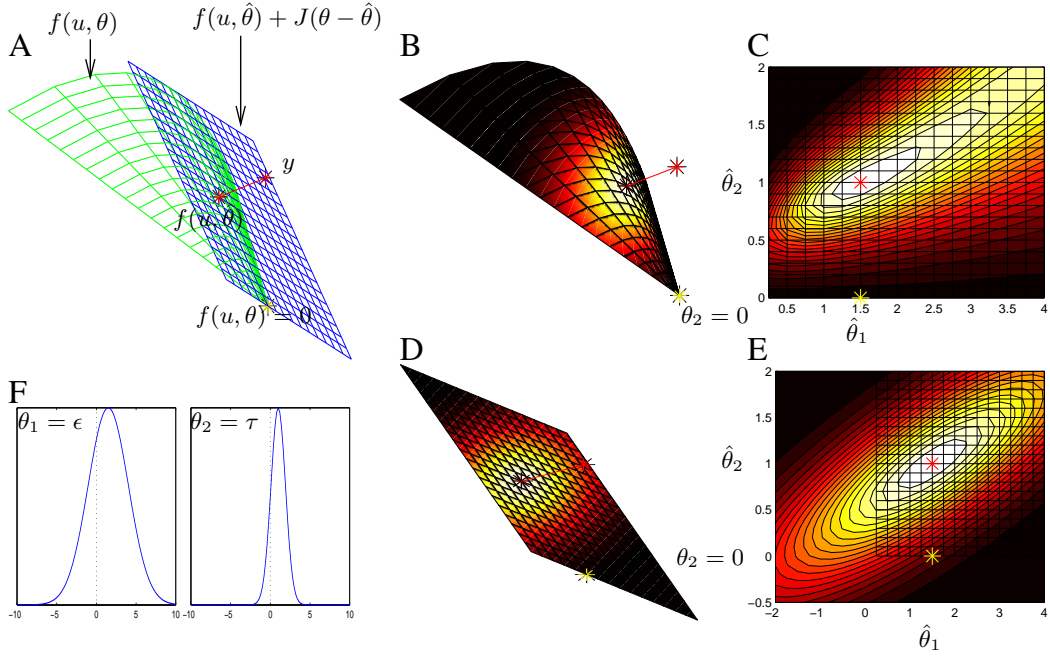


Figure 4.4: Geometrical interpretation of a-posteriori parameter distributions. (A) The set of all possible model outputs $f(u, \theta)$ when θ varies is a manifold whose dimension is the number of parameters. The surface displayed here was obtained by varying two Balloon Model parameters: ϵ and τ ; each closest point $f(u, \theta)$ is represented by its values at 3 selected instants. The maximum likelihood estimate $\hat{\theta}$ is such that $f(u, \hat{\theta})$ is the closest point on the surface to measured data y . Then, in our linearization, the manifold is approximated to its tangent surface at point $f(u, \hat{\theta})$. (B) Actual a posteriori distribution of θ on the surface. (C) The same distribution is displayed in the θ space. (D) The a posteriori distribution of θ when approximating the manifold to its tangent surface. (E) The same distribution is displayed in the θ space: its variations are similar to those in (C) around $\hat{\theta}$, but become wrong for distant values of θ . (F) Marginal a posteriori distributions of $\theta_1 = \epsilon$ and $\theta_2 = \tau$ under the approximation: they are pretty flat. Also, they allow negative values for the parameters, whereas actually, $\epsilon < 0$ is much less probable and $\tau < 0$ is impossible.

inference. It is important however to understand what this linearization will mean for the a posteriori probabilities. Figure 4.4 shows that the obtained probabilities might be wrong for θ distant from $\hat{\theta}$, and, as a consequence, so might be the normalization constant.

We notate $\tilde{y} = y - f(u, \hat{\theta})$. A posteriori distribution can be obtained from Bayes' rule:

$$\begin{aligned}
 p(\theta|y) &= \frac{p(y|\theta)p(\theta)}{p(y)} \\
 &\propto p(y|\theta)p(\theta) \\
 &\propto e^{-\frac{1}{2\sigma^2}(\tilde{y}-J(\theta-\hat{\theta}))^T\Sigma^{-1}(\tilde{y}-J(\theta-\hat{\theta}))} \\
 &\propto e^{-\frac{1}{2\sigma^2}(\theta-\hat{\theta}-(J^T\Sigma^{-1}J)^{-1}J^T\Sigma^{-1}\tilde{y})^T(J^T\Sigma^{-1}J)(\theta-\hat{\theta}-(J^T\Sigma^{-1}J)^{-1}J^T\Sigma^{-1}\tilde{y})}.
 \end{aligned}$$

We note that $J^T\Sigma^{-1}\tilde{y}$ is actually zero, by definition of $\hat{\theta}$. Indeed, if derivate the energy $\mathcal{E}(\theta)$ in (4.7) we get:

$$\frac{\partial \mathcal{E}(\theta)}{\partial \hat{\theta}} = 2(y - f(\hat{\theta}))^T \Sigma^{-1} J;$$

and since $\hat{\theta}$ minimizes \mathcal{E} , this derivative equals zero when evaluated on $\hat{\theta}$.

We thus found that under the linearization procedure, θ follows a Gaussian distribution:

$$(\theta|y) \sim \mathcal{N}(\hat{\theta}, \sigma^2(J^T \Sigma^{-1} J)^{-1}). \quad (4.10)$$

We are now interested in the marginal distribution of parameter θ_i . It is still Gaussian with mean and variance obtained in a straightforward manner from the above equation:

$$(\theta_i|y) \sim \mathcal{N}(\hat{\theta}_i, (\sigma^2(J^T \Sigma^{-1} J)^{-1})_{ii}).$$

But it can be computed too by integrating the θ distribution with respect to the other parameters (we suppose that $\Sigma = \sigma^2 I$ for convenience' sake):

$$\begin{aligned} p(\theta_i|y) &= \int p(\theta_i, \theta_2|y) d\theta_2 \\ &\propto \int e^{-\frac{1}{2\sigma^2}(\theta - \hat{\theta})^T J^T J(\theta - \hat{\theta})} d\theta_2 \\ &\propto \int e^{-\frac{1}{2\sigma^2}(J_i(\theta_i - \hat{\theta}_i) + J_2(\theta_2 - \hat{\theta}_2))^T (J_i(\theta_i - \hat{\theta}_i) + J_2(\theta_2 - \hat{\theta}_2))} d\theta_2 \\ &\propto \int e^{-\frac{1}{2\sigma^2}(\theta_2 - \hat{\theta}_2 + (J_2^T J_2)^{-1} J_2^T J_i(\theta_i - \hat{\theta}_i))^T J_2^T J_2(\theta_2 - \hat{\theta}_2 + (J_2^T J_2)^{-1} J_2^T J_i(\theta_i - \hat{\theta}_i))} \\ &\quad e^{-\frac{1}{2\sigma^2}(\theta_i - \hat{\theta}_i)^T J_i^T (I - J_2(J_2^T J_2)^{-1} J_2^T) J_i(\theta_i - \hat{\theta}_i)} d\theta_2 \\ &\propto e^{-\frac{1}{2\sigma^2}(\theta_i - \hat{\theta}_i)^T J_i^T (I - J_2 J_2^+) J_i(\theta_i - \hat{\theta}_i)} \\ (\theta_i|y) &\sim \mathcal{N}(E(\theta)_i, \sigma^2(J_i^T (I - J_2 J_2^+) J_i)^{-1}) \\ &\sim \mathcal{N}(E(\theta)_i, \sigma^2 \pi^{-2}). \end{aligned}$$

This way, it appears that our sensitivity factor π_i is proportional to the a posteriori precision (which is the variance inverse) of parameter θ_i in a Bayesian framework when we assume a white measure noise. The difference is that the sensitivity factor describes dynamical system properties, and as such does not depend upon the quantity of noise in measured data. We also proved the formula mentioned above:

$$\pi_i = \frac{1}{\sqrt{((J^T J)^{-1})_{ii}}}.$$

4.4 Experimental Results

4.4.1 Experimental Design

We conducted fMRI experiments in order to question the validity of the Balloon model and the estimation and sensitivity analysis described above. The stimulus consisted of a full screen binocular flashing checkerboard (12 Hz). A red cross fixation point was used throughout the experiment. Resting condition consisted of a grey screen with the fixation cross. Eight volunteers were used for this study (6 males and 2 females, from 19 to 25 years old, with no vision problem). Brain anatomy and fMRI images were acquired in the La Timone Hospital, Marseille, France, on a 3T scanner with surface

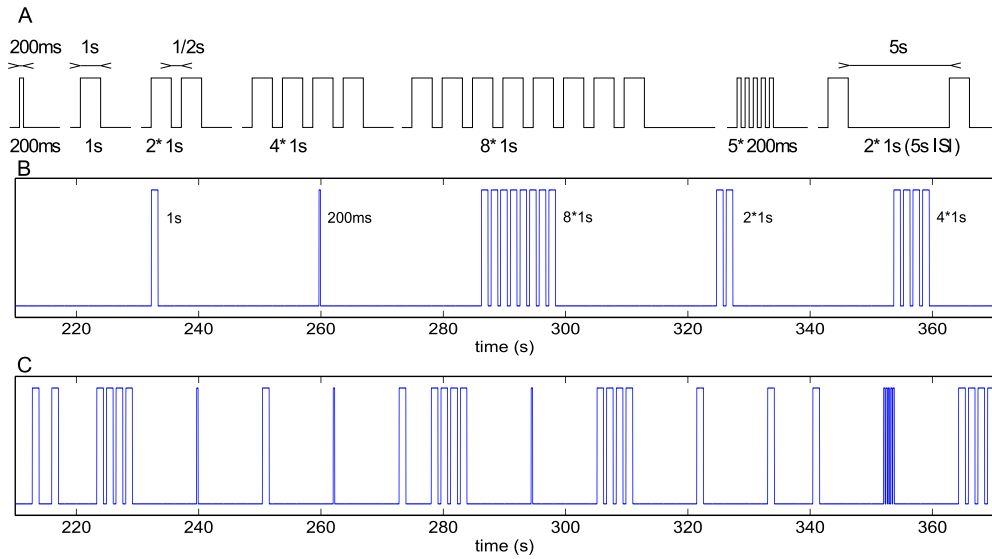


Figure 4.5: Experimental design. (A) 7 stimulation designs (B) first paradigm (stimulations are separated by 25s' rest periods). (C) second paradigm (stimulation are separated by 10s' rest periods or less).

coil. The functional scans consisted in 11 coronal occipital slices, each voxel being $2 \times 2 \times 2$ mm, with inter-scan interval $TR = 825$ ms.

In order to test the validity of the Balloon Model, we wanted to use an experimental design that operates in the nonlinear regimes of the BOLD fMRI response with respect to stimulus duration and inter-stimulus-spacing. Moreover, we wanted these nonlinearities to be due mostly to vascular effects, and minimize neuronal nonlinearities such as habituation. For this purpose, we varied the stimulus durations by using repetitions of a 1s checkerboard presentation. We used from 1 up to 8 successive presentations separated by half a second. We preferred such block repetitions to prolonged stimulation to prevent as much as possible neural habituation. Indeed, if there is a strong transient activity at the start of the stimulation, there is more chance that this transient be replicated at each repetition, whereas it would only appear once in the case of a longer stimulation. We also used one 200ms presentation, 5 successive 200ms presentations spaced by 200ms blank, and a sequence of two 1s blocks spaced by 5s (figure 4.5A). These seven designs provide complementary information that will be discussed below.

We combined the stimuli in two different paradigms. The first one consisted of two 15-minute' runs, each one containing 5 repetitions in random order of 6 different designs (200ms - 1s - 2×1 s - 4×1 s - 8×1 s - 5×200 ms (first run) or 2×1 s with 5s ISI (second run)), followed by a 25-second return to baseline (figure 4.5B). The second paradigm consisted of one 10-minute run containing the 7 designs described above, but separated only by 10s or less. It allows to compare the results when responses are overlapping (figure 4.5C).

Five subjects endured the two paradigms explained above (the first and third runs were dedicated to the first paradigm, and the middle one to the second). In a preliminary experiment, one additional subject endured the first paradigm but the data had to be discarded due to the weakness of the response, and two endured the second paradigm. The functional images were corrected for time delays. One subject needed to be motion-

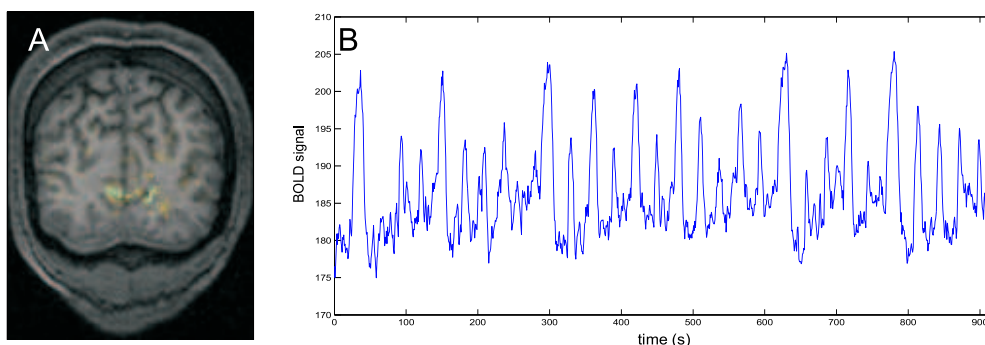


Figure 4.6: Extracted data on 40 voxels in one subject primary visual cortex. (A) Voxels selected (cyan) according to SPM F-test (light yellow) and anatomical information. (B) First eigenvector of extracted time courses in the left hemisphere.

corrected. For other subjects, motion was small relative to voxel size (2mm), and according to recommendations in [55] we did not apply motion correction to their time series.

4.4.2 Data Analysis

We first applied our methods on averaged responses to the seven stimulation patterns. In order to achieve this averaging, a first classical SPM study was done, using the stimulus convolved with three basis functions (HRF, HRF time derivative and HRF dispersion). For each subject, a cluster of voxels above a statistical threshold was selected, focusing on V1 since it is probably the region where neurons respond the most linearly to visual stimulations (V1 was selected through locating the calcarine sulci on anatomical images in both cortices, see figure 4.6A). Then a mean time course was extracted from every cluster (using the SPM first PCA eigenvalue selection, see figure 4.6B). For each subject, responses to each of the seven designs in the first paradigm were time-locked averaged (figure 4.7A), and a global mean over all subjects was calculated as well (figure 4.7B). We did not apply any high-pass filter, to preserve possible physiological low-frequency components. The baseline signal was estimated through averaging the signal over the 4s before each stimulation.

Once average responses to the stimulation patterns were calculated, we fitted to them through linear regression a simple linear model with three regressors (the stimulus convolved with a default HRF, the HRF time derivative and the HRF time dispersion, as defined in SPM), and seven different variations of the Balloon Model by using our nonlinear least square algorithm. These Balloon Model variations were obtained from equations (4.2) or (4.4) by fixing some of the parameters to physiologically plausible values and estimating the remaining ones. Results are shown in figure 4.9 and table 4.4. We assigned for convenience letters to these different models: A for the linear one and B-I for the others.

We also applied our sensitivity analysis to three models (E, G and I), to investigate the uncertainty of estimated parameters.

Secondly, we ran linear model regression and nonlinear model identifications on every voxel time course in masked brains, and applied our sensitivity analysis to these signals.

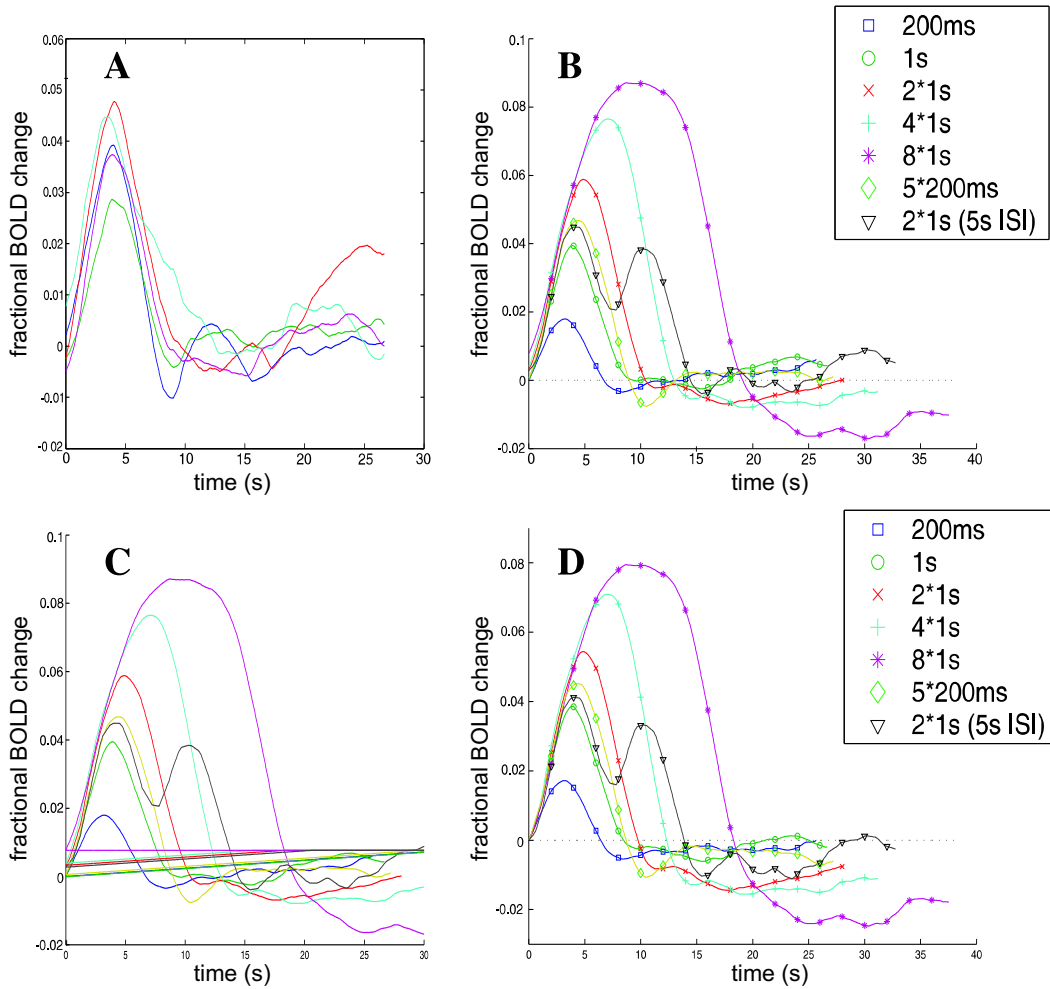


Figure 4.7: Estimated responses to the 7 designs - mean on 5 subjects. (A) inter-subjects variability (response to the 1s stimulation). (B) mean responses. (C) estimation of an overlapping return to baseline from previous responses. (D) corrected responses.

4.4.3 Qualitative Description of the Estimated Responses

We comment here the estimated responses to the experimental conditions shown in figure 4.7B, and in particular we discuss the presence of nonlinearities.

First let us note that the intensity of the signal changes is quite large: in the part B of figure 4.6, we can observe 11% signal changes in a cluster of 40 voxels. BOLD signal changes are usually smaller in the literature. However, some signal changes by 10% were reported in [121] in the monkey brain at 4.7 T, and even 20% in a single voxel. The amplitude of the signals we measured could be attributed to three reasons: they were acquired at 3T, with surface coil, and our stimulation pattern (a full-screen flashing checkerboard) was particularly intense.

We observed an ascending trend in the estimation of the responses to short duration stimuli, and the signal level before stimulus presentation seems to increase with stimulus duration. This is probably because responses to long stimulations last more than 25s after stimulation ends, so that responses to short stimulations are meddled with return

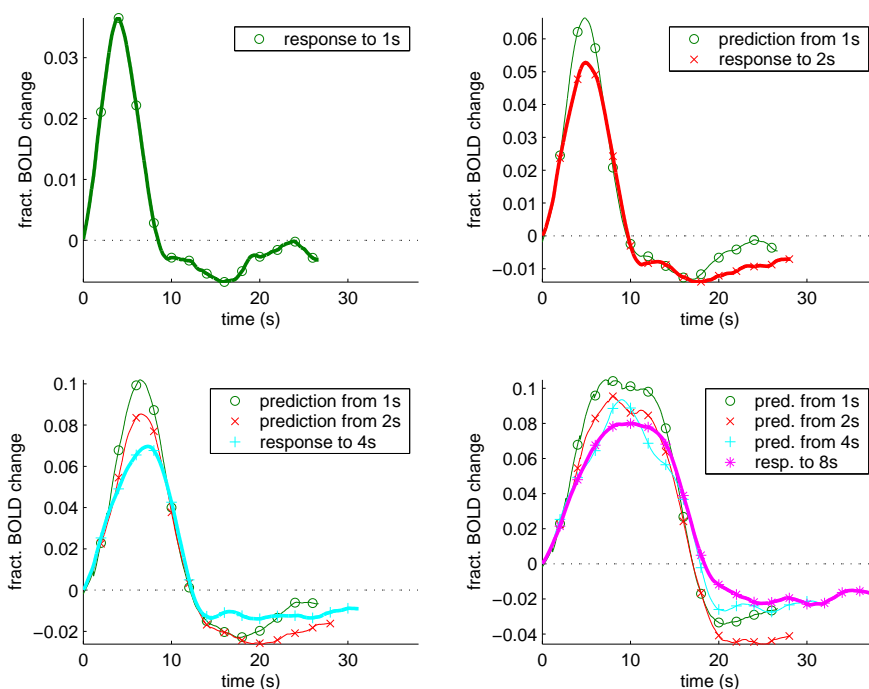


Figure 4.8: Fit between responses to long durations and their prediction by shorter durations responses. We used the corrected responses in 4.7D.

to baseline of the previous ones. We tried to correct this defect through estimating and removing a linear return to baseline (figure 4.7C,D). Nevertheless, subsequent remarks are robust to this trend removing: the analysis in figure 4.8 shows the same qualitative behaviours whether the trend has been subtracted or not.

Nonlinearities are clearly present in the short durations range: responses to 1s or $5 \times 200\text{ms}$ stimulations are much smaller than 5 times the response to the 200ms stimulation. We shall call this a *sub-linearity* effect below. Moreover, the response to the 1s stimulation is itself smaller than that to the $5 \times 200\text{ms}$ stimulation. For longer stimulations, we study linearity as shown in figure 4.8: the response to $k \times n$ repetitions is predicted by the sum of k shifted responses to n repetitions. We observe that the response to the 1s stimulus overpredicts that to the $2 \times 1\text{s}$ stimulus (see figures 4.8A and B, upper right-hand corners), which itself overpredicts that to the $4 \times 1\text{s}$ stimulus (see figures 4.8A and B, lower left-hand corners). It is not clear whether the response to the $4 \times 1\text{s}$ stimulus overpredicts that to the $8 \times 1\text{s}$ stimulus, but anyway, the shapes are different (see figures 4.8A and B, lower right-hand corners). These results are coherent with other studies [24, 41, 20, 67, 139, 173]: when comparing positive responses, the linear assumption for the BOLD response is acceptable for stimulus durations > 4 seconds, and does not hold for durations < 2 seconds.

Last, there seem to be nonlinear effects in the poststimulus undershoots too: the undershoots after longer stimulations appear to last longer than what would be predicted from shorter stimulations (even more than 25s, the time we chose to separate our presentations).

Linear Model

	SNR	HRF	HRF time derivative	HRF dispersion
A	3.21	.1	.19	-.04

First Balloon Model

	SNR	ϵ	κ_s	κ_f	τ	α	τ_{visc}^+	τ_{visc}^-	E_0	b
B	3.26	.89	.49	.19	.27	0.63	[0]	[0]	.33	.16
C	6.73	.94	.76	.48	3.72	.42	11.91	34.34	.34	.21
D	6.66	.93	.77	.48	3.65	[.4]	11.42	34.94	[.4]	.22

Second Balloon Model

	SNR	κ_n	τ_I	ξ	n	τ_f	τ_m	τ	α	τ_{visc}^+	τ_{visc}^-	b
E	6.35	[0]	\square	1.22	2.13	1.96	1.90	4.48	.43	6.11	10.76	1.80
F	4.08	[0]	\square	1.26	2.69	5.12	10.24	.20	.20	[0]	[0]	.29
G	6.27	[0]	\square	1.52	[2.5]	2.24	[$=\tau_f$]	3.89	[.4]	6.84	16.15	.73
H	6.12	[0]	\square	0.89	4.58	4.13	3.47	.20	.48	50	50	[.3]
I	6.82	2.83	1.17	2.18	[2.5]	2.34	[$=\tau_f$]	3.86	[.4]	7.90	17.57	.95

Table 4.4: Parameter values and signal to noise ratios corresponding to estimations in figure 4.9. Values inside brackets were fixed a priori (the corresponding parameters were not estimated). For the second Balloon Model, we observed that, when estimated, the delay parameter δ_t in equations (4.3) always came out null, so that we did not include it in our estimation. The signal to noise ratios have been calculated as $SNR = \|y_{model}\|/\|y - y_{model}\|$.

4.4.4 Fitting Models to Mean Responses

We fitted different models to the estimated responses by minimizing the least square errors over the seven juxtaposed curves. Figure 4.9 shows time courses and table 4.4 parameter values.

The results obtained with the linear model are shown in part A of the figure. Sub-linearities can be clearly observed: the best fit is to the 2s stimulation response, but the response peak after the 200ms stimulation is underestimated while those after longer stimulations are overpredicted. Moreover poststimulus undershoots are not well fitted.

The first physiological model we fit to the data (part B of the figure) is the original Balloon Model given by equation (4.2). It effectively appears to better capture some nonlinearities (peak amplitudes), but does not account for short time range nonlinearities (200ms and 5×200 ms stimulations). Poststimulus undershoot is not captured well either. Adding the two viscosity terms ($\tau_{visc}^+/\tau_{visc}^-$) in the volume dynamic results in a more prolonged poststimulus undershoot (part C of the figure). Part D of the figure shows that very little of the fit quality is lost when values of the Grubb parameter α and the extraction at rest E_0 are not estimated, but fixed to some physiologically plausible value (see the next sections on statistical tests and sensitivity analysis).

The next estimations use linear convolutions to model the flow and metabolic responses as in equation (4.4). Part E of the figure shows an estimation according to (4.4), except that no neural habituation was assumed. The fit to measured BOLD is comparable to that above, but it is interesting to see that the estimated flow time courses are

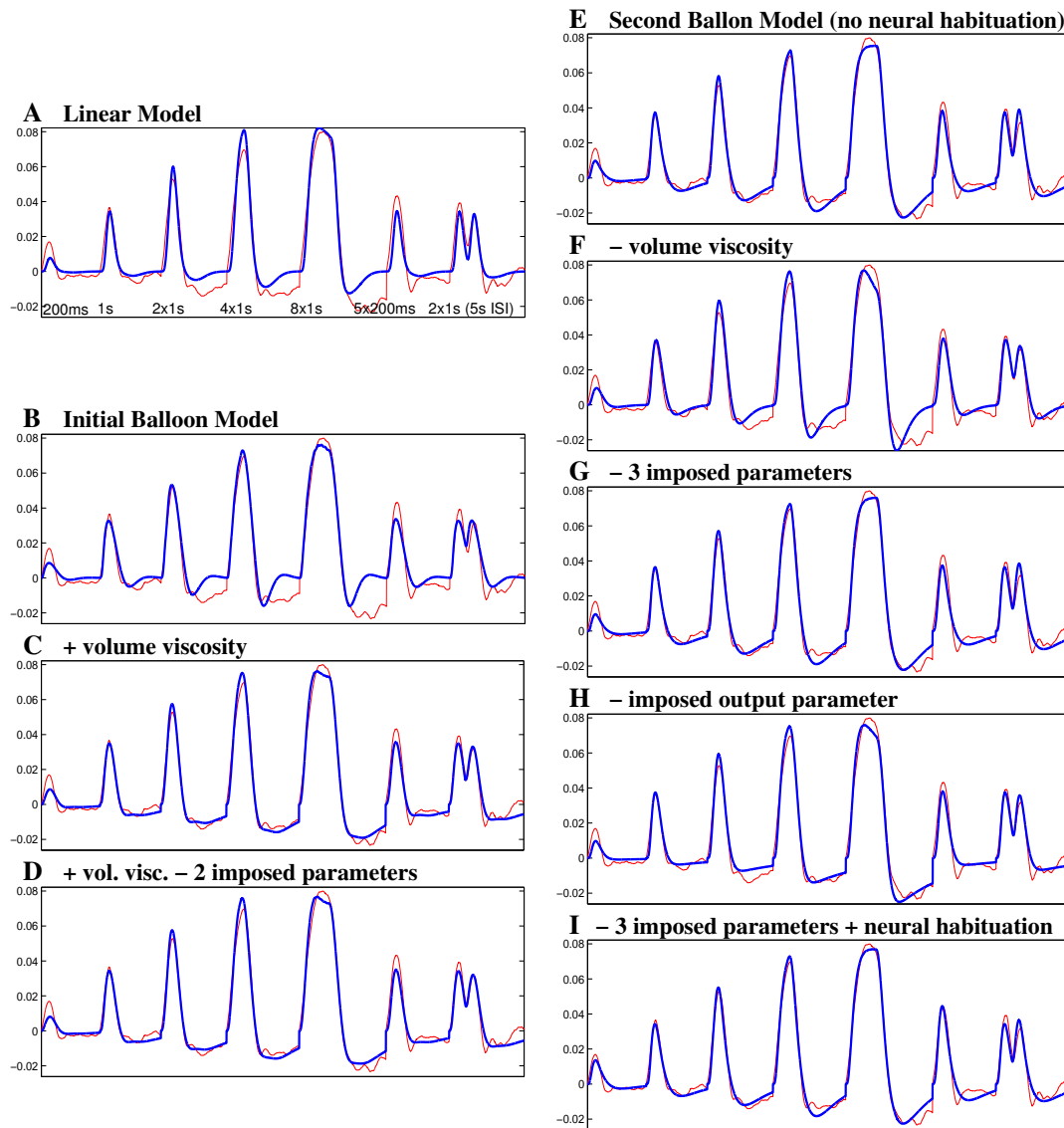


Figure 4.9: Fit of different models to the mean responses (corrected curves in figure 4.7D). Estimated parameters are shown in table 4.4.

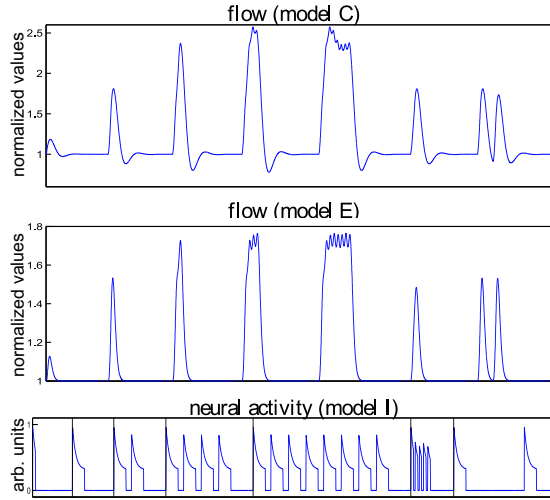


Figure 4.10: Time courses of some hidden state variables. Top and center panels show the estimated flows corresponding to models C and E in figure 4.9: they compare Friston’s model with a damped oscillator and Buxton’s convolution with a Gamma-variate function. Bottom panel show the neural activity estimated in the last model (I: second Balloon model with neural habituation).

quite different between Friston’s damped oscillator modelling and Buxton’s convolution with a Gamma-variate function (figure 4.10 top and center). We also tried some special cases of this model: first an estimation with no viscosity term in the volume dynamics (part F of the figure), but the poststimulus undershoot was not well predicted then. On the other hand, we found that some parameter values could be fixed to physiologically plausible values ($\alpha = 0.4$, $\tau_f = \tau_m$ and $n = (f_1 - 1)/(m_1 - 1) = 2.5$) without significant loss in the quality of the fit (part G of the figure). As in parts E,F,G of the figure the estimated values for the output scaling $b = V_0(a_1 + a_2)$ were a bit larger than expected, we tried a new estimation where we imposed the more physiologically plausible values $V_0 = 0.03$ and $a_1 + a_2 = 10$ (i.e. $b = .3$, part H of the figure), but it resulted in a poorer fit and less realistic values for the other parameters.

Lastly, we added to the model the simple neural habituation proposed by [32]. Including the parameters κ_n and τ_n into the estimation, i.e. allowing neural habituation, appears to be the only way to predict correctly the 200ms and 5×200 ms stimulations responses (part I of the figure). Bottom of figure 4.10 shows the corresponding neural activity time course.

4.4.5 Sensitivity Analysis of the Mean Responses

As we mentioned before, activation detection through the use of hemodynamic models and model comparison tests can be achieved successfully, even though the estimation results in some uncertainty about parameter values. Here, we applied our sensitivity analysis to estimations E, G and I in figure 4.9: figure 4.11 shows results with 2% sensitivity intervals. Let us be reminded of the definition of the sensitivity intervals: “for every value v of the i th parameter in this interval, under a linear approximation of the system output with respect to the parameters, there exists a parameter set θ' verifying $\theta'_i = v$ such that $\|y(\theta') - y(\theta)\|$ is less than 2% of $\|y(\theta)\|$ ”. To visualize this

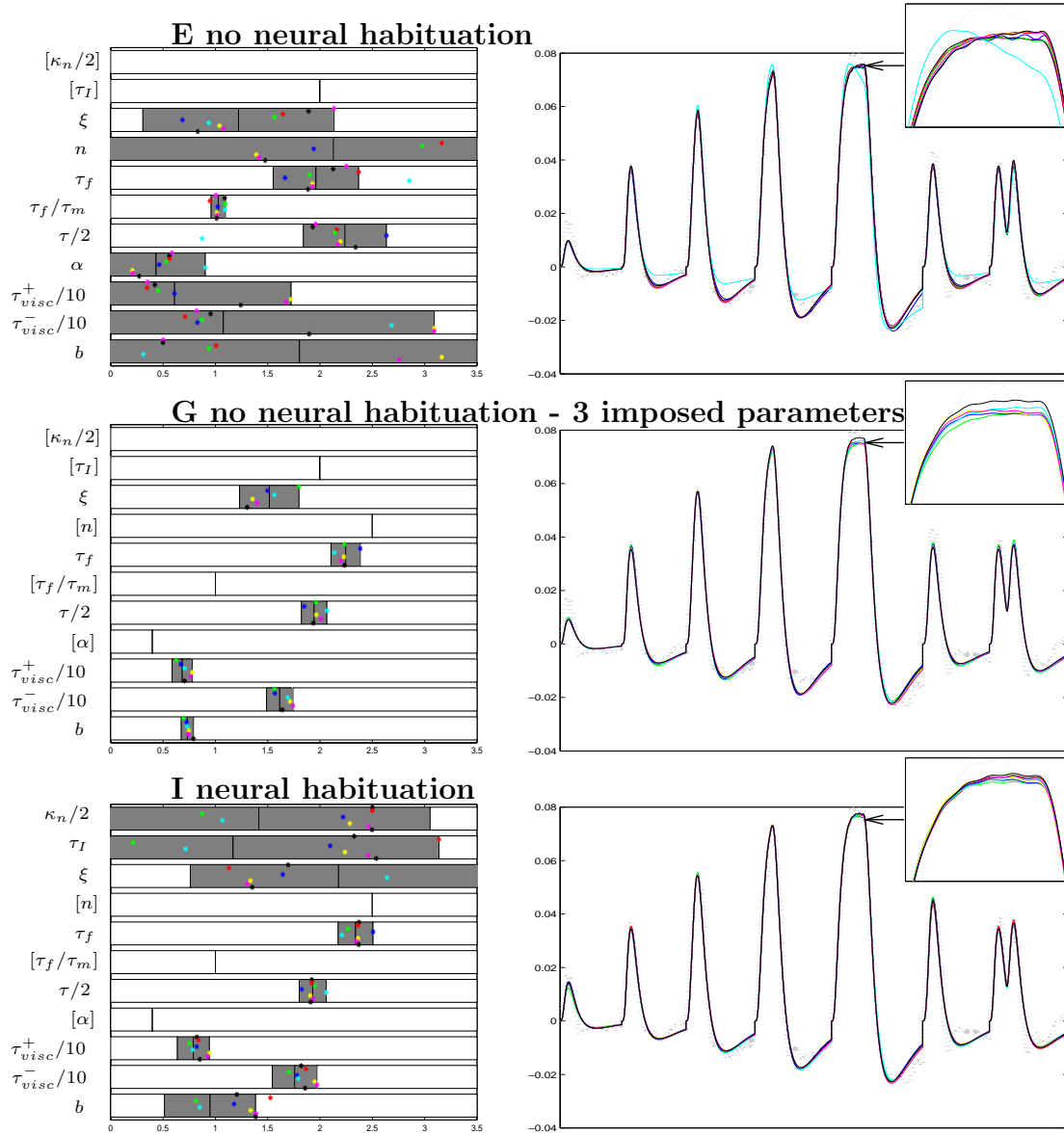


Figure 4.11: Sensitivity analysis for parameter estimation on the estimated mean responses (similar to figure 4.3). Left column: parameter sensitivity intervals for 2%-output variations; for any value of the corresponding parameter in these intervals, it is possible (in the linear approximation case) to choose the set of other parameters in such a way that the system output shall change by less than 2% with respect to the output with the middle-of-interval values; such parameter sets are represented with coloured stars. Right: measured responses (bold line), and outputs corresponding to these sets (same colour codes as for the stars). From top to bottom, graphics show different choices for which parameters are estimated or fixed to a canonical physiological value (between brackets). As expected, the less parameters are estimated, the more identifiable they are.

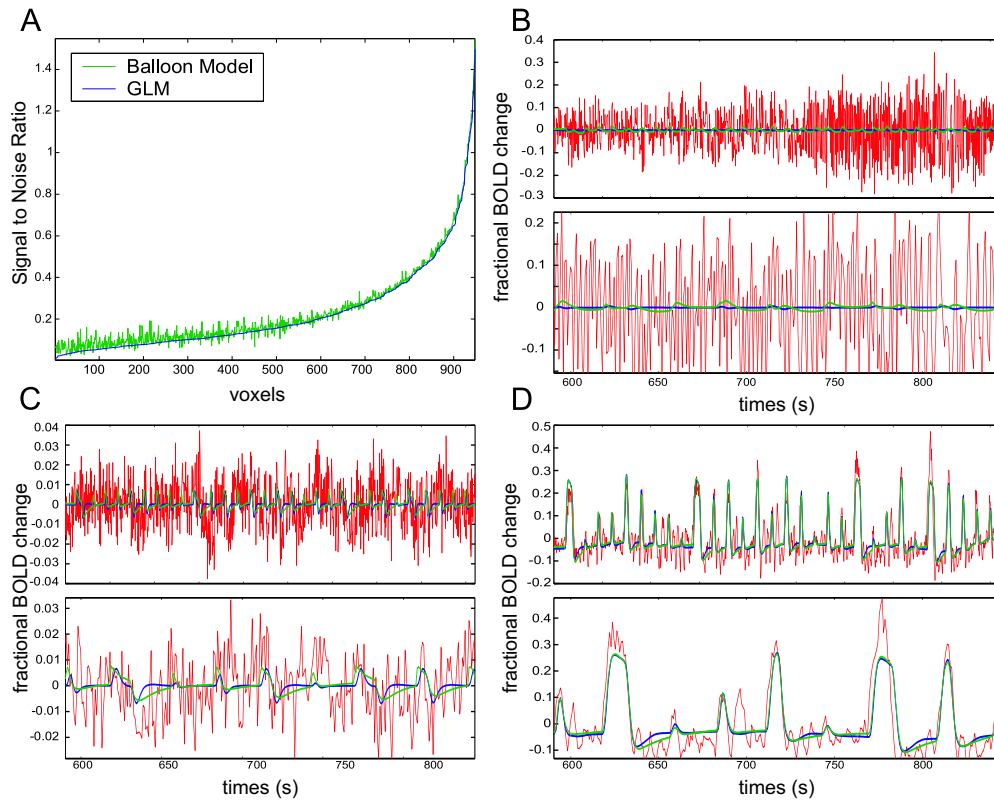


Figure 4.12: Model to data fits comparison between the GLM and the Balloon Model. (A): plot of the signal to noise ratios for the GLM and the Balloon model at all voxels ($SNR = \frac{\|y_{model}\|}{\|y - y_{model}\|}$ - voxels are sorted by their GLM SNR). (B),(C),(D): details of the fit for three voxels corresponding to the worst, medium and best SNR (red: measured signal, blue: predicted signal by the GLM, green: predicted signal by the Balloon Model).

assertion, for every estimated parameter, we give an example of a new estimation where this parameter value is fixed to that of the edge of the sensitivity interval (colour stars in the figure) and plot the new time course obtained.

It appears that we cannot estimate many parameters together correctly from the fMRI data (top and bottom rows in the figure) since many sensitivity intervals are very large. To increase the system identifiability, we must reduce the number of estimated parameters. This is the reason why, when statistical tests do validate them, we will prefer models D (for the initial Balloon Model formulation) and G (for the second one), that entail fewer parameters, for subsequent estimations on every voxels.

4.4.6 Estimation on Voxel Raw Time Course

We estimated the parameters of the model G (the second Balloon Model described in (4.4) with no neural habituation and a subset of fixed parameters) at every voxel for the three runs of each subject. As before with the mean responses, we compared the resulting fit with the one of the linear model with three regressors (canonical HRF + time derivative + time dispersion). Results on the first paradigm are shown for one subject in figure 4.12: the signal to noise ratio (SNR) is on average 22% stronger for the Balloon Model (upper left-hand corner of the figure). The figure shows details of

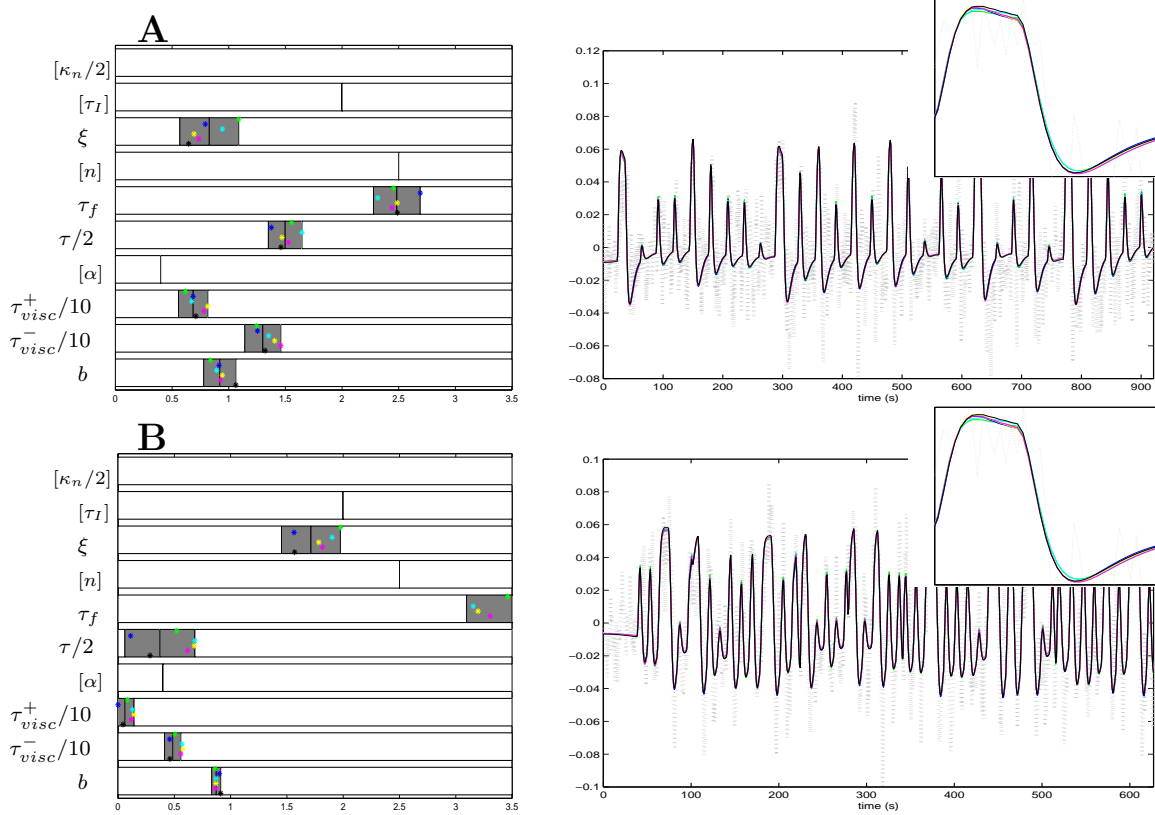


Figure 4.13: 2%-output-variations sensitivity analysis around parameters estimated on one voxel time course with the model G. (A) First paradigm. (B) Second paradigm.

the predictions of the two models for three voxels (activated, questionable and non-activated). The Balloon Model improvement is mostly in the poststimulus undershoot prediction.

Besides, it appears that parameters unidentifiability is even stronger on voxel time courses than on the means used above: figure 4.13, to be compared with the last row of figure 4.11, shows the fits and the sensitivity analysis around estimated parameters, on one activated voxel, for the first and second paradigms.

4.5 Conclusion

Our conclusion about parameter estimations in the physiological models framework is two-folds:

- On the one hand, it appears that our algorithm is able to tune model parameters in order to fit the dynamical system output to simulated or real data. And, for the latter, the Balloon Model seems to explain accurately the BOLD response, with different degrees of precision depending on which particular model version is used.
- On the other hand, it seems that the estimations lack precision on the exact values of some parameters. This is due to intrinsic system characteristics: the dynamical system will be little sensitive to variations of a given parameter as far as the changes of that parameter can be compensated by adjusting the others.

For that reason, parameter estimation cannot be used on fMRI measurements to study the values of a specific physiological or behavioural parameter - unless a very little number of parameters are left unknown, which is not the case yet in consideration of actual physiology knowledge. Please note, however, that parameter values can be investigated with more accuracy through the use of other measurement techniques, as we shall do in chapter 6 with Optical Imaging.

Nevertheless, it is still useful to use nonlinear physiological models because of their global dynamical characterization of the hemodynamic processes and of the BOLD signal. First, the methods we developed to fit a given model to fMRI measurements can be applied to the validation and selection of models, i.e. it can help the modelling efforts by deciding which particular models are the most relevant with respect to measured data. Second, if nonlinear models prove to better describe fMRI signals than linear convolution, it would be natural to use them for current fMRI analysis such as activation detection and hypothesis testing. These two points are the topic of the next chapter, where we shall propose statistical tests applicable to nonlinear models.

Chapter 5

Nonlinear Hypothesis Testing and Model Selection

In the previous chapter and the present one, we intend to prove that well-established fMRI analysis techniques, which usually rely on a simple temporal modelling of the BOLD signal with a linear convolution, can also be achieved by using more physiology-related nonlinear dynamical models.

The previous chapter was dedicated to the identification of dynamical models, i.e. the estimation of their parameters, and to the quantification of their identifiability. This was the counterpart of linear regression.

Now, we will see how hypothesis testing can be achieved in the framework of nonlinear dynamical models. It will actually have two distinct application domains. Like in the linear SPM framework, statistical tests will enable us to address cognitive questions, the simplest of them being to detect where the stimulation elicited an activation. Additionally, they will allow us to address modelling questions, by testing the significance of a model in relation to another one, which we will call *model selection*.

5.1 Statistical Test on Nonlinear Models

Nonlinear hypothesis testing is actually close to its linear counterpart, since we will again need to linearize the models with respect to their parameters. The central part of our statistical methods is the comparison between two hierarchical models, i.e. when one model is a special case of the other one.

Note that Friston et al. [58] proposed similar statistical tests in a Bayesian framework. However, these tests were relying on the marginal a posteriori distributions of model parameters, and we saw in our sensitivity analysis that individual parameter estimations lack accuracy, and in figure 4.4 that their marginal a posteriori distributions provide little information.

On the contrary, the tests we propose here directly rely on the time courses predicted by the models. Basically, they consist in Fisher tests that compare the fits to data of the two tested models, and select the most significant one, while taking into account their number of degrees of freedom.

5.1.1 Hierarchical Model Comparison

We show here how to compare two hierarchical models, i.e. when one of the two models is a particular case of the other one. Namely, if we denote by θ the whole set of parameters, a first subset θ_1 is estimated in the two models, and the remaining subset θ_2 is free in the more general model, while it is set to zero (or some other a priori value) in the simpler one. Comparing the two models means testing, under the more general assumption \mathcal{H}_1 , the null hypothesis \mathcal{H}_0 , where:

$$\begin{aligned}\mathcal{H}_1: \quad y &= f(\theta_1, \theta_2) + e_1 \\ \mathcal{H}_0: \quad y &= f(\theta_1, 0) + e_0.\end{aligned}$$

As for our Maximum Likelihood estimation, we ignore evolution noises. Also, from now on, we will suppose that the measure noises is a white Gaussian noise. It would be possible however to first whiten the signals if one has previously estimated the noise autocorrelations.

We denote by $\hat{\theta}^{\mathcal{H}_1}$ and $\hat{\theta}^{\mathcal{H}_0}$ the Maximum Likelihood estimated parameters under the two hypotheses, p_1 and p_2 the sizes of θ_1 and θ_2 and n the number of measure points. Then we linearize the model around the \mathcal{H}_1 estimate:

$$f(\theta_1, \theta_2) = f(\hat{\theta}^{\mathcal{H}_1}) + J_1(\theta_1 - \hat{\theta}_1^{\mathcal{H}_1}) + J_2(\theta_2 - \hat{\theta}_2^{\mathcal{H}_1}).$$

By doing so, the model becomes affine. We make it linear by a change of origin; let us use notations:

$$\begin{aligned}g(\theta) &= f(\theta) - f(\hat{\theta}^{\mathcal{H}_1}) + J_1\hat{\theta}_1^{\mathcal{H}_1} + J_2\hat{\theta}_2^{\mathcal{H}_1} = J_1\theta_1 + J_2\theta_2 \\ \tilde{y} &= y - f(\hat{\theta}^{\mathcal{H}_1}) + J_1\hat{\theta}_1^{\mathcal{H}_1} + J_2\hat{\theta}_2^{\mathcal{H}_1}.\end{aligned}$$

With this simplification, the two hypotheses become:

$$\begin{aligned}\mathcal{H}_1: \quad \tilde{y} &= J_1\theta_1 + J_2\theta_2 + e_1 \\ \mathcal{H}_0: \quad \tilde{y} &= J_1\theta_1 + e_0.\end{aligned}$$

We can now apply usual statistics for the General Linear Model [18]. Basically, we can obtain them by a suitable orthogonalization of the space \mathbb{R}^n where y lives in :

$$\mathbb{R}^n = X_1 \oplus X_2 \oplus X_3.$$

X_1 is the subspace defined by the columns of J_1 , X_2 is the subspace orthogonal to X_1 such that $X_1 \oplus X_2$ is the subspace defined by the columns of J_1 and J_2 , and X_3 is the orthogonal of $X_1 \oplus X_2$. Their dimensionalities are then p_1 , p_2 and $n - p_1 - p_2$, respectively. Then \tilde{y} can be decomposed as:

$$\tilde{y} = \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_3, \quad \tilde{y}_i \in X_i, \forall i.$$

These values are:

$$\begin{aligned}\tilde{y}_1 &= g(\hat{\theta}^{\mathcal{H}_0}), & \tilde{y}_2 &= g(\hat{\theta}^{\mathcal{H}_1}) - g(\hat{\theta}^{\mathcal{H}_0}) \quad \text{and} & \tilde{y}_3 &= \tilde{y} - g(\hat{\theta}^{\mathcal{H}_1}) \\ & & &= f(\hat{\theta}^{\mathcal{H}_1}) - f(\hat{\theta}^{\mathcal{H}_0}) & &= y - f(\hat{\theta}^{\mathcal{H}_1}).\end{aligned}$$

The hypothesis \mathcal{H}_0 means that \tilde{y}_2 and \tilde{y}_3 are independents and follow the distributions $\tilde{y}_2 \sim \mathcal{N}(0, \sigma^2 I_{p_2})$ and $\tilde{y}_3 \sim \mathcal{N}(0, \sigma^2 I_{n-p_1-p_2})$, where σ^2 is the unknown measure noise variance. Then it is well-known that the ratio of their square norms normalized by their dimensions follows a Fisher law with p_2 and $\nu = n - p_1 - p_2$ degrees of freedom [18]:

$$F = \frac{n - p_1 - p_2}{p_2} \frac{\|\tilde{y}_2\|^2}{\|\tilde{y}_3\|^2} \sim \mathcal{F}(p_2, n - p_1 - p_2).$$

Because of the linear approximation of the model, it is preferable for computing $\|\tilde{y}_2\|$ to use the difference of model fits:

$$\begin{aligned} \|\tilde{y}_2\|^2 &= \|\tilde{y}_2 + \tilde{y}_3\|^2 - \|\tilde{y}_3\|^2 \quad (\text{since } \tilde{y}_2 \perp \tilde{y}_3) \\ &= \|y - f(\hat{\theta}^{\mathcal{H}_0})\|^2 - \|y - f(\hat{\theta}^{\mathcal{H}_1})\|^2, \end{aligned}$$

which gives the statistic

$$F = \frac{n - p_1 - p_2}{p_2} \frac{\|y - f(\hat{\theta}^{\mathcal{H}_0})\|^2 - \|y - f(\hat{\theta}^{\mathcal{H}_1})\|^2}{\|y - f(\hat{\theta}^{\mathcal{H}_1})\|^2}.$$

We have shown that, under the null hypothesis \mathcal{H}_0 , F follows a Fisher distribution $\mathcal{F}(p_2, \nu)$. If $f_{p_2, \nu}$ is the Fisher cumulative distribution function ($f_{p_2, \nu}(z) = P(F < z)$), then comparing the two dynamical models will consist in calculating the p-value, of this statistic, that we will denote by $\mathcal{H}_0 \triangleleft \mathcal{H}_1$:

$$\mathcal{H}_0 \triangleleft \mathcal{H}_1 = 1 - f_{p_2, n-p_1-p_2}(F).$$

The hypothesis \mathcal{H}_0 will be rejected (i.e. the more general model will be selected) if and only if this p-value is less than some threshold.

5.1.2 Application to Activation Detection

It is straightforward to use this statistical test for activation detection. Indeed, it consists in comparing a hemodynamic model with a zero (no output) model:

$$\begin{aligned} \mathcal{H}_1: \quad y &= f(\theta) + e_1 \quad (\text{activation + noise}) \\ \mathcal{H}_0: \quad y &= e_0 \quad (\text{noise only}). \end{aligned}$$

Activation detection consists then in calculating in each voxel the statistic

$$F = \frac{n - p}{p} \frac{\|y\|^2 - \|y - f(\hat{\theta})\|^2}{\|y - f(\hat{\theta})\|^2}, \quad (5.1)$$

where n and p are the number of measure points and model parameters, respectively. Under the null hypothesis, F follows a Fisher law with parameters p and $n - p$, $\mathcal{F}(p, \nu)$. The test will consist in calculating the p-value $1 - f_{p, n-p}(F)$ of this statistic and declare the voxel activated if this p-value is less than some probability.

In a similar way, it is possible to take low-frequency confounds into account for activation detection, if we denote by $(C_i)_{1 \leq i \leq k}$ the used basis of low-frequency components and set the two hypotheses

$$\begin{aligned} \mathcal{H}_1: \quad y &= f(\theta) + \sum \beta_i C_i + e_1 && \text{(activation + confounds + noise)} \\ \mathcal{H}_0: \quad y &= \sum \beta_i C_i + e_0 && \text{(confounds + noise)}. \end{aligned}$$

5.1.3 Application to Other Cognitive Questions

Cognitive questions often compare the responses of a brain region to different stimulation patterns. We note that it is possible indeed for the dynamical systems to have several inputs, corresponding to these different stimulation designs; therefore there should be additional “neural efficiency” parameters, each one controlling the response to one specific stimulation. For instance, the first equation of the initial Balloon Model (4.2) would become:

$$\dot{s} = \sum_i \epsilon_i u_i - \kappa_s s - \kappa_f (f - 1).$$

Let us recall that in usual linear fMRI analysis, Student T-tests are often used to address cognitive questions, since they allow to test the positiveness of a given contrast, i.e. decide whether a linear combination of the parameters $\sum_i \alpha_i \theta_i$ (the efficiency parameters ϵ_i being of particular interest) is significantly positive. The framework we developed does not allow us to perform such Student tests. However, it is always possible to replace a Student test by a Fisher test [181]: a tested contrast is significantly positive if and only if it is declared significantly different from zero by the corresponding F-test (i.e. the hierarchical test between the model and a sub-model where the contrast is imposed to be zero) and its evaluation on the Maximum Likelihood estimate is positive.

5.1.4 Non-Hierarchical Model Comparison

We now want to design a second test to compare two models that are not hierarchical. It is not possible here to build a single statistic derived from a first model to test the second model hypothesis. However, it is possible to make the two models special cases of a third more general model, test them in relation to it, and then compare the two obtained p-values. We propose to construct this third model from the linearization of the first two; if we have the following linearization schemes:

$$\begin{aligned} \mathcal{H}_1: \quad y &= f_1(\theta_1) + e_1 \approx f(\hat{\theta}_1) + J_1(\theta_1 - \hat{\theta}_1) + e_1 \\ \mathcal{H}_2: \quad y &= f_2(\theta_2) + e_2 \approx f(\hat{\theta}_2) + J_2(\theta_2 - \hat{\theta}_2) + e_2, \end{aligned}$$

then we build the third model as a weighted sum of the two linearizations, using an additional parameter γ :

$$\mathcal{H}_3: \quad y = \gamma f(\hat{\theta}_1) + (1 - \gamma) f(\hat{\theta}_2) + J_1(\theta_1 - \hat{\theta}_1) + J_2(\theta_2 - \hat{\theta}_2) + e_3.$$

The parameters of this third model are then γ , θ_1 and θ_2 (note that the actual number of degrees of freedom may be less than $1 + p_1 + p_2$ however, since J_1 , J_2 and $(f(\hat{\theta}_1) - f(\hat{\theta}_2))$ may have linear dependencies). Since it is a linear model, its Maximum Likelihood estimate is obtained straightforward by pseudo-inverse multiplication. Then the hierarchical test (5.1.1) is applied twice to test under this third model the \mathcal{H}_1 and \mathcal{H}_2 hypothesis (using the appropriate number of degrees of freedom), and the hypothesis with the best p-value is then retained. We will denote by $\mathcal{H}_1 \diamond \mathcal{H}_2$ the ratio between these two p-values: $\mathcal{H}_1 \diamond \mathcal{H}_2 = (\mathcal{H}_1 \triangleleft \mathcal{H}_3) / (\mathcal{H}_2 \triangleleft \mathcal{H}_3)$.

Note that this non-hierarchical test is rather qualitative than quantitative: it is not easy indeed to interpret the value of the ratio between the p-values, except for the fact that it is larger or less than one.

5.2 Experimental Results

We continue here the analysis of the fMRI experiment that we began in the last chapter, by applying the statistical tests described above. These tests will be used to compare the different models we identified (figure 4.9), and to derive activation maps.

To better illustrate the model comparison technique, we will address some specific physiological questions:

- Are physiological models statistically significant compared to a simple three-regressors linear model ?
- Which methods for flow and metabolism modelling perform best among those offered by (4.2) and (4.4)?
- Does the introduction of volume viscosity significantly improve the models?
- Does the introduction of neural habituation significantly improve the models?
- Is it possible to fix the values of some parameters to physiologically plausible values with no significant loss in model precision?

5.2.1 Model Selection on Average Responses

We applied our Fisher test to compare the models and to select the one which most significantly predicts the average responses to stimulations (figure 4.9). We used the hierarchical test between two models when the first one was a special case of the second one (i.e. entailed the same equations but less free parameters). The procedure consisted then in computing the statistic (5.1.1) and its p-value, and selecting the more general model (i.e. declare the additional modelling in it statistically significant) if this p-value was less than 0.01. In other cases, we used the non-hierarchical test, which is the ratio between two hierarchical test p-values, and selected the model with the strongest p-value. Table 5.1 shows the hierarchical relations between models and the test results.

All of the test results are consistent with the hierarchical orders. This means that any increase in the model complexity was declared statistically significant, even if it required additional parameters. If we refer to the physiological questions asked above, it implies that the volume viscosity and neural habituation modelling are both statistically significant (p-values $B \triangleleft C$, $F \triangleleft E$ and $G \triangleleft I$ were all below statistical threshold), and that every parameter does contribute significantly to the model fit, and thus should be kept free in estimations procedures (p-values $D \triangleleft C$, $G \triangleleft E$ and $H \triangleleft E$ are below statistical threshold).

The non-hierarchical comparisons show the superiority of physiological models over the linear model when they entail volume viscosity ($A \diamond D < 1$ and $A \diamond G < 1$), but the original Balloon Model with no volume viscosity is considered less significant than the linear model ($A \diamond B > 1$), the improvement in SNR being not large enough to counter-balance the increased number of degrees of freedom. Finally, to fairly compare the flow and metabolism modelling methods, we need to compare model C or D against model

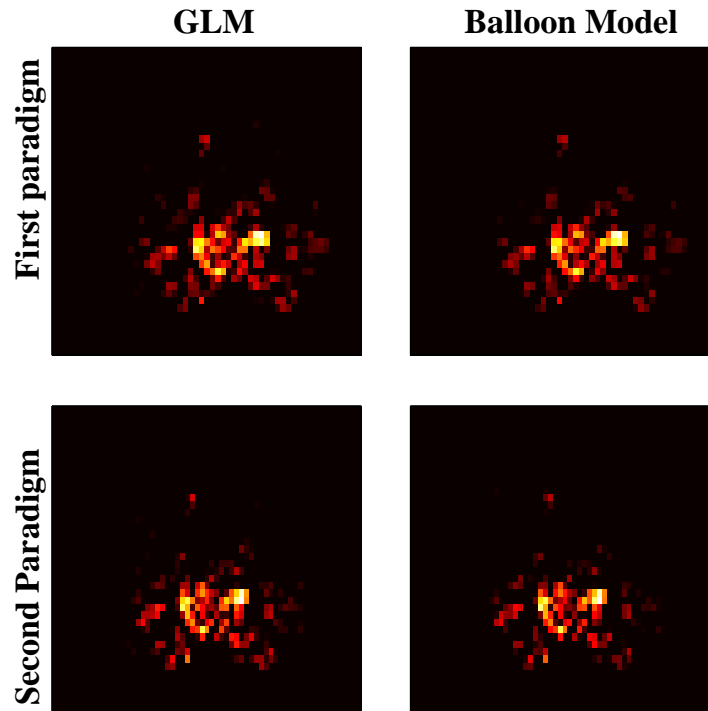


Figure 5.1: Activation maps with p-value 0.001.

E or G (since these models all include volume viscosity but no neural habituation). The four corresponding tests ($C \diamond E$, $D \diamond E$, $C \diamond G$ and $D \diamond G$) all agreed to select the original formulation (flow evolution described by a damped oscillator and flow-locked metabolism). The best model, however, remains I, which includes neural habituation.

5.2.2 Activation Maps

We applied the activation detection test in every voxel of masked brains, by selecting activated voxels with the statistical threshold $p=0.001$ for the statistic (5.1). Then we compared activation maps obtained with the different models. They actually appeared to all be very similar (see figure 5.1 for a comparison between the activation maps obtained with the linear model and model G). A possible explanation is that our stimulation was very efficient: the full-screen flashing checkerboard elicited a strong response with a large spatial extent, which was clearly identified by all the models. Another point of view is that, although it was possible to rank the different models by comparisons on average responses, their efficiency become similar when the signals become noisier, and as a result they give similar results in terms of activation detection.

5.2.3 Model Comparison on Voxel Raw Time Courses

We applied the same model comparisons as above to single voxel time courses. As a first example, table 5.2 shows SNR and comparison tests for two voxels from one subject in the first paradigm (one with a high and the other with a medium SNR) and the right column in figure 5.2 shows their time courses. It appears that the more


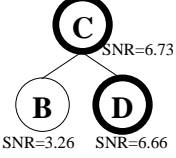
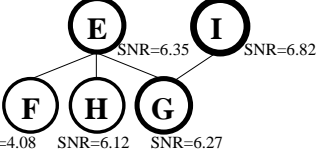

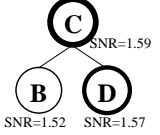
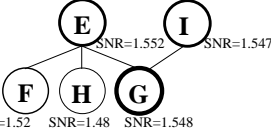
Linear model	First Balloon model	Second Balloon model
		
Hierarchical tests: $B \triangleleft \underline{C} = 0$ $F \triangleleft \underline{E} = 0$ $D \triangleleft \underline{C} = 0.0024$ $G \triangleleft \underline{E} = 0.00044$ $H \triangleleft \underline{E} = 4.6 \cdot 10^{-13}$ $G \triangleleft \underline{I} = 0$		Nonhierarchical tests: $\underline{A} \diamond B = 40.6$ $\underline{C} \diamond E = 78.4$ $\underline{D} \diamond E = 169.2$ $A \diamond \underline{D} = 0$ $\underline{C} \diamond G = 12.5$ $\underline{D} \diamond G = 47.8$ $A \diamond \underline{G} = 0$ $\underline{C} \diamond I = 0.12$ $\underline{D} \diamond I = 0.36$

Table 5.1: Model selection tests applied to models in figure 4.9. The top figure shows the dependencies between models, along with the signal to noise ratio of each model fit. Circles are thicker for preferred models. Below, hierarchical and non-hierarchical test results are presented. For each test, the selected model is underlined. Selection is applied as follows: hierarchical tests produce p-values of the first model hypothesis tested in relation to the second one, we thus chose the first model if this p-value is above 0.01; non-hierarchical tests are the ratio between the p-values of the two models in relation to a more general one, we thus chose the first model if this ratio is above 1, see text.

Linear model	First Balloon model	Second Balloon model
		
Hierarchical tests: $B \triangleleft \underline{C} = 1.9 \cdot 10^{-15}$ $F \triangleleft \underline{E} = 2.5 \cdot 10^{-7}$ $D \triangleleft \underline{C} = 3.6 \cdot 10^{-5}$ $\underline{G} \triangleleft E = 0.19$ $H \triangleleft \underline{E} = 3.9 \cdot 10^{-10}$ $\underline{G} \triangleleft I = 0.87$		Nonhierarchical tests: $A \diamond \underline{B} = 5.8 \cdot 10^{-4}$ $\underline{C} \diamond E = 2.7 \cdot 10^7$ $\underline{D} \diamond E = 2.9 \cdot 10^4$ $A \diamond \underline{D} = 0$ $\underline{C} \diamond G = 9.1 \cdot 10^6$ $\underline{D} \diamond G = 9402$ $A \diamond \underline{G} = 1.5 \cdot 10^{-9}$ $\underline{C} \diamond I = 3.4 \cdot 10^7$ $\underline{D} \diamond I = 4.8 \cdot 10^4$


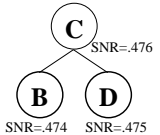
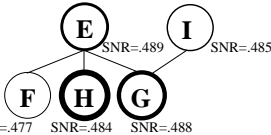
Linear model	First Balloon model	Second Balloon model
		
Hierarchical tests: $\underline{B} \triangleleft C = 0.013$ $F \triangleleft \underline{E} = 0.005$ $\underline{D} \triangleleft C = 0.86$ $\underline{G} \triangleleft E = 0.18$ $\underline{H} \triangleleft E = 0.44$ $\underline{G} \triangleleft I = 0.88$		Nonhierarchical tests: $\underline{A} \diamond B = 6659$ $\underline{C} \diamond \underline{E} = 0.003$ $\underline{D} \diamond \underline{E} = 0.012$ $\underline{A} \diamond \underline{D} = 65.1$ $\underline{C} \diamond \underline{G} = 0.0014$ $\underline{D} \diamond \underline{G} = 0.0057$ $A \diamond \underline{G} = 0.19$ $\underline{C} \diamond \underline{I} = 0.008$ $\underline{D} \diamond \underline{I} = 0.036$

Table 5.2: Model selection tests applied to two voxels (one with high SNR and one with medium SNR), see text.

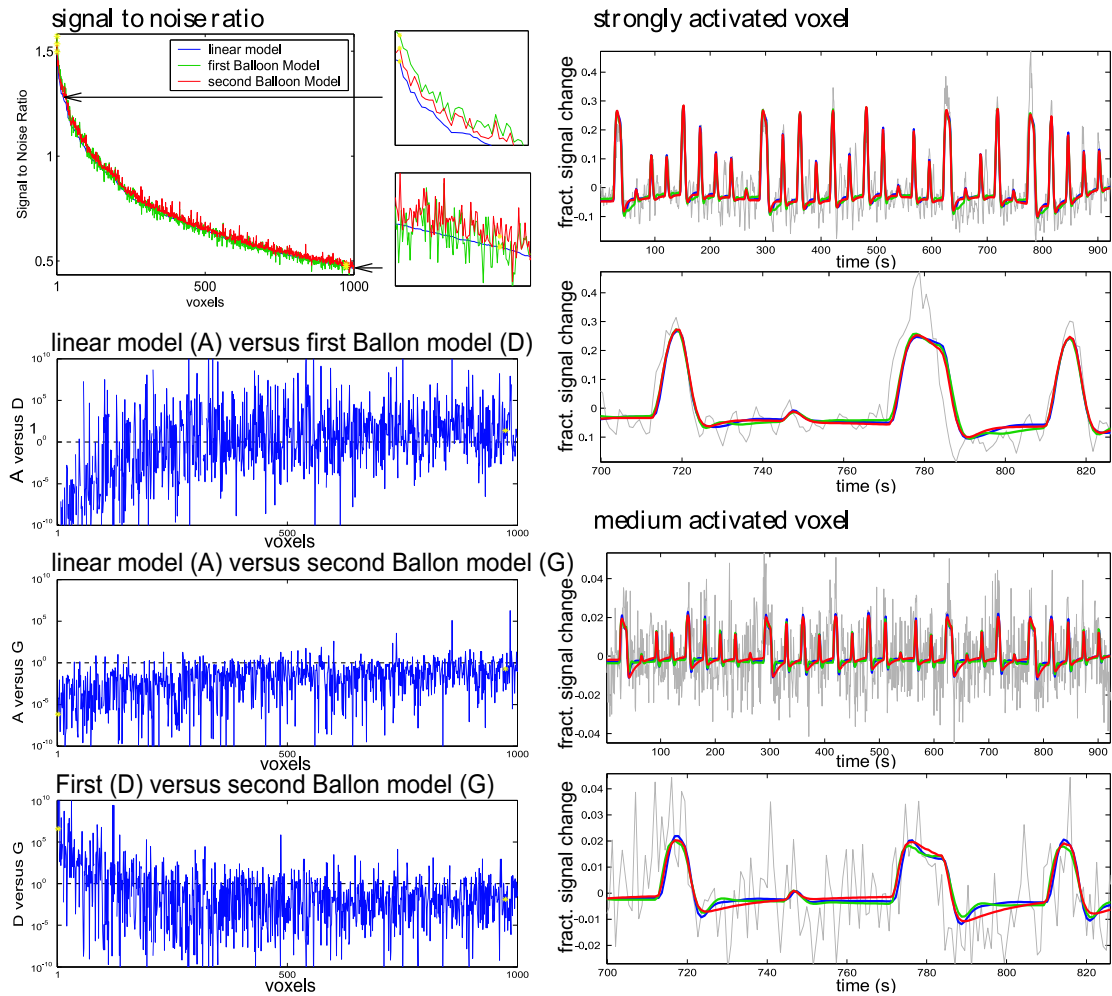


Figure 5.2: Models identification and comparisons for three models on the 1000 most activated voxels for one subjects, sorted in decreasing order of linear model Signal to Noise Ratio (top left - $SNR = \|y_{model}\|/\|y - y_{model}\|$). Comparison tests are plotted in the left column: first model is considered more significant than the second one when the test result is above 1. The right column shows details of the fit to measure for two voxels with strong and medium SNR (corresponding to the yellow starts in the left column graphs).

noisy the data, the more simpler models are preferred, because the fit improvement obtained with more complex models becomes less statistically significant. Indeed, in both cases the most complex model with neural habituation (I) is not selected any more, whereas it was selected on the average responses in the last section. However, the volume viscosity significance remains unchanged (p-values $B < C$ and $F < E$ are below the threshold). For the first voxel, the preferred model is C (first Balloon Model formulation + viscosity with all 9 parameters estimated). For the second voxel, it is H (second Balloon Model formulation with imposed output factor), but the simple linear model with three regressors also has a good rank. Model G appears to be a good compromise of significance between the two voxels.

Figure 5.2 also compares the linear model (A) and two Balloon Model formulations (D and G) in the 1000 most activated voxels for the same run. For voxels with higher SNR, the selection order is the same as the SNR order (D - G - linear model). For voxels with lower SNR, the SNR order is on average: G - D - linear model, while the selection order (which gives more importance to models with few degrees of freedom) is on average: G - linear model - D.

5.3 Discussion

5.3.1 Choice of Hemodynamic Model

The physiological questions we raised above deal with the statistical significance of several features in the models: What is the significance of different dynamical models against a simple linear model? Which flow and metabolism is the most significant between equations (4.2) and (4.4)? What is the significance of volume viscosity modelling and neural habituation? What is the significance of keeping all parameters free against fixing the values of some of them?

We actually found that these significances, and thus the model selection result, was highly dependent on the data quality: complex physiological models are validated by the measure when the signal to noise ratio is high, while simpler models are preferred on noisier or less activated voxels. On one hand, it means that in order to ask precise questions about hemodynamic models it is necessary to work with high-quality signals; this is obvious. On the other hand, it implies that when analyzing noisier fMRI signals, the model to use does not need to be the most precise with respect to hemodynamic features, but the one that finds a good compromise between precision and a reasonable number of parameters. A good candidate appears to be the model denoted by the letter G: using the model (4.4) proposed by [32], but assuming no neural habituation, fixing Grub's parameter ($\alpha = .4$), and fixing the flow-metabolism coupling in terms of amplitudes ($n = (f_1 - 1)/(m_1 - 1) = 2.5$), time constants ($\tau_f/\tau_m = 1$) and delay ($\delta_t = 0$), which leads to six degrees of freedom. However, since the quality of fMRI signals is likely to improve in the future, it argues in favour of pursuing detailed modelling efforts.

We now address the questions above and discuss their physiological meaning, relying mostly on the analysis of the average responses (figure 4.9, and tables 4.4,5.1). We recall that significant improvement brought by a new modelling does not prove that the latter is physiologically grounded; for that reason we will often recommend further investigations with other brain imaging modalities to confirm some of our answers.

First, it appeared that the major improvement brought by dynamical models over linear models was their ability to fit the poststimulus undershoot, thanks to the in-

roduction of viscosity parameters for the volume dynamic τ_{visc}^+ and τ_{visc}^- and their hysteresis behaviour (equation (4.3)). For that matter, the statistical test validated this hypothesis, even on signals with lower SNR. We thus think it would be very interesting to confirm it with other imaging modalities like ASL or Optical Imaging on animals. Besides, we observed in our data that the return to baseline after the poststimulus undershoot can be very slow, resulting in drifts in the signal that may need further modelling if they are confirmed: for example, [9] proposed a decline of a tissue oxygen buffer during sustained or repeated activations to explain these long-lasting decreases of the BOLD signal.

The question of neural habituation belongs to the more general issue of the nonlinearities in the fMRI signal with respect to stimulation. Comparisons in figure 4.8 show that nonlinearities appeared in our data. And only the modelling of a neural habituation was able to capture nonlinearities in the short time ranges: model I in figure 4.9 was the only one to correctly fit responses to 200ms and 5×200 ms stimulations. This improvement in the fit was judged significant by our statistical test. When no neural habituation was supposed, however, the other Balloon Model variations did account for nonlinearities between responses to 1, 2, 4 and 8 repetitions of the 1s stimulation, thus giving a vascular explanation to them: the blood vessel cannot inflate linearly with the flow, and the deoxyhemoglobin content cannot decrease until it is negative. These results suggest that vascular nonlinearities occur for stimulations longer than 1s, while nonlinearities occurring for shorter stimulations should have a neural origin. However we think this conclusion should be considered cautiously, because models are susceptible to be improved. In particular we notice that in the two Balloon Model formulations the flow response is considered linear with respect to neural activity, whereas a nonlinear term could be introduced, as proposed by [67]. We will investigate these questions about the flow dynamic more thoroughly in chapter 6 with the Optical Imaging (see also [139, 37, 15] for studies on blood flow).

We now endeavour to discuss the comparison between flow and metabolism modelling methods. As shown in figure 4.10, the estimated flow time courses for example are very different depending on which model is used: the damped oscillator proposed by [59] allows flow oscillations while there is no flow undershoot in the convolution with a gamma-variate function proposed by [32]. Our comparison tests applied on average signals selected the first modelling. We should keep in mind that the flow and metabolism mechanisms are probably more complex than either of these models, and are still under investigation.

At last, it appeared that some parameters (Grub's parameter α , and extraction at rest E_0 for the first Balloon Model or metabolic-vascular coupling ratios n and τ_f/τ_m for the second Balloon Model) could be fixed to some physiological a priori values without much decrease in the Signal to Noise Ratio. This is consistent with our sensitivity analysis, which shows that when many parameters are estimated together, their values can be varied much with only small changes of the system output. But all the same, the statistical tests selected the full models with every parameters estimated, meaning that every parameter carries at least a bit of relevant information. Note also that it would be of great help to be able to fix the output scaling parameters of the BOLD measure, i.e. the volume fraction at rest V_0 (estimation may be achieved using the VASO technique [125, 126]) and the scanner-dependent terms a_1 and a_2 [33, 148, 44, 32], which would improve the sensitivity of other parameters, in particular, the neural efficiency ϵ .

5.3.2 Using Dynamical Models in fMRI

We have shown that physiological models expressed as dynamical systems can be used in fMRI analysis to fit predicted responses to the data, and replace the linear regression with empirical basis functions. In particular, it is possible to obtain activation maps that rely on a Fisher test computed after a linearization of the model with respect to parameters around the maximum likelihood estimate. The activation maps obtained using linear and nonlinear methods are very similar, actually, and this can be regarded as a justification for using linear convolution today. Another argument for using linear models would be that nonlinear methods are still expensive in terms of computation time: our algorithm was implemented in C++ and took up to 10 seconds (100 iterations) to perform parameter estimation at one voxel.

But the advantage of using dynamical models is that they are more related to physiology. Also, they can handle nonlinear effects of the hemodynamic processes. Lastly, dynamical model offer more flexibility than linear ones, which in mathematical terms is expressed by the fact that the manifold of all possible outputs of a dynamical model does not need to be a vector space. As a result, dynamical models can more easily account for precise characteristics of the BOLD dynamic, even with a reduced number of parameters, and, at the same time, forbid dynamics that are not physiologically plausible.

Therefore it is probable that new advances in modelling and acquisition techniques will make their use highly valuable to detect activations, and also for other kinds of analysis; in particular those which require a detailed characterization of the temporal aspects of fMRI, such as estimation of neural time courses or fusion with EEG [153, 152, 47]. Moreover, advances in modelling and validation with other modalities will enable us to increase the constraints by imposing some parameter values, or assuming a priori distributions which can be used for estimation in a Bayesian framework [58].

5.3.3 Assumptions Used

For our parameter estimation and statistical test, we have used several approximations and assumptions.

We have already discussed the problems of the linearization with respect to the parameters in the last chapter. Our statistical test has the advantage of being quite robust to imprecision in parameter estimations. However it may be biased by the linearization: the actual distribution of this statistic may be much more complicated than a Fisher law.

There is another concern with our assumption on the noise: since we ignored evolution noises, it is probable that the actual resulting measure noise is coloured, whereas we considered it white. Further studies should try to use a regressive noise, relying on estimation of the autocorrelations in the data, and then whiten the signals before applying the statistical tests.

5.4 Conclusion

In the last chapter and the present one, we have shown methods that enable the use of dynamical models to analyze fMRI signals: a scheme to examine evolution equations and their stability, a new Maximum Likelihood procedure for system identification, a

sensitivity analysis to study the estimation precision, and a statistical tests that can be used to compare models or detect activation.

By applying the statistical model comparison methods to the results of a flashing checkerboard experiment, we could determine which one, among a small set of hemodynamic models, was the most statistically significant. We found it to be the Balloon Model proposed by [32], mostly on account of its volume viscosity and neural habituation models. However, its flow and metabolism models were judged slightly less significant than those found in [59]. More generally, our study raised many questions about some hemodynamic features, many of them concerning the flow response to brain activity. These questions led us to investigate the flow dynamic in more detail using Optical Imaging, which is the topic of the next chapter.

On the other hand, we found that, when working on noisier signals, it was necessary to use simpler models; an example choice is to ignore the neural habituation process and fix a chosen subset of parameters from the former model. Of course, it will be interesting to apply our comparison tests to other existing or future models.

Also, we were able to obtain activation maps by using these dynamical models. They were actually very similar to those obtained with a linear model, thus indicating that for detecting activation, linear models may be sufficient today. Nevertheless we think that, because of the rapid improvements in acquisition and modelling, realistic physiological models will become necessary for a better temporal characterization of the fMRI signals.

Part III

Investigation on the Cerebral Blood Flow in Optical Imaging

Chapter 6

Dynamics and Nonlinearities of the Flow Response

This chapter presents an experimental study in Optical Imaging which addresses certain questions raised by our work with hemodynamic models in fMRI. We chose to focus on the dynamics and the modelling of the cerebral blood flow (CBF). Indeed, although the CBF-induced wash-out in activated brain regions is the main cause of the BOLD response observed in fMRI, today there is little knowledge on the mechanisms responsible for it, and the models used for the CBF dynamics appear to have less physiological background than others and are eventually oversimplified.

This study was performed with Ivo Vanzetta at the INCM-CNRS Marseille. There, he and Frédéric Chavane, who both belong to the team “Equipe Dynamique de la Vision et de l’Action” (headed by Guillaume Masson), set up the first Optical Imaging facility in France, operative since end 2004 as the first European Optical Imaging facility for the awake monkey.

6.1 Introduction

The CBF increase, which has been observed since the pioneering studies of Roy and Sherrington in 1890 [154], is the main cause of the overshoot consecutive to neural activation observed in fMRI signals. Indeed, there is an additional blood flow through the cerebral vasculature, providing an increased oxygen supply in a much larger proportion than the actual increase of the oxygen metabolism; thus the total deoxyhemoglobin content in the vessels decreases, resulting in an increased BOLD signal.

As a consequence, the blood flow that enters the brain vasculature is of particular importance in hemodynamic models: it is generally assumed to completely determine the dynamic of blood volume changes, and has a major influence on that of oxygenation. Therefore, an accurate modelling of the blood flow response is critical.

However, we saw in chapter 2 that the regulation of the blood flow by the neural activity depends on an extensive number of factors, and that detailed knowledge about these factors is limited. In consequence, models are based on empirical observations rather than on physiology, with different precision levels depending on the needs for temporal resolution. In most studies with long stimulation periods or low temporal resolution (above one second), the flow response has been modelled with a trapezoidal shape [33, 9]. In the previous chapters we used more precise modelling approaches:

the damped oscillator proposed by Friston et al.[59], and the convolution with an appropriate gamma-variate function proposed by Buxton et al.[32]. However, not enough systematic experimental studies have been conducted to adequately question either of these models.

Furthermore, all mentioned models assume a linear relation between neural activity and flow response. Yet, our fMRI study showed that, under such an assumption, among nonlinear effects reported by fMRI studies [24, 41, 173, 17, 67, 20, 139], those in the short time ranges (i.e. for which stimulation repetition times are below one second) cannot be ascribed to vascular effects and hence should probably already exist at the neuronal level. This would imply that fMRI may be used to investigate the nonlinearities of the cortical activity. On the other hand, it is also likely that nonlinear effects exist in the mechanisms that regulate blood flow. In this context, a first nonlinear model was proposed by Glover et al.[67], with a trapezoidal response shape whose ramp up/down times were assumed to be independent on the stimulus duration and whose plateau duration was a nonlinear Fermi function of the stimulus.

For these reasons, the main objective of our study was to question the possible nonlinearities of the CBF response with respect to neural activity. This was made possible by the recourse to Optical Imaging, with which hemodynamic as well as electrical time courses can be monitored - at high spatio-temporal resolution - on the same animal. We recorded electrical (i.e. neuronal) signals with the fluorescent voltage-sensitive dye (VSD) technique, blood flow with a laser-Doppler probe, and blood volume in intrinsic optical imaging. We found that, indeed, such nonlinearities exist.

Additionally, our experimental data allowed us to study the interactions between the dynamics of CBF and of the cerebral blood volume (CBV). This is another critical issue for modelling the BOLD response, since these interactions result in the characteristic undershoot of the late BOLD response. Indeed, this undershoot is generally attributed to a delayed return to baseline of the venous volume compared to the flow, which results in an increased deoxyhemoglobin content in the veins. In particular, in our fMRI study we observed that an additional delay (and hence a more prolonged BOLD undershoot) could be modelled by a compliance of the volume. More generally, in every vascular compartment, one should expect the volume dynamics to follow that of the flow, for the simple reason that volume dynamics can be expressed as the difference between incoming and outgoing flows. However, our Optical Imaging data appears to be in contradiction with it: in our measurements, the volume time courses preceded those of the flow. We shall present these results and discuss how they could be explained.

6.2 Methods

6.2.1 Experimental Paradigm

Our experimental paradigm was inspired by numerous studies on linearity and non-linearity of the BOLD signal, as well as by the fMRI experiment we presented in the previous chapter. A way to introduce nonlinearities in the neural and/or hemodynamic responses is to use stimulations with a wide range of possible durations. That way the nonlinearities of each measured quantity with respect to the stimulation length can be easily checked by comparing the responses to these different conditions. The question then is to know whether nonlinearities observed at the hemodynamic level are already present at the neuronal level.

In our experiment, we used visual stimulations of various durations: 100ms, 200ms, 300ms, 600ms, 1s, 1.2s and 2s. We also used an additional “blank” condition, i.e., with no stimulation. The choice of these lengths was motivated by the following considerations:

- On the one hand, we wanted to explore a large span of durations, to increase the nonlinear effects.
- On the other hand, it was necessary to find a compromise between the number of stimulated conditions and the number of repetitions of each (and hence the signal to noise ratio of estimated responses).
- Also, the stimulation lengths were limited by acquisition, which should not exceed 6 or 7 seconds, due to a limited fixation time (after which the monkey breaks the fixation).
- Luckily, we were mostly interested in short durations, since in our fMRI study they produced nonlinearities which could not be explained by vascular effects when both the neural and the flow responses were supposed linear.
- Lastly, to easily compare predicted and measured responses it is convenient to use stimulus durations that are multiples of each other.

The visual stimulation consisted in drifting vertical and horizontal sinusoidal gratings, with temporal and spatial frequency known to reliably and optimally elicit responses in the primary visual cortex (V1) of the macaque.

The experiment was split into sixteen acquisition days: there were twelve acquisition sessions for simultaneous blood flow and volume recordings, for an overall total of 400 repetitions of each stimulation condition, and four acquisition sessions with the voltage-sensitive dye, for an overall total of 85 repetitions of each stimulation condition.

6.2.2 Optical Imaging Setup

Animal

A male adult macaque monkey was used for these experiments (*M. mulatta*, 6-7 Kg), with a 2cm diameter circular trepanation above the right visual cortex (V1/V2 border, at about 2.5° eccentricity of the visual field). All surgical procedures were performed in agreement with European and NIH guidelines (see [8] for detail).

Acquisition

The macaque was awake during experiments while his head is maintained stable via a head-holder. He was trained to fixate on the center of the screen, marked by a small red fixation spot appearing at the beginning of each trial. If he kept the fixation until the fixation spot disappeared (i.e. during the whole trial acquisition time) he was rewarded with a drop of juice. Data acquisition was synchronized with the heartbeat, which was recorded independently by a pulse-oximeter. A trial started when the monkey began fixating on the fixation point ($0.08-0.1^\circ$), displayed on a CRT screen. After a variable (1-2s) delay, the stimulus (the drifting grating) appeared and the monkey had to continue to fixate until the fixation point disappeared, for a total fixation period of

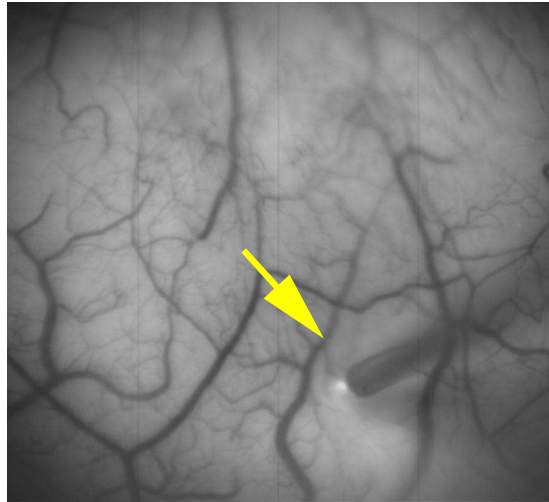


Figure 6.1: Cortex vasculature image of the monkey, as it is observed by intrinsic Optical Imaging at wavelength $\lambda = 570nm$ (used for blood volume recording) and the laser-Doppler probe (arrow) used for blood flow recording.

about 8s. “Blank” trials were identical with stimulated ones, except that no stimulus was presented. An isoluminant uniform, grey screen, was shown between the trials. The inter-trial intervals were at least 10s. To avoid visually-evoked noise, data from trials when the monkey broke fixation were rejected.

Optical Recordings

The intrinsic measure of blood volume consisted in monitoring the brain surface with a high-precision CCD camera, while it is illuminated at a specific, so called “isosbestic” wavelength (green light, $\lambda = 570nm$). Images were acquired at 200Hz using a commercial, CCD-based, imaging system (Imager 3001, Germantown, USA).

Laser-Doppler Recordings

We use a Periflux 5001 (Perimed, Stockholm; wavelength of 780 nm) laser Doppler flowmeter to which a needle probe (Perimed PF 411; outer diameter, 450 micron; fiber separation, 150 micron - estimated sample volume: $\sim 1mm^3$) was attached. The probe was positioned ~ 1 mm from the brain surface.

Voltage-Sensitive Fluorescence Dye

The stimulation paradigm for the VSD recordings was the same as for optical imaging of intrinsic signals, with the exception of data acquisition times being much shorter (4s), due to the fast decay of the neuronal responses. All details concerning VSD imaging are given in Slovin and Grinvald, 2002 [160].

6.2.3 Signal Processing

For the CBV, CBF and VSD acquisitions, the signals were averaged according to experimental conditions, and then responses to the different stimulations were normalized by

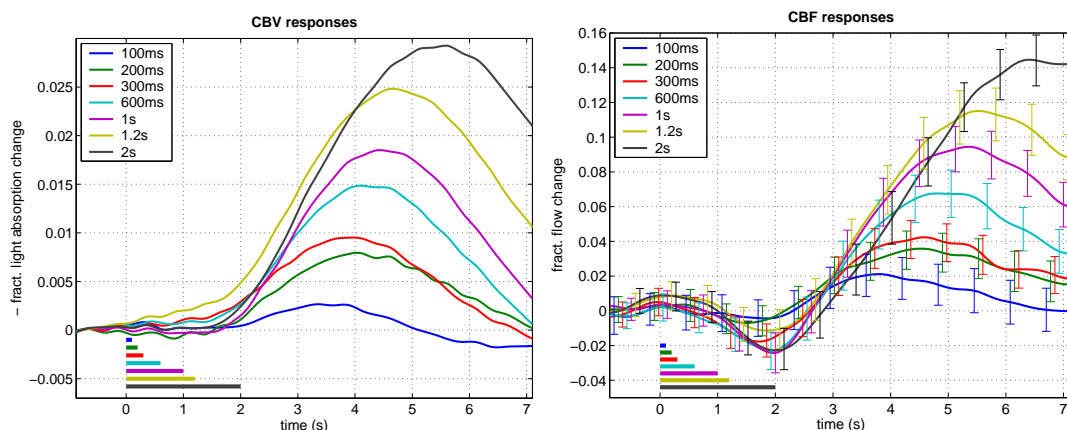


Figure 6.2: Hemodynamic responses to the seven stimulation durations: blood volume (left) and blood flow(right).

those recorded during the blank condition (see chapter 1, section 1.4, for more details). We also stress that the CBV and VSD signals were acquired as two series of images, while the CBF signal was acquired as a single-point time course (the laser-Doppler probe actually measures an average flow in a volume of about 1mm^3). Hence, we averaged CBV and VSD signals over space as well.

However, below we shall also display CBV responses at single locations. To obtain these signals, it was necessary to register the images from different acquisition days. This was achieved by selecting control points in the images and finding the linear transformations which minimized the distances between corresponding points in two images.

6.3 Results

6.3.1 Responses to the Stimulations

The figure 6.2 displays the average responses of CBV and CBF to the seven stimulation durations. Determining the precise link between blood volume and light absorption presupposes an accurate knowledge of the average photon path length in tissue, which is largely unknown in monkey cortex (although several studies have tried to estimate it in rat: [137]). For this reason, the variations reported here are defined up to an unknown multiplicative constant. Moreover, the CBF responses we show have been previously low-pass filtered at a cut-off frequency of 2Hz, to eliminate the remaining heart beat oscillations. Whereas a clear stimulus-induced increase of both CBV and CBF was observed, in several cases we could also observe a small initial undershoot in the flow responses, which, however, on the average remained below the error tolerance. Currently we don't have an explanation for this initial negative deflection of the laser Doppler signal, which has to be further investigated with experimental and stimulation paradigms especially designed for this purpose.

The figure 6.3 displays the electrical responses measured with the VSD technique. The recorded fluorescence variations reflect variations in the polarization state of the neuronal membranes, and are thus commonly thought to reflect mostly dendritic potential changes, caused by synaptic activity. Even after averaging, these responses were

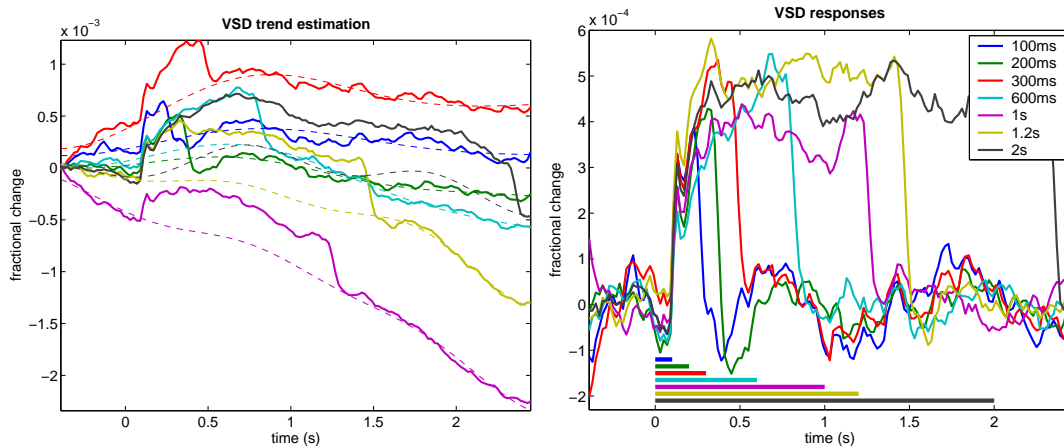


Figure 6.3: Electrical responses to the seven stimulation durations. There were large low-frequency drifts in these responses that needed to be carefully removed (left; see figure 6.4 also), in order to correctly estimate the responses (right).

endowed with strong, artifactual, low-frequency drifts. We thus had to establish a procedure to remove these drifts without altering the shape of the responses, since a basic high-pass filtering would have necessarily smoothed the steps up and down of the responses. This procedure is explained in figure 6.4 and caption.

6.3.2 Non-linearities

We first tested the linearity of the different responses with respect to the stimulation length. We used comparisons between responses to a given stimulation length and the linear prediction obtained by adding shifted responses to a smaller stimulation length. The volume and flow responses clearly exhibit non-linearities shown in figure 6.5: the responses to long stimulations are smaller than their predictions from responses to shorter stimulations. We refer to this phenomenon as *sub-linearity*. Let us however stress that the contrary happens for the shortest stimulation lengths: the response to 200ms is stronger than its prediction from response to 1s. We call this *super-linearity*. Altogether, the results suggest that the mechanisms linking electrical to hemodynamic responses are quite complex. Besides, the electrical responses either do not show such non-linearities, or they are much less perceptible. Although, on account of a slight overlap between the VSD responses, the predictions do not fit exactly the actual measurements, the global amplitudes of these measurements are not significantly different from those of the predictions.

We also used response amplitudes to explore linear and nonlinear behaviours (figure 6.6). Amplitudes of the electrical (VSD) responses were calculated as the area below the responses. For the hemodynamic measurements, however, it was not possible to use the area below the responses, since the acquisition time was not long enough for the signals to return to baseline. We thus used the peak values instead. This choice appears reasonable because stimulation lengths are relatively short. The underlying reasoning here is that, although even in the linear case the peak values do not increase linearly with the stimulation length (they rather saturate at a certain stimulus duration), in the present case stimulation was short, and only the response to the 2s stimulation is possibly

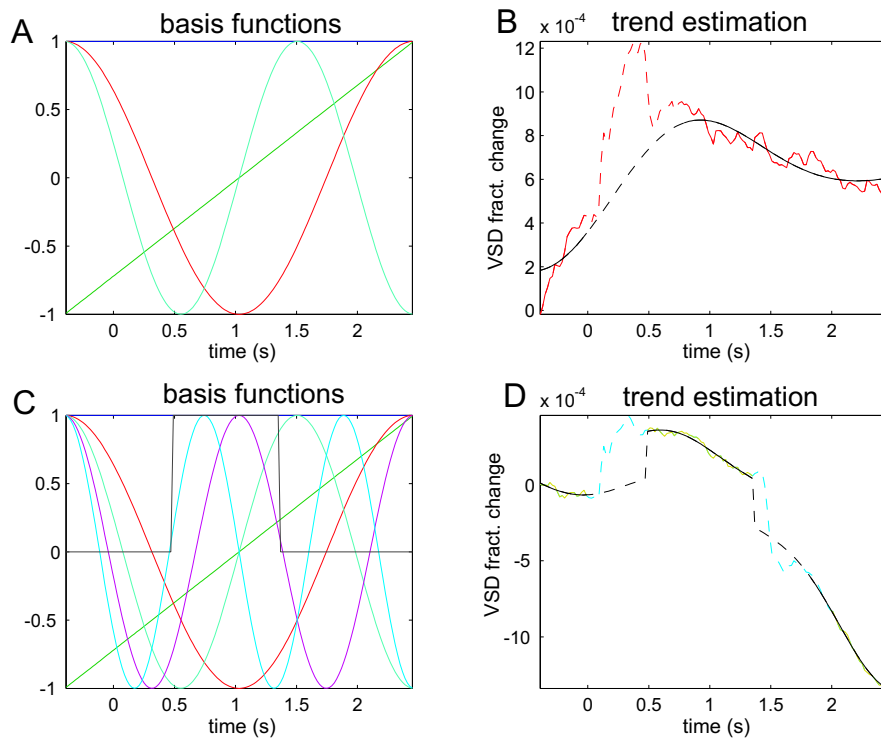


Figure 6.4: Estimation of the drift to be removed from the Voltage Sensitive Dye responses. The upper row shows how the drift is estimated for the responses to 100ms, 200ms, 300ms and 600ms stimulations. The bottom row shows how the drift is estimated for the responses to the 1s, 1.2s and 2s stimulations. (A) The drift is assumed to be a linear combination of appropriate basis functions. (B) This combination is estimated to minimize the distance between the drift and the measure. However, this distance is calculated only using time points where the electrical activity is assumed to be zero (solid lines in the upper right display), otherwise the part of the signal with the response (dashed lines) would have been altered. (C) For longer responses, it is necessary to estimate the slow variations even during the response. So, in addition to the former assumptions, the response is assumed to reach a plateau 500ms after the stimulation begin (depicted by the box car function). (D) The value of this plateau is estimated at the same time as the drift coefficients, by considering only time points assumed to have zero or plateau values (solid lines in the bottom right display).

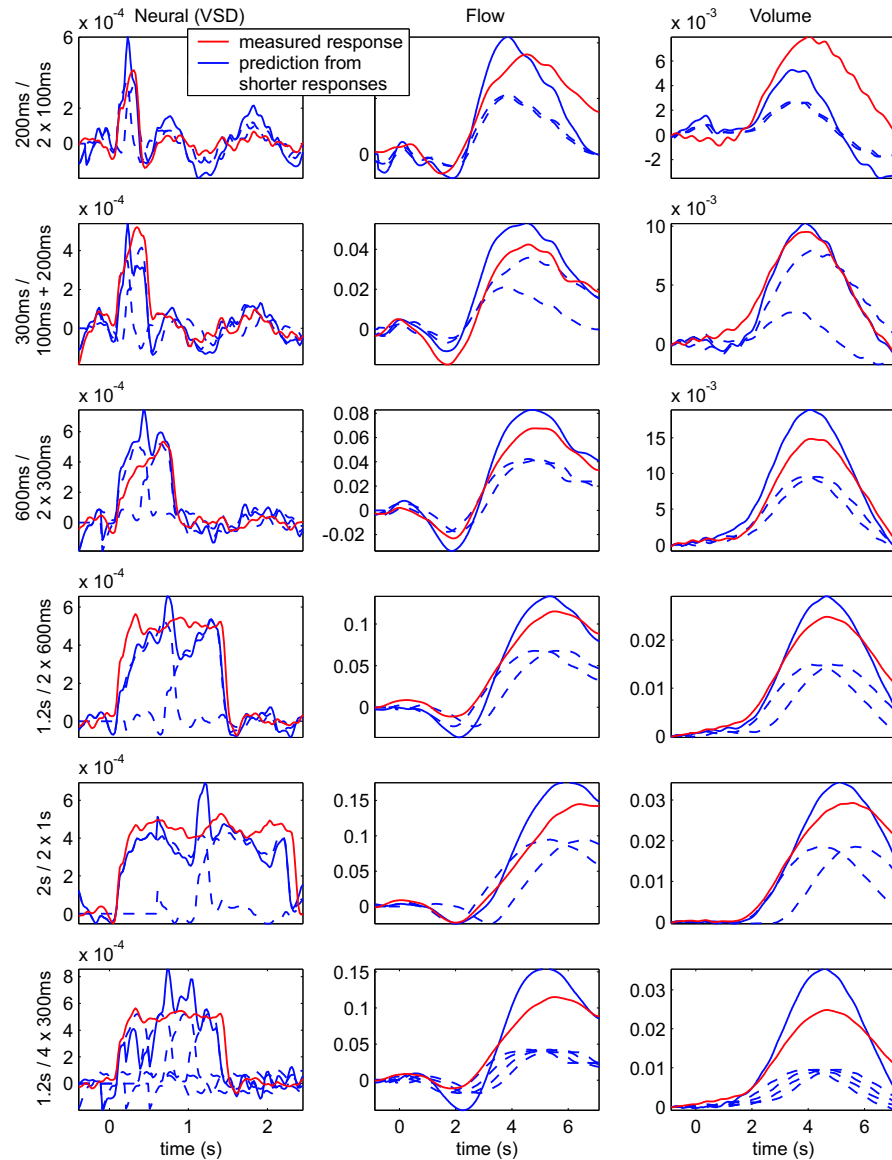


Figure 6.5: *Non-linearity comparisons*: responses to the longer stimulations (red lines in the displays) are compared to their predictions from responses to shorter stimulations (blue solid lines), obtained as the sum of several shifted responses (blue dashed lines). We observe *sub-linearities* for the hemodynamic responses (volume and flow): predictions are larger than actual responses, except for the comparison between the 200ms response and its prediction by the 100ms response, in which case the contrary occurs (*super-linearity*). These non-linearities of the volume and flow responses appear to be similar. On the other hand, the electrical (i.e. VSD) responses have a much more linear behaviour: in spite of some differences, the areas below the curves were identical between prediction and actual measure.

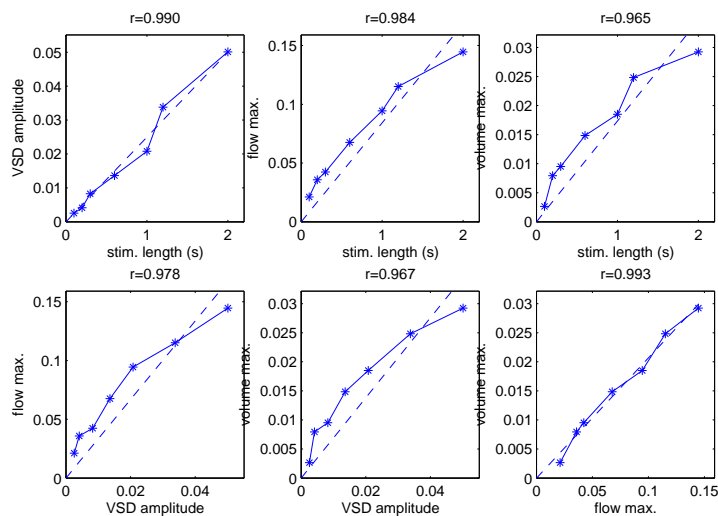


Figure 6.6: Amplitudes of the different response are plotted against stimulation lengths or against each other and linear regression results are displayed. For these amplitudes we used the peak values of the hemodynamic responses, and the area below the responses for electrical VSD responses. Stimulation length / electrical responses and flow / volume couples are in linear relation, whereas the couples electrical responses (or stimulation lengths) / either hemodynamic responses exhibit non-linear slope-decreasing patterns.

underestimated when using the peak value. We also tried using other estimations of the amplitude, such as the area under the response of extrapolated time courses, or the area under the response until the response peak, and obtained similar results (not shown). In addition, a linear regression was applied to each of the curves. As a result, the VSD amplitude appears to be a linear function of the stimulation length (correlation coefficient is 99% and the regression errors look random), whereas the CBF and CBV amplitudes show a nonlinear dependence on the stimulation length (smaller correlations, and the residuals exhibit more significant patterns). As a consequence, CBF and CBV also appear to be nonlinear with respect to the neural activity (or at least with respect to its VSD measure). Interestingly, CBF and CBV amplitudes are proportional.

At last, we emphasized these nonlinearities by fitting different dynamical models to the measured CBF responses. We used the neural activity measured with the VSD technique as the input of these models, denoted by N . Results are shown in figure 6.7. The first two rows show the fit of linear models: the damped oscillator proposed by Friston and his colleagues [59], and the convolution with a Gamma-variate function proposed by Buxton and his colleagues [32] (see chapter 2, section 2.1.3). In the two subsequent models, we added nonlinearities in the neurovascular coupling using the same equations as in the neural habituation model proposed by Buxton and colleagues (please note however that these equations do not model neural habituation any more, since the neural activity measured with the VSD technique appeared to be linear with the stimulation). For that purpose, we introduce a new intermediary variable, denoted by w , and additional parameters τ_w and κ_w . In addition, for the four models we added the possibility of a delay in the responses, denoted by δ_t . For clarity's sake, below are written the three equation systems, displaying the estimated parameter values:

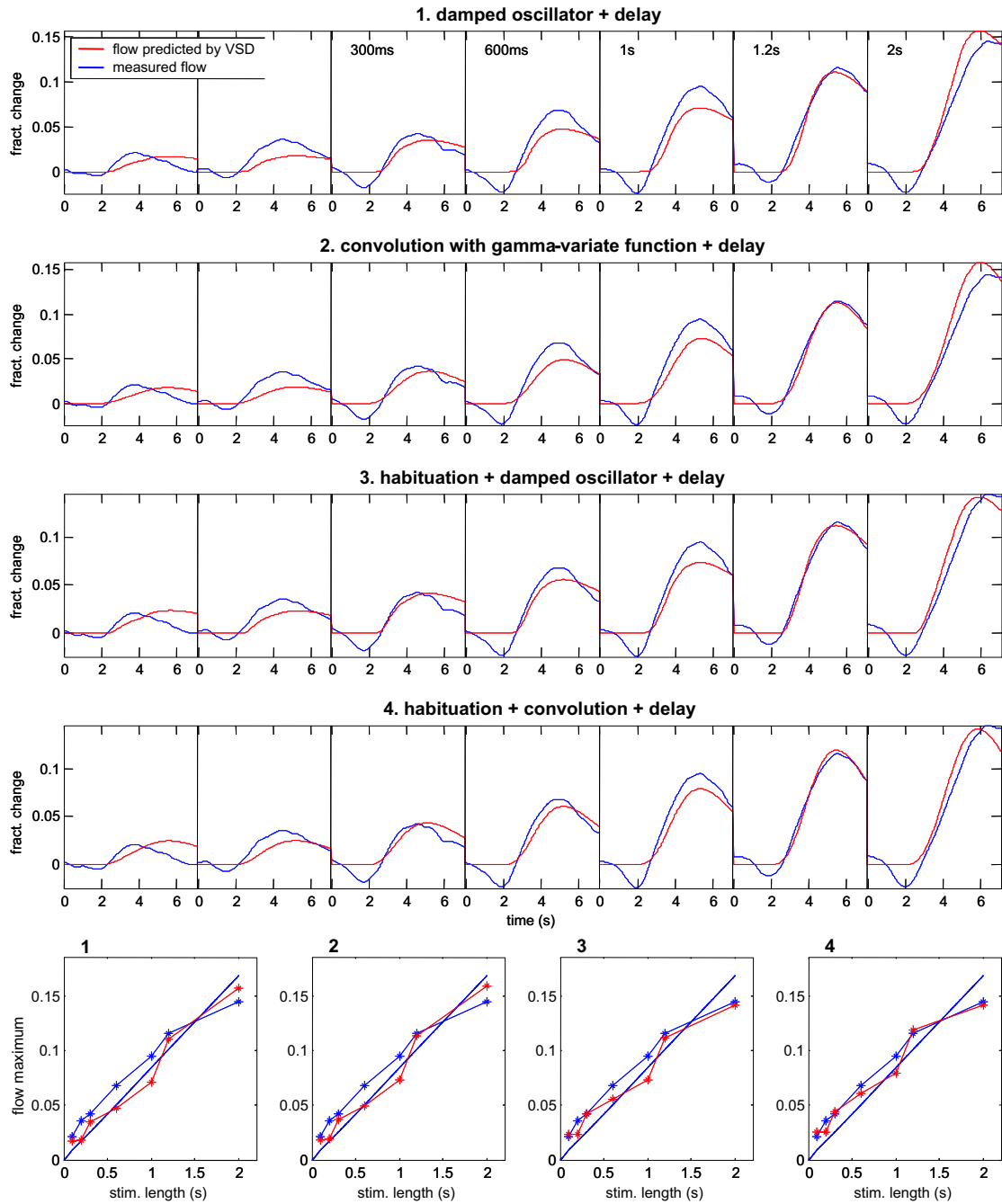


Figure 6.7: Fit of different models (in red) to the measured CBF responses (in blue). The two first rows show results with linear models, and the two next with nonlinear models. The graphs on the right compare the response-peak values of model and measurements. Note that for the linear models, the small departures from linearity are the same as those of the voltage sensitive dye, since VSD measures were used as the input of models.

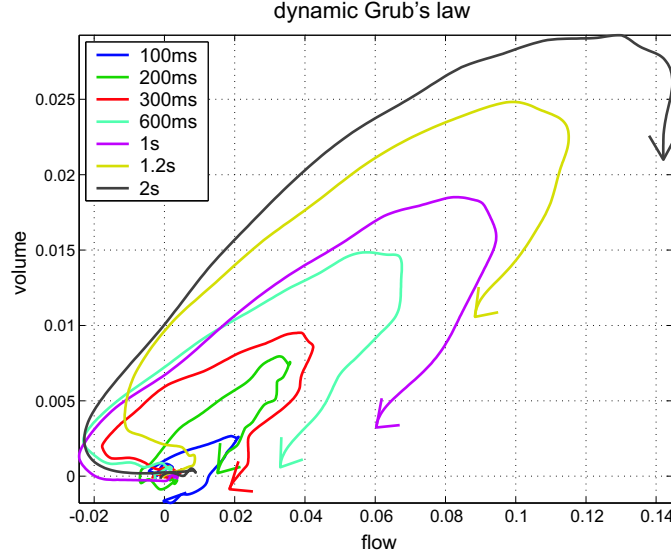


Figure 6.8: Temporal dynamic of the relation between CBF and CBV.

$$\begin{aligned}
 \textcircled{1} \quad \dot{f} &= \epsilon N(t - \delta_t) - \kappa_s \dot{f} - \kappa_f (f - 1) & \delta_t &= 2.1s, \\
 & & \epsilon &= .22, \kappa_s = 1.05s^{-1}, \kappa_f = .22s^{-1} \\
 \\
 \textcircled{2} \quad \begin{cases} f &= 1 + \xi h_f * N \\ h_f(t) &= \frac{1}{6\tau_f} \left(\frac{t - \delta_t}{\tau_f} \right)^3 e^{-\frac{t - \delta_t}{\tau_f}} \end{cases} & \delta_t &= 1.2s, \\
 & & \xi &= .76, \tau_f = 4.2s \\
 \\
 \textcircled{3} \quad \begin{cases} \ddot{f} &= \epsilon w(t - \delta_t) - \kappa_s \dot{f} - \kappa_f (f - 1) \\ w &= \max(h_w * N, 0) \\ h_w(t) &= D(t) - \frac{\kappa_w}{\tau_w} e^{-\frac{\kappa_w + 1}{\tau_w} t} \end{cases} & \delta_t &= 2s, \kappa_w = 1.4, \tau_w = 1.5s^{-1}, \\
 & & \epsilon &= .29, \kappa_s = .82s^{-1}, \kappa_f = .2s^{-1} \\
 \\
 \textcircled{4} \quad \begin{cases} f &= 1 + \xi h_f * w \\ w &= \max(h_w * N, 0) \\ h_f(t) &= \frac{1}{6\tau_f} \left(\frac{t - \delta_t}{\tau_f} \right)^3 e^{-\frac{t - \delta_t}{\tau_f}} \\ h_w(t) &= D(t) - \frac{\kappa_w}{\tau_w} e^{-\frac{\kappa_w + 1}{\tau_w} t} \end{cases} & \delta_t &= 1.35s, \kappa_w = 3.1, \tau_w = 4.5s^{-1}, \\
 & & \xi &= 1.19, \tau_f = 4.1s.
 \end{aligned}$$

The results show some similarities with those of our fMRI study. The two linear models fail to accurately capture the amplitude variations of the different responses, and the compromise that is found under-fits responses to short stimulations and over-fits responses to longer stimulations. The nonlinear models do fit the amplitude better, but do not exactly account for the temporal dynamics of the flow yet, in particular for the responses to short stimulations.

6.3.3 Flow and Volume Dynamics

A close look at figure 6.2 shows that CBV responses begin between 500ms and 1s earlier than those of CBF. Figure 6.8 underscores this fact by plotting volume against the flow for the seven responses, which constitutes a dynamic equivalent for the usual graphs that establish the Grubb's law at steady state. It clearly appears in this figure that the dynamics of the volume precedes that of the flow. As we shall see in the discussion, this is in contradiction with Buxton's Balloon model, at least when considering a single vascular compartment. In this model indeed the volume dynamics inside a given vascular compartment cannot be in advance on the flow either entering or leaving the compartment. A possible explanation for the detected delay might be that our measurements do not originate from the same spatial location or from the same vascular compartment. However, we compared the CBF measures to CBV time courses at different locations in the image, and we always observed the same delay. Actually, the CBV responses were very similar at different locations, and we were not able to isolate responses specific from the different compartments. This is probably because the different compartments have been mixed together, because of the motion of the vasculature, which have not been registered.

6.4 Discussion

6.4.1 Non-Linearity of the Hemodynamic Response

In this experimental study, we found a nonlinear relation between the responses to a visual stimulation of the cerebral blood flow and the neural electrical activity as measured by the voltage sensitive dye technique. These results argue against existing models of the flow response to brain activity. This is of particular importance concerning the application of such models in brain imaging, in fMRI for example. Indeed, it implies that nonlinearities observed in the short time ranges in fMRI are not necessarily present at the neuronal level.

These nonlinearities were manifested by the fact that the amplitude of the hemodynamic responses does not increase linearly with that of the VSD measurements; rather, we found a complex relation between the responses to different stimulation lengths, where the response intensity increased faster than linearly for stimulus durations shorter than 200ms, and slower than linearly for longer durations.

Let us discuss the possible origins of these nonlinearities, although our knowledge of the mechanisms driving the flow response is limited. First, it cannot be excluded that they correspond to some neuronal nonlinear effects that do not appear in the VSD measurements. Both the flow response and the VSD measurements are thought to be caused mostly by the synaptic activity of neurons and to correlate with the LFP. It remains possible however that the flow response is also partly influenced by a transient spiking activity which does not affect the voltage-sensitive dye.

The nonlinearities may also arise in the multiple vasoactive signals which control the flow response through the dilatation of arterial vessels. In such case, it may be possible to model a *flow-inducing signal*, in the spirit of the one proposed in [59], which would respond non-linearly to the neural activity and metabolism needs, with a response function looking like the fourth graph in figure 6.6.

Lastly, the nonlinearities may have vascular causes, through threshold and ceiling

effects in the vessel dilatation, or in the upstream pressure. In such case, new modelling should require a precise understanding of the mechanic properties of the vessels, and of the relations between flow and volume in general.

6.4.2 Blood Flow Dynamics

We observed additional surprising facts on the cerebral blood flow dynamics. First, we observed an initial undershoot before signal increase in the CBF responses. These variations are below the estimation error and might be attributed to noise effects. However they may also have a relation with our further observations. Next, it appears that the delay of the positive response increases with stimulation lengths, which constitutes a second nonlinear effect that is not captured by the predictions with simple nonlinear habituation models.

Besides, our major observation is that the CBF responses are delayed with respect to the CBV responses. This delay has been observed in other studies [129, 95]. However, it is in contradiction with the different versions of the Balloon Model, which all require that the flows entering and leaving a vascular compartment should be in advance on the compartment's volume.

Let us recall the Balloon model equation: it is based on the conservation of matter inside a given compartment:

$$\dot{v} = \frac{1}{\tau}(f_{in} - f_{out}),$$

where f_{in} and f_{out} are the flow entering and leaving the compartment, respectively.

The simpler of these models was proposed by Buxton et al.[31] and defines f_{out} with its steady-state value:

$$\dot{v} = \frac{1}{\tau}(f_{in} - v^{1/\alpha}), \quad (6.1)$$

where α is the Grubb's parameter. Since this equation models a return to steady state, it involves in particular that the volume cannot depart from this steady state unless the inflow changes first. Hence the inflow should precede the volume, whereas the outflow, in this expression, is locked to the volume.

Furthermore, the additional compliance effect introduced in [30] models an increased delay of the volume with respect to the flow:

$$\dot{v} = \frac{1}{\tau}(f_{in} - v^{1/\alpha} - \tau_{visc}\dot{v}).$$

We saw in chapter 4 that this equation is equivalent to:

$$\dot{v} = \frac{1}{\tau + \tau_{visc}}(f_{in} - v^{1/\alpha}).$$

Like for equation (6.1), the volume changes should follow those of the inflow. Concerning the outflow, is:

$$f_{out} = v^{1/\alpha} + \tau_{visc}\dot{v} = v^{1/\alpha} + \frac{\tau_{visc}}{\tau + \tau_{visc}}(f_{in} - v^{1/\alpha}),$$

and, since f_{in} is in advance on $v^{1/\alpha}$, f_{out} should also precede v .

Last, similar considerations can lead to the same conclusion using modelling of compliance with an additional differential equation proposed by Kong et al.[107].

To explain the delay we observed between volume and flow responses, we could argue that the various vascular compartments do not contribute with the same weights to our measurements. E.g. it may be possible that the flow leaving the venous balloon is delayed with respect to the capillaries' volume. Actually, it is our intention to test our data with respect to the three-compartment model proposed by Zheng et al.[184].

We also intend to reconsider the equations for the flow dynamic, in particular in the arterial compartment. We think it would be instructive to test models where the control on the volume dynamics by the outflow is directly modulated by signals in relation with neuronal activity. As proposed already at the end of the previous section, this would lead to a volume dynamics that depends directly on brain activity, rather than through the variations of inflow. Such new models may account for early volume responses.

6.5 Conclusion

The study we have presented raises several questions on the vascular dynamics and their relations to neuronal activity, in terms of nonlinear effects and delays between the different responses. To adequately address those questions, we are planning both additional measurements and modelling efforts.

On the other hand, in order to address the questions concerning the blood flow dynamics in the different vascular compartment, high-spatial resolution data are needed. With this respect, the laser-Doppler technique has the disadvantage of giving only single-point measurements (in contrast to intrinsic optical imaging recordings that allows to record time courses from every pixel in the image). Hence, the flow measured by the laser Doppler probe is an unknown combination of the flow in different parts of the vasculature. In the next chapter we shall thus present a new method using intrinsic signals to estimate the flow time courses in individual vessels.

Chapter 7

Bi-dimensional Flowmetry with Intrinsic Recordings

Accurate investigations on hemodynamic physiological models require a high spatio-temporal resolution. In particular, a good understanding of the interplays between CBF, CBV, and the electrical activity of the brain should rely on separate measurements of the flow and volume in the different vascular compartments. So far, however, CBF has been measured mostly using the Laser Doppler technique, which only provides single-point measurements.

In this chapter, we shall present a new method which allows recording CBF at micro-vascular resolution with intrinsic optical imaging. Indeed, with high spatio-temporal resolution optical imaging recordings, it is possible to track erythrocytes moving inside individual vessels. We developed an automatic method which extracts the erythrocyte motion information from the noisy optical signals. Validation was obtained by comparison with simultaneous Laser Doppler CBF measurements, using the heart-beat pulsation as a test signal. We also detected visually-evoked CBF responses, in several individual elements of the cortical micro-vasculature.

7.1 Introduction

Baseline blood flow and its sensory-evoked changes are measured mostly using the Laser Doppler technique, which takes advantage of the Doppler shift in back-reflected light when a monochromatic laser beam illuminates the moving red blood cells (RBC). In spite of the high signal-to-noise ratio (SNR) obtainable with Laser Doppler flowmetry, only single-point measurements are possible with this technique; moreover, its spatial resolution is limited, since the recorded signals are averaged over $\sim 1\text{mm}^3$ of cortical tissue, thus containing several different vessel types.

Recently, Laser-speckle flowmetry, a method based on the same principle underlying the Laser Doppler technique [26] has become available. This allows studying CBF and its activity-evoked responses with high temporal resolution in two dimensions over comparatively large samples of cortex (in the order of tens of mm^2 [50]). However, to our knowledge, this new technique provides estimations of the cerebral flow inside the capillary bed rather than in the micro-vessels.

Grinvald and his colleagues [73] proposed to use the hemoglobin packed in RBCs as an intrinsic contrast agent giving information about the blood velocity in individual

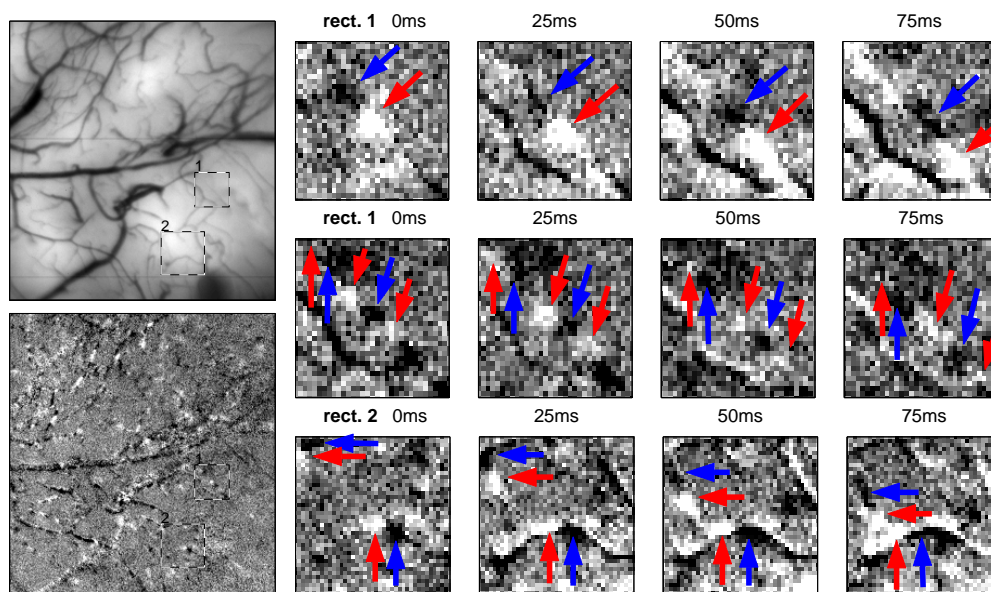


Figure 7.1: Left: (top) Raw image of the brain vasculature; (bottom) The same frame “enhanced” by the pre-processing step and displayed with an increased contrast. Black (white) dots correspond to loci of increased (decreased) local RBC concentration in the single frame with respect to average. Right: Zooms onto individual small vessels (rectangles in the left images). For each time-sequence (row), the four frames (time shown on top) show the motion of RBCs through a small blood vessel: note the shift with time of the blue and red arrows, respectively pointing to clusters of RBCs and RBC-free blood plasma regions in the vessel.

vessels, across the entire imaged area (figure 7.1 highlights examples of RBCs motions in vessels). They could measure the blood flow in several single vessels distributed over large portions of the retina in humans and anesthetized cats. However, baseline CBF values were reported, but not its temporal modulations, neither due to the heart-pulsation, nor due to a sensory stimulus.

Using two-photon fluorescence microscopy, two groups (Kleinfeld et al., 1998 [103] and Chaigneau et al., 2003 [34]) have succeeded to record CBF as well as its activity-evoked changes in single vessels. Indeed, the ultra high spatial resolution of two-photon fluorescence enabled us to separately measure the velocity and the quantity of moving RBCs, both of which contribute to blood flow (flow = speed \times RBC density). The two-photon fluorescence approach is very useful for exploring the neuro-vascular interactions at the micro-mechanism scale, especially in view of the possibility to add important anatomical details (e.g., proximity of astrocytes, pericytes etc.). On the other hand, the technique involves finite scanning times and the typical dimensions of the imaged samples are much smaller than a cortical column; its applicability is thus essentially limited to the observation of one vessel at a time. This shortcoming hampers the investigation of relations between neuronal activity and the vascular dynamics in the various vascular compartments.

Our technique extends the two approaches above, in the sense that it provides an automatic estimation of the motion of RBCs in individual vessels and the temporal variations of this motion, and that it can be applied to optical imaging signals, (though the latter have less spatial resolution and signal to noise ratio than two-photon fluorescence). It relies on several data processing features, part of them being evoked by the

computer vision domain. Our two major contributions were first a robust estimation and correction of movement of the brain, and second the use of the *tensor product* to estimate directions in bi-dimensional images.

In a first experiment, we confirmed with Laser Doppler measurements performed simultaneously with optical imaging that CBF modulations by the cardiac pulsation can be clearly detected by this method. In a second experiment, we obtained sensory-evoked responses, which appeared to have different SNR and different temporal characteristics, according to which vascular compartment the vessels belonged.

7.2 Method

Blood is not a homogeneous fluid; rather, hemoglobin travels through blood vessels, packed in red blood cells (RBCs). When the vasculature is imaged at sufficiently high resolution, these RBCs (or clusters thereof) are visible in the recorded images as small contrast changes which shift along the vasculature with time. Example are displayed in figure 7.1. Since, however, the large differences in reflected light intensity between blood vessels and parenchyma are much larger than the small intra-vascular light-intensity differences due to the presence or not of RBCs, raw images (upper left-hand panel) do not allow to appreciate single RBCs' contrasts. These large differences can be eliminated by adequate pre-processing of the light-intensity images, including correction of movement and spatial and temporal smoothing (bottom left hand panels shows the pre-processing result and zooms highlight some RBC motions). We shall now present the different steps of our estimation procedure, which are summarized in figure 7.2.

7.2.1 Data Pre-processing

Co-Registration of Images

Since the flow estimation focuses on vessels, it is critical that their position in the images does not change during the totality of the acquisition sequence. Besides, since our data was acquired on the awake monkey, we could not avoid shifts in the images, mainly due to brain movements within the skull (which can be quite large, considered the scale is $10\mu m$ per pixel). Therefore, it is necessary to register the various acquisitions, on a frame by frame basis, in order to estimate and correct these movements.

We estimated rigid motions in the images, defined by a translation vector \mathbf{t} and a rotation angle θ . Please note that ideally a non-rigid estimation should be used, since the vessels (and especially the larger ones) do move slightly with respect to the cortical surface because of the blood-pressure variations. In our case, however, rigid transformations were enough in most (99%) of the cases and thus we did not try the much more complex non-rigid transformations, also because of the massive user input they require.

The estimation was achieved by an energy minimization technique: the energy we used was the square distance between a reference frame I_1 (which was chosen as the very first frame of the whole experimental acquisition) translated and rotated according to \mathbf{t} and θ and the considered frame I_2 :

$$(\hat{\theta}, \hat{\mathbf{t}}) = \underset{\theta, \mathbf{t}}{\operatorname{argmin}} \|I_2 - \mathcal{T}_{\mathbf{t}} \circ \mathcal{R}_{\theta}(I_1)\|^2$$

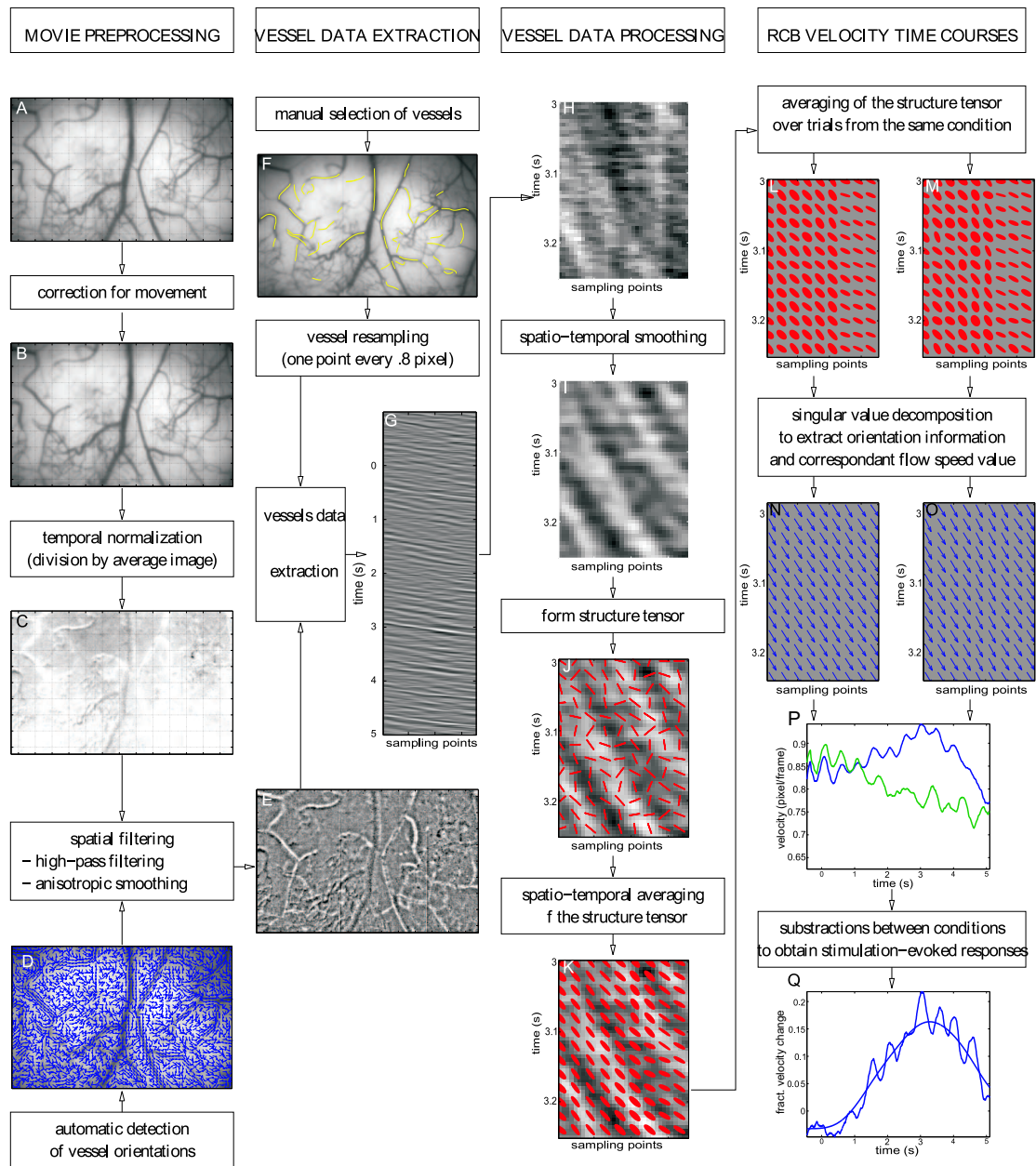


Figure 7.2: Summary of the flow estimation procedure (see text).

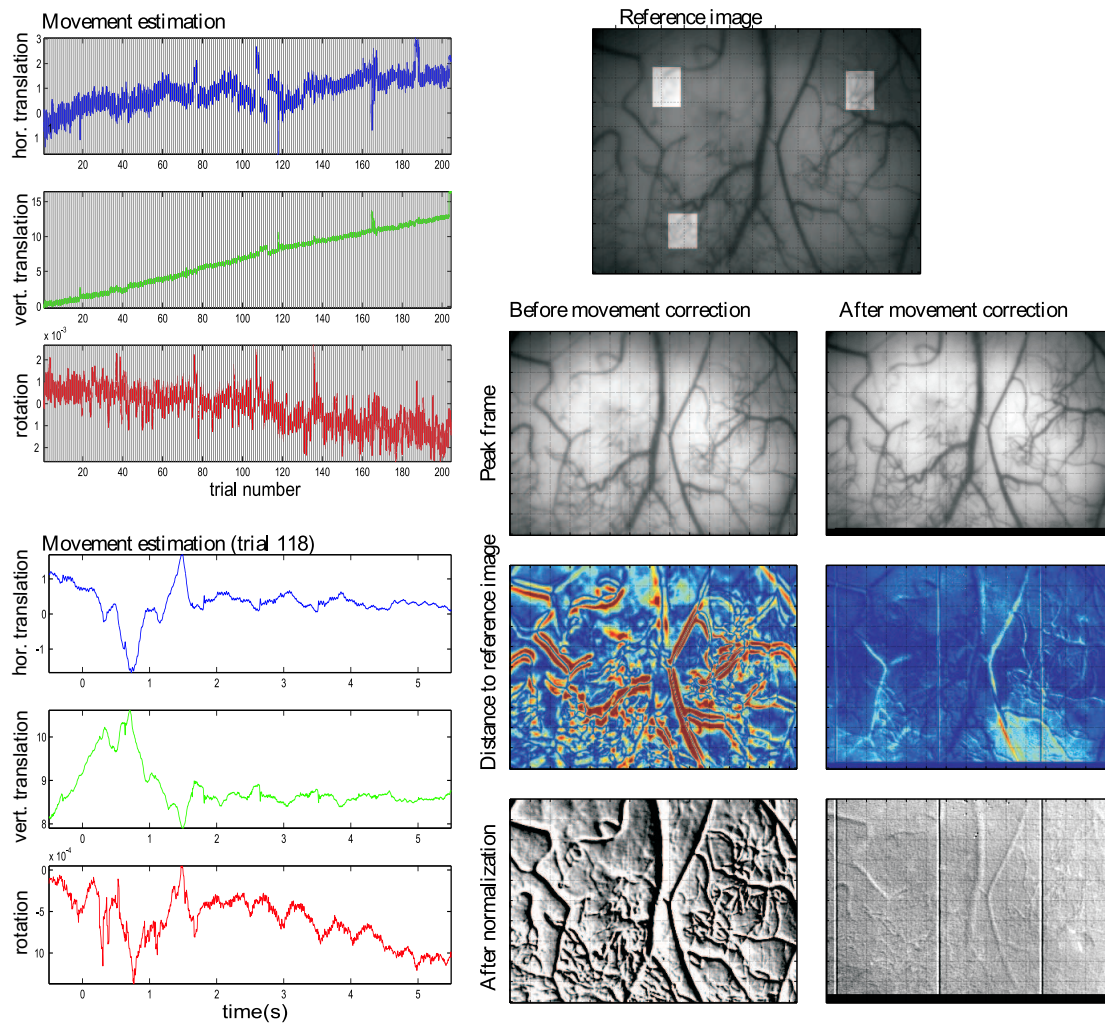


Figure 7.3: The image co-registration procedure (see text).

Please note that we used a linear interpolation to calculate the transformed images $\mathcal{T}_t \circ \mathcal{R}_\theta(I_1)$. Then, we used the Matlab function `fminsearch`, which implements the simplex search method [110], for minimizing this energy. Once the parameter had been estimated, the frame was re-calculated using a linear interpolation.

The figure 7.3 shows the co-registration results. First please note that in order to speed up the estimation procedure (which needed to be applied for nearly 250.000 frames), we actually used only a small part of the points in the image in the energy computation, which appears brightest on the display of the reference image. On the left, the estimated translation and rotation parameters are displayed for all trial repetitions and details are shown for the trial which contained the most movements. On the right we display the details for the frame, which corresponds to the peak of movement (at 0.8s): the frame, its distance to the reference, and the result of subsequent normalization (see below for details on this pre-processing step) are displayed, before and after movements have been corrected for. The registration appeared to work accurately, and the large corrections needed for the last images underscore the importance of this step. Incidentally, the increase in the images' shift is most likely due to a slow readjustment of the intracranial pressure. Indeed, the cranial chamber had been opened for cleaning just before imaging.

Filtering

Let us be reminded that hemoglobin - the main absorber in the blood - is contained in RBCs, which travel along vessels either alone or grouped in small clusters. At their specific location in the vessel, they cause local decreases in reflected light intensity. When the vasculature is imaged at appropriate wavelengths and at sufficiently high resolution, these RBCs (or clusters thereof) thus appear in the recorded images as local contrast variations which drift along the vasculature with time. In the raw images, those small contrast gradients are obviously masked by the much larger contrast differences existing between blood vessels and parenchyma (figure 7.2, displays A before and B after the re-alignment step). Since those latter patterns are, however, nearly constant in time, they can be eliminated by normalizing for the static light intensity differences among different parts of the image, dividing, on a pixel by pixel basis, each frame by its mean value, calculated over an acquisition sequence (trial), allowing us thereby to "flatten the images" (figure 7.2, display C).

Afterwards, the intravascular contrasts can still be reinforced by adequate spatial filtering: first the global light-intensity variations inside the images can be eliminated using a high-pass filter. Second, the signal to noise ratio can be increased by a low-pass filter that smoothes the images in small neighbourhoods, and thus cancels the random independent noises of the pixels in such neighbourhoods. Please note that we used an anisotropic smoothing that preferentially operated orthogonally to the vessels, so as not to decrease the RBC contrasts in the vessel directions. The vessel directions were obtained by the same tensor structure technique as that which we shall describe below for the estimation of RBCs' speed (display D), and the anisotropic smoothing was achieved through using a family of directional Gaussian kernels (filter half-widths were 3 pixel in the principal dimension and 1 pixel in the orthogonal dimension). As a result, the intravascular contrasts are largely enhanced (display E), and it becomes possible to track individual clusters of RBCs moving along single vessels (see the zooms in figure 7.1).

Extraction of Vessel Data

Although common image segmentation algorithms allow automatic detection of the vasculature, here, we preferred to select individual vessels by hand, mostly to be able to avoid RBC motion tracking ambiguities arising at the crossings of different vessels or at vascular branching points (display F in figure 7.2).

Since RBCs travel inside blood-vessels, the direction of their velocity-vector in each individual vessel is obviously determined by the vascular anatomy and is assumed to be parallel to the vessel's axis. Next, the amplitude of the RBCs' velocity (i.e. their speed in the vessel) must be extracted from the spatio-temporal data, at each position x along the particular vessel and at each time point t . For each vessel and for each individual trial, we thus cast the data into a 2D array (display G and zoom in display H), containing the recorded light intensity values $I(x, t)$ for each of the n_t time points (i.e. the number of frames) and for each of the n_x spatial points along the vessel.

In this image, thin oblique lines can be observed. They represent the spatio-temporal trajectories of RBCs in the vessel and contain the relevant motion information. Indeed, they indicate how the spatial RBC contrasts move along the vessel during time. The slope of these lines is directly related to the RBCs' speed: if a RBC moves from position x to position x' between times t and $t + 1$, then its speed along the vessel direction is $(x' - x)/(t + 1 - t) = x' - x$. The higher the speed, the more these lines approach a horizontal direction.

As a consequence, we aim at estimating the directions of these lines. We first enhanced them by 2-D low-pass filtering, to enhance the signal to noise ratio (display I). Please note however that this filtering could have been applied equivalently on the images, before vessel data extraction (it would have consisted in a temporal low-pass filtering and an additional anisotropic spatial smoothing). The reason why we applied them on these 2-D matrices is that the filter parameters could be more easily chosen through the visualization of the obtained lines.

7.2.2 Detecting Hemoglobin Trajectories with the Structure Sensor

We present here how to estimate these line directions. Let's notate I as the image representing intensity time courses in the vessel, I_t and I_x as its derivatives with respect to time and space. Then the gradient vector at point (x, t) is $(I_x(x, t), I_t(x, t))$, and, if we ignore noises, the direction of trajectories lines is its orthogonal vector $(-I_t(x, t), I_x(x, t))$. RBC speed is then given by the inverse of this trajectories slope, $v(x, t) = -I_t(x, t)/I_x(x, t)$.

Formally, we can prove this formula by writing the optical flow equation [28], which states that reflected light intensity does not change along particle trajectories: if we notate $x(t)$ the position of a given particle at time t , we have

$$\begin{aligned} \frac{d}{dt}I(x(t), t) &= 0 \\ vI_x + I_t &= 0 \\ v &= -I_t/I_x. \end{aligned} \tag{7.1}$$

This formula for RBC speed, however, cannot be applied directly because of noise in the image. It is necessary to stabilize the algorithm by using a neighbourhood averaging system. The goal is to find an average direction of trajectories over a small neighbourhood (we used a neighbourhood of about 30 time points by 5 pixels length), which can

be achieved with the structure tensor. The structure tensor is known for enabling the determination of a local mean orientation for intensity gradient in an image [19].

Let us define the following degenerate symmetric matrix:

$$A(x, t) = \begin{pmatrix} I_x(x, t) \\ I_t(x, t) \end{pmatrix} \begin{pmatrix} I_x(x, t) & I_t(x, t) \end{pmatrix} = \begin{pmatrix} I_x^2 & I_x I_t \\ I_x I_t & I_t^2 \end{pmatrix}.$$

The first eigenvector of A is the gradient direction (I_x, I_t) and its second eigenvector (with zero eigenvalue) is our trajectories direction $(-I_t, I_x)$. The structure tensor is defined by averaging matrix A over a small neighbourhood

$$S(x, t) = \sum_{(x', t') \in \mathcal{V}(x, t)} A(x', t').$$

Then the first eigenvector of S gives the direction in which image intensity varies the most ; the second one gives the direction in which image intensity varies the least (which is precisely the direction we are looking for) and the inverse of its slope is the estimated RBCs velocity.

The displays J and K in figure 7.2 show the variations of these matrices A and S in the image. To better understand them, let us recall that each of these symmetric matrices defines a scalar product and a distance in \mathbb{R}^2 . The ellipses shown in the displays represent scaled unit bowls for these distances. Their major and minor axis are respectively the second and first eigenvectors of the matrices. Since the matrix A is degenerate, the corresponding ellipses (display J) are actually segments pointing in the direction orthogonal to the image gradient. The average direction of these segments is that of the lines, but there are large random fluctuations due to the noise in the image. On the other hand, the spatial averaging included in the structure tensor (display K) allows detecting an average direction. Please note that the noisier the data, the more ellipses tend to become circles, which eventually would preclude the estimation of direction.

At this stage, RBC-motion time courses can be extracted from the line direction information, given by the singular value decomposition of the structure tensor. One time course can be estimated for each point of the vessel (i.e. for each column of the lines image), and a mean time course can be computed for the whole vessel. The left part of figure 7.4 shows an example of a signal estimated from a single trial. A first way to average over all the trials from a same experimental condition would be to extract these RBCs motion time courses for every trial, and then average them, in order to produce an estimated response to the condition. However, it happened in our dataset that the estimation procedure failed for a few trial repetitions: at some spatio-temporal points, the lines were not clearly visible, and the structure tensor was much closer to a circle, resulting sometimes in estimated orientations that approached the horizontal direction and hence to values of the flow approaching infinity and causing divergences, even in the average responses. As a consequence, it turned out to be valuable to average the structure tensor itself over trials (instead than averaging the inverse of its principal direction slope). The displays L and M show these averaging for the two conditions in our experiment (L: stimulated condition, M: blank), and the displays N and O show the line directions obtained through the singular value decomposition of these average structure tensors. As a result, the responses to the stimulated and blank conditions, estimated from every trial and averaged on every points of the vessel, are shown in the

display P, and their subtraction (stimulated - blank), which gives the sensory-evoked response, is shown in display Q.

7.2.3 Conducted Experiments

Experimental Setup

The experimental setup for simultaneous optical imaging and laser-Doppler acquisitions was similar to that described in the previous chapter, except for the resolution of the optical recordings. As a matter of fact, higher spatial and temporal resolutions were required: spatial precision was needed to focus on the vessels, and temporal precision was needed so that RBC motions inside the vessels remained at acceptable levels (not much faster than one pixel per frame). We imaged a square $\sim 3.5 \times 3.5$ mm (10μ m per pixel) of cortex at 200Hz. Additionally, to allow for an optimal identification of vascular activity, the optical system [150] was focused onto the cortical vasculature, rather than 300μ m below the surface as traditionally done.

Experimental Paradigm

We conducted two experiments to assess the validity of our technique. The first experiment aimed at testing the technique on the detection of CBF modulations by the heart beat. We thus simultaneously recorded optical and laser-Doppler signals. The monkey endured no particular stimulation during this experiment. However, he was asked to fixate a small red fixation point to avoid visual responses due to eye-movements, and, moreover, recordings were synchronized with the heart beat recorded independently by a pulse-oximeter. Around one hundred trials were acquired.

In the second experiment, we used a visual stimulation, consisting in a drifting square grating that was displayed for 0.6s. Stimulated trials were alternated randomly with “blank conditions (no grating). There were 136 repetitions of the stimulated condition, and 68 repetitions of the blank condition.

7.3 Results

7.3.1 Extraction of RBC-Motion Information

The figure 7.4 shows results from our first experiment, through comparing the extracted motion signal (blue traces) with the simultaneous CBV (green traces) and laser-Doppler CBF measurements. The left part of the figure displays these time courses for a single vessel and a single trial. The right part of the figure superimposes the estimations and measurements for several vessels, calculated over every trial as explained in the Method section. The same cardiac pulsation frequency was obtained with all three measurements (as well as by a commercial pulse oximeter applied on the monkey’s ear, not shown), strongly supporting the reliability of our method. Please note that the decay of oscillations in the right part of the figure is due to that the following: since the acquisition is synchronized with heart-beat pulsation, the oscillating signals for every trial are in phase at the start of acquisition, and thereafter this synchronization decreases with time due to small physiological variations in the cardiac rhythm.

Interestingly, the phase of the cardiac pulsation differed in the three measurements. To test whether the different spatial localization of the measurements might explain the

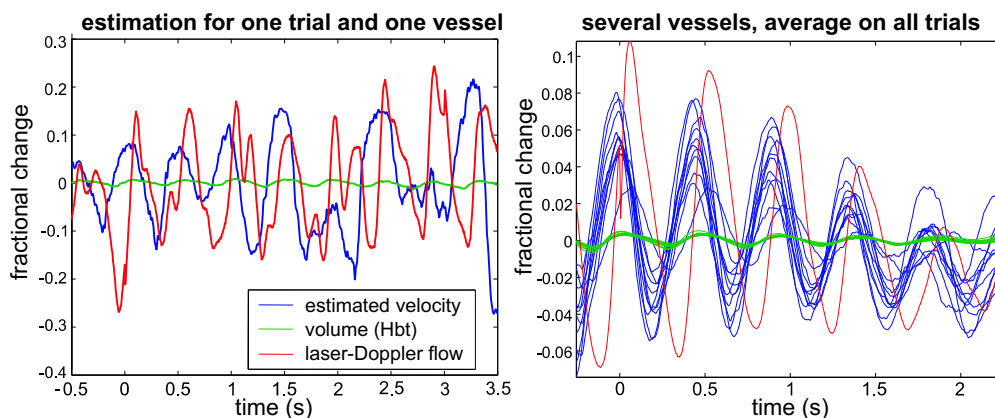


Figure 7.4: *Heart beat estimation* The flow estimations conducted for the first experiment (no stimulation) shows the fluctuation due to the heart pulsation. Left: result for one trial repetition in one vessel (blue line). Right: average response for all trial repetitions; the responses in several vessels are superimposed (blue lines). They are to be compared with the volume time courses (green), obtained directly from the optical signals, and the laser-Doppler measurements (red). The frequency of the heart pulsation was found accurately by our method. The small phase shift between our estimation and laser-Doppler measures is most likely due to different locations of the laser-Doppler probe and the vessels considered.

observed phase-lag, in figure 7.4 we compared the RBC speed time-courses obtained in 6 different vessels, averaged over trials to increase the SNR. Part of these phase differences can be explained by the different spatial localization of the measurements, and, as a matter of fact, we also found phase differences between the RBC-motion signals in different vessels, suggesting different delays in blood transfer. However, the observed phase-lags were only small and are thus unlikely to explain the large dephasings, in particular between Laser-Doppler and RBC-speed signals. Its precise origin thus remains to be understood.

7.3.2 Sensory-Evoked Changes in RBCs' Speed

We calculated sensory-evoked responses to the visual presentations in 45 vessels. Figure 7.5 shows these responses for 7 representative vessels: the RBCs-motion responses (blue traces) are super-imposed with CBV responses (in green) in the same vessels. Error bars were calculated by gathering the trials in groups of 16 repetitions of the stimulated condition and 8 repetitions of the blank condition. The qualitative aspects of the results were quite different depending on the nature of the vessel. The arteries showed the best-quality responses. These responses occurred early on and returned to baseline while the volume was still increasing. Conversely, no sustainable response was found in the veins, possibly because of lower RBCs contrasts (lines appeared less clearly in the images). In smaller vessels, response patterns could be seen too, although they were noisier, and additional acquisitions would probably be necessary to bring more evidence to these responses. In particular, in these vessels, the RBC-speed often initially dropped after the stimulus onset, before it began to increase. Currently, we do not have a clear explanation of this phenomenon, which has not been reported by Laser Doppler flowmetry studies. One possible explanation for the discrepancy is that, here, we specifically measure RBC speed, whereas blood flow as measured by the Laser Doppler technique depends both

on RBC speed and density.

7.4 Discussion

7.4.1 RBCs-Motion Estimation

We proved that, by tracking erythrocytes moving through the vasculature, it is possible to use intrinsic optical imaging to estimate the fluctuation of cerebral blood flow in micro-vessels. Our method relies upon two facts: first of all, it is possible to extract the data from individual vessels in order to produce 2D images where the motion of RBCs can be casted into oblique lines. Secondly, it is possible to automatically estimate the directions of these lines by the use of standard image analysis methods, and hence estimate the speed of RBCs inside the vessels.

As far as data extraction is concerned, it is essential to spatially co-register all the recorded frames. As a matter of fact, *in vivo*, the cortical vasculature is not static, but moves under the effects of respiration and heart-beat pulsation and, in the case of awake preparations, as a result of the animal's movements. We found that estimating rigid transformations was sufficient to account for the largest motions observed in these images and extract reliably information from the vessel data. However there are not only global motions, but also local deformation that would require more sophisticated, nonlinear image registrations. These local effects could be observed on large vessels, as artefacts related to the movements of these vessels affected by strong heart-beat pulsation.

Although beyond the scope of the present study, the development of efficient optimization procedures for estimating non-rigid image deformation is clearly of primary importance. It would also enable the co-registration of datasets acquired in different imaging sessions, where the position and angle of the camera with respect to the cortex is often not exactly reproducible.

Next, we proposed a method based on the structure tensor to estimate directions in the lines' images. The structure tensor provided a natural way to average directions in an image, which were calculated in every pixel as the direction orthogonal to image gradient. It thus enables us to overcome noises in the lines' images. To produce accurate results, high temporal and spatial resolutions were needed: our acquisitions were sampled at 200Hz, and pixels were 10x10 microns.

The methods proved adequate to accurately estimate the value of RBCs' speed, and even its temporal fluctuations. As a matter of fact, we could capture the heart pulsation in the vessels, and the estimated oscillations were in agreement with those observed on CBV signals and CBF acquired with the laser-Doppler technique. Additionally, we could observe sensory-evoked responses, with estimation qualities depending on the nature of the vessels. As a caveat, please note that additional work will be necessary to definitively rule out stimulation-induced artefacts (such as vessel movements under an increased blood pressure) as contributions to the responses we found. .

Since the tensor structure is a quite standard image analysis tool, many improvements in the method are probable. For instance, a disadvantage of this method is that it is not robust to the presence of several distinct directions in a same image, which can occur because of vessel crossing or artefacts due to local deformations of the vasculature. Other methods such as wavelet analysis and Gabor filter may be better suited to choose a maximum likelihood direction rather than averaging the several directions present in

the image. Even more, we have only considered one spatial dimension so far, i.e. the direction of the vessels. However, more accurate optical flow techniques could operate on the full width of the vessels rather than extracting signals along one-dimensional lines. This may be decisive to obtain reliable signals in vessels (veins in particular) for which the line directions did not appear clearly in the extracted spatio-temporal 2D images.

7.4.2 Application for Physiological modelling

As opposed to standard methods providing only point-like blood flow measurements, our method allows us to determine blood flow and its modulations across a 2 dimensional image. This should prove useful to study the neurally-evoked blood flow responses, and in particular, the blood-flow in different vascular compartments.

Still, the responses we obtained in small arteries, which were the ones estimated with the highest reliability, already lead to interesting physiological observations. In these vessels, the blood velocity reaches a maximum and then returns to baseline while at the same time the blood volume is still increasing. Conversely, the signals in the venules appear to be more synchronized with volume variations, or even slightly delayed. This can be put in relation with the results of the study we presented in the previous chapter on CBV and its relation with CBF, where we observed that the CBV response preceded that of CBF. It could suggest that the measurements acquired with laser-Doppler originated more from the non-arterial rather than from the arterial compartments. Moreover, in the two studies we could observe initial drops in the flow responses. Though currently we do not have an explanation for these observations, they suggest complex interactions between the various hemodynamic parameters (CBV, CBF, RBC speed, RBC density), which constitute an interesting subject for further investigations.

In conclusion, it appears that in order to address modelling questions, it is highly valuable to separate the signals issued from the different vascular compartments. Lastly, we point out that the method we presented here (concerning blood flow in individual micro-vessels) is complementary to the recently developed laser-speckle flowmetry [35, 179] which appears to be sensitive mainly to cerebral blood flow in parenchyma, i.e. in the capillary bed.

7.4.3 Conclusion

The method we have presented appears promising for recording the blood velocity in the cerebral vasculature with standard intrinsic optical recording. We would like first to consolidate the method with robust image processing techniques. Then we will address modelling questions on the microvascular blood circulation.

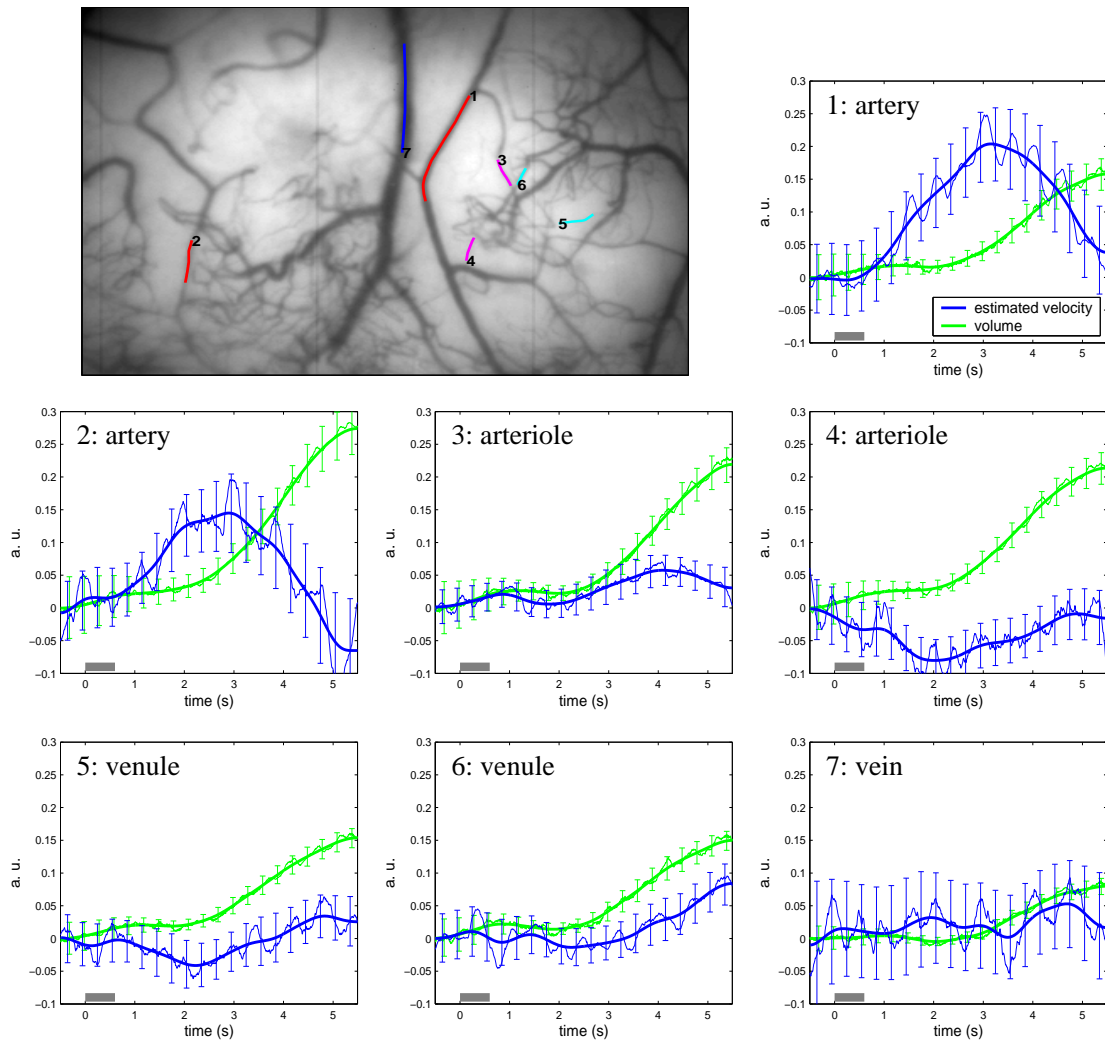


Figure 7.5: *Sensory-evoked responses* The response of CBF (blue) and CBV (green) are displayed superimposed for a few chosen vessels. The quality of sensory-evoked flow responses depend on the vascular compartment they originate from: there are significant responses in the arteries, which appear to be much in advance on those of the volume. Responses in arterioles and venules seem to exhibit some characteristic patterns that could be sensory-evoked responses, but the level of noise is more important. Lastly, no response is detected in the veins.

Part IV

Using Nonlinear Models in EEG-fMRI Fusion

Chapter 8

EEG-fMRI Fusion using Kalman Filtering

We have seen how the framework of stochastic dynamical equations, which we used for parameter estimation in fMRI, can be extended to measures from other modalities. In this chapter, we extend our models to the EEG measure by adding the spatial dimension (i.e. considering all cortical locations as a whole instead of studying each point independently), and propose to use the Kalman filtering technique to estimate the unknown cortical activity.

This method can be seen as a possible answer to a very large inverse problem, where fMRI and EEG are noisy and partial measures of the neuronal signals. As a result, it can also be applied to fMRI or EEG alone, in which case it could be considered as a deconvolution method [67] for fMRI analysis, or as an inverse problem that ensures temporal smoothness for EEG analysis.

8.1 Methods

8.1.1 Why Choose the SDE Formulation ?

We would like to present the reflection that led us to use the SDE formulation and temporal filtering techniques, because it may be instructive to better understand the problems underlying the EEG-fMRI fusion, and maybe imagine new methods.

The principle of EEG-fMRI fusion relies on the assumption that there exists some kind of neural activity, denoted by N , which is seen by the two modalities. Our aim is to estimate this activity. This N has a very high dimensionality since it gathers the time courses of brain activity at all cortical locations, the spatial and temporal resolution being those of fMRI and EEG, respectively. Basically, one could consider the two measures as noisy functions of this neural activity:

$$\begin{cases} z &= f_{\text{eeg}}(N) + \eta_{\text{eeg}} \\ y &= f_{\text{fmri}}(N) + \eta_{\text{fmri}}, \end{cases} \quad (8.1)$$

where the noises η_{eeg} and η_{fmri} would be sums of physiological effects and measure artefacts. Please note that in the physiological effects we include activities seen by only one of the two modalities. All the variables in this system are matrices whose spatial dimensions are the number of cortical sources (for N , y , η_{fmri}) or the number of EEG electrodes (for z , η_{eeg}), and whose temporal dimensions are the number of samples of the

EEG (for N , z , η_{eeg}) or of the fMRI (for y , η_{fmri}). The figure 8.1, which represents the physiological models that we will describe below, shows these matrices. If the functions f_{eeg} and f_{fmri} are known, estimating N becomes an inverse problem that one could try to solve by minimizing an adequate energy containing fit to data and smoothness terms.

But minimization over such a large unknown would be very long. It is worth then to consider that the system (8.1) has some structure. In particular, the EEG measure at a given instant t only depends on the neural activity at the same instant over the whole cortex, thanks to the quasi-static approximation of Maxwell's equations. At the same time, we could consider that, in a first approximation, the BOLD signal measured at a given spatial location s only depends on the neural time course at the same location. Hence we write:

$$\begin{cases} z(t) &= f_{\text{eeg}}^{(t)}(N(t)) + \eta_{\text{eeg}}(t) \\ y(s) &= f_{\text{fmri}}^{(s)}(N(s)) + \eta_{\text{fmri}}(s). \end{cases}$$

Now the variables in this system are vectors, represented by the blue sections in figure 8.1. EEG measure function $f_{\text{eeg}}^{(t)}$ and the distribution of the noise $\eta_{\text{eeg}}(t)$ are actually the same for all t . On the other hand, fMRI measure function $f_{\text{fmri}}^{(s)}$ and noise $\eta_{\text{fmri}}(s)$ distribution may vary across regions.

There appears to be some symmetry between time and space, and EEG and fMRI. However, there is a major difference between them: while the EEG measure in one electrode really depends on neural activity at all cortical locations, the fMRI signal at a given time t only depends on neural signals in the about last thirty seconds. As a consequence, estimations of neural activity at two sufficiently distant moments should be independent. This is the reason why we came to temporal-filtering and belief-propagation techniques. It consists in considering only all locations together instead of all locations and all instants together; and then running the estimation for increasing times, while the filtering technique ensures that the temporal interactions are respected.

Temporal filtering methods rely on a stochastic differential equations formulation, where this space/time symmetry is broken:

$$\begin{cases} \dot{X}(t) &= \mathcal{F}(X(t)) + \xi(t) \\ Y(t) &= \mathcal{G}(X(t)) + \eta(t). \end{cases} \quad (8.2)$$

- X denotes the system hidden-states: the neural activity N and the other quantities that participate to the EEG and fMRI signals, like hemodynamic variables. Note that when we identified dynamical systems in fMRI in the chapter 4, we approximated the neural activity by the stimulation sequence u , which was an external input to the system. Now there is no more system input and the neural activity N belongs to the hidden states X . The temporal evolution of X is characterized by the evolution function \mathcal{F} and the evolution noise ξ .
- Y embeds the EEG and fMRI measures, characterized by the measure function \mathcal{G} and the measure noise η . \mathcal{G} operates directly on the neural activity at time t , or on the other hidden-states.

We will first describe the physiological equations used for this system, and then the Kalman filter and Kalman smoother used to estimate X given the measures Y .

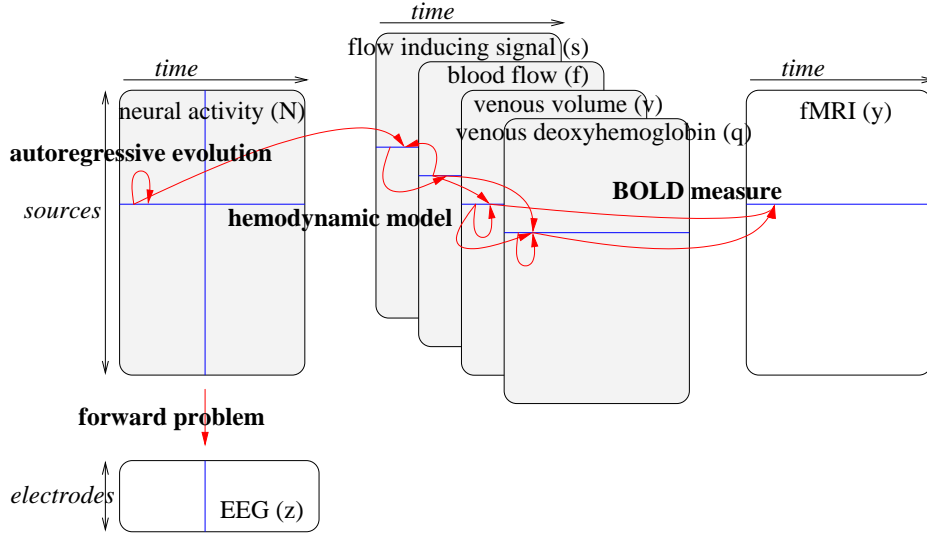


Figure 8.1: *Fusion model*. EEG measure is described by the forward problem, while a physiological dynamical system involving neural activity and hidden hemodynamic quantities underlies fMRI measure. Our fusion algorithm estimates the whole spatio-temporal time course of neural activity, but also of all hemodynamic variables.

8.1.2 Physiological Models

We denote by n_s the number of cortical sources considered, n_e the number of EEG electrodes, T the number of EEG samples and n_i the number of fMRI samples.

At the heart of our model, we suppose that the cortical activity N , which is seen by both the EEG and fMRI measures, can be described by the electrical activity of synchronous assemblies of neurons. We thus suppose it is equal to the electrical activity on the cortical surface ($N = j$).

In order to assume temporal and spatial smoothness of this sources activity, we model a simple autoregressive evolution:

$$\dot{N}(t) = -\lambda N(t) + \xi_N(t),$$

where $\lambda > 0$ is a feedback term (return to zero) ξ_N is a Gaussian evolutive noise $\xi_N(t) \sim \mathcal{N}(0, Q_N)$. Q_N is constructed according to geometrical constraints to ensure the spatial smoothness of the cortical activity. The temporal smoothness can be tuned through the value of λ and the scaling of Q_N .

Since N denotes electrical activity, the EEG measure is obtained directly by the forward problem:

$$z(t) = GN(t) + \eta_{\text{eeg}}(t),$$

where G is the $n_e \times n_s$ Lead-Field matrix for the EEG forward problem, and $\eta_{\text{eeg}}(t) \sim \mathcal{N}(0, \Sigma_{\text{eeg}})$.

Let us consider the BOLD responses now. It is described by the hemodynamic models studied in the previous chapters. We consider that there is an independent hemodynamic process in every cortical location s :

$$\begin{cases} \dot{x}(s, t) &= F(x(s, t), \phi(N(s, t))) + \xi_{\text{fmri}}(s, t) \\ y(s, t) &= G(x(s, t)) + \eta_{\text{fmri}}(s, t). \end{cases}$$

We recognize in this system the set of hemodynamic variables x , the evolution and measure functions F and G , the evolution and measure noises ξ_{fmri} and η_{fmri} that are supposed white and Gaussian. The model input should not be the electrical activity N directly, but rather a metabolic activity derived from N according to a - possibly nonlinear - function $\phi(N)$. For the simulations shown below, we used the Balloon model (4.2) formulated by Friston et al. in [59], and we supposed a simple linear interaction between electrical and metabolic activities $\phi(N) = \epsilon N$.

To clarify our notations, we re-write these equations together, to fit the general system formulation in (8.2):

$$\begin{cases} \dot{N}(t) &= -\lambda N(t) + \xi_N(t) \\ \dot{s}(t) &= \epsilon N(t) - \kappa_s s(t) - \kappa_f (f(t) - 1) + \xi_s(t) \\ \dot{f}(t) &= s(t) + \xi_f(t) \\ \dot{v}(t) &= \frac{1}{\tau} (f(t) - v(t)^{1/\alpha}) + \xi_v(t) \\ \dot{q}(t) &= \frac{1}{\tau} (f(t)^{\frac{1-(1-E_0)^{1/f(t)}}{E_0}} - v(t)^{1/\alpha-1} q(t)) + \xi_q(t) \\ z(t) &= GN(t) + \eta_{\text{eeg}} \\ y(t) &= V_0(k_1(1 - q(t)) + k_2(1 - q(t)/v(t)) + k_3(1 - v(t))). \end{cases} \quad (8.3)$$

Hidden states $N(t)$, $s(t)$, $f(t)$, $v(t)$, $q(t)$ are all vectors of dimension n_s . They are concatenated in a unique hidden-state vector $X(t)$ which dimension is thus $5n_s$. All operations above (multiplications, divisions, power laws) are performed element-wise. The EEG and fMRI measure vectors z and y have respectively n_e and n_s elements. Note also that fMRI has much fewer sampling points than EEG. However, we observed that the Kalman filter we present now was more robust if we interpolated the fMRI measure at every EEG sampling point. At last, we suppose that at time $t = 0$, all variables are at rest state (i.e. $N = s = z = y = 0$ and $f = v = q = 1$).

8.1.3 Kalman Filter and Kalman Smoother

Given the measure Y , we can estimate the hidden state X recursively for increasing times, by using an extended Kalman filter [180, 36]: if both evolutive and measure noises are Gaussian, and if at each sampling instant the nonlinear evolution and measure functions are approximated up to first order, then the a posteriori distribution of the hidden states is also Gaussian; the Kalman filter calculates the mean and variance of this distribution.

We chose to expose here mathematical details of the Kalman methods. However, the most important point of this section is that the Kalman filter iteratively calculates a posteriori means and variances of the hidden-states given the measures in the past and the present, and that the Kalman smoother is a second iterative filter that operates backward (from the future toward the past) and calculates a posteriori means and variances given all the measures.

We re-write the system equations in their discretized form (please pay caution we use the same notations \mathcal{F} and \mathcal{G} as in (8.2), but only \mathcal{G} is really the same function):

$$\begin{cases} X_{k+1} &= \mathcal{F}(X_k) + \xi_{k+1} \\ Y_k &= \mathcal{G}(X_k) + \eta_k. \end{cases} \quad (8.4)$$

The Kalman filter computes the distributions $p(X_k|Y_1..Y_k)$, and the Kalman smoother, the distributions $p(X_k|Y_1..Y_T)$. We will use the following notations for means and variances:

$$\begin{aligned}\hat{X}_k^l &= E(X_k|Y_1..Y_l) \\ P_k^l &= V(X_k|Y_1..Y_l) = E((X_k - \hat{X}_k^l)(X_k - \hat{X}_k^l)^T|Y_1..Y_l).\end{aligned}$$

Kalman Filter

The Kalman filter starts with an a priori distribution on X_1 : $p(X_1) \sim \mathcal{N}(\hat{X}_1^0, P_1^0)$. Then it iteratively applies the two update steps:

- *Measure update*: update the a posteriori distribution $p(X_k|Y_1..Y_{k-1})$ by the measure Y_k , to obtain the a posteriori distribution $p(X_k|Y_1..Y_k)$; during this step, the a posteriori variance decreases since additional information is brought by the measure.
- *Evolution update*, also called *belief propagation*: use the known distribution $p(X_k|Y_1..Y_k)$ to predict that of hidden-states at the next instant, $p(X_{k+1}|Y_1..Y_k)$; during this step, the a posteriori variance increases on account of the unknown evolution noise.

Evolution Update

We suppose known the a posteriori distribution of state X_{k-1} given measures until instant k , $p(X_{k-1}|Y_1..Y_{k-1}) \sim \mathcal{N}(\hat{X}_{k-1}^{k-1}, P_{k-1}^{k-1})$, and want to predict that of X_k .

First, the extended Filter algorithm needs to linearize the evolution equation around the a posteriori mean:

$$X_k \approx \mathcal{F}(\hat{X}_{k-1}^{k-1}) + A(X_{k-1} - \hat{X}_{k-1}^{k-1}) + \xi_k,$$

$$\text{with } A = \frac{\partial \mathcal{F}}{\partial X}(\hat{X}_{k-1}^{k-1}).$$

$(X_k|Y_1..Y_{k-1})$ is the sum of Gaussian variables, hence it is Gaussian too with mean and variance

$$\begin{aligned}E(X_k|Y_1..Y_{k-1}) &= \mathcal{F}(\hat{X}_{k-1}^{k-1}) + A(E(X_{k-1}) - \hat{X}_{k-1}^{k-1}) + E(\xi_k) \\ V(X_k|Y_1..Y_{k-1}) &= V(AX_{k-1}) + V(\xi_k).\end{aligned}$$

$$\boxed{\begin{aligned}\hat{X}_k^{k-1} &= \mathcal{F}(\hat{X}_{k-1}^{k-1}) \\ P_k^{k-1} &= AP_{k-1}^{k-1}A^T + Q.\end{aligned}}$$

Measure Update

During the measure update, a posteriori mean and variance of X_k are adjusted to take into account the new information from the measure Y_k .

Again, we need to linearize the measure equation:

$$Y_k \approx \mathcal{G}(\hat{X}_k^{k-1}) + C(X_k - \hat{X}_k^{k-1}) + \eta_k,$$

with $C = \frac{\partial \mathcal{G}}{\partial X}(\hat{X}_k^{k-1})$.

The a posteriori distribution of X_k given measures until Y_k is determined using the Bayes rule:

$$p(X_k|Y_1..Y_k) = \frac{p(Y_k|X_k, Y_1..Y_{k-1})p(X_k|Y_1..Y_{k-1})}{p(Y_k|Y_1..Y_{k-1})}$$

The denominator is an uninteresting constant, and thanks to the Markov chain property of the system, we have $p(Y_k|X_k, Y_1..Y_{k-1}) = p(Y_k|X_k)$ (conditionally to X_k , Y_k is independent of measures before the instant k). For convenience, we denote $\tilde{X}_k = X_k - \hat{X}_k^{k-1}$ and $\tilde{Y}_k = Y_k - \mathcal{G}(\hat{X}_k^{k-1})$.

$$\begin{aligned} \log(p(X_k|Y_1..Y_k)) &= \log(p(\eta_k = \tilde{Y}_k - C\tilde{X}_k|X_k)) + \log(p(X_k|Y_1..Y_{k-1})) \\ &= -\frac{1}{2}((\tilde{Y}_k - C\tilde{X}_k)^T R^{-1}(\tilde{Y}_k - C\tilde{X}_k) + \tilde{X}_k^T [P_k^{k-1}]^{-1} \tilde{X}_k) + \text{cte} \\ &= -\frac{1}{2}(\tilde{X}_k - (C^T R^{-1} C + [P_k^{k-1}]^{-1})^{-1} C^T R^{-1} \tilde{Y}_k)^T (C^T R^{-1} C + [P_k^{k-1}]^{-1}) \\ &\quad (\tilde{X}_k - (C^T R^{-1} C + [P_k^{k-1}]^{-1})^{-1} C^T R^{-1} \tilde{Y}_k) + \text{cte}. \end{aligned}$$

We found that $p(X_k|Y_1..Y_k)$ is a Gaussian law with mean $\hat{X}_k^{k-1} + (C^T R^{-1} C + [P_k^{k-1}]^{-1})^{-1} C^T R^{-1} (Y_k - \mathcal{G}(\hat{X}_k^{k-1}))$ and variance $(C^T R^{-1} C + [P_k^{k-1}]^{-1})^{-1}$. Actually, it is possible to show by matrices transformations [180] that they are equal to

$$\begin{array}{l} \hat{X}_k^k = \hat{X}_k^{k-1} + K(Y_k - \mathcal{G}(\hat{X}_k^{k-1})) \\ P_k^k = (I - KC)P_k^{k-1} \\ K = P_k^{k-1} C^T (R + CP_k^{k-1} C^T)^{-1}. \end{array}$$

Also, Max Welling in [180] argued that in practice the obtained matrix P_k^k may not be symmetric positive-definite due to numerical imprecision, and recommends to use the following form instead:

$$P_k^k = (I - KC)P_k^{k-1}(I - KC)^T + KRK^T.$$

Kalman Smoother

The Kalman filter does not use all the information from the measures to derive a posteriori probabilities, since it only takes the past and present into account. The Kalman smoother operates backward to include the missing information from measures in the future, starting with the last Kalman filter estimate, $p(X_T|Y_1..Y_T)$.

Indeed, it is possible, once we know the a posteriori distribution of X_{k+1} , $p(X_{k+1}|Y_1..Y_T)$, to deduce that of X_k . However, the derivation is quite complex and we only give the results here. Again, please refer to [180] for a complete demonstration.

The mean and variance of the distribution $p(X_k|Y_1..Y_T)$ are given by:

$$\boxed{\begin{aligned}\hat{X}_k^T &= \hat{X}_k^k + J(\hat{X}_{k+1}^T - \hat{X}_{k+1}^k) \\ P_k^T &= P_k^k + J(P_{k+1}^T - P_{k+1}^k)J^T \\ J &= P_k^k A^T [P_{k+1}^k]^{-1},\end{aligned}}$$

with $A = \frac{\partial \mathcal{F}}{\partial X}(\hat{X}_k^k)$.

Because more observations are included in the estimation of the state, the result will be less noisy compared to the Kalman filter result, hence the name ‘‘smoother’’.

8.1.4 Implementation

8.3.2

We already mentioned that, for numerical reasons, we interpolated the fMRI time course to the sampling of the EEG (20Hz for our simulations). We noticed on preliminary simulations with small datasets that the linear approximations of the model could be used without significant change in the results:

$$\begin{cases} \mathcal{F}(X) &= \left(\frac{\partial \mathcal{F}}{\partial X} \right)_{X=0} X \\ \mathcal{G}(X) &= \left(\frac{\partial \mathcal{G}}{\partial X} \right)_{X=0} X \end{cases}$$

We notice then that under this linearization the evolution of the a posteriori variances P_k^k is independent of the measurements (neither the measures, neither the estimated means do appear in the update steps), and it ultimately converges to a limit variance matrix. It is also the case for P_k^T . They can thus be pre-computed until the convergence occurs (in less than two hundred iterations). Then only the evolution of a posteriori means needs to be computed by filtering the measures forward and backward, which is much faster.

8.2 Results on Synthetic Data

8.2.1 Cortex Mesh Downsampling

Our simulations are based on a realistic head model based on the segmentation of an anatomical MRI image, which was acquired by Anne-Lise Paradis at the CEA-SHFJ in Orsay. The number of sources on the cortex is currently limited by memory constraints: indeed, the hidden-states dimension in the model (the length of vector \hat{X}_k) is equal to the number of sources n_s times the number of variables for one source ($5n_s$ for the model we used), and the main limitation comes from the square matrices of this size used by the algorithm (P_k^{k-1} , P_k^k , J). Thus we used a model with 1000 sources on the cortex.

Besides, the forward problem matrix G was computed using a realistic head model, with the symmetric boundary elements method implemented in the OpenMEEG software developed by Geoffroy Adde in the Odyssee Team [1], after the different tissues had been segmented on the anatomical MRI image. To preserve the quality of the forward problem, it was necessary to work with a high-resolution mesh (about 45000 vertices) rather than a decimated mesh. Hence the need to design a decimation procedure which allowed us to interpolate between the 45000 vertices of the full mesh and

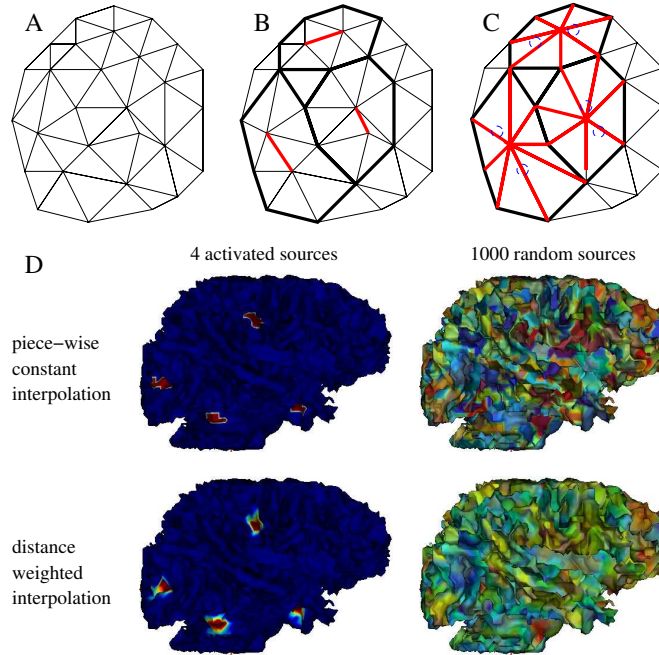


Figure 8.2: *Mesh decimation by contractions of edges*: the decimation iteratively chooses a set of independent edges (B), i.e. edges whose points are not neighbours in the graph; then these edges are contracted into single points, and connections with other neighbours are updated (C); the algorithm also updates an interpolation function between the edges which have been deleted (blue rings) and the other old and new ones. (D) shows the result of interpolation on the high-resolution cortex of activities on the decimated cortex: in the left column, four sources have been set to one and the others to zero; in the right column, the thousand sources have a random activity; the two rows compare two types of interpolation (we used the nearest neighbour interpolation for our synthetic data).

the 1000 vertices of the decimated one. The figure 8.2 shows the decimation algorithm by contractions of edges we used [51, 61], and results of a piece-wise constant compared to a smooth interpolation.

8.2.2 Artificial Data

We simulated cortex activity according to the auto-regressive model (8.1.2):

$$\dot{N}(t) = -\lambda N(t) + \xi_N(t).$$

To this random activity, we added a large sinusoidal impulse, to test the algorithm ability to handle fast and strong variations. Then the EEG and fMRI signals were simulated according to the models summarized in system (8.3), including evolution and measure noise. Figure 8.3 shows the stimulated sources activity, and EEG and fMRI measures.

Then we ran the Kalman filter and smoother. We did not estimate the physiological parameters.

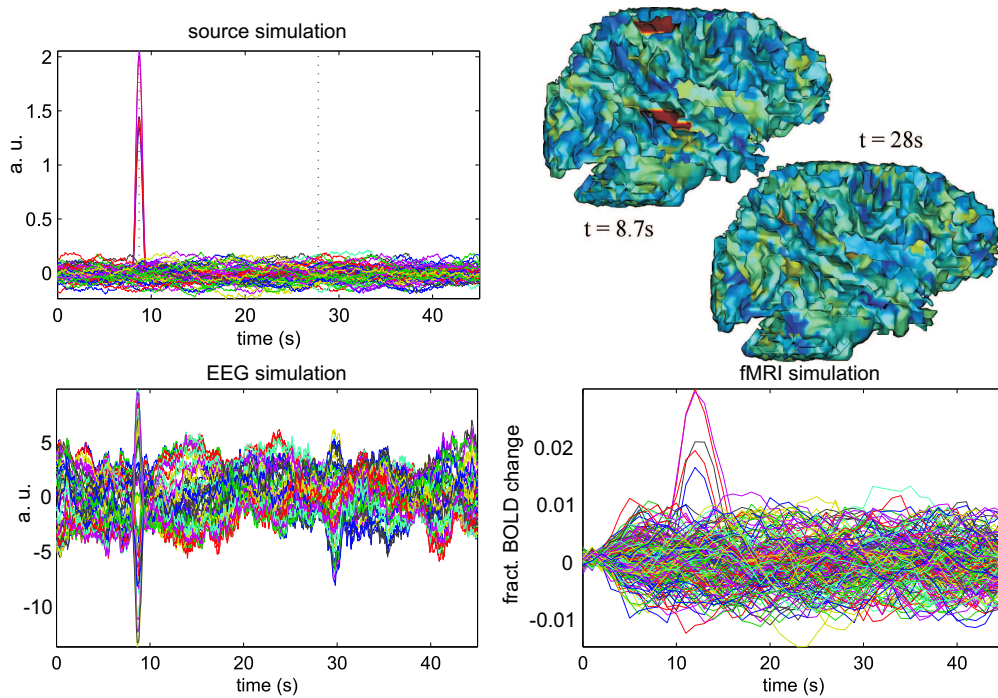


Figure 8.3: *Artificial data*. Top: (left) plot of the time courses of all cortical sources; neural activity was generated according to the evolutive model (8.1.2) and a rapid burst (1s half width) was added to a few sources; (right) activity mapped on the cortex at instants represented by dotted vertical lines. Bottom: EEG and fMRI measures generated by the forward problem (8.1.2) and the Balloon Model (8.1.2).

8.2.3 Results

Our algorithm can work with EEG or fMRI measures alone, or with both of them simultaneously. To assess the improvement brought by the fusion, we ran the algorithm on each modality separately. Figures 8.4 and 8.6, and table 8.4 show the results for EEG, fMRI and EEG+fMRI estimations. Besides, figure 8.5 shows details of hidden-states estimates after the filter and smoother steps.

We first describe the latter figure, which allows a better understanding of the specific roles of the two filters. There are two peaks in the neural activity estimated by the Kalman filter (upper lefthand display). The first one is due to the EEG measure updates and occurs at the correct instant. The second one is due to the fMRI measures updates,

	source estimation accuracy	EEG prediction	fMRI prediction
EEG estimation	25.0 %	100 %	24.9 %
fMRI estimation	88.0 %	86.1 %	97.2 %
EEG+fMRI estimation	88.8 %	100 %	97.1 %

Table 8.1: Correlation coefficients between actual and estimated/predicted time courses. Through the shape of the neural time course obtained with the EEG+fMRI estimation significantly changed from the shape of those obtained with the fMRI-alone estimation, there is not an obvious superiority in terms of estimation accuracy.

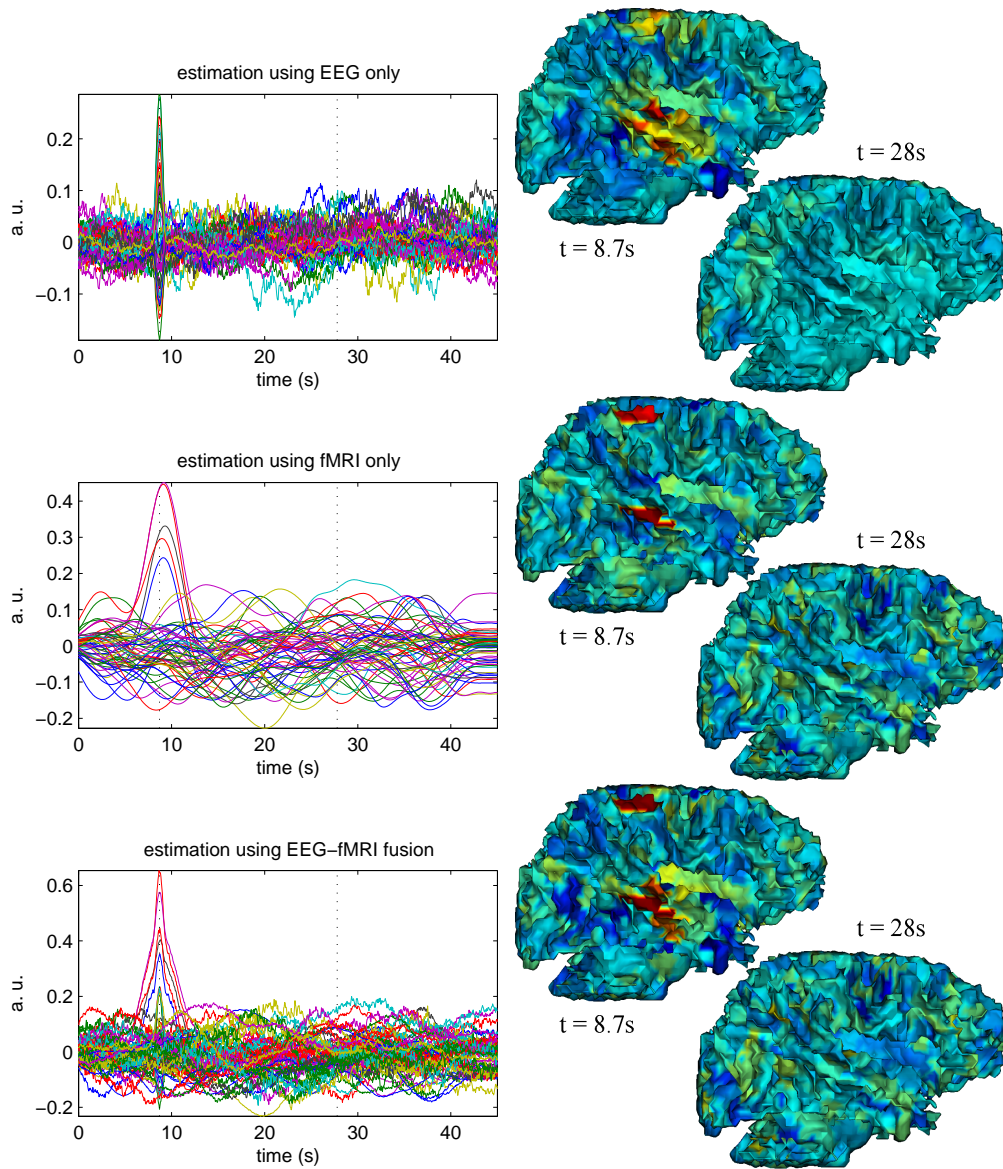


Figure 8.4: *Estimated cortical activity* when using the Kalman filter estimation on EEG and fMRI measures, both separately and together. Left: estimated cortical-source time courses. Right: estimated activities at two specific instants are mapped onto the cortex.

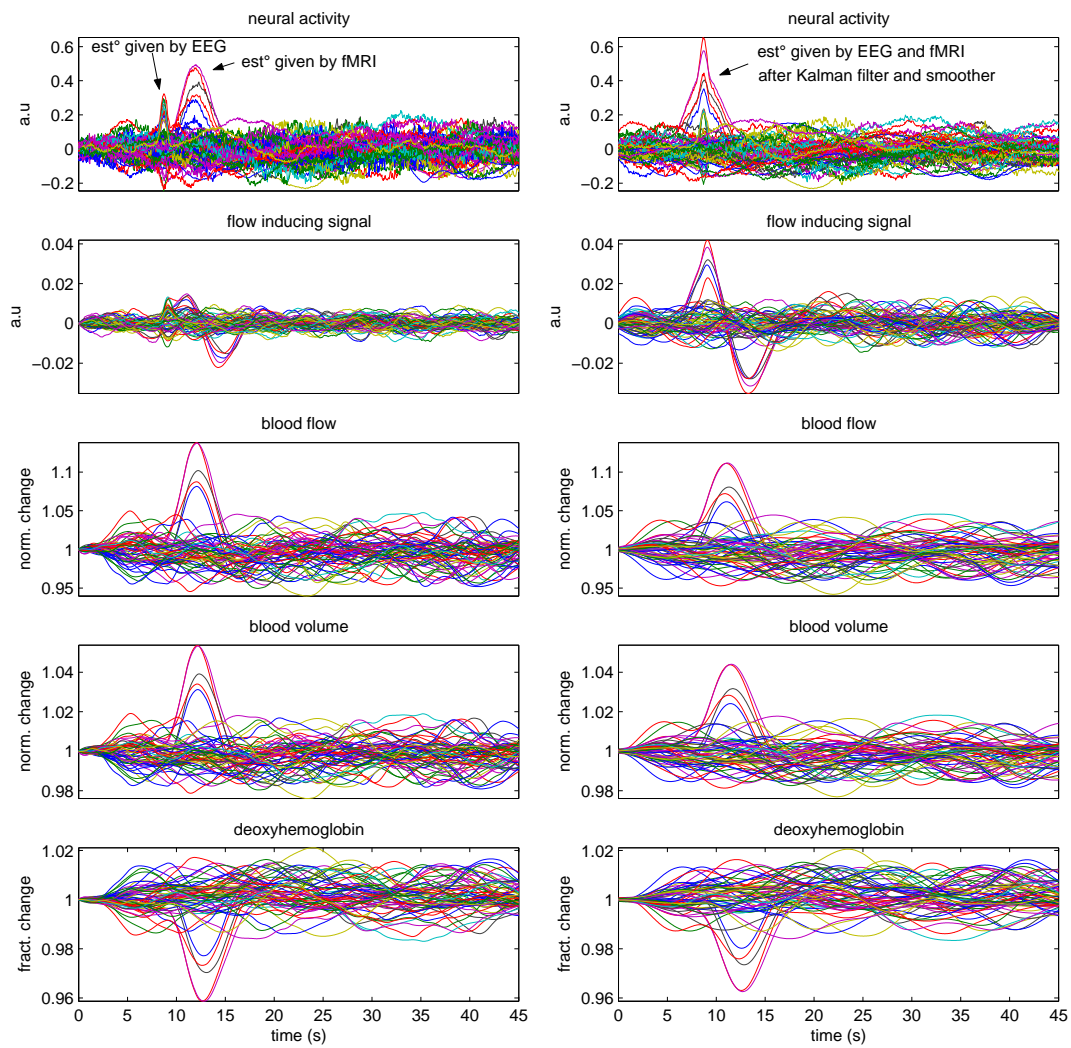


Figure 8.5: *Kalman filter and smoother comparisons*: we show here all the hidden-states estimations after the Kalman filter step and after the subsequent Kalman smoother step. This illustrates why it is necessary to apply the smoother, which includes information coming from the future in the estimation of hidden-states at a given instant (see text).

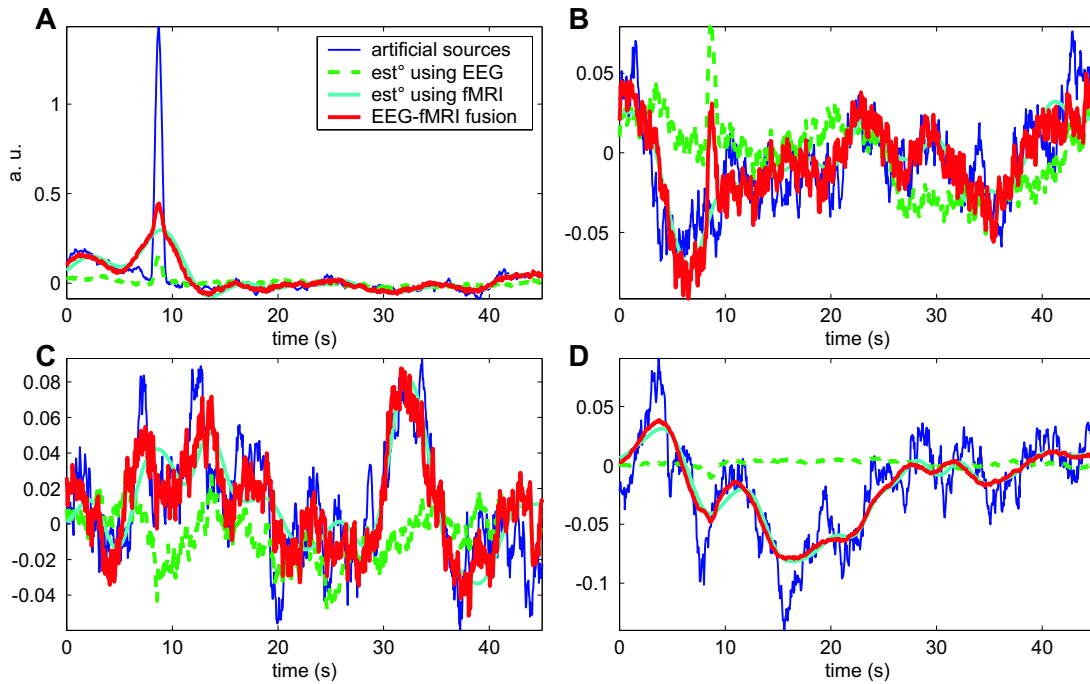


Figure 8.6: *Estimated time courses* of a few cortical sources are shown, and compared to their true values. (A) fMRI alone was able to detect a spread out activity of an actually activated source; the fusion estimation is slightly more focal, but cannot violate more the smoothness constraint. (B) The inverse situation: EEG found activation in a non-activated source, and fMRI turns down this estimation. (C) A typical example of how fMRI alone recovers the low frequency fluctuations, whereas EEG brings a complementary information on fast variations. (D) Here, EEG was unable to give any information on the source activity; fusion estimation is then the same as that of fMRI alone.

and occurs later, because of the delay in the BOLD response. After the application of the Kalman smoother however, since the information coming from the future propagates these estimations back in time, the two peaks are combined into a single one which accounts for the two measurements.

Now, in figure 8.4, the respective qualities of EEG and fMRI in term of temporal and spatial resolutions are clearly illustrated. Indeed, in the EEG sole estimation, the temporal pattern of activation peak around $t=8.5s$ is recovered exactly, but its localisation is quite diffuse (more sources are activated than in the original data, and the activation powers are reduced). On the other hand, in the fMRI sole estimation, it is the estimated time course that is diffuse, whereas the method found the right activation foci. The fusion algorithm then finds a compromise between the two estimations, and the smoothness of neural sources time courses. However, it was not able to find the exact amplitude of the activity peak as in figure 8.3, but this is not surprising since the algorithm used an a priori auto-regressive model for sources activity (equation (8.1.2)), which supposes a temporal smoothness.

Figure 8.6 shows details of the estimated time courses of a few cortical sources with different characteristics. We can observe that, when they are applied to fMRI only, the Kalman filter and smoother perform a deconvolution of the BOLD signal for each source location, that lead to a quite accurate estimation already (center row in figure 8.4 and

cyan time courses in figure 8.6). In comparison, the EEG estimation (top row in 8.4 and green time courses in 8.6) is much worse, due to the ill-posedness of the inverse problem. In the fusion context (bottom row in 8.4 and red time courses in 8.6), EEG brings a complementary information in terms of rapid fluctuations, leading to a better estimation of ongoing neural activity. In particular, it brings more precision in the temporal shape of the activity peak. However, the changes it brings in the estimation of the rest of the ongoing neural activity are not much significant, as shown also in table 8.1, where the increase in the correlation coefficients between estimated and actual time courses appears to be limited.

8.3 Discussion

Our work illustrates how information coming from simultaneous EEG and fMRI measures can be integrated together for a more efficient estimation of the ongoing neural activity. At the same time, we are aware that many hypothesis used in our simulation may not be verified in the reality. This is why we shall summarize first, the ideas brought by our method, and second, the improvements it would probably require.

8.3.1 Combining EEG and fMRI Measures with Filtering Techniques

The major interest of our fusion method and simulation results is probably to show that, in spite of their pronounced differences in terms of temporal and spatial resolutions, EEG and fMRI can be used in the same framework to infer on the underlying cortical activity. We proposed the stochastic differential equations framework, however others have already been studied [77].

In particular, both measurements can bring temporal and spatial information, whereas in numerous existing methods, the two modalities are supposed to share an information on only one of the two aspects (see chapter 3). More precisely, we observed that the effect of adding fMRI information into an estimation achieved with EEG only is to reduce the extent of activation patterns; and conversely, the effect of the EEG measurements when achieving fusion is that the estimated time courses are more focused on the activity periods, compared to those obtained with fMRI alone.

Besides, it appears that the SDE framework is useful for the analysis of each modality alone: applied to the EEG alone, the Kalman filter and smoother technique solves an inverse problem that constraints the cortical sources to be smooth in time. Applied to fMRI alone, it performs a signal deconvolution.

At last, it has the advantage of easily taking into accounts physiological models thanks to its use of dynamic models.

8.3.2 Open Questions

Regarding algorithmic and implementation, the filters described in this chapter allow to greatly simplify a very large inverse problem. However, their execution time increases drastically with the number of cortical sources used, mainly because of the size of variance matrices and other matrices involved in their computation, which is the square of the number of sources time the number of hidden-states per source. In our implementation in C++, the convergence of these variances described in section takes less than one second when using 30 cortical sources, about ten minutes with 300 sources, and several

hours with 1000 sources. These expensive calculation times are just the reason why we could not additionally estimate the noise parameters using a proper EM algorithm [65]. However, there may be possibilities to find some structure in these matrices which would allow a sparse representation and therefore a reduction of memory usage and computation times.

Considering physiological models, we supposed a proportionality relation between electrical and metabolic processes. It is very simplified, and we would rather desire a dependence of the metabolism on some powers of the electrical activity in the frequency domain. But this would introduce strong nonlinearities in the equations, and it is likely that the local linearizations needed by the extended Kalman filter would lead to errors in the estimation. It would be interesting, then, to try other filtering methods that are more robust to nonlinearities (particle filters [165], unscented Kalman filter...).

At last, we would like to discuss how our method could be applied to real data. Let us first note that, in the form as it was presented, it estimates an ongoing activity rather than an event-related activity. As a matter of fact, the model does not include any external input such as stimulation patterns. For that reason, it may be indicated to estimate activity directly from EEG and fMRI signals, rather than on evoked signals. This would have applications on experiment that study a non-triggered activity (for example, the activity of the resting brain), or the activities elicited by single trials, which is the case, for example, in the experiments on epileptic subjects that we will describe in the next section.

However, non-evoked signals are obviously much noisier than evoked signals, since in the latter, the averaging over repetitions of the same stimulation reduces all the signals and noises that are not related to that stimulation. Our method, of course, is still applicable on such evoked signals. It would suppose to first estimate an average response to the stimulation at each electrode for the EEG measurements and in each voxel for the fMRI measurements. Then use these responses as the measures of our model and estimate the neural activity evoked by the stimulation.

8.4 Conclusion

The EEG-fMRI fusion method we have presented in this chapter shows that the hemodynamic models which we have investigated in this thesis can be very useful to build a general framework for the analysis of measurements issued from very different modalities.

The algorithm presented in this chapter is likely to be used on real data since it is able to handle high dimensionalities at usual EEG and fMRI resolutions.

However, its use on real data will probably require new investigations in two directions. First a better modelling of the link between electrical and metabolic activities in the cortex. Actually, this point is a topic on its own, which is currently under investigation by several teams [123, 12, 170]. Secondly, more robust algorithm such as particle filters [165] should be explored in order to handle stronger nonlinearities in the models.

General Conclusion

As a whole, the PhD work addressed two kinds of questions about hemodynamic models:

“Why use biologically plausible models in brain imaging analysis?”

This was the topic of chapters 4, 5 and 8.

Nowadays, most fMRI analysis rely on the linear convolution with an empirical hemodynamic response function (HRF). We proved in chapters 4 and 5, using statistical tests which were designed for that purpose, that linear convolution models cannot predict the BOLD response as well as nonlinear dynamical models do. They obviously cannot predict nonlinear aspects of the BOLD response due to neural habituation or vascular saturation. Also, they are less efficient at handling some variabilities of the response among regions or subjects, such as the post-stimulus undershoot, whereas dynamical models can capture this variability with a reduced number of parameters.

Moreover, we proved in the same chapters that it is possible to use dynamical models for the analysis of fMRI data. The counterpart of linear regression from the general linear model (GLM) framework is the model identification in the dynamical model framework (chapter 4), and the counterpart of statistical tests from the GLM framework is the model selection in the dynamical model framework (chapter 5). Please note that our model selection algorithm was based on Fisher tests applied to linearized models; other algorithm could be proposed however, such as the Akaike Information Criterion (AIC).

However, this does not mean that dynamical models should always be preferred to linear models. On the contrary, linear models, which entail the smaller number of parameters, should be used to address simple cognitive questions, such as localizing activation in a sensory or motor evoked experiment. All the more since calculation time is still heavy with nonlinear methods. On the other hand, precise hemodynamic models become necessary when addressing more complex questions on the temporal aspect of the BOLD responses, such as the presence or absence of neural habituation.

Dynamical models have a second advantage on linear models. Since they are a more natural way to quantify physiological processes, they provide a convenient framework for the combination of different imaging modalities. We proved in chapter 8 that, using biological models, it is possible to consider the EEG and fMRI measures as noisy outputs of a global dynamical system relying on a hidden cortical activity, and to use them together to better estimate this activity, i.e. estimate the time courses of every cortical source.

“Which additional modelling is required?”

This was the topic of chapters 6 and 7.

Experiments described in chapter 6 were motivated by the probably over-simplistic models used for the flow response to neural activity. Indeed, they suppose that the flow responds linearly to the neural activity, and then that the volume response is driven solely by the flow. On the contrary, our data suggests a nonlinear interaction between neural electrical activity and the flow response (e.g. including saturation effects), and a more subtle interaction between the flow and volume responses. A first improvement in such models would be to consider several vascular compartments, at least three with the arteries, capillaries and veins, instead of the venous compartment alone.

These results are confirmed by our new technique, presented in chapter 7, to estimate the blood flow in single blood vessels. Indeed, using this technique, we can observe quite different responses in the different compartments: we observe early responses in the arteries, and we observe later responses in the small vessels, accompanied occasionally with a first early dip, whereas we cannot observe responses in the larger veins.

Please note that the technique itself is a very interesting fruit of the connexions between computer science and neuroscience communities, since it is based on the application of image processing techniques on biological image sequences.

As a result, the future directions that should be investigated in hemodynamic modelling are the mechanical dynamics of the flow (its variations related to those of volume in the different compartments), the processes involved in the flow induction, and to which part of the electrical activity they are related (membrane potential or spikes? how does it depend on amplitudes, frequencies?...).

As a final word, this work shows how biophysiological modelling and analysis techniques in imaging are connected and question each other. New models call for new methods to take them into account in the analysis of the data, and progress in analysis techniques and imaging performances appeal new improvements in the models. It is probable that in the coming years, the analysis of data acquired using fMRI or other imaging techniques dependent on hemodynamic effects, will rely more and more on an acute understanding of the underlying processes and their variability among cortical regions and among subjects.

Publications

Journal articles

Thomas Deneux and Olivier Faugeras **Model selection and activation detection in nonlinear fMRI** *NeuroImage* (in Press), 2006 [chapters 4 and 5]

O. Faugeras, G. Adde, G. Charpiat, C. Chefd'Hotel, M. Clerc, T. Deneux, R. Deriche, G. Hermosillo, R. Keriven, P. Kornprobst, J. Kybic, C. Lenglet, L. Lopez-Perez, T. Papadopoulo, J.-P. Pons, F. Ségonne, B. Thirion, D. Tschumperlé, T. Viéville and N. Wotawa. **Variational, geometric, and statistical methods for modeling brain anatomy and function.** *NeuroImage*, 23S1:S46-S55, 2004. [chapter 4]

Conferences

Thomas Deneux, Ivo Vanzetta, Frédéric Chavanne, Olivier Faugeras and Guillaume Masson **Linearity and non-linearity of sensory-evoked neuronal and hemodynamic responses in awake monkey V1** *Human Brain Mapping, Florence*, 2006. [chapter 6]

Thomas Deneux and Olivier Faugeras **Physiological Model selection in fMRI** *Human Brain Mapping, Florence*, 2006. [chapter 5]

Ivo Vanzetta, Thomas Deneux, Guillaume Masson and Olivier Faugeras **Cerebral blood flow recorded at high sensitivity in two dimensions using high resolution optical imaging** *IEEE Symposium on Biomedical Imaging, Washington*, 1264-1267, 2006. [chapter 7]

Thomas Deneux and Olivier Faugeras **EEG-fMRI Fusion of non-triggered data using Kalman filtering** *IEEE Symposium on Biomedical Imaging, Washington*, 1068-1071, 2006. [chapter 8]

Thomas Deneux, Olivier Faugeras, Daniel J. Jacobsen and Lars Kai Hansen **Neural activity vs. fMRI : Parameters Estimation for the Balloon Model** *Human Brain Mapping, Budapest*, 2004. [chapter 4]

Daniel J. Jacobsen, Lars Kai Hansen, Thomas Deneux and Olivier Faugeras **Non-linear hemodynamics : The spatial distribution of Balloon model parameters** *Human Brain Mapping 2004, Human Brain Mapping, Budapest*, 2004. [chapter 4]

Bibliography

- [1] G. Adde, M. Clerc, O. Faugeras, R. Keriven, J. Kybic, and T. Papadopoulo. *Symmetric BEM formulation for the M/EEG forward problem*. In Proceedings of IPMI, Lecture Notes in Computer Science, pages 524–535. Springer Verlag, 2003.
- [2] Geoffray Adde, Maureen Clerc, and Renaud Keriven. *Imaging methods for MEG/EEG inverse problem*. In Symposium on Noninvasive Functional Source Imaging, 2005.
- [3] R. J. Adler. *The geometry of random fields*. Wiley, New-York, 1981.
- [4] S. P. Ahlfors, G. V. Simpson, A. M. Dale, J. W. Belliveau, A. K. Liu, A. Korvenoja, J. Virtanen, M. Huotilainen, R. B. Tootell, H. J. Aronen, and R. J. Ilmoniemi. *Spatiotemporal activity of a cortical network for processing visual motion revealed by MEG and fMRI*. *J Neurophysiol*, 82(5):2545–2555, 1999.
- [5] Seppo P Ahlfors and Gregory V Simpson. *Geometrical interpretation of fMRI-guided MEG/EEG inverse estimates*. *Neuroimage*, 22(1):323–332, May 2004.
- [6] B M Ances, E Zarahn, J H Greenberg, and J A Detre. *Coupling of neural activation to blood flow in the somatosensory cortex of rats is time-intensity separable, but not linear*. *J Cereb Blood Flow Metab*, 20(6):921–930, Jun 2000.
- [7] A Andrade, F Kherif, J F Mangin, K J Worsley, A L Paradis, O Simon, S Dehaene, D Le Bihan, and J B Poline. *Detection of fMRI activation using cortical surface mapping*. *Hum Brain Mapp*, 12(2):79–93, Feb 2001.
- [8] Amos Arieli, Amiram Grinvald, and Hamutal Slovin. *Dural substitute for long-term imaging of cortical activity in behaving monkeys and its clinical implications*. *Journal of Neuroscience Methods*, 114:119–133, March 2002.
- [9] Agnès Aubert and Robert Costalat. *A model of the coupling between brain electrical activity, metabolism, and hemodynamics : Application to the interpretation of functional neuroimaging*. *NeuroImage*, 17:1162–1181, 2002.
- [10] F Babiloni, F Cincotti, C Babiloni, F Carducci, D Mattia, L Astolfi, A Basilisco, P M Rossini, L Ding, Y Ni, J Cheng, K Christine, J Sweeney, and B He. *Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function*. *Neuroimage*, 24(1):118–131, Jan 2005.
- [11] Fabio Babiloni, Donatella Mattia, Claudio Babiloni, Laura Astolfi, Serenella Salinari, Alessandra Basilisco, Paolo Maria Rossini, Maria Grazia Marciani, and Febo

- Cincotti. *Multimodal integration of EEG, MEG and fMRI data for the solution of the neuroimage puzzle*. *Magn Reson Imaging*, 22(10):1471–1476, Dec 2004.
- [12] A P Bagshaw, C Hawco, C G Bénar, E Kobayashi, Y Aghakhani, F Dubeau, G B Pike, and J Gotman. *Analysis of the eeg-fmri response to prolonged bursts of interictal epileptiform activity*. *Neuroimage*, 24(4):1099–1112, Feb 2005.
- [13] S Baillet and L Garnero. *A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem*. *IEEE Trans Biomed Eng*, 44(5):374–385, May 1997.
- [14] S. Baillet, J. C. Mosher, and M. Leahy. *Electromagnetic brain mapping*. *IEEE Sig. Proc. Mag.*, 18(6):14–30, 2001.
- [15] Yashar Behzadi and Thomas T. Liu. *An arteriolar compliance model of the cerebral blood flow response to neural stimulus*. *NeuroImage*, 25(4):1100–1111, May 2005.
- [16] R Beisteiner, M Erdler, C Teichtmeister, M Diemling, E Moser, V Edward, and L Deecke. *Magnetoencephalography may help to improve functional MRI brain mapping*. *Eur J Neurosci*, 9(5):1072–1077, May 1997.
- [17] G. S. Berns, A. W. Song, and H. Mao. *Continuous functional magnetic resonance imaging reveals dynamic nonlinearities of "dose-response" curves for finger opposition*. 19(14):RC17, Jul 15 1999.
- [18] Peter J. Bickel and Kjell A. Doksum. *Mathematical statistics basic ideas and selected topics 1*. New York London Sydney Prentice-Hall, 2001.
- [19] Joseph Bigun, Gösta H. Granlund, and Johan Wiklund. *Multidimensional orientation estimation with applications to texture analysis and optical flow*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790, August 1991.
- [20] Rasmus M. Birn, Ziad S. Saad, and Peter A. Bandettini. *Spatial heterogeneity of the nonlinear dynamics in the fMRI BOLD response*. *NeuroImage*, 14:817–826, 2001.
- [21] Christian Bénar, Yahya Aghakhani, Yunhua Wang, Aaron Izenberg, Abdullah Al-Asmi, Francois Dubeau, and Jean Gotman. *Quality of EEG in simultaneous EEG-fMRI for epilepsy*. *Clin Neurophysiol*, 114(3):569–580, Mar 2003.
- [22] G. E. P. Box and G. C. Tiao. *Bayesian Inference in Statistical Analysis*. New York: Wiley, 1992.
- [23] J L Boxerman, P A Bandettini, K K Kwong, J R Baker, T L Davis, B R Rosen, and R M Weisskoff. *The intravascular contribution to fMRI signal change: Monte Carlo modeling and diffusion-weighted studies in vivo*. *Magn Reson Med*, 34(1):4–10, Jul 1995.
- [24] Geoffrey M. Boynton, Stephen A. Engel, Gary H. Glover, and David J. Heeger. *Linear systems analysis of functional magnetic resonance imaging in human v1*. *The Journal of Neuroscience*, 16(13):4207–4221, 1996.

- [25] C. A. Brebbia. *The boundary element method for engineers*. Pentech Press, London, 1984.
- [26] J D Briers. *Laser Doppler, speckle and related techniques for blood perfusion mapping and imaging*. *Physiol Meas*, 22(4):35–66, Nov 2001.
- [27] G Brinker, C Bock, E Busch, H Krep, K A Hossmann, and M Hoehn-Berlage. *Simultaneous recording of evoked potentials and T2*-weighted MR images during somatosensory stimulation of rat*. *Magn Reson Med*, 41(3):469–473, Mar 1999.
- [28] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. *High accuracy optical flow estimation based on a theory for warping*. In Springer-Verlag., editor, *Proceedings of the 8th European Conference on Computer Vision*, Prague, Czech Republic, pages 25–36, May 2004.
- [29] R B Buxton. *The elusive initial dip*. *Neuroimage*, 13(6 Pt 1):953–958, Jun 2001. Comment.
- [30] R. B. Buxton, K. Miller, L. R. Frank, and E. C. Wong. *BOLD signal dynamics: the balloon model with viscoelastic effects*. In *Sixth Meeting, International Society for Magnetic Resonance in Medicine*, page 1401, Sydney, Australia, 1998.
- [31] Richard B. Buxton and L. R. Frank. *A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation*. *J. of Cerebral Blood Flow and Metabolism*, 17:64–72, 1997.
- [32] Richard B. Buxton, K. UludaNg, D. J. Dubowitz, and T. T. Liu. *Modelling the hemodynamic response to brain activation*. *NeuroImage*, 23:220–233, 2004.
- [33] Richard B. Buxton, E. C. Wong, and L. R. Frank. *Dynamics of blood flow and oxygenation changes during brain activation: the balloon model*. *Magn. Reson. Med.*, 39:855–864, 1998.
- [34] Emmanuelle Chaigneau, Martin Oheim, Etienne Audinat, and Serge Charpak. *Two-photon imaging of capillary blood flow in olfactory bulb glomeruli*. *Proc Natl Acad Sci U S A*, 100(22):13081–13086, Oct 2003.
- [35] Haiying Cheng, Qingming Luo, Shaoqun Zeng, Shangbin Chen, Jian Cen, and Hui Gong. *Modified laser speckle imaging method with improved spatial resolution*. *J Biomed Opt*, 8(3):559–564, Jul 2003. *Evaluation Studies*.
- [36] C.K. Chui and G. Chen. *Kalman Filtering with Real-Time Applications*. Springer Ser. Info. Sci., Vol. 17. Springer, Berlin, Heidelberg, 1987.
- [37] Eric R Cohen, Kamil Ugurbil, and Seong-Gi Kim. *Effect of basal conditions on the magnitude and dynamics of the blood oxygenation level-dependent fMRI response*. *Journal of Cerebral Blood Flow and Metabolism*, 22(9):1042–1053, 2002.
- [38] L B Cohen. *Changes in neuron structure during action potential propagation and synaptic transmission*. *Physiol Rev*, 53(2):373–418, Apr 1973.
- [39] L B Cohen, R D Keynes, and B Hille. *Light scattering and birefringence changes during nerve activity*. *Nature*, 218(140):438–441, May 1968.

- [40] O. Creutzfeld. *Neurophysiological correlates of different functional states of the brain*. In *Brain work. The coupling of function, metabolism and blood flow in the brain*, pages 21–47, Copenhagen: Munksgaard, 1975. Ingvar DH, Lassen NA, eds.
- [41] Dale and Anders M. *Selective averaging of rapidly presented individual trials using fMRI*. *Human Brain Mapping*, 5:329–390, 1997.
- [42] A.M. Dale and M.I. Sereno. *Improved localization of cortical activity by combining EEG and MEG with mri cortical surface reconstruction: A linear approach*. *Journal of Cognitive Neuroscience*, 5:162–176, 1993.
- [43] Jean Daunizeau. *Localisation et dynamique des sources d’activité cérébrale par fusion d’informations multimodales EEG/IRMf*. *PhD thesis, Université Paris XI, France, October 2006*.
- [44] T. L. Davis, K. K. Kwong, R. M. Weisskoff, and B. R. Rosen. *Calibrated functional mri: mapping the dynamics of oxydative metabolism*. *Proc. Natl. Acad. Sci. U.S.A.*, 95:1834–1839, 1998.
- [45] T. L. Davis, R. M. Weisskoff, K. K. Kwong, J. L. Boxerman, , and B. R. Rosen. *Temporal aspects of fMRI task activation: Dynamic modeling of oxygen delivery*. *Proc. Int. Magnetic Resonance in Medicine*, 2:69, 1994.
- [46] Valeria Della-Maggiore, Wilkin Chau, Pedro R Peres-Neto, and Anthony R McIntosh. *An empirical comparison of SPM preprocessing parameters to the analysis of fMRI data*. *NeuroImage*, 17(1):19–28, 2002.
- [47] Thomas Deneux and Olivier Faugeras. *EEG-fMRI fusion of non-triggered data using kalman filtering*. In *International Symposium on Biomedical Imaging*, pages 1068–1071. *IEEE, NIH*, 2006.
- [48] Anna Devor, Andrew K Dunn, Mark L Andermann, Istvan Ulbert, David A Boas, and Anders M Dale. *Coupling of total hemoglobin concentration, oxygenation, and neural activity in rat somatosensory cortex*. *Neuron*, 39(2):353–359, Jul 2003.
- [49] U Dirnagl, B Kaplan, M Jacewicz, and W Pulsinelli. *Continuous measurement of cerebral cortical blood flow by laser-Doppler flowmetry in a rat stroke model*. *J Cereb Blood Flow Metab*, 9(5):589–596, Oct 1989.
- [50] Andrew K Dunn, Anna Devor, Anders M Dale, and David A Boas. *Spatial extent of oxygen metabolism and hemodynamic changes during functional activation of the rat somatosensory cortex*. *Neuroimage*, 27(2):279–290, Aug 2005.
- [51] Matthias Eck, Tony DeRose, Tom Duchamp, Hugues Hoppe, Michael Lounsbery, and Werner Stuetzle. *Multiresolution analysis of arbitrary meshes*. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 173–182. *ACM Press New York, NY, USA*, 1995.
- [52] A S Ferguson, X Zhang, and G Stroink. *A complete linear discretization for calculating the magnetic field using the boundary element method*. *IEEE Trans Biomed Eng*, 41(5):455–460, May 1994.

- [53] Guillaume Flandin, Ferath Kherif, Xavier Pennec, Denis Riviere and Nicholas Ayache, and Jean-Baptiste Poline. *Parcellation of brain images with anatomical and functional constraints for fMRI data analysis*. In Proceedings of the IEEE Symposium on Biomedical Imaging, pages 907–910, 2002.
- [54] P. T. Fox, M. E. Raichle, M. A. Mintun, and C. Dence. *Nonoxidative glucose consumption during focal physiologic neural*. Science, 241:462–464, 1988.
- [55] Luis Freire, Alexis Roche, and Jean-Francois Mangin. *Motion correction algorithms may create spurious brain activations in the absence of subject motion*. NeuroImage, 14(3):709–722, September 2001.
- [56] Luis Freire, Alexis Roche, and Jean-Francois Mangin. *What is the best similarity measure for motion correction in fMRI time series?* IEEE Transactions on Medical Imaging, 21(5):470–484, 2002.
- [57] P Fries, J H Reynolds, A E Rorie, and R Desimone. *Modulation of oscillatory neuronal synchronization by selective visual attention*. Science, 291(5508):1560–1563, Feb 2001.
- [58] Karl J. Friston. *Bayesian estimation of dynamical systems: an application to fMRI*. NeuroImage, 16:513–530, 2002.
- [59] Karl J. Friston, A. Mechelli, R. Turner, and C. J. Price. *Nonlinear responses in fMRI : the balloon model, volterra kernels, and other hemodynamics*. NeuroImage, 12:466–477, 2000.
- [60] K.J. Friston, O. Josephs, G. Rees, and R. Turner. *Non-linear event-related responses in fMRI*. Magnetic Resonance in Medicine, 39:41–52, 1998.
- [61] Michael Garland, Andrew Willmott, and Paul S. Heckbert. *Hierarchical face clustering on polygonal surfaces*. In Proceedings of the 2001 symposium on Interactive 3D graphics, pages 49–58. ACM Press New York, NY, USA, 2001.
- [62] J S George, C J Aine, J C Mosher, D M Schmidt, D M Ranken, H A Schlitt, C C Wood, J D Lewine, J A Sanders, and J W Belliveau. *Mapping function in the human brain with magnetoencephalography, anatomical magnetic resonance imaging, and functional magnetic resonance imaging*. J Clin Neurophysiol, 12(5):406–431, Sep 1995.
- [63] J. S. George, D. M. Schmidt, D. M. Rector, and C. C. Wood. *Dynamic functional neuroimaging integrating multiple modalities*. In SM Smith P Jezzard, PM Matthews, editor, Functional MRI: An introduction to Methods, pages 353–382. Oxford Univ. Press, York, 2001.
- [64] D. B. Geselowitz. *On the magnetic field generated outside an inhomogeneous volume conductor by internal current sources*. IEEE Transactions on Magnetics, 6:346–347, 1970.
- [65] Zoubin Ghahramani and Geoffroy Hinton. *Parameter estimation for linear dynamical systems*. Technical Report CRG-TR-96-2, University of Toronto, 1996.

- [66] A. Gjedde. *The relation between brain function and cerebral*. In *Cerebrovascular Disease*, pages 23–40. H. H. Batjer, Ed., Lippincott-Raven, Philadelphia, 1997.
- [67] G. H. Glover. *Deconvolution of impulse response in event-related BOLD fMRI*. *NeuroImage*, 9(4):416–29, Apr 1999.
- [68] Robin I Goldman, John M Stern, Jerome Jr Engel, and Mark S Cohen. *Simultaneous EEG and fMRI of the alpha rhythm*. *Neuroreport*, 13(18):2487–2492, Dec 2002.
- [69] Jean Gotman, Christian-G Bénar, and François Dubeau. *Combining EEG and fMRI in epilepsy: Methodological challenges and clinical results*. *Journal of Clinical Neurophysiology*, 21(4), August 2004.
- [70] A Grinvald, R D Frostig, E Lieke, and R Hildesheim. *Optical imaging of neuronal activity*. *Physiol Rev*, 68(4):1285–1366, Oct 1988.
- [71] A Grinvald, E Lieke, R D Frostig, C D Gilbert, and T N Wiesel. *Functional architecture of cortex revealed by optical imaging of intrinsic signals*. *Nature*, 324(6095):361–364, Nov 1986.
- [72] A. Grinvald, D. Shoham, A. Shmuel, A. Arieli, and et al. *In-vivo optical imaging of cortical architecture and dynamics*. In U. Windhorst and H. Johansson, editors, *Modern techniques in neuroscience research*. Springer, Umea, Sweden, 1999.
- [73] Amiram Grinvald, Tobias Bonhoeffer, Ivo Vanzetta, Ayala Pollackand Eyal Aloni, Ron Ofri, and Darin Nelson. *High-resolution functional optical imaging: from the neocortex to the eye*. *Ophthalmol Clin N Am*, 17:53–67, 2004.
- [74] C Grova, J Daunizeau, J-M Lina, C G Benar, H Benali, and J Gotman. *Evaluation of EEG localization methods using realistic simulations of interictal spikes*. *Neuroimage*, 29(3):734–753, Feb 2006.
- [75] R L Jr Grubb, M E Raichle, J O Eichling, and M M Ter-Pogossian. *The effects of changes in PaCO₂ on cerebral blood volume, blood flow, and vascular mean transit time*. *Stroke*, 5(5):630–639, Sep 1974.
- [76] Yaroslav O. Halchenko, Stephen José Hanson, and Barak A. Pearlmutter. *Multi-modal integration: fMRI, mri, EEG, MEG*. In Luigi Landini, Vincenzo Positano, and Maria Filomena Santarelli, editors, *Advanced Image Processing in Magnetic Resonance Imaging*. Dekker, book series on *Signal Processing and Communications*, 2005.
- [77] Y.O. Halchenko, B.A. Pearlmutter, S.J. Hanson, and A. Zaimi. *Fusion of functional brain imaging modalities via linear programming*. *Biomedizinische Technik (Biomedical Engineering)*, 48(2):102–104, 2004.
- [78] M. Hamalainen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa. *Magnetoencephalography: Theory, instrumentation, and applications to noninvasive studies of the working human brain*. *Rev. Modern Phys.*, 65:413–497, 1993.
- [79] D. J. Heeger, G. M. Boynton, J. B Demb, E. Seidemann, and W. T. Newsome. *Motion opponency in visual cortex*. *Journal of Neuroscience*, 19:7162–7174, 1999.

- [80] D. J. Heeger, A. C. Huk, W. S. Geisler, and D. G. Albrecht. *Spikes versus BOLD, what does neuroimaging tell us about neuronal activity?* *Nature Neuroscience*, 3(2):631–633, 2000.
- [81] D. J. Heeger and D. Ress. *What does fMRI tell us about neuronal activity?* *Nature Reviews Neuroscience*, 3:142–151, Feb 2002.
- [82] Bertil Hille. *Ion Channels of Excitable Membranes*. Sinauer Associates, Inc., Sunderland, Massachusetts, 1940.
- [83] A Hirota, K Sato, Y Momose-Sato, T Sakai, and K Kamino. *A new simultaneous 1020-site optical recording system for monitoring neural activity using voltage-sensitive dyes*. *J Neurosci Methods*, 56(2):187–194, Feb 1995.
- [84] Silvina G Horovitz, Bruno Rossion, Pawel Skudlarski, and John C Gore. *Parametric design and correlational analyses help integrating fMRI and electrophysiological data during face processing*. *Neuroimage*, 22(4):1587–1595, Aug 2004.
- [85] Silvina G Horovitz, Pawel Skudlarski, and John C Gore. *Correlations and dissociations between BOLD signal and P300 amplitude in an auditory oddball task: a parametric approach to combining fMRI and ERP*. *Magn Reson Imaging*, 20(4):319–325, May 2002.
- [86] Barry Horwitz and David Poeppel. *How can EEG/MEG and fMRI/PET data be combined?* *Hum Brain Mapp*, 17(1):1–3, Sep 2002.
- [87] F.R. Huang-Hellinger, H.C. Breiter, G. McCormack, M.S. Cohen, K.K. Kwong, J.P. Sutton, T.L. Davis, R.L. Savoy, R.M. Weisskoff, J.W. Belliveau, and B.R. Rosen. *Simultaneous functional magnetic resonance imaging and electrophysiological recording*. In *Human Brain Mapping*, 1996.
- [88] Costantino Iadecola. *Neurovascular regulation in the normal brain and in Alzheimer’s disease*. *Nat Rev Neurosci*, 5(5):347–360, May 2004.
- [89] T Iijima, G Matsumoto, and Y Kidokoro. *Synaptic activation of rat adrenal medulla examined with a large photodiode array in combination with a voltage-sensitive dye*. *Neuroscience*, 51(1):211–219, Nov 1992.
- [90] C. Janz, S. P. Heinrich, J. Kornmayer, M. Bach, and J. Hennig. *Coupling of neural activity and BOLD fMRI response: new insights by combination of fMRI and vep experiments in transition from single events to continuous stimulation*. *Magnetic Resonance in Medicine*, 46(3):482–6, Sep 2001.
- [91] B Jeffs, R Leahy, and M Singh. *An evaluation of methods for neuromagnetic image reconstruction*. *IEEE Trans Biomed Eng*, 34(9):713–723, Sep 1987.
- [92] P. Jezzard. *Physiological noise: strategies for correction*. Springer-Verlag, Heidelberg, 1999.
- [93] Peter Jezzard and Stuart Clare. *Principles of nuclear magnetic resonance and mri*. In *SM Smith P Jezzard, PM Matthews, editor, Functional Magnetic Resonance Imaging: An Introduction to Methods*, pages 67–92. Oxford Univ. Press, York, 2001.

- [94] Wenjun Jin, Ren-Ji Zhang, and Jian-young Wu. Voltage-sensitive dye imaging of population neuronal activity in cortical tissue. *J Neurosci Methods*, 115(1):13–27, Mar 2002.
- [95] M Jones, J Berwick, D Johnston, and J Mayhew. Concurrent optical imaging spectroscopy and laser-Doppler flowmetry: the relationship between blood flow, oxygenation, and volume in rodent barrel cortex. *Neuroimage*, 13(6 Pt 1):1002–1015, Jun 2001.
- [96] D. Kelen. On cytochrome, a respiratory pigment, common to animals, yeast, and higher plants. *Proc. R. Soc. B*, 98:2602–2612, 1925.
- [97] J Kershaw, B A Ardekani, and I Kanno. Application of Bayesian inference to fMRI data analysis. *IEEE Trans Med Imaging*, 18(12):1138–1153, Dec 1999.
- [98] Stefan J Kiebel and Karl J Friston. Statistical parametric mapping for event-related potentials: I. Generic considerations. *Neuroimage*, 22(2):492–502, Jun 2004.
- [99] Stefan J Kiebel and Karl J Friston. Statistical parametric mapping for event-related potentials (II): a hierarchical temporal model. *Neuroimage*, 22(2):503–520, Jun 2004.
- [100] Stefan J Kiebel, Jean-Baptiste Poline, Karl J Friston and Andrew P Holmes, and Keith J Worsley. Anatomically informed basis functions. *NeuroImage*, 11(6):656–667, Jun 2000.
- [101] J M Kilner, J Mattout, R Henson, and K J Friston. Hemodynamic correlates of EEG: a heuristic. *Neuroimage*, 28(1):280–286, Oct 2005.
- [102] Dae-Shik Kim, Itamar Ronen, Cheryl Olman, Seong-Gi Kim, Kamil Ugurbil, and Louis J Toth. Spatial relationship between neuronal activity and BOLD functional MRI. *Neuroimage*, 21(3):876–885, Mar 2004.
- [103] David Kleinfeld, Partha P. Mitra, Fritjof Helmchen, and Winfried Denk. Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex. *PNAS USA*, 95:15741–15746, December 1998.
- [104] H Kober, C Nimsky, M Moller, P Hastreiter, R Fahlbusch, and O Ganslandt. Correlation of sensorimotor activation with functional magnetic resonance imaging and magnetoencephalography in presurgical functional imaging: a spatial analysis. *Neuroimage*, 14(5):1214–1228, Nov 2001.
- [105] Christof Koch and Idan Segev. *Methods in Neuronal Modeling: from Ions to Networks*. MIT Press, 1989.
- [106] Matthias Kohl, Ute Lindauer, Georg Royl, Marc Kühl, Lorenz Gold, Arno Villringer, and Ulrich Dirnagl. Physical model for the spectroscopic analysis of cortical intrinsic optical signals. *Phys. Med. Biol.*, 45:3749–3764, 2000.

- [107] Y. Kong, Y. Zheng, D. Johnston, J. Martindale, M. Jones, S. Billings, and J. Mayhew. A model of the dynamic relationship between blood flow and volume changes during brain activation. *Journal of Cerebral Blood Flow and Metabolism*, 24:1382–1392, 2004.
- [108] A Korvenoja, J Huttunen, E Salli, H Pohjonen, S Martinkauppi, J M Palva, L Lauronen, J Virtanen, R J Ilmoniemi, and H J Aronen. Activation of multiple cortical areas in response to somatosensory stimulation: combined magnetoencephalographic and functional magnetic resonance imaging. *Hum Brain Mapp*, 8(1):13–27, 1999.
- [109] F Kruggel, C S Herrmann, C J Wiggins, and D Y von Cramon. Hemodynamic and electroencephalographic responses to illusory figures: recording of the evoked potentials during functional MRI. *Neuroimage*, 14(6):1327–1336, Dec 2001.
- [110] J.C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9(1):112–147, 1998.
- [111] P.-J. Lahaye, S. Baillet, J.-B. Poline, and L. Garnero. Fusion of simultaneous fMRI/EEG data based on the electro-metabolic coupling. In *Proc. 2th Proc. IEEE ISBI*, pages 864–867, Arlington, VA, Apr. 2004.
- [112] N. Lange and S. L. Zeger. Non-linear fourier time series analysis for human brain mapping by functional magnetic resonance imaging. *Appl. Stat.*, 46(1):1–29, 1997.
- [113] H Laufs, A Kleinschmidt, A Beyerle, E Eger, A Salek-Haddadi, C Preibisch, and K Krakow. EEG-correlated fMRI of human alpha activity. *Neuroimage*, 19(4):1463–1476, Aug 2003.
- [114] H Laufs, K Krakow, P Sterzer, E Eger, A Beyerle, A Salek-Haddadi, and A Kleinschmidt. Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proc Natl Acad Sci U S A*, 100(19):11053–11058, Sep 2003.
- [115] Martin Lauritzen and Lorenz Gold. Brain function and neurophysiological correlates of signals used in functional neuroimaging. *J Neurosci*, 23(10):3972–3980, May 2003.
- [116] F Lazeyras, I Zimine, O Blanke, S H Perrig, and M Seeck. Functional MRI with simultaneous EEG recording: feasibility and application to motor and visual activation. *J Magn Reson Imaging*, 13(6):943–948, Jun 2001.
- [117] L Lemieux, K Krakow, and D R Fish. Comparison of spike-triggered functional MRI BOLD activation and EEG dipole model localization. *Neuroimage*, 14(5):1097–1104, Nov 2001.
- [118] A. D. Letgatt, J. Arezzo, and H. G. Jr. Vaughan. Averaged multiple unit activity as an estimate of phasic changes in local neuronal activity: effects of volume-conducted potentials. *Journal of Neuroscience Methods*, 2:203–217, 1980.

- [119] A K Liu, J W Belliveau, and A M Dale. *Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations*. Proc Natl Acad Sci U S A, 95(15):8945–8950, Jul 1998.
- [120] A K Liu, J W Belliveau, and A M Dale. *Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations*. Proc Natl Acad Sci U S A, 95(15):8945–8950, Jul 1998.
- [121] N K Logothetis, H Guggenberger, S Peled, and J Pauls. *Functional imaging of the monkey brain*. Nat Neurosci, 2(6):555–562, Jun 1999.
- [122] Nikos K Logothetis. *The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal*. Philos Trans R Soc Lond B Biol Sci, 357(1424):1003–1037, Aug 2002.
- [123] Nikos K Logothetis. *The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal*. Philos Trans R Soc Lond B Biol Sci, 357(1424):1003–1037, Aug 2002.
- [124] Nikos K Logothetis and Brian A Wandell. *Interpreting the BOLD signal*. Annu Rev Physiol, 66:735–769, 2004.
- [125] Hanzhang Lu, Xavier Golay, James J. Pekar, and Peter C.M. van Zijl. *Functional magnetic resonance imaging based on changes in vascular space occupancy*. Magnetic Resonance in Medicine, 50(2):263–274, Aug 2003.
- [126] Hanzhang Lu, Meng Law, Glyn Johnson, Yulin Ge, Peter C. M. van Zijl, and Joseph A. Helpert. *Novel approach to the measurement of absolute cerebral blood volume using vascular-space-occupancy magnetic resonance imaging*. Magnetic Resonance in Medicine, 54(6):1403–1411, Oct 2005.
- [127] David J.C. MacKay. *Information theory, inference, and learning algorithms*. Cambridge University Press, 2003.
- [128] P J Magistretti and L Pellerin. *Cellular mechanisms of brain energy metabolism and their relevance to functional brain imaging*. Philos Trans R Soc Lond B Biol Sci, 354(1387):1155–1163, Jul 1999.
- [129] D Malonek, U Dirnagl, U Lindauer, K Yamada, I Kanno, and A Grinvald. *Vascular imprints of neuronal activity: relationships between the dynamics of cortical blood flow, oxygenation, and volume changes following sensory stimulation*. Proc Natl Acad Sci U S A, 94(26):14826–14831, Dec 1997.
- [130] D Malonek and A Grinvald. *Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping*. Science, 272(5261):551–554, Apr 1996.
- [131] J B Mandeville, J J Marota, C Ayata, G Zaharchuk, M A Moskowitz, B R Rosen, and R M Weisskoff. *Evidence of a cerebrovascular postarteriole windkessel with delayed compliance*. J Cereb Blood Flow Metab, 19(6):679–689, Jun 1999.

- [132] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11:431–441, (see <http://mathworld.wolfram.com/Levenberg-MarquardtMethod.html>) 1963.
- [133] Eduardo Martinez-Montes, Pedro A Valdes-Sosa, Fumikazu Miwakeichi, Robin I Goldman, and Mark S Cohen. Concurrent EEG/fMRI analysis by multiway Partial Least Squares. *Neuroimage*, 22(3):1023–1034, Jul 2004.
- [134] S J Matcher, C E Elwell, C E Cooper, M Cope, and D T Delpy. Performance comparison of several published tissue near-infrared spectroscopy algorithms. *Anal Biochem*, 227(1):54–68, May 1995.
- [135] C Mathiesen, K Caesar, N Akgoren, and M Lauritzen. Modification of activity-dependent increases of cerebral blood flow by excitatory synaptic activity and spikes in rat cerebellar cortex. *J Physiol*, 512 (Pt 2):555–566, Oct 1998.
- [136] J Mayhew, D Johnston, J Martindale, M Jones, J Berwick, and Y Zheng. Increased oxygen consumption following activation of brain: theoretical footnotes using spectroscopic data from barrel cortex. *Neuroimage*, 13(6 Pt 1):975–987, Jun 2001.
- [137] J Mayhew, Y Zheng, Y Hou, B Vuksanovic, J Berwick, S Askew, and P Coffey. Spectroscopic analysis of changes in remitted illumination: the response to increased neural activity in brain. *Neuroimage*, 10(3 Pt 1):304–326, Sep 1999.
- [138] B W Mel. Synaptic integration in an excitable dendritic tree. *J Neurophysiol*, 70(3):1086–1101, Sep 1993.
- [139] Karla L. Miller, Wen-Ming Luh, Thomas T. Liu, Antigona Martinez, Takayuki Obata, Eric C. Wong, Lawrence R. Frank, and Richard B. Buxton. Nonlinear temporal dynamics of the cerebral blood flow response. *Human Brain Mapping*, 13:1–12, 2001.
- [140] G. A. Millikan. Experiments on muscle hemoglobin in vivo; the instantaneous measurement of muscle metabolism. *Proc. Roy. Soc. B*, 123:218–241, 1937.
- [141] U. Mitzdorf. Properties of the evoked potential generators: current source-density analysis of visually evoked potentials in the cat cortex. *International Journal of Neuroscience*, 33:33–59, 1987.
- [142] J C Mosher, R M Leahy, and P S Lewis. EEG and MEG: forward solutions for inverse methods. *IEEE Trans Biomed Eng*, 46(3):245–259, Mar 1999.
- [143] Christoph Mulert, Lorenz Jager, Robert Schmitt, Patrick Bussfeld, Oliver Pogarell, Hans-Jurgen Moller, Georg Juckel, and Ulrich Hegerl. Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *Neuroimage*, 22(1):83–94, May 2004.
- [144] Dinesh G Nair. About being BOLD. *Brain Res Brain Res Rev*, 50(2):229–243, Dec 2005.

- [145] P L Nunez and R B Silberstein. *On the relationship of synaptic activity to macroscopic measurements: does co-registration of EEG with fMRI make sense?* Brain Topogr, 13(2):79–96, Winter 2000.
- [146] T. Obata, T. T. Liu, K. L. Miller, W. M. Luh, E. C. Wong, and R. B. Buxton. *Discrepancies between BOLD and flow dynamics in primary and supplementary motor areas : application of the balloon model to the interpretation of BOLD transients.* NeuroImage, 21:144–153, 2004.
- [147] S Ogawa, T M Lee, R Stepnoski, W Chen, X H Zhu, and K Ugurbil. *An approach to probe some neural systems interaction by functional MRI at neural time scale down to milliseconds.* Proc Natl Acad Sci U S A, 97(20):11026–11031, Sep 2000.
- [148] S. Ogawa, R. S. Menon, D. W. Tank, S.-G. Kim, H. Merkle, J. M. Ellerman, and K. Ugurbil. *Function brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging: a comparison of signal characteristics with a biophysical model.* Biophys. J., 64:803–812, 1993.
- [149] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Nonlinear models.* Numerical Recipes in C, pages 681–688, 1992.
- [150] E H Ratzlaff and A Grinvald. *A tandem-lens epifluorescence microscope: hundred-fold brightness advantage for wide-field imaging.* J Neurosci Methods, 36(2-3):127–137, Feb 1991.
- [151] G. Rees, K. Friston, and C. Koch. *A direct quantitative relationship between the functional properties of human and macaque v5.* National Neuroscience, 3:716–723, 2000.
- [152] Jorge Riera, E. Aubert, K. Iwata, R. Kawashima, X. Wan, and T. Ozaki. *Fusing EEG and fMRI based on a bottom-up model: inferring activation and effective connectivity in neural masses.* Phil. Trans. R. Soc., B 360:1025–1041, 2005.
- [153] Jorge Riera, J. Watanabe, I. Kazuki, M. Naoki, E. Aubert, T. Ozaki, and R. Kawashima. *A state-space model of the hemodynamic approach: nonlinear filtering of BOLD signals.* NeuroImage, 21:547–567, March 2004.
- [154] C. S. Roy and C. S. Sherrington. *On the regulation of the blood supply of the brain.* Journal of Phisiology, 11:85–108, 1890.
- [155] A Salek-Haddadi, Friston K J, Lemieux L, and Fish D R. *Studying spontaneous EEG activity with fMRI.* Brain Res Brain Res Rev, 43(1):110–133, 2003.
- [156] Matthias Schulz, Wilkin Chau, Simon J Graham, Anthony R McIntosh, Bernhard Ross, Ryouhei Ishii, and Christo Pantev. *An integrative MEG-fMRI study of the primary somatosensory cortex using cross-modal correspondence analysis.* Neuroimage, 22(1):120–133, May 2004. *Clinical Trial.*
- [157] Sameer A Sheth, Masahito Nemoto, Michael Guiou, Melissa Walker, Nader Pouratian, and Arthur W Toga. *Linear and nonlinear relationships between neuronal activity, oxygen metabolism, and hemodynamic responses.* Neuron, 42(2):347–355, Apr 2004.

- [158] E Shtoyerman, A Arieli, H Slovin, I Vanzetta, and A Grinvald. Long-term optical imaging and spectroscopy reveal mechanisms underlying the intrinsic signal and stability of cortical maps in V1 of behaving monkeys. *J Neurosci*, 20(21):8111–8121, Nov 2000.
- [159] Manbir Singh, Sungheon Kim, and Tae-Seong Kim. Correlation between BOLD-fMRI and EEG signal changes in response to visual stimulus frequency in humans. *Magn Reson Med*, 49(1):108–114, Jan 2003.
- [160] Hamutal Slovin, Amos Arieli, Rina Hildesheim, and Amiram Grinvald. Long-term voltage-sensitive dye imaging reveals cortical dynamics in behaving monkeys. *J Neurophysiol*, 88(6):3421–3438, Dec 2002.
- [161] I Tasaki, A Watanabe, R Sandlin, and L Carnay. Changes in fluorescence, turbidity, and birefringence associated with nerve excitation. *Proc Natl Acad Sci U S A*, 61(3):883–888, Nov 1968.
- [162] Bertrand Thirion. fMRI data analysis: statistics, information and dynamics. *PhD thesis, Ecole Doctorale Télécom Paris*, October 2003.
- [163] N J Trujillo-Barreto, E Martínez-Montes, L Melie-García, and P A Valdés-Sosa. A symmetrical bayesian model for fMRI and EEG/MEG neuroimage fusion. *International Journal on Bioelectromagnetism*, 3(1), 2001.
- [164] D Y Ts’o, R D Frostig, E E Lieke, and A Grinvald. Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science*, 249(4967):417–420, Jul 1990.
- [165] R. van der Merwe. Sigma-point kalman filters for probabilistic inference in dynamic state-space models. In *Workshop on Advances in Machine Learning*, Montreal, June 2003.
- [166] B. D. van Veen, van Dronglen. W., M. Yuchtman, and A. Suzuki. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng*, 44:867–880, 1997.
- [167] B.D. van Veen and K.M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2):4–23, April 1988.
- [168] S Vanni, J Warnking, M Dojat, C Delon-Martin, J Bullier, and C Segebarth. Sequence of pattern onset responses in the human visual areas: an fMRI constrained vep source analysis. *Neuroimage*, 21:801–817, 2004.
- [169] I Vanzetta and A Grinvald. Increased cortical oxidative metabolism due to sensory stimulation: implications for functional brain imaging. *Science*, 286(5444):1555–1558, Nov 1999.
- [170] Ivo Vanzetta, Thomas Deneux, Guillaume Masson, and Olivier Faugeras. Linearity and non-linearity of sensory-evoked neuronal and hemodynamic responses in awake monkey v1. In *INCM workshop, Marseille, October 2005*.

- [171] Ivo Vanzetta, Rina Hildesheim, and Amiram Grinvald. *Compartment-resolved imaging of activity-dependent dynamics of cortical blood volume and oximetry*. *J Neurosci*, 25(9):2233–2244, Mar 2005.
- [172] Ivo Vanzetta, Hamutal Slovin, and Amiram Grinvald. *Spatio-temporal characteristics of neurovascular coupling in the anesthetized cat and the awake monkey*. In *International Symposium Brain Activation and CBF Control*, pages 145–153, Tokyo, 2001.
- [173] A. L. Vazquez and D. C. Noll. *Nonlinear aspects of the BOLD response in functional mri*. *NeuroImage*, 7(2):108–118, 1998.
- [174] D. Vitacco, D. Brandeis, R. Pascual-Marqui, and E. Martin. *Correspondence of event-related potential tomography and functional magnetic resonance imaging during language processing*. *Hum Brain Mapp*, 17(1):4–12, 2002.
- [175] Maxim Volgushev, Joachim Pernberg, and Ulf T Eysel. *A novel mechanism of response selectivity of neurons in cat visual cortex*. *J Physiol*, 540(Pt 1):307–320, Apr 2002.
- [176] T. D. Wager, A. Vazquez, L. Hernandez, and D. C. Noll. *Accounting for nonlinear BOLD effects in fMRI: parameter estimates and a model for prediction in rapid event-related studies*. *NeuroImage*, 25(1):206–18, Mar 2005.
- [177] Anthony B Waites, Marnie E Shaw, Regula S Briellmann, Angelo Labate, David F Abbott, and Graeme D Jackson. *How reliable are fMRI-EEG studies of epilepsy? A nonparametric approach to analysis validation and optimization*. *Neuroimage*, 24(1):192–199, Jan 2005. *Case Reports*.
- [178] S Warach, J R Ives, G Schlaug, M R Patel, D G Darby, V Thangaraj, R R Edelman, and D L Schomer. *EEG-triggered echo-planar functional MRI in epilepsy*. *Neurology*, 47(1):89–93, Jul 1996. *Case Reports*.
- [179] B. Weber, C. Burger, M. T. Wyss, G. K. von Schulthess, F. Scheffold, and A. Buck. *Optical imaging of the spatiotemporal dynamics of cerebral blood flow and oxidative metabolism in the rat barrel cortex*. *European Journal of Neuroscience*, 20:2664–2670, 2004.
- [180] Max Welling. *Max welling’s classnotes in machine learning*. http://www.cs.toronto.edu/~welling/classnotes/papers_class/KF.ps.gz.
- [181] K. J. Worsley. *Statistical analysis of activation images*. In *SM Smith P Jezzard, PM Matthews, editor, Functional Magnetic Resonance Imaging: An Introduction to Methods*, pages 251–270. Oxford Univ. Press, York, 2001.
- [182] K.J. Worsley, S. Marrett, P. Neelin, and A.C. Evans. *Searching scale space for activation in pet images*. *Human Brain Mapping*, 4:74–90, 1996.
- [183] K.J. Worsley, S. Marrett, P. Neelin, A.C. Vandal, K.J. Friston, and A.C. Evans. *A unified statistical approach for determining significant signals in images of cerebral activation*. *Human Brain Mapping*, 4:58–73, 1996.

-
- [184] Ying Zheng, David Johnston, Jason Berwick, Danmei Chen, Steve Billings, and John Mayhew. *A three-compartment model of the hemodynamic response and oxygen delivery to brain*. *NeuroImage*, 28:925–939, 2005.
- [185] Ying Zheng, John Martindale, David Johnson, Myles Jones, Jason Berwick, and John Mayhew. *A model of hemodynamic response and oxygen delivery to brain*. *NeuroImage*, 16:617–637, 2002.