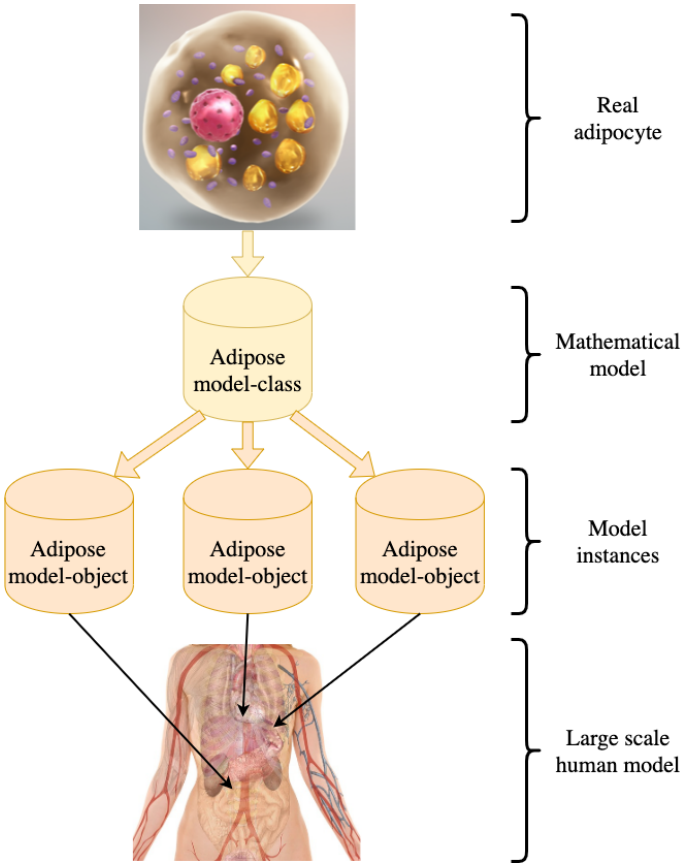
CHALMERS
UNIVERSITY OF TECHNOLOGY



Mathematical modelling of the human physiology

Developing a new framework for creating mathematical models adapted to the digital twin project and modelling the neurovascular coupling

Master's thesis in Master Program Physics

Gustav Magnusson

Department of Physics
Chalmers University of Technology
Gothenburg, Sweden 2020

Master's thesis 2020

Mathematical modelling of the human physiology

Developing a new framework for creating mathematical models adapted to the digital twin project and modelling the neurovascular coupling

Gustav Magnusson

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Physics
Chalmers University of Technology
Gothenburg, Sweden 2020

Mathematical modelling of the human physiology
Developing a new framework for creating mathematical models adapted to the digital twin project and modelling the neurovascular coupling
Gustav Magnusson

Supervisor: Gunnar Cedersund, Department of Biomedical Engineering, Linköping University
Examiner: Peter Apell, Department of Physics, Chalmers University of Technology

Master's Thesis 2020:NN
Department of Physics
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: The new modelling framework, developed to meet the new requirements of the digital twin project

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2020

Mathematical modelling of the human physiology
Developing a new framework for creating mathematical models adapted to the digital twin project and modelling the neurovascular coupling
Gustav Magnusson
Department of Physics
Chalmers University of Technology

Abstract

This master's thesis is carried out within the field of systems biology and aims to contribute to the advancement of mathematical modelling for human physiology. This is important for many reasons, e.g. the potential reduction of animal testing. The work consists of two related parts: firstly, the development of a new framework for creating mathematical models to fulfill requirements of the Digital twin project, an undertaking currently being performed by Gunnar Cedersund's group at Linköping University. The Digital twin project aims to produce a detailed computer model of an individual person which can be tuned to their unique physiology, i.e. a digital twin. This digital twin could then be used to improve patient compliance and/or understanding by simulating the likely effects of a medical treatment, or to make risk predictions for various diseases. The second part of the thesis revolves around the development of a neurovascular coupling model. This model is used to test the hypothesis that the post-stimulus response seen in a typical blood oxygen level dependent signal, measured by a magnetic resonance imaging camera, is due to a change in relative excitatory and inhibitory neural activity. This neurovascular coupling model also serves as a test for the new modelling framework described above.
A first version of a digital twin software, where the new modelling framework plays a central role, was successfully completed and provides a solid foundation for further development in the digital twin project. The neurovascular coupling model was successful in showing the capabilities of this new modelling framework, however due to concerns regarding its validity it cannot definitively support the neural hypothesis. Instead, the model demonstrates the importance of thoroughly understanding the underlying physiology in order to assess the legitimacy of a neurovascular coupling model. This thesis has provided that understanding, and therefore can be viewed as a good starting point for future efforts to model the neurovascular coupling phenomenon.

Keywords: mathematical modelling, systems biology, object-oriented modelling, neurovascular coupling, fMRI, BOLD.

Acknowledgement

Thanks to the people in Gunnar Cedersund's research group who welcomed me and made life in Linköping quite pleasant. Extra thanks to Sebastian Sten who would come to serve as my closest mentor and help me with my research. I also want to thank my friends Gesche Siebke and Evan Mitchell who, to my knowledge, were the only ones willing to read the thesis from page to page and at the same time look for grammatical errors and give feedback, true friends indeed. And last, thanks to my supervisor Gunnar Cedersund and examiner Peter Apell for assisting me along the way.

Gustav Magnusson, Gothenburg, November 2020

Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

The next list describes several symbols and abbreviations that will be later used within the body of the document.

$CMRO_2$ Cerebral Metabolic Rate of Oxygen

f Relative change of cerebral blood flow

n Coupling factor

n_E Relative change of excitatory neural activity

n_I Relative change of inhibitory neural activity

r Relative change of cerebral metabolic rate of oxygen

y Relative change of MRI signal, also called the BOLD signal

BOLD Blood Oxygen Level Dependent

CBF Cerebral Blood Flow

CBV Cerebral Blood Volume

CVODE Numeric ODE solver package

DAE Differential Algebraic Equation

dHb Deoxyhemoglobin

fMRI Functional Magnetic Resonance Imaging

IDA Numeric DAE solver package

IQM IntiQuan MATLAB

MH Metropolis-Hastings

MRI Magnetic Resonance Imaging

NVC Neurovascular coupling

ODE Ordinary Differential Equation

1
Introduction

1.1 Animal testing

Modern medicine has seen an accelerating development with revolutionary discovery such as anesthesia, antibiotics, imaging techniques and vaccine (Hajar, 2015). With the scientific revolution taking place at the end of the 18th century humans learned how to make careful observation and draw empirical conclusions about biological, chemical and physical laws and processes that governs the world. This principle of carefully observing is very evident in medical studies. Before any new drug or medical procedure is allowed to be delivered to humans it has to undergo many safeguarding experiments to validate its function. One of these safeguards is animal testing (Hajar, 2011). Today the standard procedure for drug testing is a several stage process where animal testing is an essential link in the chain of validating a new drug, see figure 1.1.

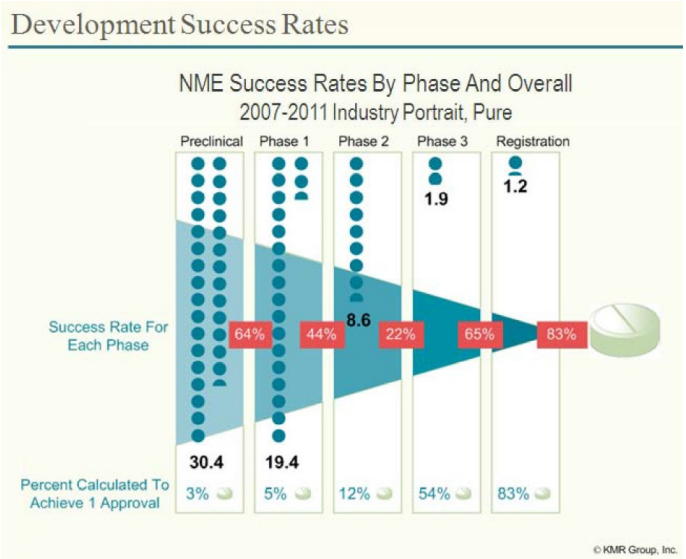


Figure 1.1: Statistics from the US Food and Drug Administration (FDA). The different phases when assessing the validity of a New Molecular Identity (NME) as a potential drug. The red boxes show the percentage of drugs that reaches the next phase and the number at the bottom shows the probability for a drug in this phase to succeed in all remaining phases and becoming approved by the FDA. Animal testing are included in the preclinical phase and human clinical trials includes phase 1-3. Reprinted from Lovell-Badge, 2013.

Even though this system of testing and validating is standard across the world it is not very effective. Over 90% of all drugs that pass through the preclinical stage of non-human testing, which includes animal testing, are stopped at one of the later human trial stages because of unforeseen effects on humans. The infamous 'Elephant man' trials from 2006 is one example. Six volunteers faced life threatening conditions after testing the drug TGN1412 which caused multi-organs failure and swelling of the head, giving name to the trials (Attarwala, 2010). The drug, which was develop for treatment of autoimmune diseases, had shown no serious side effects in the preclinical studies and had been tested on primates in much higher doses than given to the six volunteers. It is no secret that animal models are far from perfect models of the human physiology. In fact, the use of animals is motivated by the notion that animals are different from humans. Research in Alzheimer's disease is one example where the use of animal models have been questioned on how useful they are at modelling the human pathology (McDonald and Overmier, 1997; Drummond and Wisniewski, 2017). In order to study Alzheimer's disease animals, often mice, are artificially given the same symptoms which are seen in a typical Alzheimer's patient (Philipson et al., 2010), often an increase of β-amyloid plaques and neurofibrillary tangles which are believed to be good biomarkers (Reitz, 2012). But drugs successfully treating these symptoms in animals have so far not been able to stop the progression of the disease in humans (Kolata, 2020). The core problem is that Alzheimer's disease is poorly understood and whether amyloid plaques and neurofibrillary tangles are the source or just symptoms of the disease is not known (Markou et al., 2009). Nevertheless, animal testing is the standard benchmark when it comes to assessment and validation of new drugs and medical procedures. However, the high percentage of stopped drugs and the known dissimilarities in human and animal physiology makes one question the justification of the current system. The suffering and horrors experienced by millions of animals used every year in various animal testing are motivated by the essential need for these tests (Ferdowsian and Beck, 2011). Not just that most drugs passing animal trials are latter stopped in human trials, but what about all the drugs that potentially would work on humans but are stopped because of negative outcome on animals? These arguments should perhaps not be used to discredit the current system, but instead motivate the search for alternative methods.

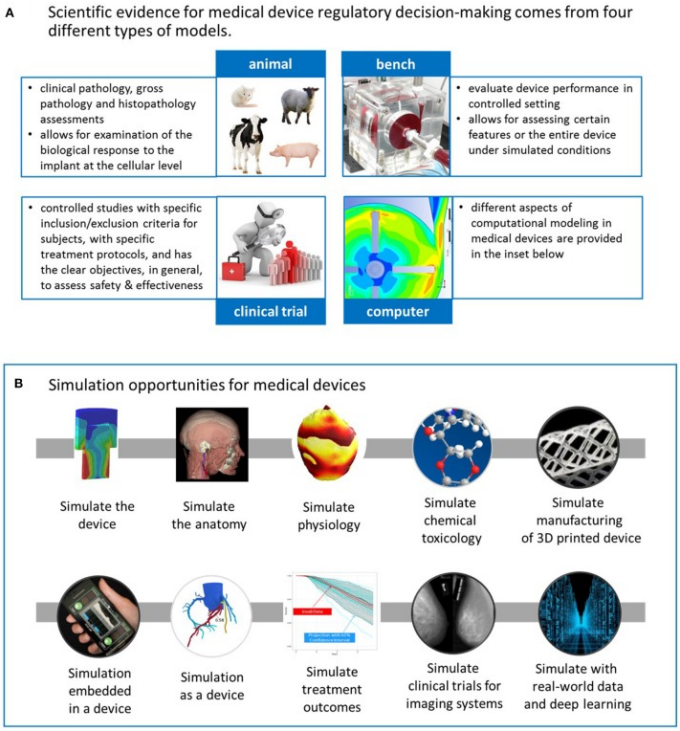


Figure 1.2: The top figure shows different ways of gaining scientific evidence for the functionality of a new drug or medical device. Traditionally animal studies, bench testing and clinical trials have been used, but now with improved computer power an alternative of using computer models has opened up new opportunities. Some of these opportunities are shown in the bottom figure. Reprinted from Morrison et al., 2018.

There is now a growing amount of research trying to find more effective alternatives to the current system of medicine development (Doke and Dhawale, 2015). The 3Rs is an initiative to Reduce, Refine and Replace animal testing and multiple 3Rs centers has been established, for example here in Sweden, to assist and fund research with the long-term goal to face out animal testing. One successful example comes from the US in the development of insulin pumps. Patient with diabetes type 1 need to have insulin given to them externally in response to a meal. Failing in meeting the rise in blood glucose with an appropriate increase in blood insulin can cause hyperglycemia and be life threatening to a person. Insulin pumps are designed to measure the blood glucose level and regulate it by releasing insulin into the blood stream. These insulin pumps used to be validated using animals, usually dogs (Cobelli, Renard, and B. Kovatchev, 2011). But since 2008, the Food and Drug Administration (FDA) has accepted the use of simulation program to validate insulin pumps (B. P. Kovatchev et al., 2009). This has made the use of animals obsolete as companies both save time and money by using simulation instead of animals (Morrison et al., 2018). This is an example where facing out animal testing has benefited the industry as well as animal welfare.

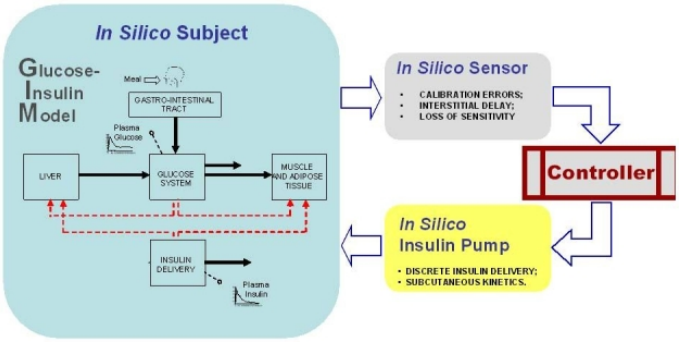


Figure 1.3: An In Silico simulation environment, where the functionally of the insulin controller, shown in red text above, can be tested by computer simulations. Since 2008, FDA has approved these types of alternative methods for validating insulin pumps for patient with type 1 diabetes. This has caused animal testing, which used to be the standard validation procedure, to become obsolete. Reprinted from B. P. Kovatchev et al., 2009.

1.2 Aim of thesis

This master's thesis is taken place at Gunnar Cedersund's group at Linköping University where work is being done with in the field of systems biology to model various biological systems (Cedersund, 2020). The group is partly funded by the Swedish 3Rs center as it aims to develop mathematical models of the human physiology that has the potential to replace animal models. One major goal of the group is the digital twin project, a computer platform which would allow users, such as a patient, to interact with the computer models describing their own physiology, i.e. a digital twin. The work of this thesis is within the digital twin project and has two main focuses:

- Create a new modelling framework which is adapted for the digital twin project. The models need to be made interactive, both with other models in order to be simulate larger and more complex systems and with a user who are using the digital twin software to understand his or her physiology.

- Study and develop a model for the neurovascular coupling between neural activity, oxygen consumption and blood flow. The main goal is to test the claims made by Mullinger et al., 2017 about the origin of the post-stimulus response, seen in a typical Blood Oxygen Level Dependent (BOLD) signal. The developed model also serves to test the new modelling framework.

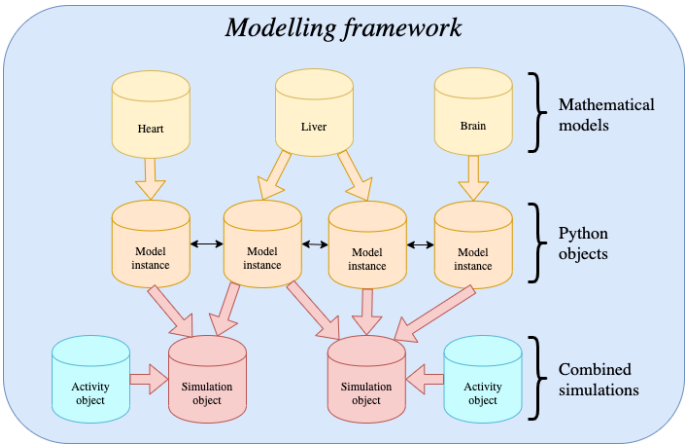


Figure 1.4: The new modelling framework, which is one of the main focuses of this thesis. The goal is to develop a more flexible modelling framework which is adapted for the new challenges of building a digital twin software. The main idea is to think of the models as autonomous entities, objects, that be combined with other models to perform more complex simulations. These models, or objects, should also be communicative and able to provide information about themselves to other parts of the digital twin software.

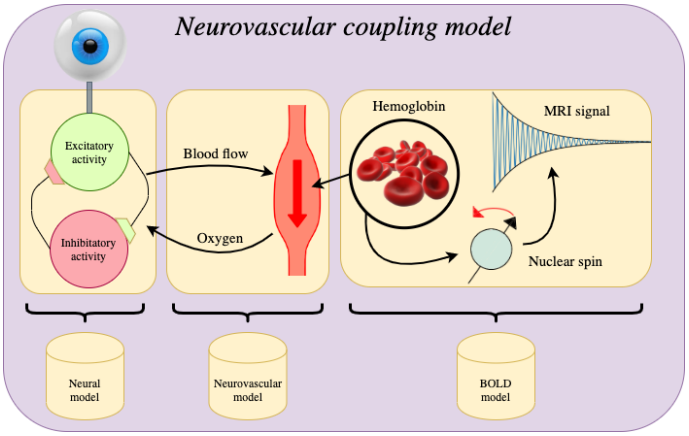


Figure 1.5: The second focus of this thesis is to build a neurovascular coupling model which relates a change in neural activity to a corresponding change in oxygen consumption and blood flow which gives rise to changed oxygenation of the blood, which in turn is captured by the Blood Oxygen Level Dependent (BOLD) signal, measured by a Magnetic Resonance Imaging (MRI). The goal is to test the claims made by Mullinger et al., 2017 about the origin of the post-stimulus response, seen in a typical BOLD signal, being causes by a relative change of neural activity. The developed model would also serve to test new modelling framework.

1.2.1 Limitation

The digital twin project is a large endeavor and there are multiple hurdles in the way before a functional platform is at place. One such hurdle is to make the computer models interactive, both with each other and with a user, which is the main problem

addressed in this thesis. The aim is not to develop or adapt all the models that are thought to be incorporated into the digital twin. Also, the actual user and model interface are not developed here, their development is the main focus of a bachelor thesis group, doing their work within the field of computer science. There is however a close collaboration between this master's thesis work and the work done in the bachelor group, but with the clear separation with the user and model interface being developed by the bachelor group and the modeling framework being developed here. The neurovascular coupling has been extensively studied and many different models have been developed which address everything from the actual neurotransmitters and proteins involved to the imaging accusation (Richard B. Buxton, 2013). The model developed in this master's thesis is building, to a large extent, on previous work and is not aimed at extensively model all the different aspect of the neurovascular coupling. Instead, it focuses on the observations made by Mullinger et al., 2017 which suggest that the post-stimulus response in the BOLD signal is due to a relative change in excitatory and inhibitory neural activity.

2

Background and theory

2.1 Systems biology

This work is done within the field of systems biology. Systems biology seeks to explain and model biological processes through the use of mathematics and computer simulations, even though the exact definition of systems biology is not crystal clear. This has opened a new way of understanding the vast complexity of biological processes. The main difference from classical biological studies is that instead of being a purely linear process, where the focus is on either proving or disproving a given hypothesis, systems biology allows for a more circular approach when developing new knowledge. In the light of experimental data, a model is hypothesized which describes the underlying biological process. From model simulations predictions are produced, these can then, by the use of statistical tools, be compared with the experimental data to validate the model as acceptable or not (Johansson, 2017). If not accepted this gives, apart from an updated belief in your hypothesis, insight how the actual model should look like. When a sufficiently good model is developed, it can be used to find prediction that further distinguish the model from others, these predictions are called 'core predictions' (Cedersund and Roll, 2009). The core predictions give clues where to look next for more experimental data to further validate the model and drive the development forward. This approach is at heart a circular process where models are being developed according to existing experimental data and new experimental data is gathered through the insight gained from the models, see figure 2.1.

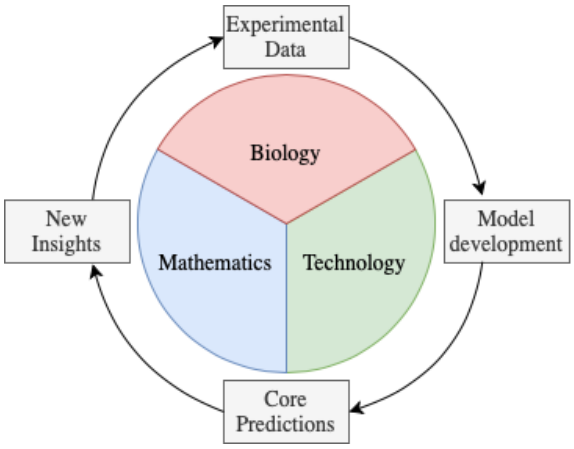


Figure 2.1: Showing the intersections of systems biology, being dependent on both biology, mathematics and technology. The figure also shows the circular process at the heart of systems biology. Experimental data give rise to a model hypothesis. The model, if accepted, can be used to guide the experimentalist in gathering new data by making core predictions that distinguish the model. This will hopefully lead to new insight and further drive the development forward.

The above figure also illustrates the intersection of systems biology between biology, technology and mathematics.

2.1.1 Differential equation and computer models

The types of computer models studied in this thesis are so called dynamical casual models (DCM), based on ordinary differential equations (ODE). Differential equation describes changes in a systems over time with in the realm of differential calculus. Biological systems are often identified through the changes in its internal states and interaction with other systems, for example metabolic flows or neural activity due to stimulus, why ODE:s are a useful tool. The main difference to other scientific fields, where mathematics and computer simulations have been used for a long time, is the great complexity found in biological systems (Johansson, 2017). As an example, consider the gene network shown in figure 2.2. A certain gene codes for a certain protein and the transcriptions rate is controlled by other proteins, which can both be increased or decreased. The transcripted protein can in turn affect the transcription rate of other proteins and thus allowing for a complicated nested network.

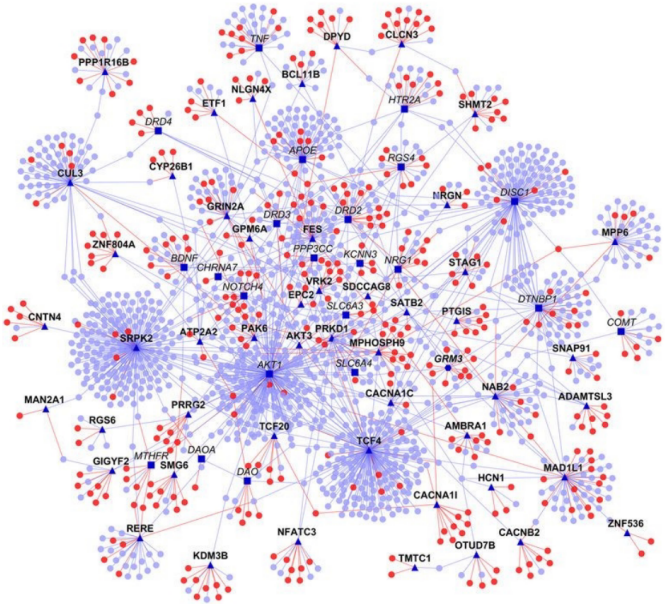


Figure 2.2: A gene network, illustrating the great complexity found in biological systems. The transcription rate of certain protein, showing as a triangle node in the above figure, depends on other proteins, the connected circular nodes. These other proteins are themselves also dependent on other proteins, which allows for complicated networks with feedback loops. Reprinted from Johansson, 2017.

Proteins, such as those shown in figure 2.1, would in a computer model be represented by a state variable $\vec{\mathbf{P}}$, perhaps describing the number density of the individual proteins $\vec{\mathbf{P}}_i$. The dynamics would then be described as coupled ODE:s:

$$\frac{\mathrm{d}\vec{\mathbf{P}}}{\mathrm{d}t} = \vec{\mathbf{f}}(t, \vec{\mathbf{P}}(t), u(t)) \tag{2.1}$$

where $\vec{\mathbf{f}}$ describe the production and clearance of proteins as a function of time t, the current protein density $\vec{\mathbf{P}}(t)$ and some input function $u(t)$, could for example be an experiment defined function. The actual modeling goes into to the choice of the function $\vec{\mathbf{f}}(t, \vec{\mathbf{P}}(t), u(t))$, which in principle could be any sort of function, but is chosen given some hypothesis about the underlying biological system. The strength of a computer model lies in its ability to be easily changed and adapted to test different hypothesis, simply by changing the function $\vec{\mathbf{f}}$. Often, when building a computer model, the shape of the function is fixed but parameters, $\vec{\mathbf{p}}$, determining the exact behaviour of the function are free to change. These parameters could describe timescales, coupling strengths or physical quantities that are not exactly known, but often bounded to some interval. Thus, we should really consider the model function to be a function of the parameters as well:

$$\vec{\mathbf{f}} = \vec{\mathbf{f}}(t, \vec{\mathbf{P}}, u(t), \vec{\mathbf{p}}). \tag{2.2}$$

2.1.2 Bayesian inference

There are multiple ways to obtain optimized parameters that try to make the model mimic the data as close as possible, often done by minimizing a cost function, such as the mean square error. But these different methods will usually only produce a point estimate of the 'best parameters' and not give a measure of the uncertainty in the parameters. An alternative approach, which lets you incorporate prior knowledge and will produce a joint probability distribution over all the parameters, is Bayesian inference. In Bayesian inference, we are interested in computing the conditional probability function $P(\vec{\mathbf{p}} \mid \text{Data}, \vec{\mathbf{f}})$, which is the probability for a given set of parameters $\vec{\mathbf{p}}$ being true, or correct, given the data at hand and the assumption that the model function $\vec{\mathbf{f}}$ is valid. This distribution is often not possible to directly calculate but instead, it can be retrieved by the use of Bayesian theorem, see equation (2.3).

$$P(\vec{\mathbf{p}} \mid \text{Data}, \vec{\mathbf{f}}) = \frac{P(\text{Data} \mid \vec{\mathbf{p}}, \vec{\mathbf{f}}) \cdot P(\vec{\mathbf{p}} \mid \vec{\mathbf{f}})}{P(\text{Data} \mid \vec{\mathbf{f}})} \tag{2.3}$$

Here the probability $P(\vec{\mathbf{p}} \mid \text{Data}, \vec{\mathbf{f}})$ is called the posterior distribution and is related to the likelihood function $P(\text{Data} \mid \vec{\mathbf{p}}, \vec{\mathbf{f}})$, prior probability $P(\vec{\mathbf{p}} \mid \vec{\mathbf{f}})$ and the evidence $P(\text{Data} \mid \vec{\mathbf{f}})$. What makes Bayes theorem useful is that the expression on the left hand side of equation (2.3) can be computed given some assumptions about the model and data statistics. Important to remember is that it is assumed that the model function $\vec{\mathbf{f}}$ is correct and are only computing the posterior distribution for the parameters under this assumption. It is really this assumption that is to be tested: can the model function $\vec{\mathbf{f}}$ satisfactory describe the data and what are the corresponding parameter values. To save space and make the equation more readable I will not write out this assumption explicitly and instead write Bayes theorem as in equation (2.4).

$$P(\vec{\mathbf{p}} \mid \text{Data}) = \frac{P(\text{Data} \mid \vec{\mathbf{p}}) \cdot P(\vec{\mathbf{p}})}{P(\text{Data})} \tag{2.4}$$

Then it is important to remember the implicit assumption that model function $\vec{\mathbf{f}}$ is valid.
In equation (2.4), the denominator $P(\text{Data})$ is generally not necessary to compute as it only a normalization factor and do not depend on the parameters $\vec{\mathbf{p}}$. However, $P(\text{Data})$can be used when comparing two or more competing models, which is not done here. Thus, to compute the posterior distribution $P(\vec{\mathbf{p}} \mid \text{Data})$ it is only the two terms in the nominator, namely the prior probability and the likelihood function, that needs to be computed. The first of these, the prior, is simply chosen, the idea is that it will incorporate prior knowledge about the parameters. This may be knowledge from previous studies or simply just reasonable values for the parameters. The posterior probability can be seen as the updated version of the prior, reflecting the new knowledge gained from the data. For simplicity it is assumed that the prior distribution describes independent parameters so that it can be separated into the individual parameter prior distributions, see equation (2.6).

$$P(\vec{\mathbf{p}}) = \prod_{i=1}^{p} P(p_i) \tag{2.5}$$

I will assume normal distributions for the priors, with parameter p_i being described by mean μ_i^{p} and standard deviation σ_i^{p} according to equation (2.6).

$$P(p_i) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{p_i - \mu_i^{\mathrm{p}}}{\sigma_i^{\mathrm{p}}}\right]^2\right) \tag{2.6}$$

The likelihood function measures how likely the data was given a specific set of parameters. To compute the likelihood function, assumption needs to be made about the underlying system being modelled. Often, by taking several samples of the data D_i^j, where the index i is the data point index and the index j is the sample index, the central limit theorem is invoked to argue that the sample mean D_i is normally distributed, see equation (2.7-2.10).

$$D_i = \frac{1}{s}\sum_{j=1}^{s} D_i^j \tag{2.7}$$

$$\sigma_i^{\mathrm{D}} = \sqrt{\frac{1}{s-1}\sum_{j=1}^{s}(D_i^j - D_i)^2} \Big/ \sqrt{s} \tag{2.8}$$

$$\mu(D_i^j) = \mu_i^{\mathrm{D}} \tag{2.9}$$

$$D_i \sim \mathcal{N}(\mu_i^{\mathrm{D}}, \sigma_i^{\mathrm{D}}) \tag{2.10}$$

Here the mean μ_i^{D} is the mean value for the individual samples D_i^j and σ_i^{D} is the unbiased estimate of the sample mean standard deviation. Notice in equation (2.8) the denominator $\sqrt{s}$, this is because this is the standard error of the mean, SEM, and not the standard deviation, SD, see equation (2.11).

$$\mathrm{SEM} = \frac{\mathrm{SD}}{\sqrt{\text{Number of samples}}} \tag{2.11}$$

In terms of the model, it is the mean μ_i^{D} that is usually being predicted. The likelihood function then looks like in equation (2.12).

$$P(\text{Data} \mid \vec{\mathbf{p}}) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left[\frac{D_i - M_i}{\sigma_i^{\mathrm{D}}}\right]^2\right) \tag{2.12}$$

Where M_i is the model predicted value for data point i, perhaps resulting from a time series with data collected at time points t_i, and D_i and σ^{D} are the sample mean and standard deviation from equation (2.7) and (2.8).
We can now compute the posterior distribution as a product of individual factors using equation (2.5), (2.12) and (2.6), however this product tends to be a very small number so instead the logarithmic of the posterior is often computed, see equation (2.13).

$$\begin{aligned}\log(P(\vec{\mathbf{p}} \mid \text{Data})) &= \log\left(\frac{P(\text{Data} \mid \vec{\mathbf{p}}) \cdot P(\vec{\mathbf{p}})}{P(\text{Data})}\right) \\ &= \sum_{i=1}^{d} -\frac{1}{2}\left(\frac{D_i - M_i}{\sigma_i^{\mathrm{D}}}\right)^2 + \sum_{i=1}^{p} -\frac{1}{2}\left[\frac{p_i - \mu_i^{\mathrm{p}}}{\sigma_i^{\mathrm{p}}}\right]^2 (+ \text{ constant terms})\end{aligned} \tag{2.13}$$

Here the constant terms are the prefactors $\frac{1}{\sqrt{2\pi}}$ from equations (2.5) and (2.12) and the denominator P(Data from equation (2.4), which can all be seen as normalization factors which generally don't have to be computed.

The sum in equation (2.13) is also, a part from a factor -2, a chi-2 distributed variable. This fact will be used when setting a threshold value on the sampled posterior distribution, see section 3.2.3.

2.2 Digital twin project

For systems biology to be a serious alternative to existing animal testing, it needs to be able to model large complex systems on multiple different time and length scale. This is the goal of the digital twin project being developed in Gunnar Cedersund's group at Linköping University. The aim of the digital twin project is to develop a platform in which models describing the human physiology can be combined and provide an interface from which they can be simulated. The models parameters would be adapted to a user's medical data such that the models mimic that specific user's physiology, i.e. a digital twin. The digital twin would assist to answer question such as risk factors for various diseases or help patient to stay on medical plan by illustrating the effect of a certain treatment. The project is so far in its cradle and the work of this thesis is focused on developing a modelling framework for combining models in a standardized way and making them interactive with a user and model interface.

2.2.1 Module based modelling

The human physiology is very complex and a process taking place in one part of body will unavoidable affect other, perhaps remote, parts as well. An example is the glucose homeostasis which maintain the blood sugar level within acceptable values. It works through a close feedback system between food intake, the pancreas, liver, brain and other peripheral tissues such as fat and muscle. An increase in glucose level of the blood from eating food will signal the pancreas to start producing insulin which in turns enables the uptake glucose by other organs in the body. In a similar way will a decrease in glucose blood level, perhaps due to an increased activation of muscles, signal to the pancreas to start producing glucagon which tells the liver to secrete glucose, stored in the liver as glycogen, into blood stream. Also, the brain is believed to play a central part through the brain-centered glucoregulatory system (BCGS) (Scarlett and Schwartz, 2015). Thus, a complete model needs to, in varying detail, incorporate models of all of these different organs and tissues. An overview of glucose homeostasis is shown in figure 2.3.

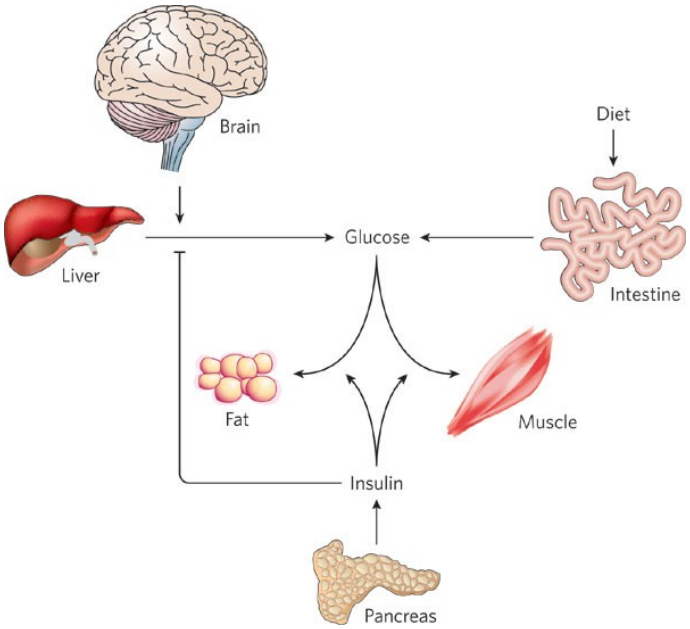


Figure 2.3: An overview of glucose homeostasis which incorporates many different organs and tissues which all plays a role in regulating the glucose levels of the blood. Thus, to model the glucose homeostasis there needs to be, in varying detail, models of all these different organs and tissue. Reprinted from Rosen and Spiegelman, 2006.

However, different organs and tissues are often studied and modelled independently, as separate parts of a larger system. The reasons for this are many, but the most obvious is perhaps the sheer complexity of accounting for the entire system in one big model, which would make the modelling task unfeasible. What then is needed is a way to combine smaller models into larger ones, such that a complicated system can first be understood by its integral parts, before being modelled in its entire complexity. This is a module-based approach where models are seen as being composed by several smaller modules. The modules can be replaced or refined, without affecting the other modules in the model, they may also be reused in other modelling task, which perhaps also requires models for fat or muscle tissue.

2.2.2 Current modelling framework

The current modelling framework used in Gunnar Cedersund's group is based on the IntiQuan MATLAB (IQM) toolbox developed for MATLAB (IntiQuan, 2017). It's a simple to use toolbox that transform differential equation defined in a text-based environment to a compiled MATLAB executable (MEX) file, which then can be run from within MATLAB to perform simulations. Once the file is compiled it is not possible to do any further adjustment apart from changing parameters and initial state values. Most importantly, this means that models can't communicate with other models or share variables as they are being simulated, a big limitation of the current toolbox. Models developed independently needs to have the ability to be incorporated into a larger model, where they can influence each other, just as they would do in real experiment. An example is taken from the work done in Gunnar Cedersund's group where three different models describing weight loss (Hall and

Jordan, 2008), food intake (Dalla Man, Rizza, and Cobelli, 2007) and adipose tissue (Brännmark et al., 2013) are combined into a larger whole body model, see figure 2.4. This new model can simultaneously model weight changes as well as cellular changes of adipocytes in response to a meal.

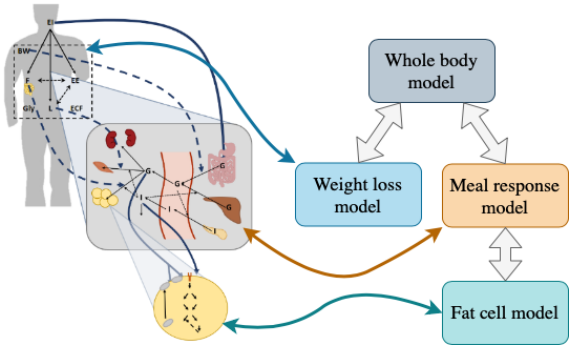


Figure 2.4: An example of how models are being combined into larger more complicated models. Here three models describing weight loss (Hall and Jordan, 2008), food intake (Dalla Man, Rizza, and Cobelli, 2007) and adipose tissue (Brännmark et al., 2013) are combined into a whole body model. One of the limitations of the old framework is however that each new combination of smaller models needs to compiled into its own MEX-file. Meaning if anything changes in a smaller model all bigger models where the smaller model is used needs to be updated and recompiled.

However, in the current framework an entirely new model needs to be created and compiled into a new MEX-file, separate from the individual models. This is a very static solution since any new combination of models needs to be compiled in this way, further, if anything changes in any of the smaller models the changes needs to be updated in every bigger model where the smaller model is used, a logistic long-term nightmare. These current limitation of the toolbox needs to be eradicated in order to achieve a more flexible and functional digital twin software, which is one of the main focuses of this thesis.

2.2.3 User and model interface

For the models to be useful in a clinical application there needs to be an interface that enables users, such as patient, to easily interact with the models and perform simulation. Such an interface is being developed in parallel by a bachelor thesis group doing their bachelor in computer science. They are using the interactive program MeVisLab that has been developed for medical imaging and procession, see figure 2.5. MeVisLab allows developers to incorporate their own software into the program, (also) called modules, which makes it easy to adapt to specific needs.

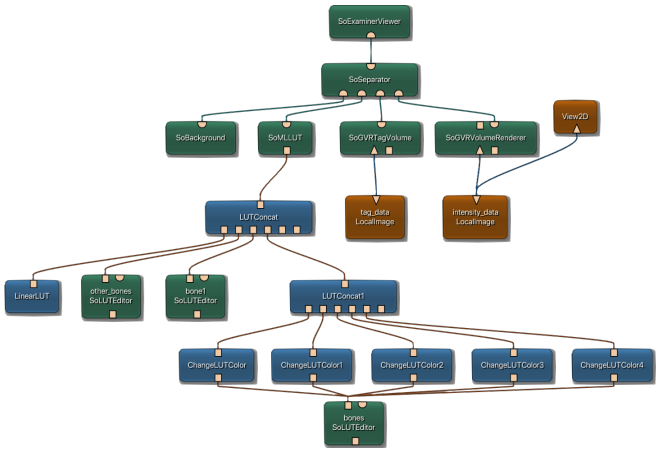


Figure 2.5: An illustration of development inside of MeVisLab where new functionality can be added by creating new modules, boxes shown in the picture. The idea is to use MeVisLab to create an interactive interface from which the models can be simulated by a user, this is done in parallel by bachelor thesis group. Reprinted from MeVisLab, 2020.

The idea is then to build a platform from which the models can not only communicate with each other but also with an interface provided by MeVisLab. The models would supply information about them self and then, through choice being made by the user, be simulated with other models to provide an interactive simulation environment.

2.3 Neurovascular coupling

Magnetic resonance imaging (MRI) has given us the capability to non-invasive look inside the brain to investigate the physiological processes taking place there. One technique extensively used since its initial development is the Blood Oxygen Level Dependent (BOLD) response (Ogawa et al., 1990). It relies on the fact that when different parts of the brain are stimulated, the local change in neural activity alters the metabolic rate and blood flow in an unequal way. The cell metabolic consumption of oxygen ($CMRO_2$), which is closely related to the energy expenditure of activated neurons (Attwell and Laughlin, 2001), is increased less than the accompanied increase in cerebral blood flow (CBF) (Davis et al., 1998). The driving mechanism is believed to be a feed forward network from neural activity to blood flow, thus the name 'Neurovascular coupling' (Hillman, 2014), see figure 2.6. The unequal change in CBF and $CMRO_2$ causes the relative concentration of oxyhemoglobin to deoxyhemoglobin to increase, which in turn changes the local magnetic susceptibility since oxyhemoglobin is diamagnetic whereas deoxyhemoglobin is paramagnetic (Ogawa et al., 1990). This change in local magnetic susceptibility is captured by an MRI camera which allows for a functional mapping of the brain due to various types of stimulus. But even though the technique has been extensively used, the question still stand as to exactly what physiological process are measured in the BOLD-response (Mishra et al., 2016). The evidence points towards CBF and

$CMRO_2$ being controlled in parallel by neural activity but haemodynamic effects such as change in cerebral blood volume (CBV) makes the casual relation between the BOLD-response and neural activity obscured (Richard B. Buxton et al., 2014). To make a heuristically connection between the underlying physiological processes and the BOLD-response computer models are used. Building realistic models of the neurovascular coupling is valuable for many reason, for example to understand different neurological diseases which often are accompanied with an altered hyperemia (Petzold and Murthy, 2011).

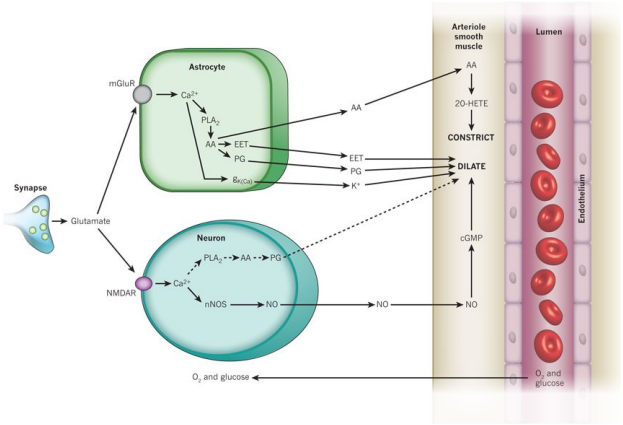


Figure 2.6: Showing a feed forward network from synaptic activity to blood flow with the believed important metabolites (not discussed here). Neurons constantly signal to local arterioles their current energy demand and the arterioles respond by either dilating or constricting to increase or decrease the local supply of oxygen and glucose. The communication is believed to be mediated through different substrates such as nitrogen oxide (NO) and glial cells such as astrocytes. As the oxygen consumption increase there is a relative larger increase in blood flow which alters the local concentration of hemoglobin to deoxyhemoglobin, which can be captured by MRI camera. This is used to map local brain activity in the so-called BOLD-response. Reprinted from Attwell, Buchan, et al., 2010.

2.3.1 Physics behind the BOLD-response

Magnetic resonance imaging relies on one thing, nuclear spin, or more specifically the spin of hydrogen nucleus. Human are largely made up by water, about 60%, which means that there is a large amount of hydrogen atoms spread out of the human body (Plewes and Kucharczyk, 2012). Nuclear spin is a fundamental property of matter and has with it an associated magnetic moment, which is being probed by the MRI technique (Richard B. Buxton, 2013). In MRI, a magnetic field causes the magnetic moment of the hydrogen nucleus to align with the field, then a short radio frequency (RF) pulse causes the magnetic moment to flip with an angle to the applied magnetic field, the flip angle. When the RF pulse then is turned off the magnetic moment will precess around applied magnetic field with a frequency given by the Larmor frequency, see equation (2.14).

$$\omega = \gamma |B| \tag{2.14}$$

Here ω is the precession frequency around the applied magnetic field B and γ is the gyromagnetic constant, which for a hydrogen nucleus has the value $2.675 \times 10^8\,\mathrm{rad\,T^{-1}}$. A typical magnetic field used in MRI is 3 T which gives a precession frequency $f = \omega\ 2\pi \approx 128\,\mathrm{MHz}$.

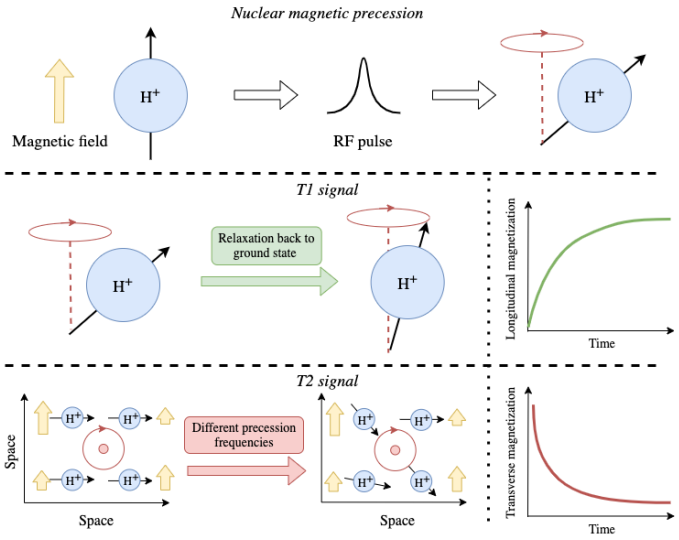


Figure 2.7: The top part of the figure shows hows the magnetic moment of the hydrogen nucleus, H^+, is first aligned with an applied magnetic field in the low-energy state. A radio frequency pulse, tuned to the hydrogen nucleus, is then applied to flip the magnetic moment away from the direction of the magnetic field, this will in turn cause the magnetic moment to precess with a frequency given by the Larmor equation (2.14). The middle part of the above figure shows how the nuclear moments of the hydrogen nucleus return, or decay, back to pointing along the magnetic field, which give rise to the T1 signal measuring the longitudinal magnetization. The decay time is different for different tissue why the T1 signal can be used to visualize tissue content. The bottom part shows a top view with the magnetic field pointing out of the paper. Magnetic field inhomogeneities will cause hydrogen nucleus at the different spatial location to precess at slightly different frequency, causing them to come out of phase and reducing the transverse magnetization which gives rise to the T2 signal.

This precession will in turn cause a electromagnetic signal that is being measured by the MRI camera. The key aspect that makes MRI useful is that the magnetic moment returns, or decays, to its original state in different time depending on the environment the nucleus is in. In this way the MRI scan can differentiate between fat, bone, muscle and other types of tissue by comparing the decay time of the induced precession signal. There are two different types of signal, the T1 and T2 signal, which is the signal along the magnetic field and the signal perpendicular to the magnetic field, respectively. The signal decay measured in the T1 signal is caused by nuclear spins returning to aligning with the magnetic field after initially being flipped by the RF pulse, see picture 2.7. The signal decay measured in the T2 signal is largely due to the nuclear spins coming out of phase from each other and pointing in random directions in the plane perpendicular to the magnetic field,

see figure 2.7. This dephasing is caused by several effects such as local variation in magnetic field strength causing a slightly different Larmor frequency and diffusion of atoms and molecules (Richard B. Buxton, 2013).
It is the sensitivity to local magnetic field variation that makes the T2 signal useful to detect changes in relative concentration of oxyhemoglobin and deoxyhemoglobin. Oxyhemoglobin and deoxyhemoglobin have different magnetic properties, oxyhemoglobin being diamagnetic and deoxyhemoglobin being paramagnetic. Therefore, a change in the relative concentration of oxyhemoglobin/deoxyhemoglobin will alter the local magnetic field and thereby affecting the decay of the T2 signal. Many factors plays in to the exact response such as blood vessel diameter and orientation, blood volume fraction and magnetic field strength, but in general since brain tissue is mostly diamagnetic, like oxyhemoglobin, deoxyhemoglobin will cause a greater magnetic field strength variation and thus a larger decay of the T2 signal (Richard B. Buxton, 2013), see figure 2.8.

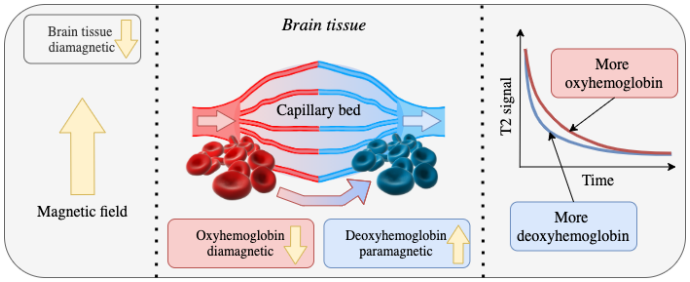


Figure 2.8: Schematic overview of the effect of change of the relative concentration of oxyhemoglobin and deoxyhemoglobin to the T2 signal. The T2 signal is sensitive to local magnetic field variation. Oxyhemoglobin is diamagnetic like the surrounding tissue whereas deoxyhemoglobin is paramagnetic. Blood coming from the lungs contains almost 100% oxygenated hemoglobin, illustrated as red blood cells in the above figure, as the blood vessel branch into smaller vessel in the capillary bed oxygen diffuse to the surrounding tissue which consumes it to produce energy for its cells. Blood returning to the lungs in the veins has then a lower amount of oxyhemoglobin and a larger amount of deoxyhemoglobin, illustrated as blue blood cells in the above figure. A larger extraction of oxygen, and thus a larger amount of deoxyhemoglobin, will cause greater magnetic variation which in turn will cause a faster decay of T2 signal. The opposite is also true, a smaller extraction of oxygen will increase the amount oxyhemoglobin and cause a slower decay of T2 signal.

The amount of oxygenated hemoglobin and deoxygenated hemoglobin further depend on the local extraction of oxygen from the blood to the surrounding tissue, happening at the capillary bed. What is surprising with the BOLD-response is that increased neural activity, and thus increased consumption of oxygen, results in an increased concentration of oxyhemoglobin and thus an increased T2 signal. This is counter-intuitive since this means the local extraction of oxygen has decreased as the consumption of oxygen increase. To understand this phenomenon, we have to understand the interplay between blood flow, oxygen concentration and energy consumption.

2.3.2 Physiology behind BOLD-response

The brain constantly consumes energy, and even though it only comprise 2% of the total body mass it accounts for around 20% of the total rest energy consumption (Raichle and Gusnard, 2002). It mainly gets its energy from oxidative metabolism of glucose and therefore rely on arterial blood supply of both glucose and oxygen (Raichle and Gusnard, 2002). Oxygen has very low solubility in water, which is the main component of blood. Evolution has therefore equipped human with red blood cells containing hemoglobin molecules which binds oxygen and increase the solubility of oxygen by a factor 30-50 (Richard B. Buxton, 2013). The binding of oxygen to hemoglobin follows an approximate sigmoidal curve with a near saturation of hemoglobin at oxygen partial pressure (pO_2) larger than 100 torr, which is the typical pressure in lungs, see figure 2.9. The partial pressure in brain tissue is around $pO_2 = 20$ torr, close to the believed pO_2 of the atmosphere about two billion years ago when the first oxidative metabolism developed (Richard B. Buxton, 2010).

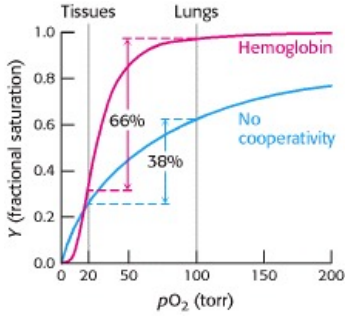


Figure 2.9: Binding of oxygen to hemoglobin as function of partial pressure of oxygen. The figure shows the typical partial pressure in both lungs and tissue, it also illustrates how the delivery of oxygen from the lungs to tissue is increased from about 38% to 66% by the cooperation between the subunits of hemoglobin. Reprinted from Berg, Tymoczko, and Stryer, 2002.

The above figure shows how hemoglobin increase the delivery of oxygen from the lungs to tissue compared to a hypothetical non-cooperative protein, about 1.7 times more of oxygen can be delivered by the fact that hemoglobin's subunits cooperate (Berg, Tymoczko, and Stryer, 2002).
When a certain brain region is activated, due to some type of stimulus, you would expect the local concentration of oxygenated hemoglobin to decrease, because of the increased consumption of energy and therefore the increased consumption of oxygen. But the opposite is seen in the BOLD-response, an increase in the consumption rate of oxygen, $CMRO_2$, measured per tissue volume and time with unit $\mathrm{mM\,min^{-1}}$, is accompanied with an increase in local oxygenated blood. This counter intuitive reaction is due to the fact that cerebral blood flow, CBF, defined as the amount arterial blood delivered per volume tissue and time with unit $\mathrm{min^{-1}}$, increase about two times more than the $CMRO_2$ in response to increased brain activity. These two quantities are related by a simple mass balance equation:

$$CMRO_2 = E \cdot CBF \cdot [O_2]_a \tag{2.15}$$

where E is the local oxygen extraction factor and $[O_2]_a$ is the arterial oxygen concentration, measured in mM and largely reflecting the saturation level of hemoglobin and the hematocrit. Assuming that the delivered arterial blood contain a near constant concentration of oxygen, about 98% saturation, we see from equation (2.15) that a smaller quotient $CMRO_2/CBF$ is accompanied with an decreased extraction factor E. A decrease in local extraction fraction leads to an increased average blood oxygen concentration and thus resulting in the counter intuitive BOLD-response.

2.3.2.1 Oxygen diffusion hypothesis

The question then arises what physical aspect demands for an increased blood oxygen concentration in response to an increase oxygen consumption? One possible answer can be understood by considering a simple oxygen diffusion model as done by Buxton (Richard B. Buxton, 2010). In this model the local delivery of oxygen from arterial blood to brain tissue can increase due to two things: recruitment of new capillaries or an increased oxygen gradient between blood and tissue leading to greater diffusion. The first options is believed to be negligible or have very little effect (Göbel, Theilen, and Kuschinsky, 1990). For the later options, either the average pO_2 in the tissue must drop or the average pO_2 in the blood must increase. Richard B. Buxton, 2010 hypothesized that in response to an increased oxygen consumption the tissue partial pressure, pO_2^{tissue}, is prevented from dropping by instead increasing the blood partial pressure, pO_2^{blood}. They provide two main reason why it would be beneficial for pO_2^{tissue} to remain close to baseline. The first has to do with reaction kinetics and the fact that many reactions in the brain has Michaelis constants, K_m, close to baseline values of pO_2^{tissue} and therefore sensitive to changes in partial pressure. The other reason has to do with thermodynamics for oxidative metabolism. The change in Gibbs free energy, ΔG, for oxidative metabolism of glucose depends on the local concentration of oxygen, if the concentration is decreased so is also the available energy to produce ATP from ADP and Pi, which is thermodynamic uphill process.

Believing the constant pO_2^{tissue} hypothesis, the oxygen diffusion is increased by increasing pO_2^{blood}. For the pO_2^{blood} to increase there needs to be a decrease of the local extraction fraction E, which, as we pointed out earlier, means a decreased ratio $CMRO_2/CBF$. In this way, the mismatch between the change in $CMRO_2$ and CBF is explained by the need for increased oxygen diffusion at the same time as pO_2^{tissue} is kept constant or close to baseline values, see figure 2.10.

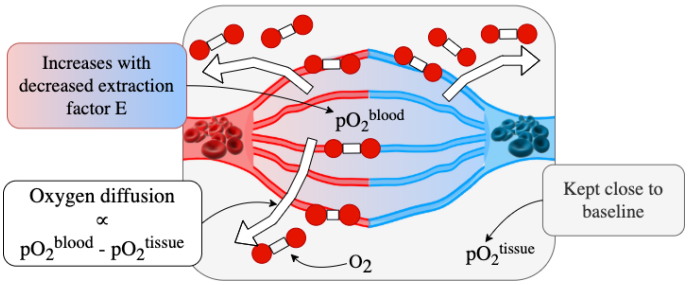


Figure 2.10: A schematic picture of the oxygen diffusion hypothesis which explain the unequal change in blood flow and oxygen consumption seen in the BOLD-response to arise from the need of a larger partial oxygen concentration gradient. It is hypothesis that partial pressure of oxygen in tissue, pO_2^{tissue}, is kept close to baseline, whereas the partial pressure of oxygen in the blood, pO_2^{blood}, can increase by decreasing the local extraction fraction E, see equation (2.15).

An important quantity when talking about changes in CBF and $CMRO_2$ is the coupling factor n, defined in equation (2.16) as the relative change in oxygen consumption divided by the relative change in blood.

$$n = \frac{\Delta_r \mathrm{CBF}}{\Delta_r \mathrm{CMRO_2}} \tag{2.16}$$

Here Δ_r indicate a relative change with respect to baseline ($\Delta_r \mathrm{CBF} = (\mathrm{CBF} - \mathrm{CBF_0})/\mathrm{CBF_0}$ where $\mathrm{CBF_0}$ is the baseline value). An increased coupling factor means a relative larger increase of CBF compared to $CMRO_2$ and thus a reduced extraction fraction E, see equation (2.15). The coupling factor has been measured in various parts of the brain with PET scans and calibrated-BOLD measurement with typical values between 2-4, see figure 2.11. The oxygen diffusion model develop by Richard B. Buxton, 2010 also predicts coupling factors in this range in order to keep pO_2^{tissue} constant, see the solid red line in figure 2.11.

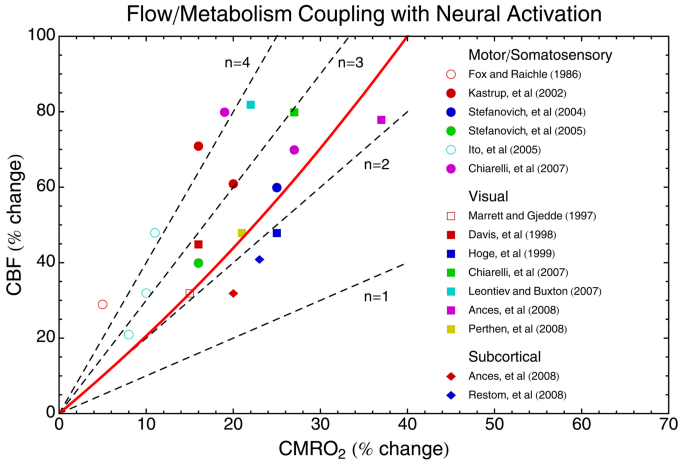


Figure 2.11: Experimental measurement of the change in blood, CBF, and oxygen consumption, $CMRO_2$. PET measurements are indicated with open symbols and calibrated-BOLD measurements are indicated by filled symbols, different symbol shapes are used for different brain areas, as indicated. The dashed lines shows constant values for coupling factor n from equation (2.16) and the solid red curve shows the contour for constant pO_2^{tissue}. All reported values lies above $n = 1$ which indicate a proportional larger response of CBF compared to $CMRO_2$, which is the cause of the BOLD-response. The oxygen diffusion hypothesis suggests that this to avoid a large drop in tissue pO_2. Reprinted from Richard B. Buxton, 2010.

But even though a oxygen diffusion model can provide a physiological arguments for the unequal change in $CMRO_2$ and CBF, it doesn't provide the full picture of the underlying neural activity causing the change in the first place. For example, the coupling ratio n is seen to vary for different types and duration of stimulus (Richard B. Buxton, 2010). This variation has been suggested to reflect a varying degree of excitatory and inhibitory neural activity, with excitatory and inhibitory neural activity affecting $CMRO_2$ and CBF in an unequal way (Richard B. Buxton et al., 2014). To understand why this might be the case we must understand how neurons consume energy.

2.3.2.2 Neural signalling

Neurons in the brain primarily consume energy as they respond to and transfer electrical signals (Attwell and Laughlin, 2001). Neurons maintain a negative potential across their membrane, around $-70\,mV$, mainly through pumping out sodium ions, Na^+, and pumping in potassium ions, K^+, in a ratio 3 to 2 (Richard B. Buxton, 2013). The membrane potential is further regulated by opening and closing a variety of ion channels which lets ions such as sodium, potassium, calcium and chlorine pass in and out of the cell. For example, since sodium ions is maintained at much higher concentration outside the cell, the opening of sodium channels will cause a rapid influx on these positive charge ions and quickly depolarize the cell. On the other hand, the high concentration of potassium ions inside the cell means that the opening these channels will cause K^+ ions to leave the cell and cause a hyperpolarization. These ion channels are in turn controlled by other neurons by either electrical or

chemical contact. The point of contact happens in synapses where the axon from one neuron meets the dendrite of another neuron. The most common scenario is that an action potential, travelling down the axon of the presynaptic neuron, will cause the release of neurotransmitters which then migrate over the synaptic cleft and binds to ion channels at the dendrites of the postsynaptic neuron. The combined effect of all synaptic activity at the dendrites are summed up to determine the overall membrane potential of the cell. If the cell membrane potential is pushed passed a threshold value, approximately $-55\,\mathrm{mV}$, an new action potential is generated at the axon hillock, sitting between the soma (cell body) and the axon, which propagates down the axon and influence other neurons. In this way input signals are summed up to determine if a new outgoing signal should be generated, see figure 2.12.

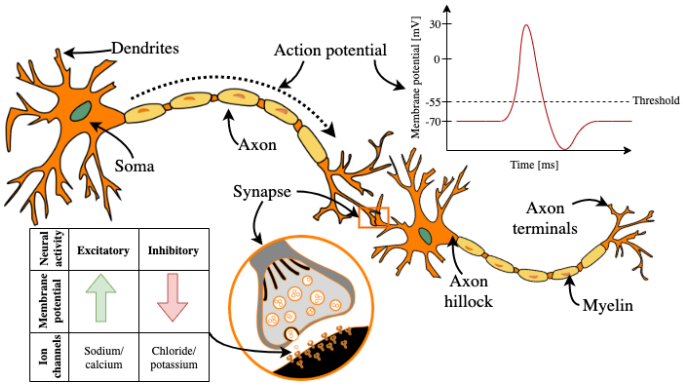


Figure 2.12: Overview of neural signaling between two neurons in the brain. A neuron comprises of a the soma (cell body), dendrites extending out from the soma, an axon covered in a myelin sheet to provide insulation and axon terminals which makes connection with other neurons in synapses. An action potential travelling down a neuron's axon, the presynaptic neuron, will upon arrival at the axon terminals cause the release of neurotransmitters. The neurotransmitters will migrate over the synaptic cleft to the dendrites of another neuron, the postsynaptic neuron, where they bind to specific ion channels. Depending on which ion channels the transmitters bind to specific ions are either let in or out of the cell which in turn either decrease or increase the membrane potential which at rest is kept at around $-70\,\mathrm{mV}$. The activity at all synapses is summed up at axon hillock, sitting between the soma and the axon, if the potential is pushed above a threshold, around $-55\,\mathrm{mV}$, a new action potential will be created. There are two large classes of neural activity, excitatory and inhibitory, with excitatory activity causing an increase of the membrane potential and inhibitory activity causing a decrease. The most common neurotransmitter associated with excitatory activity is glutamate which opens sodium and calcium ion channels, the most common neurotransmitter associated with inhibitory activity is GABA which opens chlorine and potassium channels.

Neural activity that shifts membrane potential in the positive direction, and thus closer to the threshold, is called excitatory neural activity, the most common neural transmitter associated with excitatory activity is glutamate which opens sodium and calcium ion channels (Meldrum, 2000; Lauritzen, 2005). On the other hand, neural activity that shifts the membrane in the negative direction, hyperpolariza-

tion, is inhibitory neural activity, the most common neural transmitter is gamma-Aminobutyric acid (GABA) which opens chloride and potassium ion channels (Bowery and Smart, 2006).

2.3.2.3 Energy cost for neural activity

The energy cost of neural activity comes largely from clearing neurotransmitter and repacking them from the synaptic cleft, generating and propagating the action potential along the axon and restoring electrochemical gradient across membranes (Richard B. Buxton, 2013). Most energy is believed to be associated to postsynaptic activity (Richard B. Buxton, 2013), which involves pumping out ions to restore electrochemical gradients across cell membrane, about three quarters of all energy consumed (Attwell and Iadecola, 2002). Different ions are kept closer or further from their equilibrium potential across the cell membrane in the baseline state, why restoring their gradient will cost different amount of energy. As mentioned earlier are sodium ions kept far away from equilibrium why restoring its chemical potential is a costly process (Attwell and Laughlin, 2001). Chlorine ions, Cl^-, and calcium ions, Ca^{2+}, are kept at high extracellular concentration but because of the negative potential across the membrane Cl^- ions are close to electrochemical equilibrium and Ca^{2+} ions are far from equilibrium (Richard B. Buxton, 2013). Potassium ions, K^+, are kept at a high intracellular concentration and close to equilibrium because of the negative membrane potential (Richard B. Buxton, 2013). This is why it is believed that it is excitatory neural activity that accounts for the majority of energy budget of neural activity. Inhibitory neural activity still has an energy cost for repacking neurotransmitters, but the opening of channels such as chloride tends to stabilize the membrane potential around low values why the effect of sodium and calcium current might be reduced, inhibitory neural activity might therefore even have a negative impact on the energy consumption (Richard B. Buxton et al., 2014).
Because of the unequal energy consumption of excitatory and inhibitory neural activity, is has been suggested that the two types also cause different relative changes of CBF and $CMRO_2$ and thus different coupling factors n (Richard B. Buxton et al., 2014). Excitatory activity, being a big consumer of energy, is believed to cause a large dilation of blood vessel and increased blood flow (Attwell, Buchan, et al., 2010). For inhibitory neurons the pictures is not as clear as some inhibitory neural activity cause a vasoconstriction and decreased blood flow (Uhlirova et al., 2016; Cauli et al., 2004) and others can cause vasodilation and increased blood flow (Estrada and DeFelipe, 1998). The potential variation in coupling factors n for excitatory and inhibitory neural activity could perhaps be used to determine the underlying neural activity seen in a typical BOLD-response (Richard B. Buxton et al., 2014).

2.3.3 The BOLD-response signal

The MRI signal that is important for the BOLD-response is the T2 signal, $S(t)$, discussed earlier, and decays exponentially with a decay constant R_2, see equation (2.17).

$$S(t) = \exp(-R_2 \cdot t) \tag{2.17}$$

The decay constant depends on the local concentration of deoxyhemoglobin with an increased amount deoxyhemoglobin causing a faster decay, see section (2.3.1). Another important quantity in an MRI experiment is the echo time, TE, which is the time after the RF pulse that the signal is being measured. It should be large enough, so the signal has had time to decay, but not too large to cause the signal to complete vanish. The BOLD signal, y, is then measured as the relative change of the signal S from the baseline state S_0, see equation (2.18).

$$y = \frac{\Delta S}{S_0} = \frac{S(TE) - S_0(TE)}{S_0(TE)} = \frac{\exp(-R_2 \cdot TE) - \exp(-R_2^0) \cdot TE)}{\exp(-R_2^0) \cdot TE)} = \\ = \exp\Big(-[R_2 - R_2^0)] \cdot TE) - 1 \approx -[R_2 - R_2^0] \cdot TE = -\Delta R_2 \cdot TE \tag{2.18}$$

Where R_2 is the decay rate of the system and R_2^0 is the baseline value. Notice to measure the BOLD signal y you need have measurement of the baseline state, which is usually taken as an average over some period of time when the subject, whose brain is being monitored, is assumed to be in a resting state.
The BOLD signal is usually measured due to some stimulus, although interesting investigation is also being done on the resting state of the human brain (Heuvel and Hulshoff Pol, 2010). Then the BOLD-response is often divided into two separate phases, the primary response and the post-stimulus response, see figure 2.13. The primary response is associated with an increase of oxyhemoglobin during stimulus whereas the post-stimulus response is associated with an increase of deoxyhemoglobin.

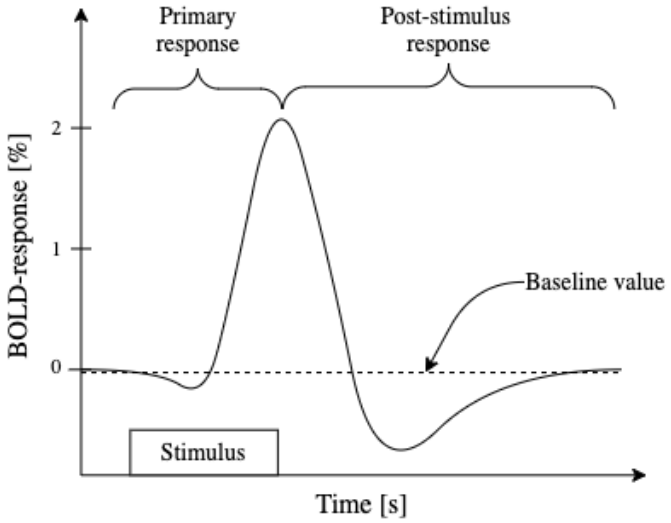


Figure 2.13: A typical BOLD-response due to some stimulus. We see the initial dip that may or may not show up due to low time resolution, the primary response peak and the post-stimulus undershoot.

The BOLD-response may look different from brain region to brain region, and also trial to trial, but some characteristic features are often seen which. includes the initial dip, the succeeding peak and the final undershoot (Richard B. Buxton, 2013),

see the above figure. These different features gives clues to the underlying physiological processes taking place (Havlicek, Roebroeck, K. J. Friston, et al., 2017). The initial dip is associated with an initial increase in deoxyhemoglobin before the blood flow has responded to the increased neural activity, it happens though on such a small timescale that it is not always observed due to poor time resolution of the MRI camera (Richard B. Buxton, 2013). The post-stimulus undershoot is believed to be attributed many different factors such as an uncoupling between blood flow and blood volume, as described in the Balloon model (Richard B Buxton, Wong, and Frank, 1998), where the blood volume decrease slower than the blood flow causing a apparent increase in deoxyhemoglobin, another effect could be a post-stimulus energy storage restoration (Richard B. Buxton, 2013).
The exact reason for post-stimulus response has been one the most debated topic for the last two century and even though many possible explanation has been given no clear consensus has been reached (Zijl, Hua, and Lu, 2012). For this reason, the post-stimulus response was the focus for Mullinger et al., 2017 in their 2017 study, which this thesis focuses on.

2.3.4 The Mullinger study

Mullinger et al., 2017 studied the BOLD-response along with other measurement of the blood flow and electromagnetic activity in a visual cortex experiment, see method section 3.2.1. They hypothesized that the post-stimulus response is primarily neural in origin and is due to relative change of excitatory and inhibitory neural activity from the primary response which can be quantified by a change in coupling factor n[1] Using the combined result from their different measurement they argued that their data supported this claim as they saw a decrease in coupling factor going from the primary to the post-stimulus response. They were also able to correlate the magnitude of the post-stimulus undershoot to an overall increase of neural activity measured by electroencephalography (EEG).
In light of these findings they suggested a qualitative explanation how the different neurological activities can explain the change in coupling factor n from the primary to the post-stimulus response, an explanation which is to be investigated in this thesis. Their qualitative explanation goes like this:

- In the primary response the stimulus drives foremost excitatory neural activity, n_E, which in turn cause an increase inhibitory neural activity, n_I, $\Delta n_E > 0$ and $\Delta n_I > 0$, see figure 2.14.
- Both excitatory and inhibitory is a positive drive for CBF and $CMRO_2$ but both cause a relative larger increase of CBF.
- This causes the coupling factor $n = CBF/CMRO_2$ to increase in primary response.
- In the post-stimulus response the stimulus is removed which causes the excitatory activity to quickly drop and become negative relative to baseline,

[1] Mullinger et al., 2017 definition of the coupling factor is actually the reciprocal of the definition given in equation (2.16), thus their definition is $n = CMRO_2/CBF$.

$\Delta n_E < 0$, the inhibitory neural activity which is not primary driven by the stimulus remains elevated, $\Delta n_I > 0$.

- Excitatory neural activity is the primary drive for both CBF and $CMRO_2$ and will cause a reduction in these which is partly offset by the inhibitory neural activity which remains elevated.

- But since inhibitory neural activity consume less energy than excitatory activity it is believed to be a relative larger drive for CBF than $CMRO_2$ when compared to excitatory neural activity. Thus the offset is greater for CBF than for $CMRO_2$. Causing a decreased in coupling constant n from the primary response.

These arguments are illustrated in figure 2.14.

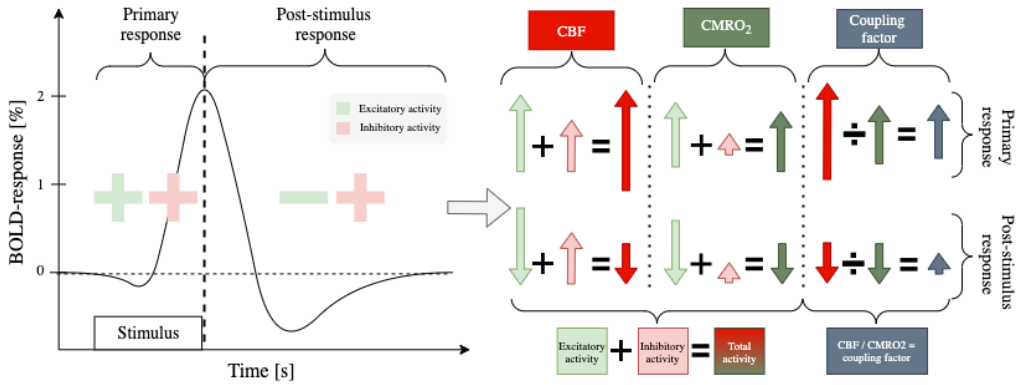


Figure 2.14: An overview of Mullinger et al., 2017 suggestive explanation for the different coupling factors n seen in the two different phases of the BOLD-response. They argue that in the primary response both the excitatory and inhibitory neural activity is positive. Since both activities are proportional larger driver for both CBF than $CMRO_2$ the coupling factor $n = CBF/CMRO_2$ is increased. In the post-stimulus phase the stimulus cessation cause excitatory neural activity to quickly reduce and become negative relative to baseline, the inhibitory activity remains elevated. The primary drive for both CBF and $CMRO_2$ is excitatory neural activity which will lead to a reduction of these, this reduction is however offset by the inhibitory neural activity remaining elevated. The crucial part is that inhibitory activity is larger drive for CBF than $CMRO_2$ when compared to excitatory activity, which cause the offset to be larger for CBF. This in turn cause a decrease of the coupling factor from the primary response.

The goal with this thesis is to build a neurovascular coupling model which incorporate the necessary physiological processes so that the claims made by Mullinger et al., 2017, summarized below, can be tested on a quantitative basis.

- The post-stimulus is primary a neural phenomenon, reflecting a relative change in excitatory and inhibitory neural activity.

- Both excitatory and inhibitory neural activity are positive drivers from CBF and $CMRO_2$ with inhibitory neural activity being a proportional larger driver for CBF than $CMRO_2$ compared excitatory neural activity.
- This cause the seen reduction in coupling factor n from the primary to post-stimulus response, explained above in figure 2.14.

3
Method

3.1 Digital twin project

The aim of the digital twin project is to develop user adapted models, digital twins, which can simulate physiology and help answer questions such as: If I do a thirty-minute work after a meal containing this and that what will happen to my glucose level? Or: If I continue with this certain lifestyle what are my chances of developing atherosclerosis when I'm fifty? These types of questions would be answered by simulating a user's digital twin doing these activities or living that particularly lifestyle. For this to be possible there needs to be an interface which enables users to simulate models and pose question such as those above. Also, models describing specific organs or tissue need to be able to be combined and simulated together in order to model a large-scale human.
In order to meet these requirements a first version of a digital twin software was developed which encompassed a user and model interface and a new modelling framework, each described below.

3.1.1 New modelling framework

The IQM toolbox, currently used for modelling development in Gunnar Cedersund's group, compiles the model ODE:s, defined in text form, to C-files which can interact with the CVODE package developed by SUNDIALS which numerically integrate the ODE:s Sundials. These C-files are written with a specific format which makes them executable by MATLAB (MEX files). Once the mathematical model, i.e. the differential equations, has been compiled into a MEX-format they cannot be further changed but only simulated. The MEX files are static, they can simulate and integrate their differential equation given a set of initial conditions and parameter values, that's it. What is needed for in the digital twin is to have models that are more dynamically, that can be simulated with other models to perform larger and more complex simulations. To achieve this the models needs to have ways of communicating with each other and inform what variables they simulate and what variables they need some other models to simulate. Also, the idea with digital twin is to have users, that have no experience with mathematical models, to be able to perform simulations and make decision about the type of stimulation to be applied, such as training, eating or sleeping (we call these stimulation *activities*). This puts entirely different demands on the modelling framework, which needs to be able to communicate to the user what things can be simulated and then from choices made by the user put together the relevant models and perform the simulation.

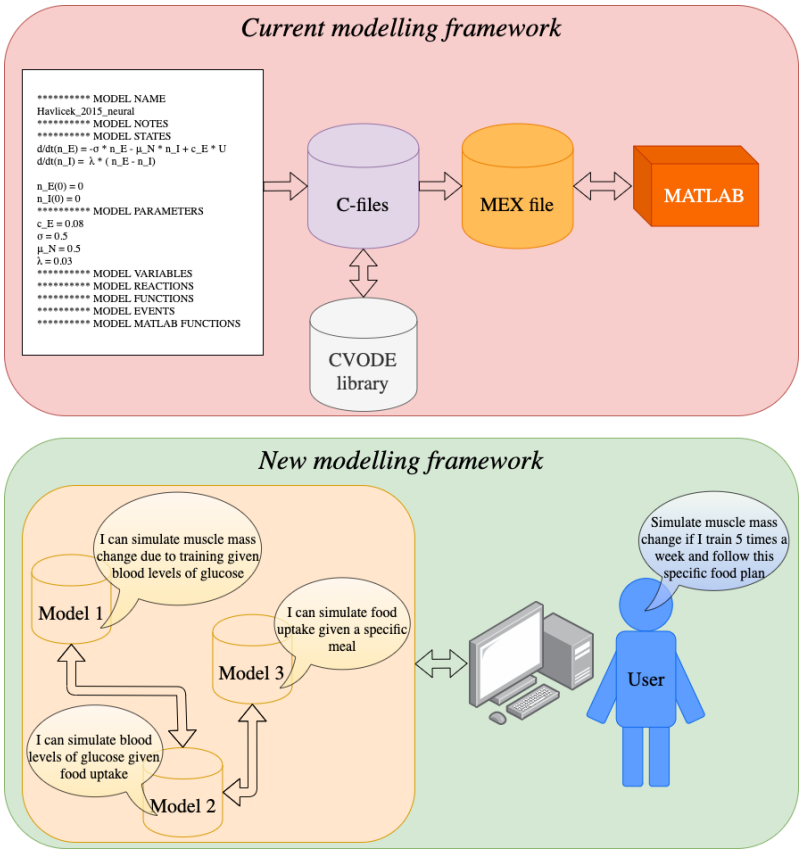


Figure 3.1: The current and new modelling framework. In the old modelling framework, based on the IQM toolbox for MATLAB, a mathematical model, defined as a set of differential equations, was compiled in two steps, first from a pure text file into C-files and then into a executable MATLAB file (MEX file), see above picture. The MEX file could then be run from MATLAB to simulate the mathematical model. This framework was built for researchers who are developing the models and know how to the models works. The digital twin project calls for a completely new way of performing simulations as the central idea is that users, with no previous experience of programming or mathematical models, should be able to make simulations and understand the result. It was therefore decided to build a new framework, see lower picture, where the mathematical models would be more autonomous and able to communicate with each other and also with the user.

As the project started it became clear that the current framework relying on MEX files was to limitating in the way models where handled and a new framework was needed to be built to handle the new demands from the digital twin. The hope was however that the new framework would be compatible with the current way of developing models in simple text from. A schematic picture of the old and new framework is shown in figure 3.1.

3.1.1.1 Model-classes and model-objects

The idea with the new framework was to think about the mathematical models as different classes which here are referred to as *model-classes*. These model-classes

would serve as blueprints from which actual *model-objects* are created which are able to communicate with other model-objects and be simulated together. An object inherits all its properties from the class it is created but is its own entity at the same time, it has its own variables and states which is separate from other objects created from the same class. In this way there is a clear separation between the mathematical model, i.e. the model-class, and the realization of the model, i.e. the model-objects. For example, it may be desirable to treat adipose tissue at different locations in the body separately when developing a large scale human model, indeed the fat cells in the liver have no direct link with fat cells in other parts of the body, except the fact that they are all fat cells. With the model-class-object framework all adipose tissue can be created from the same mathematical model, the same blueprint, of real adipocytes, but still be separate from each other and able to interact with other, different, model-objects, perhaps in the vicinity of their spatial location, see figure 3.2.

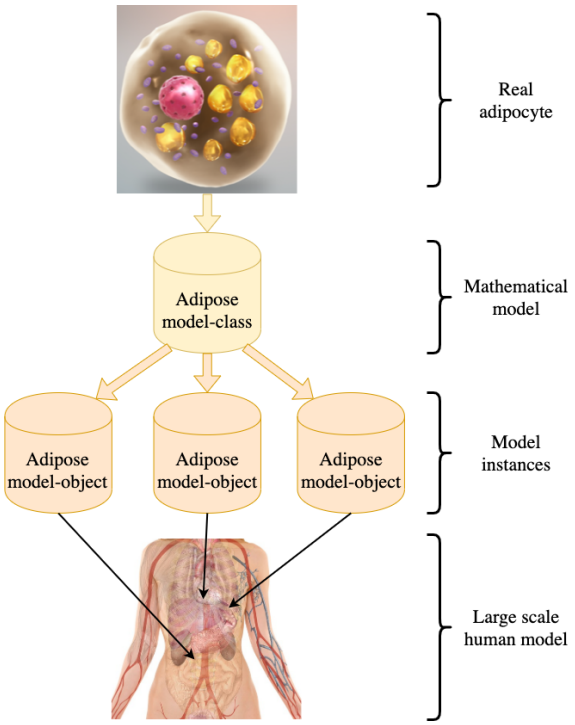


Figure 3.2: The idea with new framework where a real objects, such as adipocytes, are turned into a mathematical model, which in the new framework is represented as a model-class. From this single model of a adipocyte multiple model instances can be created. These model instances, or model-objects, can be located at different spatial location in a large scale human model and interact with different model-objects in order to perform more realistic spatially varying modelling.

This setup allows for a much more scalable modelling solution than the previous framework as there is no limit, except perhaps a computational limit, in how many model-objects you could have to describe tissue at different locations. All model-objects created from the same model-class inherits the same properties from the original mathematical model but are not bound to interact with the same model-

objects. To put it in other words: the model-class provide the rules of the game, but it is the model-objects who are doing the actual playing. Also, thinking about models as objects is a closer abstraction of the real world and allows for the models to be interactive in way that simply was not possible in the old framework.
It was decided to write the new framework in Python as it has advanced tools for integrating C-code with regular Python code. In this way new functionality, such as a user and model interface, could be written using Python but at the same time keeping the C-file structure that had been used for the mathematical models in the IQM toolbox, for example using CVODE integrator to solve the differential equations. Without going into detail in the actual code I will try to explain the new framework through its functionality.

3.1.1.2 Python-C extension modules

There are several different ways C-code can be integrated into Python, but Python-C extension modules has a straightforward way of creating new object-classes, precisely what was needed for the new framework. It was decided that new modelling framework needed to encompass three different types of objects-classes to meet the modelling and interaction demands of the digital twin software. The first type of class is model classes, discussed above, which contains the mathematical equation defining the physiological model. These classes also have attributes and method to allow them to be queried about such things as which time unit is assumed in their differential equations, how many inputs and outputs the model has, what physiological variables are simulated by the model. It also keeps tracks of its internal state-variables, variables that are not public to the rest of the program but are needed for the model to simulate. For each physiological model there will be one model-class, from which multiple model-object can be created, see figure 3.2.
The second type of object-class is a simulation-class from which simulation objects are created. Simulation objects are used to simulate various number of model-objects, created from different model-classes, simultaneously. It needs to check whether or not the model-objects are compatible which each other, for example some models need input from other models in order to simulate, which the simulation-object will check before simulating. To do the actual simulation, which basically means integrating the differential equations defined in the model-objects, it uses the CVODE integrator mentioned above. Having a distinct simulation-class makes it possible to change how the models are simulated without changing the rest of the framework. For example, there are models that uses differential algebraic equations (DAE) which cannot be solved by the CVODE integrator. Support for solving DAE:s can be added in the future simply by defining a new type of simulation-class, perhaps one that uses the IDA integrator, also developed by Sundials Sundials.
The last type of object-class is the activity-class, which enables outside information, perhaps defined by a user, to be sent to the model-objects as they are being simulated. These activities would represent various types of stimulus, for example a meal which would stimulate a model simulating the gastrointestinal tract or perhaps a cardio exercise which would change your oxygen and glucose consumption. The activities are treated as input to the various model-objects which is controlled by the simulation-object.

The different classes and objects are all shown schematically in figure 1.4 from chapter 1.

3.1.2 User and model interface

Parallel to the work focused on in this thesis, which largely revolved around building the new modelling framework discussed above, a group of bachelor students also worked on developing a user and model interface which would enable users to interact with the models and simulate them. The user interface, or *frontend*, would provide a graphical interface to the users, showing them a representation of their digital twin and a simulation panel where simulations request can be created. The simulation request would contain the physiological variables and stimulus that the user wants to simulate and how long the simulation should last. The stimulus could for example be eating, sleeping or training. The model interface, or *backend*, would communicate to the frontend information about the models such as the physiological variables they contain and which activities they can simulate. When receiving simulation requests from the frontend the backend would put together the necessary models and activities in a simulation-object to perform the simulation. The result would then be sent back to the frontend where it would be shown to the user. The digital twin software, with the relation between the frontend, backend and new modelling framework, is shown in figure 3.3.

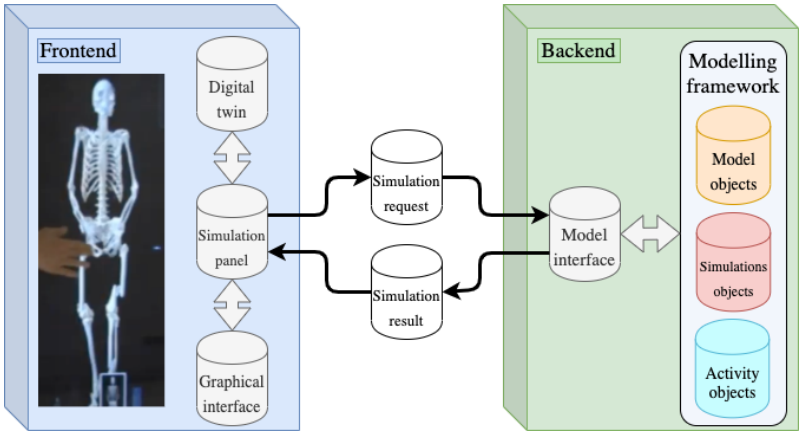


Figure 3.3: The digital twin software, showing the relationship between the frontend, backend and new modelling framework. The frontend holds a graphical interface for the user to see and interact with its digital twin and also make choices about simulation which are requested from the backend. The backend provides the model interface between the frontend and the models. Given a simulation request the backend creates a simulation-object with the appropriate models and activities that are to be simulated. The result is then sent back to the fontend where it is displayed to the user.

The frontend and backend are constantly communicating with each other as the frontend takes in instructions from the user and passes them on to the backend which in turn is providing the frontend with information about the models and possible simulations. By making a clear distinction between the frontend/user interface and

the backend/model interface these two can be potentially located on different servers. There could even be multiple types of user/model interfaces, each developed for a certain type of application.
The frontend was built using the medical image processing and visualization program MeVisLab, see figure 2.5. MeVisLab labs is a module-based program where new modules and functionality can be added on, why it was thought suited for the digital twin interface where the long term goal is to have medical scans of patient which shows organs and tissue which is linked to the different mathematical models. New modules for MeVisLab, written in Python, was developed by the bachelor thesis group. The same group also did the mayor work in developing the backend, which was also written in Python. The work of this thesis focused mainly on the development of the new modelling framework. However, the development of the frontend, backend and new modelling framework was all done in consultation of each other as these all relied on each other.

3.1.2.1 Digital twin

The digital twins themselves are nothing more than the various models with an optimized parameter set to a specific user's test data. Exactly what data is used and how the parameters are optimized was not the focus here. Instead the focus was on having a representation of the digital twin which the user can see and interact with. The interaction with the digital twin revolved around doing simulation and viewing the response that is produced. The representation would be a graphical view of the user showing different organs to which specific models are associated. The user would be able to rotate and zoom into different parts of their digital twin to inspect different organs and models. The models should be able to inform the user which physiological variables and activities they can simulate.

3.2 Neurovascular coupling model

The goal of building a neurovascular coupling model was to see if the claims made by Mullinger et al., 2017 in their 2017 paper could be supported on a quantitative basis, these claims are explained in section 2.3.4. The model would also serve as a test for the newly developed modelling framework as well as the digital twin software, explained above. The developed model builds largely on previous work done by Havlicek, Roebroeck, K. Friston, et al., 2015; Havlicek, Roebroeck, K. J. Friston, et al., 2017 but also ads some novel parts to adapt to the specific data at hand. The idea is to fit the model to the data by using Bayes inference, as explained in section 2.1.2, and by looking at the simulated variables and fitted parameter see if support can be given to Mullinger et al., 2017 about the origin of the post-stimulus response being neural. We will first describe the data.

3.2.1 Mullinger data

The data used in this thesis comes from the previously mentioned study by Mullinger et al., 2017 where they perform a visual cortex experiment, we will briefly walk

through the study here and refer to the original paper for more details.
The data set is collected by letting sixteen volunteers watch a screen displaying a checkerboard to the left eye at the same time as EEG-(electroencephalography), ASL-(arterial spin labeling) and BOLD-data was being collected of the right, contralateral, and left, ipsilateral, visual cortex brain region. ASL is a method for measuring arterial blood flow and the EEG signal allowed for a correlation between the overall brain activity to other neural responses such as those recorded in the BOLD-data. The checkerboard was displayed for 10 seconds and after followed 30 seconds long pause to record post-stimulus data before the test was repeated again. Previous work had seen different post-stimulus response depending on the stimulus being static or flickering, why the test was both conducted with the checkerboard being shown statically and flickering with a frequency of 3 Hz. The contrast of the checkerboard was reduced by 33% for the flickering signal to produce a similar primary response as the static stimulus. The test was repeated 32 times for both the static and flickering checkerboard and the average response was calculated over all subjects, see figure 3.4 showing the contralateral visual cortex ASL- and BOLD-data.

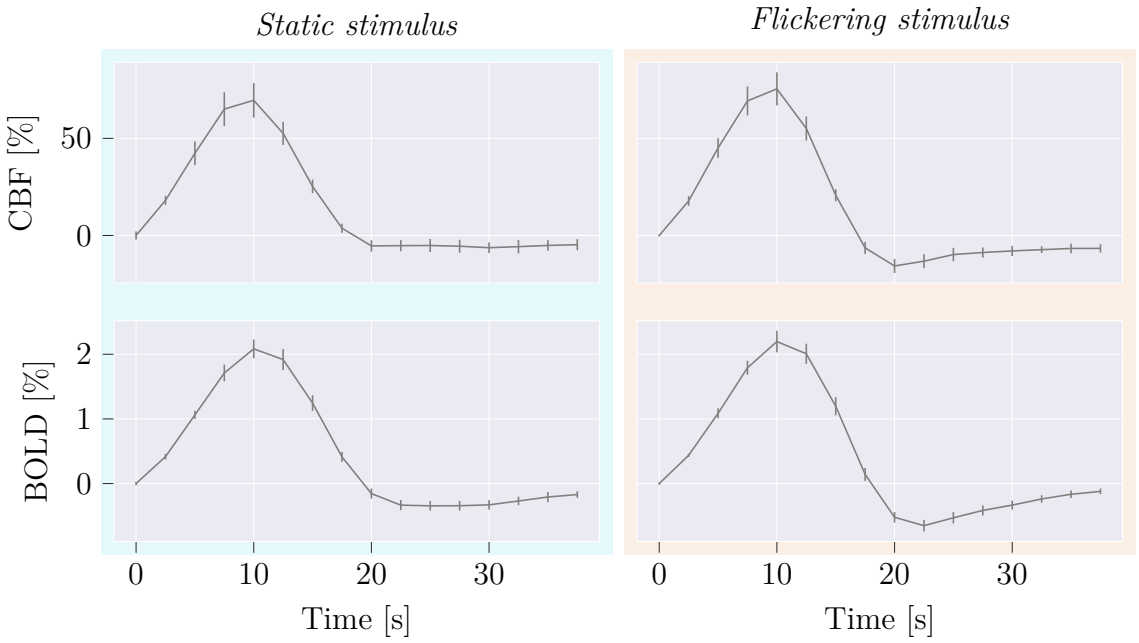


Figure 3.4: The selected data from the Mullinger study (Mullinger et al., 2017) which is used to train the developed model. The visual cortex experiment was performed with the checkerboard both showing statically and flickering with a frequency of 3 Hz. Data from EEG, BOLD and ASL where simultaneously collected for the contralateral (stimulated) and ipsilateral (non-stimulated) visual cortex region, here the focus was on the ASL-, measuring the cerebral blood flow (CBF), and the BOLD-data for the contralateral visual cortex, which is shown in the figure above.

We are focusing on the ASL- and BOLD-data, as these data are more straightforward interpret in terms of the model equations, see below. The intention was to study both the contralateral and ipsilateral visual cortex, as there is a cross-talk between

these regions (Bocci et al., 2014). Unfortunately did time not allow for studying this bilateral interplay between the left and right hemisphere, instead the focus was on only the data from the contralateral brain region (the activated side of the brain) and the study of the connection between the two hemispheres was left as a future project.

3.2.2 Model equation

The neurovascular coupling (NVC) model developed was largely based on the work done by (Havlicek, Roebroeck, K. Friston, et al., 2015; Havlicek, Roebroeck, K. J. Friston, et al., 2017), who developed a model aimed at capturing the coupling between neural activity, oxygen consumption and blood flow. The model was essentially comprised out of three different parts:

- A neural model, describing the connection between stimulus input and excitatory and inhibitory neural activity in the visual cortex.
- A neurovascular model, connecting the neural activity to changes in blood flow, CBF, and oxygen consumption, $CMRO_2$.
- A BOLD model, describing the change in MRI signal due to changes in oxyhemoglobin and deoxyhemoglobin concentrations as a result of changed $CMRO_2$ and CBF.

This clear distinction between the different parts of the model can be exploited in the new modelling framework where independent models for each part can be developed and then later combined when simulated. This way of modelling, dividing complicated models into multiple constitutive parts, is one of key features of the new modelling framework, why the neurovascular coupling model serve as a perfect test for this new approach. The original model of Havlicek, Roebroeck, K. Friston, et al., 2015 was therefore separated into three different models, each describing the different parts mentioned above, an illustration is also shown in figure 1.5 in the introduction. Each model, a neural model, a neurovascular model and a BOLD model, will be described separately along with the different adaptions made, the larger combined model will be referred to as the neurovascular coupling model.
First a small note on notation used. The models describe changes to systems and are therefore not concerned with absolute values of their system variables, instead they model the relative changes from baseline. Capital letters are used to donate the absolute variable values, such as V, and baseline values are denoted with a 0-subscript (or superscript), V_0. Relative quantities are denoted by lowercase letters and lowercase letters with a tilde:

$$v = \frac{\Delta V}{V_0} = \frac{V - V_0}{V_0} \tag{3.1}$$

$$\tilde{v} = \frac{V}{V_0} = v + 1 \tag{3.2}$$

where v is the relative change with respect to baseline and $\tilde{v}$ is the relative amplitude with respect to baseline. Since all the relative quantities are scaled by their baseline

values, and are therefore dimensionless, they can be directly compared with each other.

3.2.2.1 Neural model

The neural model is exactly the same as the one developed by Havlicek, Roebroeck, K. Friston, et al., 2015 with no adaptation. It divides neural activity based on two types of activity, excitatory and inhibitory, and describes changes in these relative to baseline. It is assumed that stimulation, from the visual field, $u(t)$ only affect the excitatory population, n_E, controlled by the parameter c. The excitatory activity in turn activates the inhibitory population n_I through the coupling constant λ, see equations (3.3) and (3.4). The inhibitory activity is further coupled to the excitatory activity through a negative feedback controlled by the parameter μ.

$$\frac{\mathrm{d}n_E}{\mathrm{d}t} = -\sigma \cdot n_E - \mu \cdot n_I + c \cdot u(t) \tag{3.3}$$

$$\frac{\mathrm{d}n_I}{\mathrm{d}t} = \lambda \cdot (n_E - n_I) \tag{3.4}$$

$$\tag{3.5}$$

This is an extremely simplified model of the neurological activation, but the hope is to capture some of the dynamics in neurological activity where research have shown that excitatory and inhibitory neurons exist in a close interplay with each other. It has been observed that a small disturbance in the relative population of excitatory and inhibitory neurons can have large impact on the overall functionality of the neural network (Isaacson and Scanziani, 2011), this motivate equation (3.4) where the term $\lambda(n_E - n_I)$ will cause the relative change in inhibitory activity to follow the relative change in excitatory activity.

3.2.2.2 Neurovascular model

The neurovascular model describe the coupling between neural activity, see equation (3.3) and (3.4), and blood flow and oxygen consumption. The coupling to blood flow is modelled as a two-stage process where the excitatory and inhibitory neural processes creates an activation signal a which in turn modulate the blood flow f:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -\rho \cdot a + \xi_E \cdot n_E + \xi_I \cdot n_I \tag{3.6}$$

$$\frac{\mathrm{d}f}{\mathrm{d}t} = -\chi \cdot f + \phi \cdot a \tag{3.7}$$

The only differences in the model equations compared with those proposed by Havlicek, Roebroeck, K. Friston, et al., 2015 are found in equation (3.6) where the parameter ξ_E is inserted in front of n_E and the term $\xi_I \cdot n_I$ is added. This is done in accordance with the claims made by (Mullinger et al., 2017) which suggested that the coupling between neural activity and blood flow is different for inhibitory and excitatory neural activity. The activation signal a in turn cause a change in the blood flow through equation (3.7). The parameters ρ and χ sets the dynamics for

the coupling and the parameters $\xi_{E/I}$ and ϕ sets the overall strength of the coupling between neural activity and arterial blood flow.
What was missing in the original model by Havlicek, Roebroeck, K. Friston, et al., 2015 is a direct link between neural activity and oxygen consumption. To model this a linear dependence was assumed between the relative change of oxygen consumption, cmrO_2, and neural activity, see equation (3.8). To allow for a small delay between the increased neural activity and increased oxygen consumption we add the state variable r which converge to the oxygen consumption cmrO_2 with time constant δ, see equation (3.9). The idea being that the change in neural activity, and therefore oxygen consumption, will not immediately affect the oxygen extraction from the blood, which is really what is being modelled, but the change has to propagate through the local tissue to the blood supply.

$$\mathrm{cmrO}_2 = \theta_E \cdot n_E + \theta_I \cdot n_I \tag{3.8}$$
$$r = \delta \cdot (\mathrm{cmrO}_2 - r) \tag{3.9}$$

It is the variable r that is of interest and describes the rate of oxygen extraction from the blood, the quantity needed to model the BOLD-response described next.

3.2.2.3 BOLD model

The actual BOLD-response signal, described in section 2.3.3, has been modelled extensively and several different models are available (Richard B. Buxton, 2013). Havlicek, Roebroeck, K. Friston, et al., 2015 uses a detailed model developed by Richard B Buxton, Wong, and Frank, 1998 which takes into account many different physiological aspects. Here a simpler model, the Davis model, has been chosen which was originally developed by Davis et al., 1998 and also used in the paper by Mullinger et al., 2017 from which the data is taken. The Davis model is one of the most common models and is appreciated for its simplicity. And even though many of the physical aspects are left out, detailed studied carried out by Griffeth and Richard B. Buxton, 2011 has shown that the Davis model is surprisingly accurate, and can be made even more accurate by using optimized parameters.
Davis et al., 1998 assumed the following relationship between the decay rate of the T2-signal R_2, the blood volume fraction V and the deoxyhemoglobin concentration [dHb]:

$$R_2 = k \cdot V \cdot [\mathrm{dHb}]^{\beta} \tag{3.10}$$

where k is proportional constant which depends on the field strength of the applied magnetic field and β is a complex parameter which reflects the size of blood vessels causing the decay rate of the MRI signal. For large vessel, radius $> 10\,\mu\mathrm{m}$, it has the value of 1 and for the smallest of capillary vessels it has the value of 2, radius $\approx 2.5\,\mu\mathrm{m}$ (Richard B. Buxton, 2013). In the original model by Davis et al., 1998 a value of 1.5 was proposed to compromise between these two values, in more recent studies a value of 1.3 has often been used (Richard B. Buxton, 2013). In the mentioned study by Griffeth and Richard B. Buxton, 2011 a value 0.9 was suggested. These type of uncertainty in exact parameter value is easily dealt with

when performing a Bayesian inference since you define a prior distribution which reflects the uncertainty in exact parameter value, see section 2.1.2.
Then, by using equation (2.18) from section 2.3.3, the following equation for the relative change of the MRI signal y is retrieved:

$$\begin{aligned} y &= -\Delta R_2 \cdot TE = (R_2^0 - R_2) \cdot TE = k \cdot TE \cdot (V_0 \cdot [\mathrm{dHb}]_0^\beta - V \cdot [\mathrm{dHb}]^\beta) \\ &= k \cdot TE \cdot V_0 \cdot [\mathrm{dHb}]_0^\beta (1 - \frac{V}{V_0} \cdot \Big(\frac{[\mathrm{dHb}]}{[\mathrm{dHb}]_0}\Big)^\beta). \end{aligned} \tag{3.11}$$

Assuming that the majority of the signal originates from the venous department the change in deoxyhemoglobin $[\mathrm{dHb}]/[\mathrm{dHb}]_0$ is equal to the change in extraction fraction E/E_0 which, by using equation (2.15), can be expressed the quotient between CBF and CMRO_2: $\mathrm{E}/\mathrm{E}_0 = \mathrm{CMRO}_2/\mathrm{CBF} \big/ \mathrm{CMRO}_2^0/\mathrm{CBF}_0 = \tilde{r}/\tilde{f}$ where $\tilde{r}$ is the relative oxygen consumption and $\tilde{f}$ is the relative blood flow with respect to baseline. Putting this into equation (3.11) we arrive at the Davis model:

$$y = M \cdot \Big[1 - \tilde{v} \cdot \Big(\frac{\tilde{r}}{\tilde{f}}\Big)^\beta\Big] \tag{3.12}$$

where the parameter M is the combined term $k \cdot TE \cdot V_0 \cdot [\mathrm{dHb}]_0^\beta$ which may vary between brain regions and different subjects. Here M will be treated as any other parameter that needs to be fitted to the data.
What is often done when applying the Davis model is to assume a fixed relationship between the cerebral blood flow, CBF, and the cerebral blood volume fraction, CBV, as first described by GRUBB et al., 1974, see equation (3.13). The value of the parameter α was originally estimated to be 0.38 but this value has come into question in more recent studies (Mark, Fisher, and Pike, 2011), here α will be treated as a parameter to be optimized.

$$\mathrm{CBV} = \mathrm{CBF}^\alpha \tag{3.13}$$

The above relationship between blood flow and volume is a static and thus neglects viscoelastic affects, such as those describe in the 'Balloon model' by Richard B Buxton, Miller, et al., 1998 where blood flow and blood volume is partly decoupled. This is one of the concerns raised by Mullinger et al., 2017 as they analyze their data using the Davis model. To overcome this shortcoming of the Davis model it was decided to model the venous volume as a mass balance between inflow, f, and outflow, f_{out}:

$$\frac{\mathrm{d}v}{\mathrm{d}t} = \frac{f - f_{out}}{t_0}. \tag{3.14}$$

as was done by Havlicek, Roebroeck, K. Friston, et al., 2015. Here t_0 is the mean transit time for the blood to pass through the blood compartment. The arterial blood inflow, f, is modelled by the neurovascular model, see equation (3.7). What is missing is an equation for f_{out}, which would have been described by $\tilde{f}_{\mathrm{out}} = \tilde{v}^{1/\alpha}$ if assuming the static assumption from equation (3.13). Instead, by adding a viscoelastic term $\tau \frac{\mathrm{d}v}{\mathrm{d}t}$ to the right hand side and using equation (3.14) we can include

viscoelastic effects, as explained by Richard B Buxton, Uludağ, et al., 2004, see equation (3.15).

$$\tilde{f}_{out} = \tilde{v}^{1/\alpha} + \tau \frac{\mathrm{d}v}{\mathrm{d}t} \overset{(3.14)}{=} \frac{1}{t_0 + \tau}(t_0 \cdot \tilde{v}^{1/\alpha} + \tau \cdot \tilde{f}) \tag{3.15}$$

Where τ is parameter setting the strength of the viscoelastic effect and which will be optimized.
By using equation (3.14) and (3.15), rather than equation (3.13), to model to the blood volume fraction v the viscoelastic effects are incorporated into to the Davis model, which, to the writer's knowledge, is the first time this has been done. Thus, by creating a rigorous mathematical model it was possible to handle one of the concerns raised by Mullinger et al., 2017 as they draw their conclusions, namely including viscoelastic effects.
Since the data used in this study is an average over all the subjects performing the visual cortex experiment, as explained above, the models and parameters should also be seen as modelling the average subject. It is most likely that the specific parameter values differ between subjects, one example is the parameter M which depends on the baseline values of deoxyhemoglobin. It is therefore important to note that this is a model of the 'average' neurovascular coupling.
All the parameters in the above models have the dimension of inverse time.

3.2.3 Bayesian inference

A common approach when building a model and fitting to data is simply to find the parameters that makes the model mimic the data as close as possible, according to some measure. In Bayesian inference the methodology is somewhat different. Instead of trying to find the parameters that most closely models the data, they are thought of as stochastic variables, whose probability distribution are to be found. The problem then boils down to computing the posterior probability for the joint distribution over all parameters: $P(\vec{\mathbf{p}} \mid \text{Data}$, see section 2.1.2. Here a Markov Chain Monte Carlo (MCMC) sampling method was used, where walkers randomly explore the parameter space and by cleverly updating their positions the random walk will mimic the posterior probability, which is of interest, see below.

3.2.3.1 Prior distribution

The first thing that one needs to do when performing Bayesian inference is choosing a prior distribution $P(\vec{\mathbf{p}})$ for the parameters $\vec{\mathbf{p}}$, see equation (2.4). Here it is assumed that the parameters are independent such as described by equation (2.5). It is also assumed that all parameters are described by a normal distribution or a log-normal distribution if there is a positivity requirement on the parameter, such as time scales or coupling parameters. Superscripts N/L will be used whenever it is needed to make a distinction between these two types. For both types of parameters, choosing a prior is then equivalent to choosing a mean, μ_i, and variance σ_i^2, for their

respective normal distribution:

$$p_i^N \sim \mathcal{N}(\mu_i, \sigma_i^2) \tag{3.16}$$
$$p_i^L \sim \exp[\mathcal{N}(\mu_i, \sigma_i^2)]. \tag{3.17}$$

The exact choice of prior, such as the type of distribution and distribution parameters, is somewhat arbitrary, mainly because the parameters in a model will never be measured directly but instead inferred from other measurement. Here, reasonable values are chosen based on previous work. For example, the parameter M depends on many factors such as magnetic field strength and deoxyhemoglobin baseline values, Mullinger et al., 2017 accounts for this in their study and assume values of M between 0.06 and 0.42 (Uludağ et al., 2004; Gauthier and Hoge, 2013). These values are used when choosing a prior for the parameter M such that they represent the 2.5 and 97.5 percentile respectively. Most of the parameters in the neural and neurovascular model had priors chosen based on values reported in Havlicek, Roebroeck, K. Friston, et al., 2015; Havlicek, Roebroeck, K. J. Friston, et al., 2017. The optimized values reported by Griffeths (Griffeth and Richard B. Buxton, 2011) was used to decide the priors for the parameters α and β. Some parameters simply did not have a clear range of values beforehand, such as $\xi_{E/I}$ and $\theta_{E/I}$, describing excitatory and inhibitory effect on arterial blood flow and oxygen consumption respectively. For these parameters the prior mean and variance was simply chosen to be zero and one respectively. The choice of zero mean was done to make minimum assumption about the relative effect that the different neural activity has blood flow and oxygen consumption, which is one of the things that is hoped to be studied with developed model. A variance of one seemed reasonable since all equations are normalized by their baseline values. The prior distribution is shown in figure 3.5 and the exact value for the distribution parameters μ_i and σ_i^2 can be found in table A.1 in the appendix.

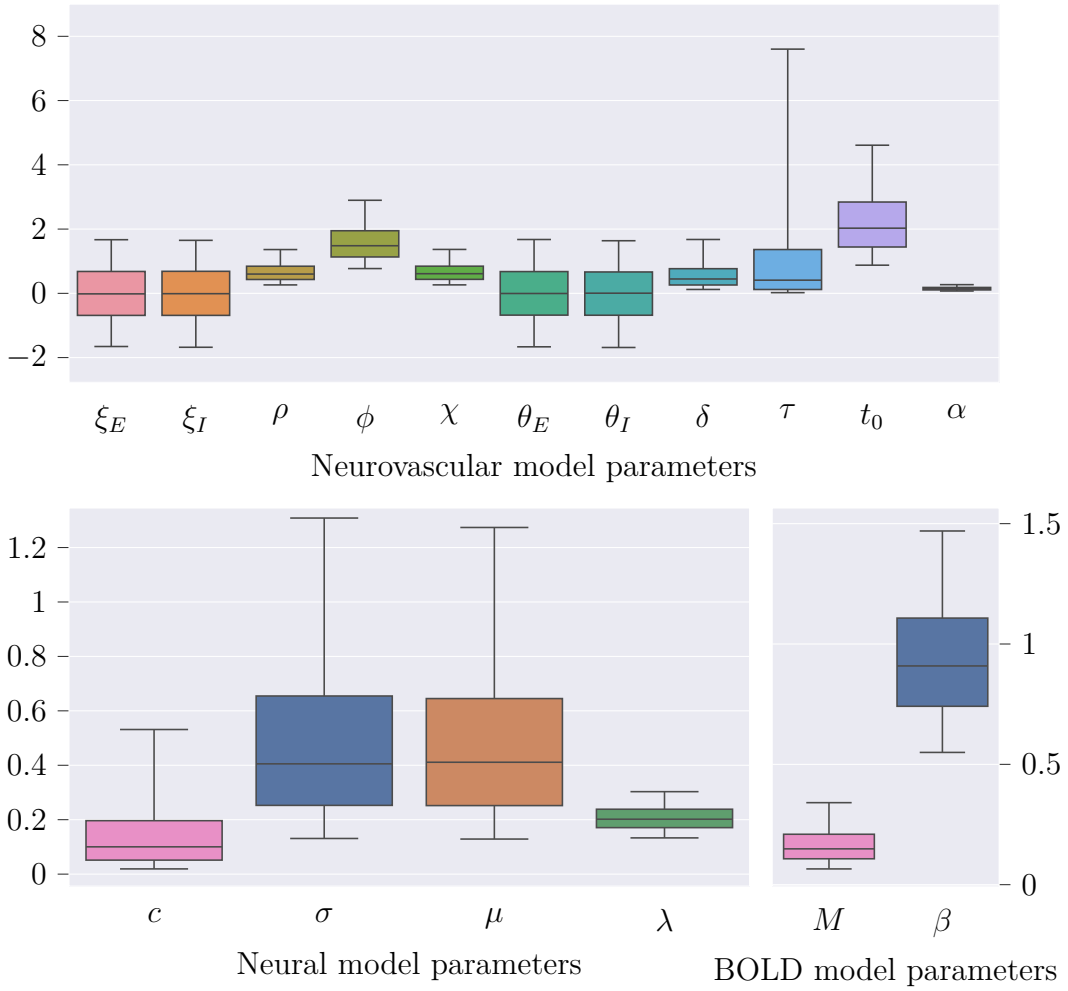


Figure 3.5: The prior distribution shown for each model, the parameters follows either a normal or log-normal distribution, see table A.1 in the appendix. The boxes include 50% of total probability and the whiskers includes 90%, the line in each box shows the median. The priors were chosen according to previously reported values and when no such values existed, they were simply chosen to be as non-informative as possible.

It can seem unscientific to simply choose a prior probability distribution, but this is not something that can be redeemed, other optimization algorithms may not to talk about prior probabilities and simply optimize the model parameters to fit the data as good as possible, perhaps with some hard bounds on the parameters space. But this is the same as having a uniform prior, with equal probability within the bounds. Not regarding a prior probability will not make the problem disappear but simply make you unaware of it, it is therefore better to make an explicit choice for a prior and be aware that the result you get are dependent on that choice.

3.2.3.2 Likelihood-function

After a prior had been chosen the next step was to identify the likelihood function, see equation (2.4). The data is shown in figure 3.4 where the cerebral blood flow and

BOLD-signal is shown for the two types of stimulus, static and flickering, giving a total of four different time series with fifteen time points each (the first time points in figure 3.4 is not modelled as it is the initial value and taken to be zero for all-time series). In the neurovascular coupling model the relative change in blood flow is modelled by the variable f from equation (3.7), the BOLD-signal is modelled by the variable y from equation (3.12) and the stimulus is represented by the function $u(t)$ from equation (3.3). The data was modelled by simulating the neurovascular coupling model twice, one time for static stimulus, S, and one time for flickering stimulus, F, and storing the values for $f_{\mathrm{S/F}}(t)$ and $y_{\mathrm{S/F}}(t)$ at the data time points t_i. This was thus the computational most costly part, all differential equation defined in the neurovascular coupling model needed to be solved twice each time the likelihood function was to be computed for a new set of parameters.
The individual data points was assumed to be independent and normal distributed such that the likelihood-function $P(\text{Data} \mid \vec{\mathbf{p}}$ could be written as product of probabilities for individual data points, see equation (2.12) from section 2.1.2. It was however the logarithm of the likelihood-function that was of interest which is given in equation (3.18).

$$\begin{aligned}\log[P(\text{Data} \mid \vec{\mathbf{p}})] = \sum_{t_i} \left(\frac{f_{\mathrm{S}}^{\mathrm{D}}(t_i) - f_{\mathrm{S}}(t_i)}{\sigma_{f,\mathrm{S}}^{\mathrm{D}}(t_i)/\sqrt{16}} \right)^2 + \left(\frac{y_{\mathrm{S}}^{\mathrm{D}}(t_i) - y_{\mathrm{S}}(t_i)}{\sigma_{y,\mathrm{S}}^{\mathrm{D}}(t_i)/\sqrt{16}} \right)^2 + \\ + \sum_{t_i} \left(\frac{f_{\mathrm{F}}^{\mathrm{D}}(t_i) - f_{\mathrm{F}}(t_i)}{\sigma_{f,\mathrm{F}}^{\mathrm{D}}(t_i)/\sqrt{16}} \right)^2 + \left(\frac{y_{\mathrm{F}}^{\mathrm{D}}(t_i) - y_{\mathrm{F}}(t_i)}{\sigma_{y,\mathrm{F}}^{\mathrm{D}}(t_i)/\sqrt{16}} \right)^2 + \text{constant terms}\end{aligned} \tag{3.18}$$

Here the first sum is for the static stimulus and the second sum for the flickering stimulus. The time series $f_{\mathrm{S/F}}^{\mathrm{D}}(t_i)$ and $y_{\mathrm{S/F}}^{\mathrm{D}}(t_i)$ are the data time series shown in figure 3.4 and $\sigma_{f,\mathrm{S/F}}^{\mathrm{D}}(t_i)$ and $\sigma_{y,\mathrm{S/F}}^{\mathrm{D}}(t_i)$ are their respective standard deviation, also shown in figure 3.4. The constant terms are all normalization factors which can be ignored when performing an MCMC sampling, see below. Notice the factor $\sqrt{16}$, where 16 is the number of subjects in the study. By adding this factor the standard deviation, SD, is turned into the standard error of the mean, SEM, see equation (2.11). Remember that it is the mean neurovascular coupling that is being modelled why the SEM should be used and not the SD.
To test the model the log-likelihood-function from equation (3.18) was first maximized by using a basin-hopping algorithm found in the Python-package SciPy. The basin-hopping algorithm takes in a number of arguments, but only a few were used which are shown in table 3.1, for more information about the algorithm see the support site (SciPy, n.d.).

Table 3.1: The options used through the project for the basin-hopping algorithm found in the SciPy-package.

Option	Choice
Local minimizer method	Nelder-Mead
Initial step-size	0.1
Temperature	1
Number of iteration	100

One thing that quickly became apparent was that the model would not be able to satisfactory describe both static and flickering stimulus with the same parameters, this was observed simply by visual inspection of the data and the optimized model predictions computed from the basin-hopping algorithm. It was first thought that allowing for two different values of the parameter c, controlling the input signal strength in equation (3.3), for the static and flickering stimulus would redeem this problem. This was believed to be appropriate since the checkerboard's contrast for the static and flickering input was different, see section 3.2.1. Denoting the two parameters $c^{\mathrm{S/F}}$, the data S/F-Data and all other parameters $\vec{\mathbf{p}}_{\mathrm{O}}$, the log-likelihood function now looked like:

$$\log[P(\text{Data} \mid \vec{\mathbf{p}})] = \log[P(\text{S-Data} \mid c^{\mathrm{S}}, \vec{\mathbf{p}}_{\mathrm{O}})] + \log[P(\text{F-Data} \mid c^{\mathrm{F}}, \vec{\mathbf{p}}_{\mathrm{O}})]. \tag{3.19}$$

This however, also turned out to be unsatisfactory as there was still large difference between the model and data, even though smaller than before. It was therefore accepted that our simple neurological model could not fully describe the two different stimulus signals with one parameter set, perhaps because the parameters are modulated in some way by the different stimulus. To handle this obstacle, it was decided to allow all the neural parameters, c, σ, μ and λ, to have different values for the static and flickering stimulus. This would turn out to be satisfactory and the final log-likelihood function thus looked accordingly:

$$\begin{aligned}&\log[P(\text{Data} \mid \vec{\mathbf{p}})] = \\ &\log[P(S\text{-Data} \mid c^{\mathrm{S}}, \sigma^{\mathrm{S}}, \mu^{\mathrm{S}}, \lambda^{\mathrm{S}}, \vec{\mathbf{p}}_{\mathrm{O}})] + \log[P(F\text{-Data} \mid c^{\mathrm{F}}, \sigma^{\mathrm{F}}, \mu^{\mathrm{F}}, \lambda^{\mathrm{F}}, \vec{\mathbf{p}}_{\mathrm{O}})].\end{aligned} \tag{3.20}$$

3.2.3.3 Markov chain Monte Carlo sampling

After that the prior distribution and likelihood function had been defined the posterior probability could be computed, see equation (2.4). To do this a Metropolis-Hasting, MH, algorithm was used with the help of the emcee-package for Python (Goodman and Weare, 2010). The MH sampling algorithm is special type of MCMC sampler which samples a function, any function, by performing many random walks where the new position of each random walker is updated in such a way that the cumulative positions visited, after many iterations, resembles the function wanted. It is a method that comes handy when you want to sample a high dimension functions where simply computing the function at enough positions is no time efficient. For example sampling a function in 10-dimensional space with 1000 steps in each dimension would call for 1000^{10} computation, a ridiculous large number of computation.
The simple version of MH-sampling algorithm goes something like this:

1. Initialize the walkers at some position in parameter space and calculate the function F at the current position.
2. Suggest a new position, $x_{suggest}$, for the random walkers.
3. Calculate the function, F, at the suggested position.

4. Update the position, x_{new}, of the walkers in the following manner: If $\frac{F(x_{suggest})}{F(x_{old})} < \gamma$ then $x_{new} = x_{old}$ otherwise $x_{new} = x_{suggest}$, where γ is random variable with a uniform distribution between zero and one.

5. Save the current position of the walkers and iterate a fix number of times by going back to step 2.

Often, is not the actual function that MH-algorithm deals with, but the logarithm of the function, this is to avoid computational error when working with very small numbers, which is usually the case if F describes a probability distribution. The update step number 4 above is then re-written as $\frac{F(x_{suggest})}{F(x_{old})} < \gamma \Leftrightarrow \log(F(x_{suggest})) - \log(F(x_{old})) < \log(\gamma)$, here it is easy to understand why any constant term in $\log(F)$ can be ignored, it will be cancelled when comparing function values. Also, when applying the MH algorithm, as described above, you can change how the sampling behaves and how fast the it converges onto the actual function by changing how new positions, the second step, are suggested. For this the standard class Stretch-Move was used with the stretch parameter a set to two, the Stretch-Move class suggest a new position for one walker by using the relative distance between two others, randomly chosen, walkers. It will not be discussed in further details, instead for more information see the article by Foreman-Mackey et al., 2012 which explains the Stretch-Move in simple terms. Also, a common approach is to do a 'burn-in' phase before the actual sampling of the function. In the burn-in phase the walkers are moving around for a certain number of iterations to settle into the function, their then updated position is used as the initial position in step 1 above.
An issue that might arise when sampling a function is that the walkers get stuck in local maximums, especially if the function is multi-modal with large 'valleys' between each extremum. This is more likely to be the case with a large parameter space, which is the case here with twenty-one parameters, two sets of four parameters for the neural model (for static and flickering stimulus), eleven parameters for the neurovascular model and two parameters for the BOLD model, see figure 3.5. If the walkers can't move from extrema to extrema you get an incorrect sampling of the target function, and thus may draw wrong conclusions. There are ways of handling this, as explained here Foreman-Mackey et al., 2012, but unfortunately time did not allow to extensively try out all different possibilities. Instead, the simplest solution of only sampling the posterior probability around its global maximum was chosen. The global maximum was found by again using the basin-hopping algorithm found in the SciPy-package by maximizing the log-posterior from equation (2.13). This is of course a big limitation as only a part of the parameter space will be sampled, but hopefully the most important part if indeed the global maximum has been found. To ensure this the basin-hoping algorithm was re-run several times with different starting points.
Once the location for the global maximum had been identified the walkers was randomly initialized around this point, with the variance from the prior distributions serving as measure of the spread for each parameter, see figure 3.5. Two-hundred walkers was used with one-hundred burn-in steps and one-thousand sampling iteration, given a total of two-hundred-thousand sampling points. The different settings used for the MH-sampling is shown in table 3.2.

Table 3.2: The settings used for the MH-sampling algorithm.

Option	Choice
Position update	Stretch-Move (scaling parameter $a = 2$)
Initial position	Global maximum found with the basin-hopping algorithm
Number of walkers	200
Number of burn-in steps	100
Number of sampling iteration	1000

When sampling, and especially when doing large samples, you will get samples with very low probability, outliers is another name. These outliers are not of interest why a threshold value on the posterior probability was set to only include samples high probability. There are multiple ways this threshold can be set, here, the fact that the logarithmic of the posterior is a chi-2 distributed variable, apart from a factor of -2, see equation 2.13. The degree of freedom for the chi-2 distribution is equal to the number of fitted data points plus the number of parameters, however, since the data had been fitted by the adjusting the parameters these are normally not counted. Thus, the degree of freedom was taken to be the number of data points which was equal to $4 \times 15 = 60$, four time series with fifteen time points each. A chi-2 test with significance level of 0.01 was done to compute a threshold value for the log-posterior and remove outliers. The remaining samples should then be a correct representation of the posterior-probability around the global maximum.
The sampled posterior distribution of the parameters was then used to see if the statements made by Mullinger et al., 2017, presented in section 2.3.4, could be supported by the neurovascular coupling model.

4

Result and Discussion

4.1 Digital twin software

The new digital twin software was successfully finished in time and was tested by implementing the neurovascular coupling model, see next section. The software comprised of a frontend (user interface), a backend (model interface) and a new modelling framework. The frontend and backend were the main focus of the bachelor thesis group, doing their work in parallel to this master thesis which main focus was the new modelling framework which is first described.

4.1.1 New modelling framework

The goal was to have a new modelling framework which allowed for a more flexible way of developing models, where multiple smaller models could be combined into a larger one. The new framework also needed to be compatible with the current way of developing models in Gunnas Cedersund's group which is based on the IQM toolbox built for MATLAB. The end result is thus a framework in which the actual mathematical models are written down in the same manner as before, in simple text format, see figure 4.1. The only difference is the added sections at the end of the text file which allow to specify how this specific model fits in with other models. There it is specified what time unit is assumed in the model, what inputs and outputs it has to interact with other models and last which variables, or *features*, that are interesting to look at for this specific model. More sections can be added as more functionality is asked for.

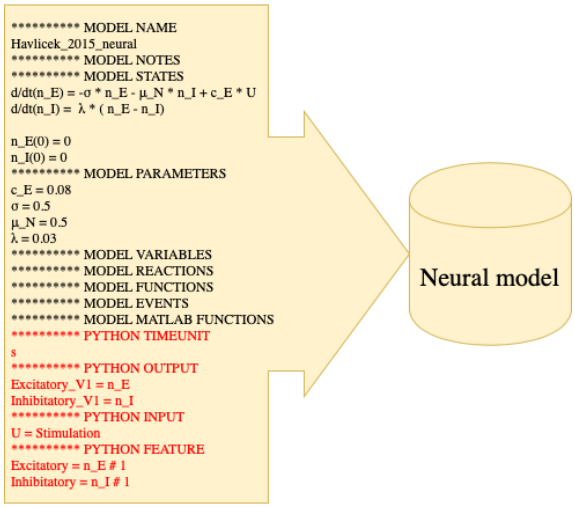


Figure 4.1: The new way of specifying models in the new modelling framework, which is built on top of the current modelling framework used in Gunnar Cedersund's group. The format of the text file specifying the differential equation looks very similar as before but new sections, marked with red above, is added to make the text file able to be compiled into a Python-class, illustrated above, instead of into a MEX-file as in the old framework, see figure 3.1.

The above figure tries to illustrate that the text file are now being compiled into a Python-class, a model-class, from which objects, model-objects can be created, see figure 3.2 from the method section. This is an entirely new way of working with models compared with the old framework were the text files were compiled into MEX-files, see figure 3.1. The actual work revolved around how to do the actual compiling of the text-file to a Python-class, not so easy to show in a figure, however.

4.1.2 Digital twin frontend and backend

The digital twin frontend, which allows users to see and interact with the models, is shown in figure 4.2. There is a representation of the digital twin to the right in which the user can zoom into and rotate. If a specific organ is clicked on, a list with all the relevant features that can be simulated for that organ is shown, this could be the glucose or insulin level of the liver for example. To the left is a panel in which choices can be made about the type of simulation that the users wish to perform on their digital twins, there are three types of choices that can currently be made: How long the simulations should run for and what features and activities should be simulated. The simulation can be repeated multiple times to perhaps simulated a regular day for an entire month. The features that are to be simulated must be connected to the activities that the users choose. For example, if the user wishes to simulate a meal (the activity) and see how the glucose in the blood (the feature) changes there must be a model which simulates the glucose of the blood and also depend, directly or indirectly, on the fact that a meal is being consumed. These types question is exactly what is handled by the backend which communicate both with the models and the frontend, an overview is shown in figure 3.3 from the

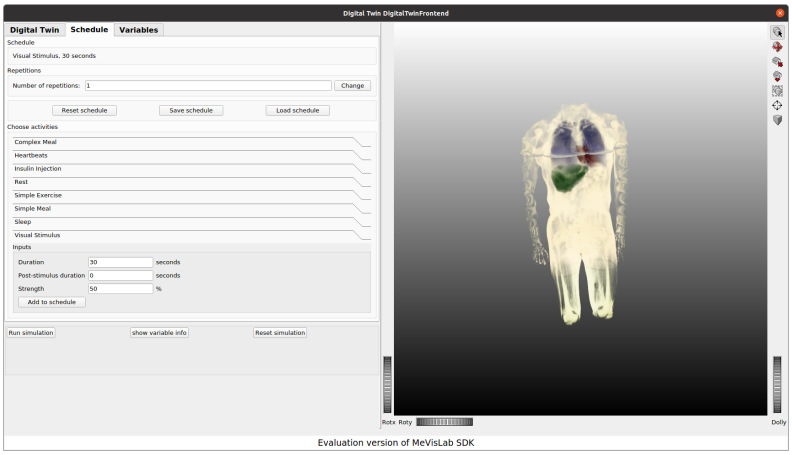


Figure 4.2: The digital twin user interface, or frontend, allowing users to see a representation of their digital twin, to the right, and a simulation panel, to the left, where they can make decision about different types of simulations they want to perform on their digital twin. The different activities showing, such as 'Simple Meal', 'Simple Exercise' etc., are not implemented but are there to show how the interface is supposed to work. What is implemented is the 'Visual Stimulus'-activity which is simulated in the neurovascular coupling model developed in this thesis.

method section.

So far the activities showing in figure 4.2 are not all implemented, but they are there to show how the digital twin software is planned to work. However, the visual stimulus activity is implemented by the use of the neurovascular coupling model explained next.

4.2 Neurovascular coupling model

The development of the NVC model had two main objective, one was to test the digital twin software and one was to test the claims made by Mullinger et al., 2017. The building of the model was successful and so was the testing of the digital twin software, which could simulate the model and make decision about the simulation. Thus, the NVC model completed one of its objective, for the claims made by Mullinger et al., 2017 the model could partly support the claims but the validity of the model can be questioned, this is discussed below.

4.2.1 Posterior distribution

After the global maximum had been identified by basin-hoping algorithm the posterior distribution for the parameters in the NVC model could be sampled around this point. For this the Metropolis-Hasting sampling method was used and the resulting marginalized posterior distribution for each parameter can be seen in figure 4.3.

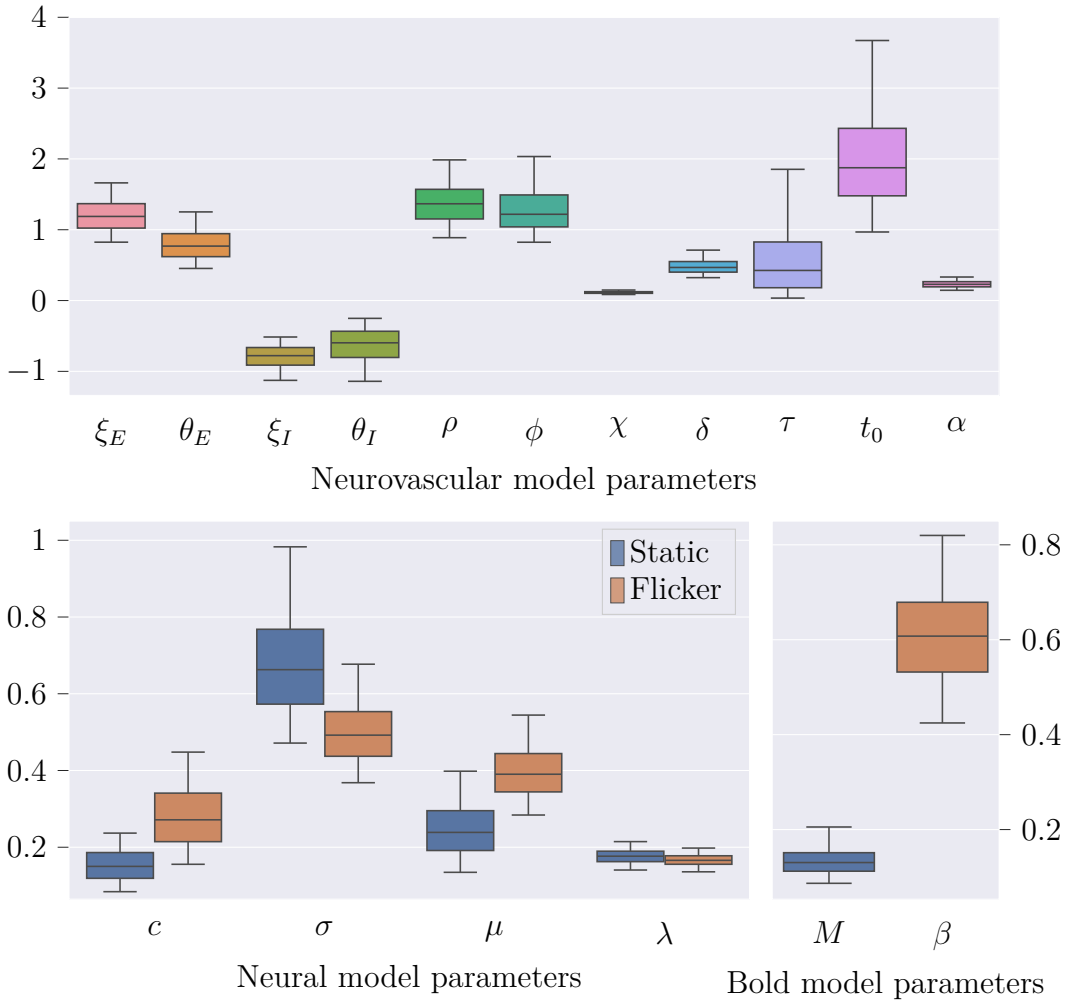


Figure 4.3: The marginalized posterior distribution sampled by the Metropolis-Hasting sampling method around the global maximum identified by the basin-hoping algorithm. The boxes contain 50% of the total probability and the wishers mark the 5 and 95 percentiles, the line in each box marks the median. The parameters are divided according to which model they belong to, the parameters for the neural model are further divided for the different types of input.

First note the distinction between the parameters $\xi_{E/I}$ and $\theta_{E/I}$ where the excitatory parameters, E, have a positive distribution and the inhibitory parameters, I, have a negative distribution. The parameters $\xi_{E/I}$ and $\theta_{E/I}$ control neural activities impact on blood flow and oxygen consumption respectively, see equations (3.7) and (3.8). Next look at parameters for the neural model, which were allowed to be different for static and flickering stimulus. First, we see that the parameter c is about double as large for the flickering stimulus as it is for the static stimulus. This make sense since the flickering stimulus is off half time compared to the static stimulus, which can be compensated by doubling the parameter c. We can also see a clear reduction of the parameter σ and an increase of the parameter μ going from static to the flickering stimulus. The parameter σ controls the overall dynamics for the excitatory neural activity, see equation (3.3), which can be viewed as low-pass filter with the cut-off

frequency set by σ. It thus seems as the model compensate for the fast dynamics of the flickering stimulus by reducing σ and thus blocking high frequency input. The parameter μ controls the negative feedback from inhibitory neural activity to excitatory neural activity, see equation (3.3). The fact that it is increased for flickering stimulus is perhaps not surprising, Mullinger et al., 2017 discussed the possibility of there being an active increase in inhibition as result of stimulus cessation, which is happening frequently in the flickering stimulus. Since the NVC model doesn't explicitly include any such effect it would seem reasonable that it would compensate by instead increasing the negative feedback μ. The possibility of an active inhibition due to stimulus cessation was not investigated further but could be interesting for a future project. We can also see that the parameter λ is similar for both static and flickering stimulus, perhaps good since it seems hard to find any good reason why it should change.

From the posterior distribution shown in figure 4.3 one-thousand samples was drawn randomly to simulate the NVC model and produce model prediction for key variables, shown in figure 4.4.

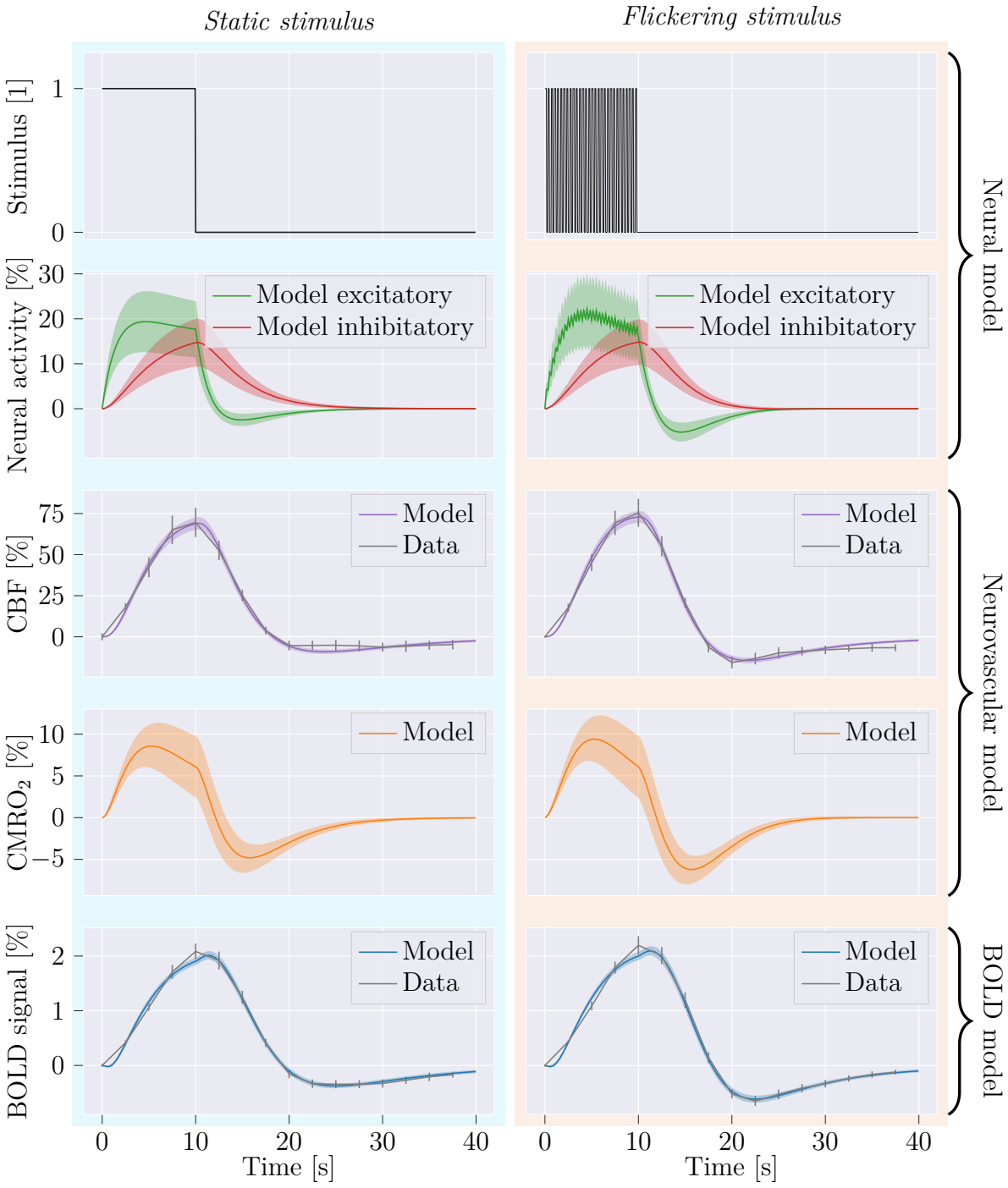


Figure 4.4: The simulation of one-thousand randomly drawn parameters from the posterior distribution shown in figure 4.3, showing the mean and plus minus one standard deviation. Characteristic features for each model are selected together with the data used to train the model, both for the static and flickering input stimulus. The topmost row shows the stimulus which is an input to the neural model and then propagates to the other models to produce the simulated result. The division of a larger model into smaller constitutive parts lies at the heart of the new modelling framework and produce a new way of developing more complex models.

The above figure shows shows key variables for the three different models, a neural model, a neurovascular mode and a BOLD-response model, comprising the full neurovascular coupling model as described in the method section. The three models are distinct from each other but at the same time dependent on each other. The input signal, first row in figure 4.4, is stimulating the neural model to produce the excitatory and inhibitory neural activities, second row. These are then used to calculate the blood flow and oxygen consumption, third and fourth row, in the neurovascular model which in turn are used by the BOLD model to calculate the resulting BOLD signal, fifth row. This it the whole idea of the new modelling framework, to be able to develop larger complex models from smaller and more simplistic constitutive parts. This opens up a whole new world of possibilities, for example, any of the smaller models shown in figure 4.4 could in turn be broken down in to multiple models or exchanged for another model, without having to interfere with the larger model as a whole. This modularity is believed to be key when going forward with the digital twin project

4.2.2 The Mullinger claims

One of the main task for the NVC model was to see if quantitative support could be given to the qualitative claims made by Mullinger et al., 2017, see section 2.3.4. Looking closer at the result shown in figure 4.4, we see that in the primary response where the input stimulus is present, 0 s–10 s, both the excitatory and inhibitory neural activities are positive and in the post-stimulus response, 10 s–40 s, the excitatory activity drops below baseline while the inhibitory activity remains elevated. Thus, our model predicts a clear change in neural activity for the primary and post-stimulus response which supports the claims made by Mullinger et al., 2017, see figure 2.14, where they argue that the cause for the post-stimulus response is due to a change in relative neural activity.

One important motivation for Mullinger et al., 2017 conclusions is the observed changed in coupling factor $n = \Delta_r \text{CBF}/\Delta_r \text{CMRO}_2$ which decreases going from the primary to the post-stimulus response. Mullinger et al., 2017 suggested that this could be caused by inhibitory neural activity having a proportional larger impact on CBF than on CMRO_2 when compared to excitatory neural activity, see figure 2.14. Unfortunately could the coupling factor for the simulations in figure 4.4 not be determined satisfactory for either the primary or post-stimulus response, mainly because of the large uncertainty in oxygen consumption. However, believing Mullinger et al., 2017 that the coupling factor is reduced in the post-stimulus response, the NVC model can be used to give an alternative explanation how this might happen due to changed neural activity.

The main assumption made by Mullinger et al., 2017 was that both excitatory and inhibitory neural activity are positive drivers for CBF and CMRO_2. This is in direct conflict with what was found by the NVC model, see figure 4.3, where it clearly shows that the inhibitory parameters ξ_I and θ_I have a negative posterior distribution. Further, by looking at the posterior-probability for the quotient $\theta_{E/I}/\xi_{E/I}$, as in figure 4.5, we can compare excitatory and inhibitory neural activities' relative impact on CBF and CMRO_2.

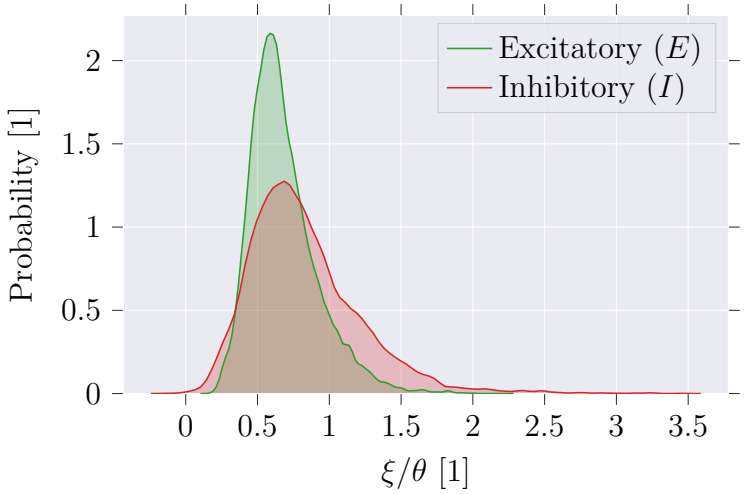


Figure 4.5: The posterior distribution for the quotient $\theta_{E/I}/\xi_{E/I}$ for both excitatory and inhibitory neural activity. The inhibitory distribution θ_I/ξ_I seem to be shifted slightly to higher values compared to the excitatory distribution θ_E/ξ_E. This would mean that inhibitory neural activity have a proportional larger impact on $CMRO_2$ than CBF compared to excitatory neural activity. However, the large overlap also suggest that the different neural activity have a similar proportional impact on $CMRO_2$ and CBF.

Though the overlap of the two distribution is large, it seems like the inhibitory neural activity has a slightly larger quotient and thus a relative larger impact on $CMRO_2$ than CBF when compared to excitatory neural activity, which is opposite to what was suggested by Mullinger et al., 2017.
In light of these observation of the parameters $\xi_{E/I}$ and $\theta_{E/I}$ the NVC model would suggest an alternative explanation for the change in coupling factor for the primary and post-stimulus response compared to the one given by Mullinger et al., 2017:

- In the primary response the stimulus drives excitatory neural activity which in turn cause an increased inhibitory neural activity, $\Delta n_E > 0$ and $\Delta n_I > 0$, see figure 4.4.
- Excitatory neural activity is a positive drive for both CBF and $CMRO_2$ while inhibitory activity is negative driver.
- In the primary response excitatory neural activity is the primary drive for both CBF and $CMRO_2$ and will cause an increase in these which is partly offset by the inhibitory neural activity.
- Since inhibitory neural activity is a proportional larger (negative) driver for oxygen consumption the offset is larger for $CMRO_2$, see figure 4.5.
- This causes the coupling factor $n = \Delta_r CBF/\Delta_r CMRO_2$ to assume a value larger than 1.
- In the post-stimulus phase the stimulus is removed which causes the excitatory activity to quickly drop and become negative relative to baseline, the inhibitory

neural activity which is not primary driven by the stimulus remains elevated, see figure 4.4.

- The changed sign of excitatory neural activity cause excitatory and inhibitory neural activity to now drive blood flow and oxygen consumption in the same direction.
- The resulting coupling factor will reflect the two population's relative effect on CBF and $CMRO_2$.
- In any case, because of excitatory and inhibitory neural activities unequal drive for CBF and $CMRO_2$ the result will be a decrease in coupling factor n from the primary response.

This if further explained in figure 4.6.

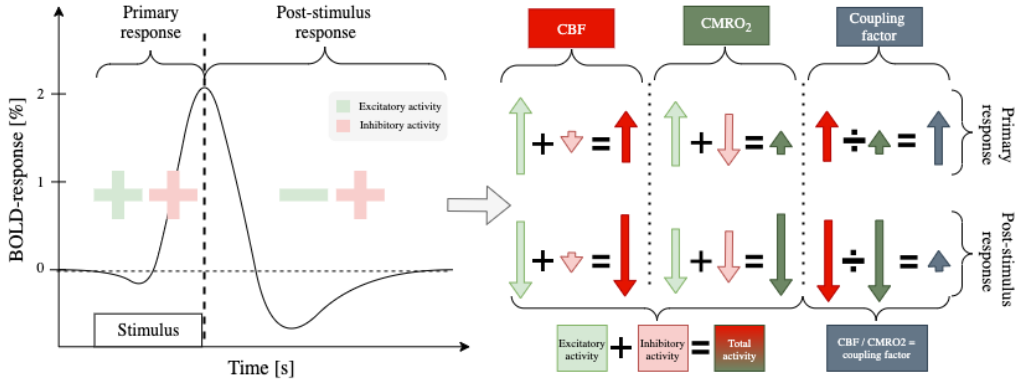


Figure 4.6: An alternative explanation for the different coupling factor n seen by Mullinger et al., 2017 in the primary and post-stimulus BOLD-response. In the primary response CBF and $CMRO_2$ is largely driven by excitatory activity which cause a positive increase of these. The increase is however offset by inhibitory neural activity having a negative impact on CBF and $CMRO_2$, the offset is largest for $CMRO_2$ since inhibitory activity has a proptional larger impact on oxygen consumption than on blood flow. This result is an increased coupling factor $n = \text{CBF}/\text{CMRO}_2$. In the post-stimulus response the removal of stimulus cause the excitatory neural activity to drop below baseline while inhibitory activity, which is not primarily driven by the stimulus, to remain elevated. The result is that the two types of neural activity now drives CBF and $CMRO_2$ in the same direction which cause the coupling factor to decrease from the primary response.

In the above figure the different relative impact on CBF and $CMRO_2$ for excitatory and inhibitory neural activity is grossly exaggerated. The large overlap in figure 4.5 suggest that excitatory and inhibitory neural activity has quite similar relative impact on blood flow and oxygen consumption.

4.2.3 The validity of the NVC model

However, there is serious reason to doubt the developed NVC model. The most obvious is the relative low $\Delta_r CMRO_2$, see figure 4.4. As stated, the coupling factor

was unable to be computed but we can get a sense for its magnitude by looking at the maximum oxygen consumption which is around 8%, see figure 4.4, compare this to the maximum blood flow which is about 70%, this tells us that the coupling factor $n = \Delta_r \text{CBF}/\Delta_r \text{CMRO}_2$ is of the order 10, a value much larger than any experimental determined value, see figure 2.11. We can further see that something is not quite right with the NVC model if we look at the calculated oxygen concentration in tissue, $\text{pO}_2^{\text{tissue}}$, see figure 4.7, which is calculated by using the oxygen diffusion model developed by Richard B. Buxton, 2010 and the simulated blood flow and oxygen consumption for static stimulus.

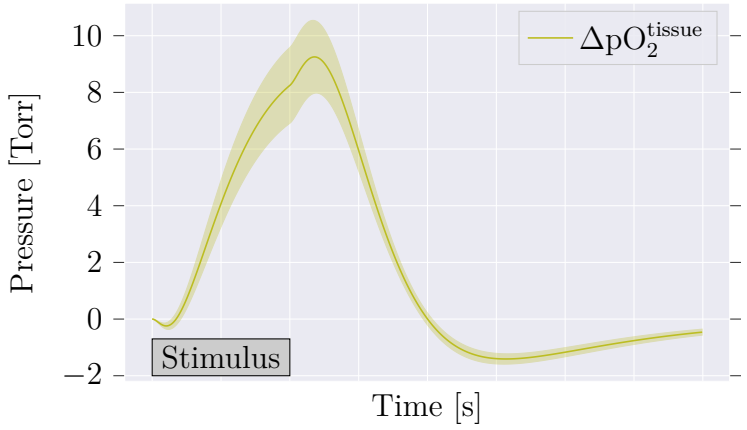


Figure 4.7: Showing the change in oxygen partial pressure in tissue, computed by using the simulate blood flow and oxygen consumption for static stimulus from figure 4.4 together with the the oxygen diffusion model developed by Richard B. Buxton, 2010, baseline value $\text{pO}_{20}^{\text{tissue}} = 25\,\text{Torr}$. We can see large changes in pressure, about 9 Torr, which is not in line with the oxygen diffusion hypothesis, see section 2.3.2, and makes one question the validity of the NVC model.

The above figure show the change of oxygen partial pressure from baseline, which in accordance with Richard B. Buxton, 2010 was set 25 Torr. We see that the NVC model predicts a large rise in $\text{pO}_2^{\text{tissue}}$, about 9 Torr, which corresponds to a relative increase of about 36%. This is not what we expect given the oxygen diffusion hypothesis, see section 2.3.2, which states the uneven change in blood flow and oxygen consumption is to keep the partial pressure of oxygen in tissue close to baseline. This artifact can also be related to the fact that the NVC model predicts a very small change in CMRO_2. The reason the model can predict a relative small change in CMRO_2 and still model the blood flow and BOLD signal reasonably well, see figure 4.4, is because it at the same time predicts a small value for the parameter $\beta = 0.6 \pm 0.1$. Compare this with the value used by Mullinger et al., 2017, $\beta = 1.3$, which is taken from literature (Mark, Fisher, and Pike, 2011).
Using the value $\beta = 1.3$ and inverting the BOLD model from equation (3.12) the oxygen consumption can be computed directly from the BOLD signal, blood flow and blood volume. This leads to more reasonable values shown in figure 4.8, also showing in the figure is the change in $\text{pO}_2^{\text{tissue}}$ if this oxygen consumption would have been used in the oxygen diffusion model.

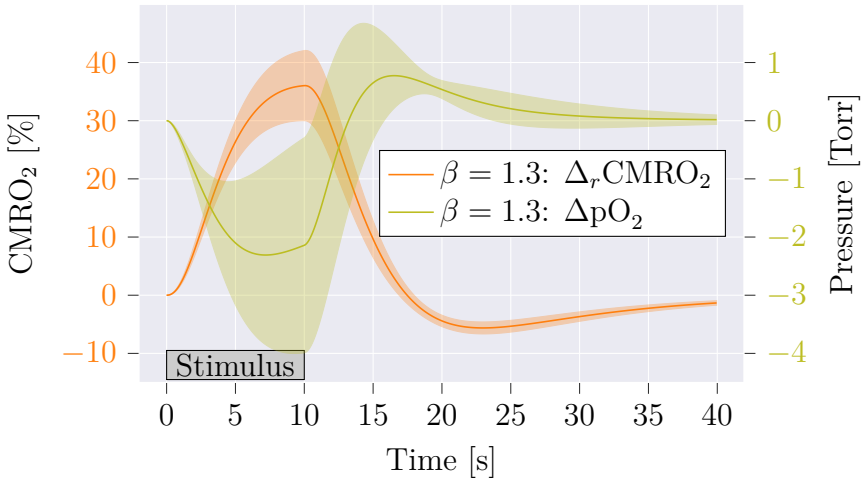


Figure 4.8: Showing the oxygen consumption, orange plot, if the parameter β would have been equal to 1.3 instead of the optimized value shown in figure 4.3 and calculating it directly by inverting the BOLD model in equation (3.12). Also showing is the change in pO_2^{tissue}, olive plot, if this oxygen consumption would have been used. These plots were computed by using the simulated values for the BOLD signal, blood flow and blood volume due to static stimulus 4.4.

We now see a much larger change in oxygen consumption, maximum change about 35%, which means a coupling factor of about two, which is much more reasonable. We can also see that the partial pressure is fairly close to baseline value, only drifts about -3 Torr in the primary response, which is expected given the oxygen diffusion hypothesis.
The reason the parameter β is allowed to assume so small values in the NVC model is because Griffeth and Richard B. Buxton, 2011 suggested to abandon the physical interpretation of β, which reflects the relative size of blood vessel impact on the local change in magnetic field, see section 3.2, and instead use an optimized value to make the Davis model more accurate. Griffeth and Richard B. Buxton, 2011 suggested a value of $\beta = 0.9$. Therefore the prior distribution for β, see figure 3.5, was chosen to include a wide range of values.
Realizing the issue with a too small value of β, a new prior was set with a much narrower band around the value of 1.3. However, this did not redeem the problem, and NVC model still predicted small changes of $CMRO_2$, now by changing the value of the parameter M to very small values, $M < 6\%$, which is smaller than reported values (Mullinger et al., 2017). Therefore it seems that the NVC model has an inherit issue, predicting to small changes in oxygen consumption, perhaps because the simple linear model assumed to relate neural activity to $CMRO_2$ is not adequate, see equation (3.8).
It could also be that the believed global maximum found by the basin-hoping algorithm is not the true maximum, in any case, the NVC model would need further investigation before taking its prediction to serious.

5
Conclusion

This project had two main goals: to develop a new modelling framework adapted to the needs of the digital twin project, and building a neurovascular coupling model which could test the new modelling framework and the claims made by Mullinger et al., 2017. Both of these goals were met to varying degrees.
The new framework for modelling proved to be a powerful tool and will be central going forward with the digital twin project, where large scale human computer models will need to be built from smaller, simpler models. The neurovascular coupling model on the other hand, was mostly successful in testing the new modelling framework and demonstrating its novel capabilities. However, the coupling model could not definitively prove the claims made by Mullinger et al., 2017. This model consistently underestimated oxygen consumption and appeared to push parameters toward nonphysical areas, which draws into question the overall validity of the results.
The work done thus far highlights the importance of understanding underlying physiology to better assess the validity of a neurovascular coupling model. Had the oxygen diffusion hypothesis not been fully understood, it would have been all too easy to erroneously overestimate the developed model's accuracy by only looking at the overall fit to data, see figure 4.4. This thesis can therefore be seen as a comprehensive review of underlying physiological and physical aspects of the described neurovascular coupling, given that immense effort was applied in researching relevant and current knowledge in the field. In conclusion, this study serves as a springboard for future researchers, to launch them in their efforts investigating human computer modelling or the neurovascular coupling phenomenon.

References

Attarwala, H. (2010). "TGN1412: From discovery to disaster". In: *Journal of Young Pharmacists* 2.3, pp. 332–336. ISSN: 09751505. DOI: 10.4103/0975-1483.66810.

Attwell, David, Alastair M. Buchan, et al. (Nov. 2010). *Glial and neuronal control of brain blood flow*. DOI: 10.1038/nature09613.

Attwell, David and Costantino Iadecola (2002). "The neural basis of functional brain imaging signals". In: *Trends in Neurosciences* 25.12, pp. 621–625. ISSN: 01662236. DOI: 10.1016/S0166-2236(02)02264-6.

Attwell, David and Simon B Laughlin (2001). *An Energy Budget for Signaling in the Grey Matter of the Brain*. Tech. rep.

Berg, JM, JL Tymoczko, and L Stryer (2002). "Hemoglobin Transports Oxygen Efficiently by Binding Oxygen Cooperatively". In: *Biochemistry*. 5th. New York: W H Freeman. Chap. 10.2. URL: https://www.ncbi.nlm.nih.gov/books/NBK22596/.

Bocci, Tommaso et al. (Feb. 2014). "Visual callosal connections: Role in visual processing in health and disease". In: *Reviews in the Neurosciences* 25.1, pp. 113–127. ISSN: 03341763. DOI: 10.1515/revneuro-2013-0025.

Bowery, N G and T G Smart (Jan. 2006). "GABA and glycine as neurotransmitters: a brief history". eng. In: *British journal of pharmacology* 147 Suppl.Suppl 1, S109–S119. ISSN: 0007-1188. DOI: 10.1038/sj.bjp.0706443. URL: https://pubmed.ncbi.nlm.nih.gov/16402094%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1760744/.

Brännmark, Cecilia et al. (2013). "Insulin signaling in type 2 diabetes: Experimental and modeling analyses reveal mechanisms of insulin resistance in human adipocytes". In: *Journal of Biological Chemistry* 288.14, pp. 9867–9880. ISSN: 00219258. DOI: 10.1074/jbc.M112.432062.

Buxton, Richard B. (2010). "Interpreting oxygenation-based neuroimaging signals: the importance and the challenge of understanding brain oxygen metabolism". In: *Frontiers in Neuroenergetics*. ISSN: 1662-6427. DOI: 10.3389/fnene.2010.00008.

– (Sept. 2013). "The physics of functional magnetic resonance imaging (fMRI)". In: *Reports on Progress in Physics* 76.9. ISSN: 00344885. DOI: 10.1088/0034-4885/76/9/096601.

Buxton, Richard B. et al. (2014). "Variability of the coupling of blood flow and oxygen metabolism responses in the brain: A problem for interpreting BOLD studies but potentially a new window on the underlying neural activity". In: *Frontiers in Neuroscience* 8 JUN. ISSN: 1662453X. DOI: 10.3389/fnins.2014.00139.

Buxton, Richard B, K Miller, et al. (1998). "BOLD signal dynamics: the balloon model with viscoelastic effects". In: *Sixth Meeting, International Society for Magnetic Resonance in Medicine, Sydney, Australia*, p. 1401.

Buxton, Richard B, Kâmil Uludağ, et al. (2004). "Modeling the hemodynamic response to brain activation". In: *NeuroImage* 23, S220–S233. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2004.07.013. URL: http://www.sciencedirect.com/science/article/pii/S1053811904003787.

Buxton, Richard B, Eric C Wong, and Lawrence R Frank (1998). *Dynamics of Blood Flow and Oxygenation Changes During Brain Activation: The Balloon Model*. Tech. rep.

Cauli, B. et al. (2004). "Cortical GABA interneurons in neurovascular coupling: Relays for subcortical vasoactive pathways". In: *Journal of Neuroscience* 24.41, pp. 8940–8949. DOI: 10.1523/JNEUROSCI.3065-04.2004.

Cedersund, Gunnar (2020). *Webpage: Integrative System Biology group*. URL: https://isbgroup.eu/research/biology.php.

Cedersund, Gunnar and Jacob Roll (2009). "Systems biology: Model based evaluation and comparison of potential explanations for given biological data". In: *FEBS Journal* 276.4, pp. 903–922. ISSN: 1742464X. DOI: 10.1111/j.1742-4658.2008.06845.x.

Cobelli, Claudio, Eric Renard, and Boris Kovatchev (Nov. 2011). *Artificial pancreas: Past, present, future*. DOI: 10.2337/db11-0654.

Dalla Man, Chiara, Robert A Rizza, and Claudio Cobelli (Oct. 2007). "Meal simulation model of the glucose-insulin system." eng. In: *IEEE transactions on biomedical engineering* 54.10, pp. 1740–1749. ISSN: 0018-9294 (Print). DOI: 10.1109/TBME.2007.893506.

Davis, Timothy L. et al. (Feb. 1998). "Calibrated functional MRI: Mapping the dynamics of oxidative metabolism". In: *Proceedings of the National Academy of Sciences of the United States of America* 95.4, pp. 1834–1839. ISSN: 00278424. DOI: 10.1073/pnas.95.4.1834.

Doke, Sonali K. and Shashikant C. Dhawale (July 2015). "Alternatives to animal testing: A review". In: *Saudi Pharmaceutical Journal* 23.3, pp. 223–229. ISSN: 1319-0164. DOI: 10.1016/J.JSPS.2013.11.002. URL: https://www.sciencedirect.com/science/article/pii/S1319016413001096.

Drummond, Eleanor and Thomas Wisniewski (Feb. 2017). "Alzheimer's disease: experimental models and reality". eng. In: *Acta neuropathologica* 133.2, pp. 155–175. ISSN: 1432-0533. DOI: 10.1007/s00401-016-1662-x. URL: https://pubmed.ncbi.nlm.nih.gov/28025715%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5253109/.

Estrada, C and J DeFelipe (Mar. 1998). "Nitric oxide-producing neurons in the neocortex: morphological and functional relationship with intraparenchymal microvasculature." In: *Cerebral Cortex* 8.3, pp. 193–203. ISSN: 1047-3211. DOI: 10.1093/cercor/8.3.193. URL: https://doi.org/10.1093/cercor/8.3.193.

Ferdowsian, Hope R. and Nancy Beck (2011). "Ethical and scientific considerations regarding animal testing and research". In: *PLoS ONE* 6.9. ISSN: 19326203. DOI: 10.1371/journal.pone.0024059.

Foreman-Mackey, Daniel et al. (Feb. 2012). "emcee: The MCMC Hammer". In: DOI: 10.1086/670067. URL: http://arxiv.org/abs/1202.3665%20http://dx.doi.org/10.1086/670067.

Gauthier, Claudine J and Richard D Hoge (May 2013). "A generalized procedure for calibrated MRI incorporating hyperoxia and hypercapnia". In: *Human Brain Mapping* 34.5, pp. 1053–1069. ISSN: 1065-9471. DOI: 10.1002/hbm.21495. URL: https://doi.org/10.1002/hbm.21495.

Göbel, U, H Theilen, and W Kuschinsky (Feb. 1990). "Congruence of total and perfused capillary network in rat brains." In: *Circulation Research* 66.2, pp. 271–281. DOI: 10.1161/01.RES.66.2.271. URL: https://doi.org/10.1161/01.RES.66.2.271.

Goodman, J and J Weare (2010). "Ensemble Samplers With Affine". In: *Communications in Applied Mathematics and Computational Science* 5.1, pp. 65–80.

Griffeth, Valerie E.M. and Richard B. Buxton (Sept. 2011). "A theoretical framework for estimating cerebral oxygen metabolism changes using the calibrated-BOLD method: Modeling the effects of blood volume distribution, hematocrit, oxygen extraction fraction, and tissue signal properties on the BOLD signal". In: *NeuroImage* 58.1, pp. 198–212. ISSN: 10538119. DOI: 10.1016/j.neuroimage.2011.05.077.

GRUBB, ROBERT L. et al. (Sept. 1974). "The Effects of Changes in Pa CO ₂ Cerebral Blood Volume, Blood Flow, and Vascular Mean Transit Time". In: *Stroke* 5.5, pp. 630–639. ISSN: 0039-2499. DOI: 10.1161/01.STR.5.5.630. URL: https://www.ahajournals.org/doi/10.1161/01.STR.5.5.630.

Hajar, Rachel (Jan. 2011). "Animal testing and medicine". eng. In: *Heart views : the official journal of the Gulf Heart Association* 12.1, p. 42. ISSN: 0976-5123. DOI: 10.4103/1995-705X.81548. URL: https://pubmed.ncbi.nlm.nih.gov/21731811%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3123518/.

– (2015). "History of medicine timeline". eng. In: *Heart views : the official journal of the Gulf Heart Association* 16.1, pp. 43–45. ISSN: 1995-705X. DOI: 10.4103/1995-705x.153008. URL: https://pubmed.ncbi.nlm.nih.gov/25838882%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4379645/.

Hall, Kevin D and Peter N Jordan (Dec. 2008). "Modeling weight-loss maintenance to help prevent body weight regain." eng. In: *The American journal of clinical nutrition* 88.6, pp. 1495–1503. ISSN: 1938-3207 (Electronic). DOI: 10.3945/ajcn.2008.26333.

Havlicek, Martin, Alard Roebroeck, Karl J. Friston, et al. (2017). "On the importance of modeling fMRI transients when estimating effective connectivity: A dynamic causal modeling study using ASL data". In: *NeuroImage*. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2017.03.017.

Havlicek, Martin, Alard Roebroeck, Karl Friston, et al. (2015). "Physiologically informed dynamic causal modeling of fMRI data". In: *NeuroImage* 122, pp. 355–372. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2015.07.078.

Heuvel, Martijn P van den and Hilleke E Hulshoff Pol (2010). "Exploring the brain network: A review on resting-state fMRI functional connectivity". In: *European Neuropsychopharmacology* 20.8, pp. 519–534. ISSN: 0924-977X. DOI: https:

//doi.org/10.1016/j.euroneuro.2010.03.008. URL: http://www.sciencedirect.com/science/article/pii/S0924977X10000684.

Hillman, Elizabeth M.C. (2014). "Coupling Mechanism and Significance of the BOLD Signal: A Status Report". In: *Annual Review of Neuroscience* 37.1, pp. 161–181. ISSN: 0147-006X. DOI: 10.1146/annurev-neuro-071013-014111.

IntiQuan (2017). *IQM toolbox*. URL: https://iqmtools.intiquan.com.

Isaacson, Jeffery S and Massimo Scanziani (2011). "How Inhibition Shapes Cortical Activity Excitation and inhibition walk hand in hand". In: *Neuron* 72.2, pp. 231–243. DOI: 10.1016/j.neuron.2011.09.027.

Johansson, Rickard (2017). *Model-based hypothesis testing in biomedicine : how systems biology can drive the growth of scientific knowledge*. Department of Biomedical Engineering Linköping University, Sweden. ISBN: 9789176854570. URL: https://login.e.bibl.liu.se/login?url=https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,uid&db=cat00115a&AN=lkp.929369&lang=sv&site=eds-live&scope=site%0Ahttp://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-141614.

Kolata, Gina (2020). *An Alzheimer's Treatment Fails: 'We Don't Have Anything Now'*. URL: https://www.nytimes.com/2020/02/10/health/alzheimers-amyloid-drug.html.

Kovatchev, Boris P et al. (2009). "In silico preclinical trials: A proof of concept in closed-loop control of type 1 diabetes". In: *Journal of Diabetes Science and Technology*. Vol. 3. 1, pp. 44–55. DOI: 10.1177/193229680900300106. URL: www.journalofdst.org.

Lauritzen, Martin (Jan. 2005). "Reading vascular changes in brain imaging: is dendritic calcium the key?" eng. In: *Nature reviews. Neuroscience* 6.1, pp. 77–85. ISSN: 1471-003X (Print). DOI: 10.1038/nrn1589.

Lovell-Badge, Robin (Jan. 2013). *Nine out of ten statistics are taken out of context*. Tech. rep. London: MRC National Institute for Medical Research. URL: www.understandinganimalresearch.org.uk.

Mark, Clarisse I, Joseph A Fisher, and G Bruce Pike (2011). "Improved fMRI calibration: Precisely controlled hyperoxic versus hypercapnic stimuli". In: *NeuroImage* 54.2, pp. 1102–1111. ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2010.08.070. URL: http://www.sciencedirect.com/science/article/pii/S1053811910011675.

Markou, Athina et al. (Jan. 2009). *Removing obstacles in neuroscience drug discovery: The future path for animal models*. DOI: 10.1038/npp.2008.173.

McDonald, Michael. P and Bruce. J Overmier (Dec. 1997). "Present Imperfect: A Critical Review of Animal Models of the Mnemonic Impairments in Alzheimer's Disease". In: *Neuroscience & Biobehavioral Reviews* 22.1, pp. 99–120. ISSN: 0149-7634. DOI: 10.1016/S0149-7634(97)00024-9. URL: https://www.sciencedirect.com/science/article/pii/S0149763497000249.

Meldrum, Brian (May 2000). "Glutamate as a Neurotransmitter in the Brain: Review of Physiology and Pathology". In: *The Journal of nutrition* 130, 1007S–15S. DOI: 10.1093/jn/130.4.1007S.

MeVisLab (2020). *MeVisLab*. URL: https://www.mevislab.de/mevislab/features/image-processing.

Mishra, Anusha et al. (Dec. 2016). "Astrocytes mediate neurovascular signaling to capillary pericytes but not to arterioles". In: *Nature Neuroscience* 19.12, pp. 1619–1627. ISSN: 15461726. DOI: 10.1038/nn.4428.

Morrison, Tina M. et al. (2018). "Advancing regulatory science with computational modeling for medical devices at the FDA's office of science and engineering laboratories". In: *Frontiers in Medicine* 5.SEP. ISSN: 2296858X. DOI: 10.3389/fmed.2018.00241.

Mullinger, K. J. et al. (Aug. 2017). "Post-stimulus fMRI and EEG responses: Evidence for a neuronal origin hypothesised to be inhibitory". In: *NeuroImage* 157, pp. 388–399. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2017.06.020.

Ogawa, S. et al. (1990). "Brain magnetic resonance imaging with contrast dependent on blood oxygenation". In: *Proceedings of the National Academy of Sciences of the United States of America* 87.24, pp. 9868–9872. ISSN: 00278424. DOI: 10.1073/pnas.87.24.9868.

Petzold, Gabor C. and Venkatesh N. Murthy (2011). *Role of astrocytes in neurovascular coupling*. DOI: 10.1016/j.neuron.2011.08.009.

Philipson, Ola et al. (Mar. 2010). *Animal models of amyloid-β-related pathologies in Alzheimer's disease*. DOI: 10.1111/j.1742-4658.2010.07564.x.

Plewes, Donald B. and Walter Kucharczyk (2012). "Physics of MRI: A primer". In: *Journal of Magnetic Resonance Imaging* 35.5, pp. 1038–1054. ISSN: 10531807. DOI: 10.1002/jmri.23642.

Raichle, Marcus E and Debra A Gusnard (Aug. 2002). "Appraising the brain's energy budget". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 99.16, pp. 10237–10239. ISSN: 0027-8424. DOI: 10.1073/pnas.172399499. URL: https://pubmed.ncbi.nlm.nih.gov/12149485%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC124895/.

Reitz, Christiane (2012). *Alzheimer's disease and the amyloid cascade hypothesis: A critical review*. DOI: 10.1155/2012/369808.

Rosen, Evan D and Bruce M Spiegelman (2006). "Adipocytes as regulators of energy balance and glucose homeostasis". In: *Nature* 444.7121, pp. 847–853. ISSN: 1476-4687. DOI: 10.1038/nature05483. URL: https://doi.org/10.1038/nature05483.

Scarlett, J and M Schwartz (2015). "Gut-brain mechanisms controlling glucose homeostasis". In: *F1000Prime Reports* 7.

SciPy (n.d.). *Basin-hopping*. URL: https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.basinhopping.html.

Sundials (n.d.[a]). *Sundials CVODE*. URL: https://computing.llnl.gov/projects/sundials/cvode.

– (n.d.[b]). "Sundials IDA". In: ().

Uhlirova, Hana et al. (2016). "Cell type specificity of neurovascular coupling in cerebral cortex". In: *eLife* 5.MAY2016, pp. 1–23. ISSN: 2050084X. DOI: 10.7554/eLife.14315.

Uludağ, Kâmil et al. (2004). "Coupling of cerebral blood flow and oxygen consumption during physiological activation and deactivation measured with fMRI". In: *NeuroImage* 23.1, pp. 148–155. ISSN: 1053-8119. DOI: https://doi.org/10.

1016/j.neuroimage.2004.05.013. URL: http://www.sciencedirect.com/science/article/pii/S1053811904002903.

Zijl, Peter C.M. van, Jun Hua, and Hanzhang Lu (Aug. 2012). *The BOLD post-stimulus undershoot, one of the most debated issues in fMRI*. DOI: 10.1016/j.neuroimage.2012.01.029.

Appendices

A
Prior distribution

Table A.1: The prior distribution for the parameters is either a normal distribution, N, or a log-normal distribution, L. In both cases are the prior distribution fully described by the mean, μ_i, and variance, σ_i^2, for their respective normal distribution, which we give here, see equations (3.16) and (3.17). Just remember for the log-normal distribution that the mean and variance given for the normal distribution is not the same as the mean and variance for the actual prior distribution[2].

Model	Parameter	Mean (μ_i)	Variance (σ_i^2)	Distribution
Neural	c	-2.3	1.0	L
	σ	-0.9	0.49	L
	μ	-0.9	0.49	L
	λ	-1.6	0.0625	L
Neurovascular	ξ_E	0.0	1.0	N
	ξ_I	0.0	1.0	N
	ρ	-0.5	0.25	L
	ϕ	0.4	0.16	L
	χ	-0.5	0.25	L
	θ_E	0.0	1.0	N
	θ_I	0.0	1.0	N
	δ	-0.8	0.64	L
	τ	-0.9	3.24	L
	t_0	0.7	0.25	L
	α	-1.97	0.16	L
BOLD	M	-1.9	0.25	L
	β	-0.1	0.09	L

[2]They are however related by a simple transformation: $\mu_L = \exp\left(\mu + \frac{\sigma^2}{2}\right)$, $\sigma_L^2 = [\exp(\sigma^2) - 1] \cdot \mu_L^2$.