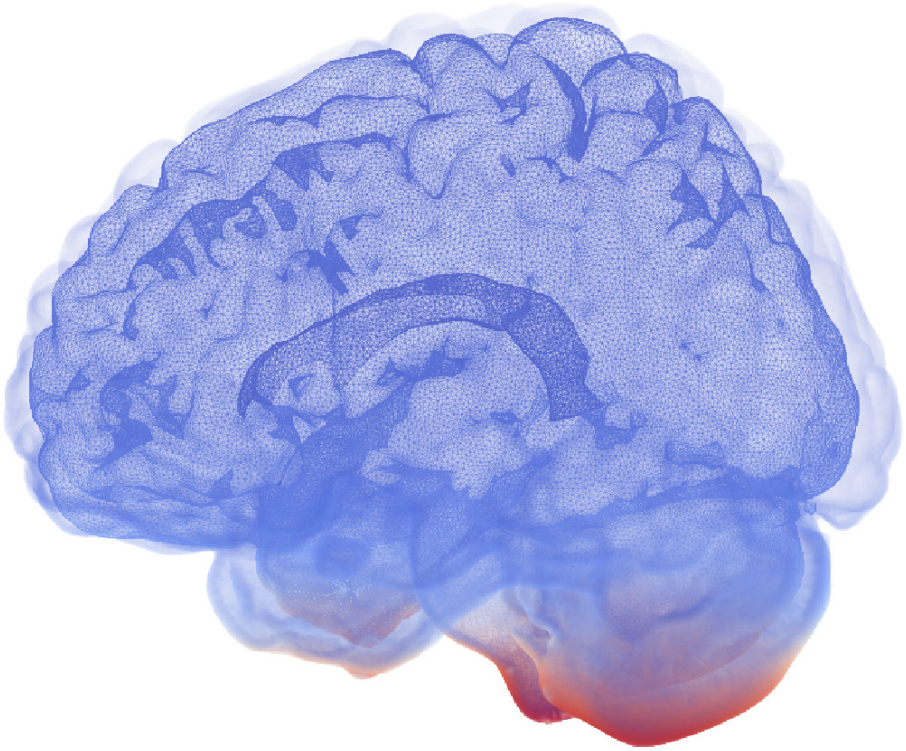
EPSRC
Engineering and Physical Sciences Research Council

InFoMM
Industrially Focused Mathematical Modelling

EPSRC Centre for Doctoral Training in Industrially Focused Mathematical Modelling



Multilevel Monte Carlo Methods for Uncertainty Quantification in Brain Simulations

Matteo Croci

UNIVERSITY OF OXFORD

simula

Contents

1 Introduction

The project motivation: brain disease

Within the Department of Numerical Analysis and Scientific Computing in Simula, Marie E. Rognes' group focuses on the simulation of physiological processes that affect human health. The expertise of the group ranges from medical research and bioengineering to mathematical modelling, numerical analysis and scientific computing. One of the group's current interests is the simulation of the fluid dynamics within the brain, which we refer to as the brain '**waterscape**'. Anomalies in the brain waterscape are believed to be related to common diseases such as **dementia** and hydrocephalus.

The most famous type of dementia is Alzheimer's disease, which affects over 44 million people worldwide.

Dementia is a generic term that describes a wide range of diseases which have the common denominator of creating a disorder in a person's mental processes. Examples include memory loss, personality changes, impaired reasoning and thinking and deterioration of social behaviour. Worldwide, about 47 million people suffer from dementia and every year 7.7 millions more cases are diagnosed, with an expected 75.6 million people being affected by 2030. This leads to an enormous health care cost : $604 billion per year in the US alone. 'More research is needed to develop new and more effective treatments and to better understand the causes of dementia' (World Health Organization).

The most well-known type of dementia is Alzheimer's disease. One of the hallmark characteristics of Alzheimer's is the formation of plaques made up of a protein called Amyloid-β. This protein can be normally found as a solute in the cellular cytosol and in the brain interstitial fluid (ISF) (present in the brain extracellular space, the space between brain cells). However, under Alzheimer's conditions, this protein accumulates in the brain forming plaques that may induce neuronal death.

Every cell in our body absorbs its nutrients from the blood and then releases its metabolic wastes in the surrounding extracellular space. From there, the metabolic waste is cleared by the lymphatic system, a network of tissues and organs specialised in toxin and waste clearance. The only part of our body where (almost) no lymphatic vessels are present is the brain, where some other unknown clearance mechanism is in action. One of the leading hypotheses in Alzheimer's research is that when this mechanism, whatever it might be, is impaired, Amyloid-β accumulates damaging the brain and causing Alzheimer's. If researchers could discover what this mechanism is, they might possibly prevent the damage and find a cure.

For this reason, over the last few years the study of the physiological mechanisms governing the movement of fluids in the brain has gained prominence: understanding how the brain waterscape works can help discover how dementia develops. However, experimenting with the human brain in vivo is extremely difficult and the subject is still poorly understood. This makes this topic a highly active interdisciplinary research area. Invasive experiments are performed on mice, but their results are not necessarily representative of the human physiology. In addition, they cannot be performed on people as that would be unethical. Non-invasive experiments, instead, do not give us enough information. For this reason, computer simulations of mathematical models describing these phenomena are gaining popularity as they represent an alternative avenue of investigation.

One example of a physical parameter subject to uncertainty is given by the brain tissue permeability which varies up to 7 orders of magnitude in computational models

Computer simulations that are based on mathematical models rely on physiological information which is obtained through *in vivo* measurements. Modern techniques allow the measurement of most physical parameters with a reasonable degree of accuracy. Nevertheless, measurement errors are still present, and, in some cases, the measurement itself can affect the quantity which is being measured. In general, some of the physiological aspects of the fluid dynamics of the brain are still not well understood.

Overall, one of the main challenges in brain simulation is the lack of accurate quantitative information on the physical input parameters needed to set up mathematical models. Quantities such as brain matter permeability, interstitial fluid flow velocity and diffusivity are only known approximately or on average. The position of the blood vessels and capillaries can be measured, but it varies from patient to patient and it is extremely difficult to resolve without significant expense. The focus of our work is estimating how uncertainty in these input parameters affects models and simulations' predictions and propagates to output quantities of clinical interest. Quantifying this uncertainty is extremely important

for medical applications. We construct surrogates for the physiological parameters that account for the uncertainty in their values through the use of **Gaussian random fields** and we use **multilevel Monte Carlo methods** to perform forward **uncertainty quantification** (UQ) on brain fluid models. We solve the mathematical problems of interest via the **finite element method** (FEM).

Biological background

The human brain is influenced by three different types of flow networks: the **blood vasculature**, the **cerebrospinal fluid** (CSF) and the **interstitial fluid** (ISF). These networks are separated by semi-permeable membranes, which only allow exchange of certain substances and liquids.

The vascular network carries the nutrients to the brain cells within the **parenchyma** (the cerebral matter) and it is strongly related to the production of CSF, which is believed to be produced from blood plasma in the brain ventricles. From the ventricles, the CSF flows in the spinal canal and in the **subarachnoid space**, the space between the arachnoid and pia meninges (see Figure 1).

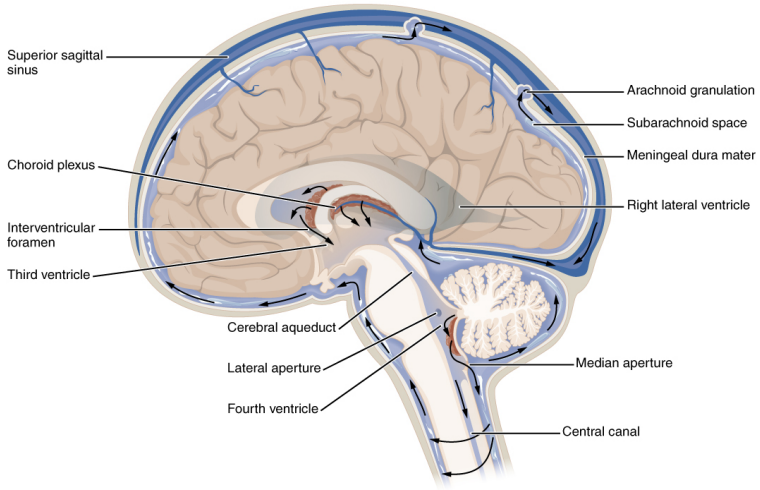


Figure 1 – A schematic of the brain CSF flow. CSF is produced within the ventricles and flows around the brain before being reabsorbed in the subarachnoid space. The parenchyma (coloured in grey in the picture) is made of brain cells and it is permeated by ISF. Picture taken from `http://www.wikiwand.com/en/Cerebrospinal_fluid`.

The cerebral matter can be divided in two parts, the gray matter and the white matter.

Brain cells absorb their nutrients from the vascular system and release metabolic waste into the interstitial space, the space between the cells. Here, the waste is absorbed and collected by the ISF, which permeates the brain tissue. The interstitial space and the ISF make the parenchyma a sponge-like solid permeated by liquid.

One of the main hypothesis about brain metabolic waste clearance is called the **glymphatic hypothesis**, according to which a waste-clearing bulk flow of ISF has been proposed to occur through the brain parenchyma. The glymphatic hypothesis is still far from being established. One of the leading alternative hypothesis, the **intramural periarterial drainage theory** (IPAD), states that solutes are instead removed from the brain along the basement membranes of capillaries and arteries. The exact mechanism is, however, still under investigation.

Overall, some aspects of ISF movement are still not well understood and discovering how the brain clears itself from metabolic waste could be one of the milestones to reach before being able to understand how Alzheimer's disease develops.

2 Mathematical model

Modelling brain tracer movement

Sugar dissolved into a glass of water spreads due to diffusion. When the water is stirred, the sugar moves with the water due to advection.

Tracers are used to highlight the presence of fluid in MRI scans.

In this project, we study the movement of a tracer within the brain ISF and we simulate on a computer the clinical study performed by Ringstad and others in [1]. In this study, a tracer is injected within the spinal canal of 8 healthy patients and its movement is monitored via **magnetic resonance imaging** (MRI). From the spinal canal, the tracer spreads upwards in the CSF and finally penetrates within the brain.

Substances dissolved within a fluid evolve according to three phenomena: **diffusion**, **advection** and **growth/decay/drainage**. Diffusion is the movement of a soluted substance from higher concentration regions to lower concentration regions and happens independently from the velocity of the liquid. If the fluid moves it also transports the substances dissolved within it. This phenomenon is called advection. Finally, a solute can react with other substances or be drained away. This causes its concentration to increase (growth) or decrease (decay/drainage).

We consider a baseline mathematical model, given by the advection-reaction-diffusion equation, that accounts for this behaviour. However, we modify it according to 5 different hypotheses present in the medical literature about ISF flow and solute clearance [2, 3]; see Figure 2. This includes investigating the relative effect of diffusion VS convection in tracer transport, and the glymphatic hypothesis mentioned earlier.

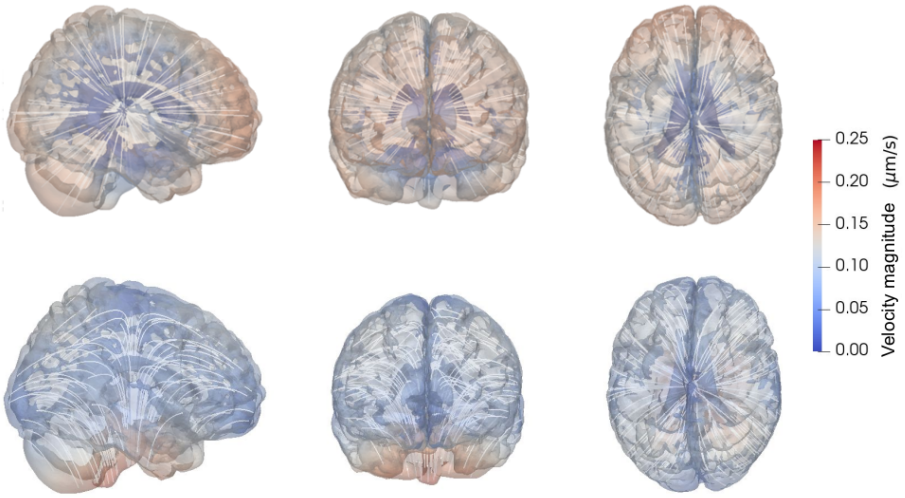


Figure 2 – Velocity magnitude and streamlines for two different models for the ISF velocity. In the top row of figures, flow is assumed to occur from the brain cortex towards the ventricles. In the bottom row, flow is assumed to follow large vascular structures upwards along the brain. From left to right: sagittal, coronal and transverse views.

We assume that no tracer or liquid penetrates through the ventricles and we impose that the concentration at the subarachnoid interface is equal to the tracer concentration in the CSF just outside the parenchyma. To compute the CSF concentration we impose that **total amount of tracer** in the brain and CSF **is preserved**, which means that

$$\text{tot. tracer in brain} \quad + \quad \text{tot. tracer in CSF} \quad - \quad \text{tot. tracer drained} \quad = \quad \text{tot. initial tracer.}$$

This results mathematically into a **non-local boundary condition**.

Quantifying the model uncertainty

Many parameters are uncertain in our model. In this project, we focus on two physical quantities: the tracer diffusivity in the ISF and the ISF velocity and we model them as **random variables** and **random fields**. A random variable is a variable whose value depends on an aleatory (or stochastic), event. For example, the outcome of a die roll is a

random variable that can have values between 1 and 6 according to how the die is cast. The way the die positions on the table is random and it affects the value of the die roll. While random variables reflect uncertainty in a single parameter value, constant in time and space, random fields model uncertainty in spatially varying quantities. A random field is a collection of random variables, one for each spatial location considered within the brain.

The win at a lottery, the value of a playing card drawn from a deck, whether it will be sunny next week... All of these can be modelled as random variables.

Simulating a stochastic event allows us to compute the related value attained by a random variable or field. This operation is called sampling. Consider a game in which we win a pound if a coin flip is heads and we lose a pound otherwise. How much are we expecting to win on average if we play many times? To compute the expected win, or better, the expected value of the coin toss, we can flip the coin a lot of times and take the average of the win and losses. For this game, we have equal probability (50%-50%) to win or lose a pound so the expected win of this game is 0 pounds. The win here is a random variable and this method of computing the expected value is called the Monte Carlo method.

In our model, the fact that diffusivity and ISF velocity are taken to be random variables or fields causes the fluid pressure and displacement to be uncertain as well. Similarly as for the coin toss, we want to be able to draw samples from the distribution of fluid pressure and of brain tissue displacement so as to compute their expected values (the expected behaviour of the brain system). By using the FEM to solve the mathematical model, we can compute the evolution of the injected tracer concentration for a given diffusivity and ISF velocity. However, the diffusivity and ISF velocity values have to be sampled as well.

Sampling random fields and variables requires the specification of their probability distribution and a sampling method. A distribution is a measure of the probability of a random variable or field to attain a given value. For example, the distribution of a coin toss is 50% probability of flipping heads and 50% probability of flipping tails. In this case the distribution is also a measure of how much do we expect a given value of a physical parameter to be likely to be the true physiological value in a living brain.

The distribution of the parameters is itself unknown, but we can provide a model for their distribution, which means that we can assume that the distribution behaves in a certain way according to some physical principles and constraints. For instance, we know that it is extremely unlikely for the parameters to be more than 3 times smaller or larger than their estimated average values. As another example, in some instances of our model we impose that no CSF/ISF leaves or enters the parenchyma through the bloodstream.

Once we have a model for the distribution, we also need to formulate a sampling method or strategy. **Sampling a random field** is a very challenging process to perform on a computer. In our work we derived two new algorithms that can sample random fields extremely efficiently (in optimal cost complexity), making them some of the fastest random field sampling methods available [4, 5].

We now have all the ingredients needed for the Monte Carlo method to compute the expected tracer movement in the brain system. However, this method is computationally costly, to the extent that a single (standard Monte Carlo) simulation would take a couple of years to run on a powerful computer. For this reason, we use instead three more advanced methods, called **multilevel Monte Carlo** (MLMC), **quasi Monte Carlo** (QMC) and **multilevel quasi Monte Carlo** (MLQMC).

Monte Carlo methods are perfectly suited for parallelization. This means that we can considerably improve efficiency by dividing the workload across different processors.

MLMC uses a hierarchy of approximation levels. Each level is characterised by an accuracy degree and computational cost. On the coarse levels, we make a rough, but cheap approximation of the samples, while on the fine levels, the accuracy is high at a high cost price. MLMC draws many (cheap) samples from the coarse levels and corrects them with samples taken from the fine levels (expensive, but we only need a small amount of them). This procedure allows a fine level approximation quality at (roughly) a coarse level cost. The accuracy of standard Monte Carlo increases slowly with the number of samples taken. For this reason, Monte Carlo simulations typically require a large number of samples. QMC improves on standard Monte Carlo by choosing the samples to be taken more carefully and thus reducing the overall number of samples needed. MLQMC combines the benefits of QMC and MLMC. Thanks to these methods, **the computational time needed for our simulations reduces from a couple of years to a couple of weeks**.

3 Results

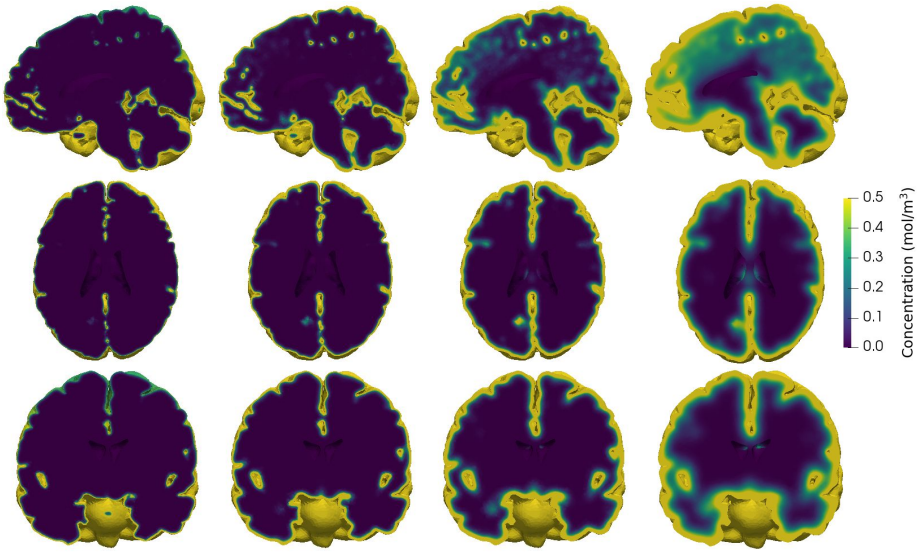


Figure 3 – The effect of pure diffusion (simulation with zero ISF velocity assumed). Brain tracer concentration after (from left to right) 1, 3, 8 and 24 hours in (from top to bottom) sagittal, transverse and coronal planes. Initially, most of the tracer is found in inferior regions. At 24 hours, tracer has penetrated substantially into the outer regions, but not into the deep, central regions.

The unit of concentration is the mole (mol). One mole corresponds to an amount of particles equal to the Avogadro number, roughly 6.022×10^{23} particles.

The investigation on the 5 different medical hypotheses revealed that the diffusion coefficient is the largest source of uncertainty in our model, showing that accurate diffusivity values are needed if models that do not account for uncertainty are to be trusted. Nevertheless, even with moderate uncertainty in the diffusivity parameter, discrepancies between simulations of diffusion (see Figure 3) and experimental data (see [1]) are too large to be attributed to uncertainties in the diffusion coefficient alone.

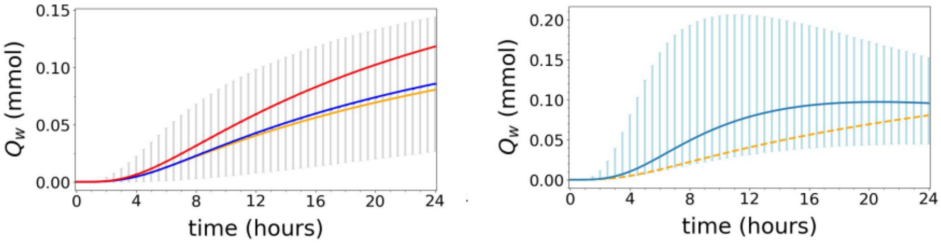


Figure 4 – Total amount of tracer Q_w expected to be found in the brain central regions (white matter), as predicted by different models. The orange line on both side represents the predictions due to pure diffusion (zero ISF velocity, see Figure 3). The red and blue lines in the plot on the left reflects a glymphatic velocity field with (see Figure 2, bottom) and without directionality respectively. The blue line in the plot on the right reflects a directional influx of fluid along brain arteries (see Figure 2, top). The errorbars represent variability due to uncertainty in the physiological parameters considered. A directional structure is needed to justify an increase in tracer penetration.

In fact, our results suggest that the presence of a non-zero ISF velocity, possibly as described by the glymphatic hypothesis, would be needed to justify the presence of tracer deep into the brain central regions (see Figure 4). However, this assumes the presence of a directional structure in ISF flow (see Figure 2) and more accurate ISF flow measurements would be needed to establish whether such a flow structure is effectively present. Overall, more quantitative experimental data are needed to obtain a deeper understanding of ISF movement.

4 Discussion, conclusions & recommendations

Simulating the fluid dynamics of the brain and brain solute movement is important for medical applications, such as disease investigation and treatment. However, these simulations are hard to set-up since some of the physiological aspects of the brain are still not well understood. This leads to uncertainty in the formulation of mathematical models and in the related parameters. To simulate the evolution of a tracer within the brain under this uncertainty, we constructed a new advection-reaction-diffusion model with random field and random variable parameters, which we used to investigate different hypotheses on ISF movement.

Some of the simulations were run on the Abel Cluster, a supercomputing facility in Norway. See `www.hpc.uio.no`.

The equation related to this model was solved by using finite elements and a combination of various advanced Monte Carlo methods (such as multilevel and quasi Monte Carlo). The computational improvement in comparison to a standard Monte Carlo method is quite significant (about 100 times faster). For both standard Monte Carlo and these more advanced methods, two new efficient random field sampling techniques were derived. These are currently among the fastest sampling methods available in the literature. The mathematical techniques and algorithms developed in this project could be applied widely, for example in mathematical finance, vehicle manufacturing, oil reservoir simulation, engineering.

The results obtained from simulating the tracer concentration model support the claim that diffusion alone is not sufficient to explain penetration of tracer deep into the central brain regions as seen in experimental data. However, a directional structure in ISF circulation might be sufficient to increase tracer transport.

For further details about this project, we refer to the author's DPhil thesis [6].

5 Potential impact

Marie E. Rognes, Chief Research Scientist and Research Professor at Simula, commented:

"Simula Research Laboratory is a limited company owned by the Norwegian Ministry of Education and Research. Our main objective is to create knowledge about fundamental scientific challenges that are of genuine value for society. At Simula Scientific Computing, we aim to develop and apply novel simulation technologies to reach new understanding of complex physical processes affecting human health. We target selected medical problems where insight from mathematical modeling can contribute to changing clinical practice.

One of our current fundamental research goals is to use modelling and simulation to better understand and predict the flow of fluid and solutes through brain tissue. These physiological mechanisms are crucial for the well-being of the brain. In spite of their importance, we understand them but little. Mathematics and numerics could play a crucial role in gaining new insight, and medical doctors express an urgent need for modelling and simulation to overcome limitations in traditional techniques. Indeed, a new avenue of investigation to understand these physiological prosesses has tremendous long-term clinical impact in association with neurological diseases such as dementia and oedema.

Croci's project "Uncertainty Quantification through multilevel Monte Carlo Simulation" is a ground-breaking first step in this direction. By applying advanced numerical techniques, he has quantified the propagation of uncertainty in material parameters through a model for solute transport in brain tissue. In particular, he has studied the effect of uncertainty in fluid velocity and diffusivity on tracer enhancement in gray and white matter. This has allowed us to reach new conclusions on the modes of solute transport in the brain with physiological implications, and gives us new numerical techniques to apply in other scientific settings. Further, the knowledge transfer between Croci and his Oxford research team and Simula Scientific Computing within the field of high-dimensional uncertainty quantification in biomedical applications has been extremely valuable."

References

[1] Geir Ringstad, Svein Are Sirirud Vatnehol, and Per Kristian Eide. Glymphatic MRI in idiopathic normal pressure hydrocephalus. *Brain*, 140(10):2691–2705, 2017.

[2] M. Croci, V. Vinje, and M. E. Rognes. Uncertainty quantification of parenchymal tracer distribution using random diffusion and convective velocity fields. *submitted to Fluids and Barriers of the CNS*, 2019. doi: 10.1101/665109.

[3] M. Croci, V. Vinje, and M. E. Rognes. Mutilevel and quasi Monte Carlo methods for uncertainty quantification of parenchymal tracer distribution. *DPhil thesis chapter - submission to a journal expected December 2019*.

[4] M. Croci, M. B. Giles, M. E. Rognes, and P. E. Farrell. Efficient white noise sampling and coupling for multilevel Monte Carlo with non-nested meshes. *SIAM/ASA Journal on Uncertainty Quantification*, 6(4):1630–1655, 2018. doi: 10.1137/18M1175239.

[5] M. Croci, M. B. Giles, M. E. Rognes, and P. E. Farrell. Multilevel quasi Monte Carlo methods for elliptic partial differential equations driven by spatial white noise. *DPhil thesis chapter - submission to a journal expected November 2019*.

[6] M. Croci. *Multilevel Monte Carlo Methods for Uncertainty Quantification in Brain Simulations*. PhD thesis, University of Oxford, 2019.