

Fast uncertainty quantification of tracer distribution in the brain interstitial fluid with multilevel and quasi Monte Carlo

Matteo Croci^{1,2}  | Vegard Vinje²  | Marie E. Rognes² 

¹Mathematical Institute, University of Oxford, Oxford, UK

²Department for Numerical Analysis and Scientific Computing, Simula Research Laboratory, Lysaker, Norway

Correspondence

Matteo Croci, Mathematical Institute, University of Oxford, Oxford, UK.
Email: matteo.croci@maths.ox.ac.uk

Funding information

Engineering and Physical Sciences Research Council, Grant/Award Number: EP/L015803/1; European Research Council, Grant/Award Number: #714892; Norges Forskningsråd, Grant/Award Number: #250731; NOTUR, Grant/Award Number: NN9316K

Abstract

Efficient uncertainty quantification algorithms are key to understand the propagation of uncertainty—from uncertain input parameters to uncertain output quantities—in high resolution mathematical models of brain physiology. Advanced Monte Carlo methods such as quasi Monte Carlo (QMC) and multilevel Monte Carlo (MLMC) have the potential to dramatically improve upon standard Monte Carlo (MC) methods, but their applicability and performance in biomedical applications is underexplored. In this paper, we design and apply QMC and MLMC methods to quantify uncertainty in a convection-diffusion model of tracer transport within the brain. We show that QMC outperforms standard MC simulations when the number of random inputs is small. MLMC considerably outperforms both QMC and standard MC methods and should therefore be preferred for brain transport models.

KEYWORDS

biomedical computing, brain fluids, brain simulation, diffusion-convection, finite element method, multilevel Monte Carlo, quasi Monte Carlo, random fields

1 | INTRODUCTION

Mathematical models in biology involve many parameters that are uncertain or in some cases unknown. Over the last years, increased computing power has expanded the complexity and increased the number of degrees of freedom of many such models. For instance, full scale modeling of the brain requires very high spatial resolution to resolve the details on the curvature of the surface. For this reason, efficient uncertainty quantification algorithms are now needed to explore the often large parameter space of a given model. Advanced Monte Carlo methods such as quasi Monte Carlo (QMC) and multilevel Monte Carlo (MLMC) have become very popular in the mathematical, engineering, and financial literature for the quantification of uncertainty in model predictions. However, applying these methods to physiologically relevant brain simulations is a difficult task given the typical complexity of the models and geometries involved.

Efficient elimination of waste products from the brain is crucial for a functional central nervous system. In humans, the brain is responsible for around 20% of the total oxygen consumption and 15% of the cardiac output.¹ Despite this high energy demand, the brain lacks lymphatic vessels, which carry waste products in the rest of the body, thus introducing the need for an effective alternative mechanism. Proposed theories of waste clearance all include a combination

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *International Journal for Numerical Methods in Biomedical Engineering* published by John Wiley & Sons Ltd.

of diffusion and convection to clear substances out of the brain. There is also consensus that interstitial fluid (ISF) within the brain can exchange with cerebrospinal fluid (CSF) surrounding the brain to transfer waste products. However, there is disagreement on where the CSF/ISF-exchange occurs, in which direction it occurs, the forces involved in transport and how much of the transport within the brain that can be explained by diffusion alone.²⁻⁷

Interchange between CSF and ISF is also an important mechanism for effective drug delivery to the brain by substance administration in the SAS. The blood-brain-barrier is one of the main obstacles for this type of drug delivery to the brain.⁸ The contrast agent gadobutrol has been used as a CSF tracer to assess transport into brain tissue, and possibly across the blood-brain-barrier.^{9,10} However, human studies are sparse, the number of subjects are limited, and MRI cannot directly capture phenomena on the microscale (eg, transport across the blood-brain-barrier).

Mathematical modeling is therefore a compelling tool to investigate CSF/ISF-exchange related to brain clearance or drug delivery. Computational studies¹¹⁻¹³ investigating CSF flow in paravascular spaces (PVS) have not been able to reproduce the average flow velocities reported in experimental studies with microspheres.^{14,15} Within the interstitium, both experimental,⁷ and computational¹⁶ studies have concluded that diffusion dominates convection. Still, tracer experiments in human beings^{9,10} show transport of gadobutrol to the brain that is unlikely to be explained by diffusion alone.¹⁷

The aforementioned mathematical models involve parameters that are usually unknown to some degree. Therefore, evaluating the models' parameter sensitivities is vital to give confidence in the model predictions and conclusions. For instance, brain tissue permeability range over several orders of magnitude in the literature.^{16,18} Recently, Fultz et al¹⁹ found coherent oscillations of electrophysiological waves, blood flow and cerebrospinal fluid flow during non-rapid eye movement sleep, suggesting that these processes are all interlinked. Full-scale modeling of CSF/ISF flow in the brain thus introduces several more interlinked parameters related to these additional processes. Previous studies in the literature that have included extensive parameter sensitivity analyses have been computationally cheap,^{20,21} allowing for parameter space exploration within a reasonable amount of time. However, when the model of interest is expensive to simulate, more advanced uncertainty quantification (UQ) methods are needed.

Standard Monte Carlo (MC) methods have successfully been applied to simulations of, for example, the cardiovascular system,²²⁻²⁴ and to brain solute transport.¹⁷ However, when working with physiologically realistic, MRI-derived geometries, standard MC methods are typically prohibitively costly and more advanced methods, for example, quasi Monte Carlo (QMC) or multilevel methods (multilevel, multilevel quasi, or multi-fidelity Monte Carlo), are to be preferred. We remark that each of these methods bring their own additional requirements: QMC requires either the output functionals of interest to have low effective dimensionality with respect to the random input and/or the input dimensions to be ordered in order of decaying importance.²⁵ Whether this assumption is satisfied is strongly problem-dependent. On the other hand for Multilevel Monte Carlo (MLMC) to work well, a good statistical coupling must be enforced between the levels and a mesh hierarchy on which a good rate of bias and variance decay can be appreciated is needed.²⁶ Even though such a hierarchy can always be obtained by refining and/or coarsening a given mesh in theory, this is far from trivial in practice. A multi-fidelity Monte Carlo (MFMC)²⁷ approach can be beneficial in this case, since it can incorporate in the hierarchy low-fidelity models that are still correlated with the fine-mesh simulations, but do not require a computational grid. However, suitable low-fidelity models are not always available in practice. Finally, multilevel quasi Monte Carlo (MLQMC) methods combine the advantages and the requirements of both QMC and MLMC and, as such, may be advantageous only in the cases in which both QMC and MLMC perform well.

In light of the above considerations, the efficiency, feasibility and performance of enhanced Monte Carlo methods such as QMC, MLMC, MFMC, and MLQMC is highly problem-dependent. Moreover, determining a-priori which method that will bring the largest performance benefits for a given problem is highly nontrivial. As the model complexity and dimension grows, choosing the most efficient UQ method becomes even more important. Additionally, knowing when QMC and MLMC both work well provides a model domain for which MLQMC methods can bring substantial additional computational improvements. Generally speaking, correctly setting up a QMC or MLMC method in complex geometries is a difficult task, and to our knowledge QMC and MLMC methods have not yet been employed for UQ in simulations of brain physiology.

In this study, we therefore investigate different enhanced Monte Carlo approaches for quantifying uncertainty in a model with random field coefficients deriving from MRI-studies in humans.^{9,10} The model, and its connection to tracer experiments in humans, has previously been described in detail,¹⁷ as well as the clinical relevance of the simulated tracer experiments in Ref. 10. We introduce and evaluate the relative performance of different Monte Carlo methods, namely standard, quasi and multilevel Monte Carlo methods^{29,30} as implemented in the FEMLMC library. The software

platform used is comprised of generally available open-source software libraries (FEniCS,³¹ libsupermesh,^{32,33} and FEMLMC³⁴).

Our findings show that QMC outperforms standard Monte Carlo, but only when the number of (in this case infinite-dimensional) random inputs is small, bringing a 10-fold improvement in the computational cost. On the other hand, MLMC always outperforms both standard Monte Carlo and QMC, and is two orders of magnitude faster than standard Monte Carlo, showing that MLMC should be the method of choice when solving brain transport problems. Our results further suggest that problems with a small number of random field inputs might be amenable to a MLQMC treatment.

This paper is organized as follows. In Section 2, we give a brief overview of mathematical, numerical and algorithmic background. In Section 3, we present the baseline stochastic model for brain tracer movement, as originally introduced in Ref. 20, and in Section 4 we formalize two UQ test problems building on this baseline model. We detail the numerical and computational solution algorithms in Section 5 before presenting numerical results in Section 6. We discuss our findings in Section 7 and conclude in Section 8.

2 | MONTE CARLO METHODS AND STOCHASTIC SAMPLING

2.1 | Preliminaries

In what follows we indicate by $\mathbf{x} \in G \subset \mathbb{R}^3$ a given spatial point in a domain G , with $t > 0$ a time point, and with ω a given probabilistic event, living in a sample space Ω . For example, we might indicate with $f(t, \mathbf{x})$ a generic function of time and space and with $z(\omega)$ a generic random variable of expected value $\mathbb{E}[z]$ and variance $\mathbb{V}[z]$. Additionally, we will consider random functions of space, varying both in space and across random realizations. These are called random fields and we use the notation, for example, $u(\mathbf{x}, \omega)$. More formally, a random field $u(\mathbf{x}, \omega)$ is a collection of random variables such that each point value $u(\mathbf{x}, \cdot)$ is a random variable for every \mathbf{x} and $u(\cdot, \omega)$ is a function of space for fixed ω .

For Gaussian fields, all point values (which are random variables) are jointly Gaussian, and a Gaussian field $u(\mathbf{x}, \omega)$ is uniquely determined by prescribing a mean function $\mu(\mathbf{x}) = \mathbb{E}[u(\mathbf{x}, \omega)]$ and a covariance function $C(\mathbf{x}, \mathbf{y}) = \mathbb{E}[(u(\mathbf{x}, \omega) - \mu(\mathbf{x}))(u(\mathbf{y}, \omega) - \mu(\mathbf{y}))]$. A ubiquitous family of Gaussian fields is the Matérn family: a Matérn field is a Gaussian field with covariance of the type

$$C(\mathbf{x}, \mathbf{y}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa r)^\nu \mathcal{K}_\nu(\kappa r), \quad r = \|\mathbf{x} - \mathbf{y}\|_2, \quad \kappa = \frac{\sqrt{8\nu}}{\lambda}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^d, \quad (1)$$

where σ^2 , ν , $\lambda > 0$ is the variance, smoothness parameter and correlation length of the field, respectively. Here $\Gamma(x)$ is the Euler Gamma function and $\mathcal{K}_\nu(x)$ is the modified Bessel function of the second kind. The σ^2 , ν and λ parameters determine the characteristics of the random field. For instance, in our model the random field represents CSF/ISF-flow and the correlation length was chosen to match approximately the distance between larger arteries and veins, to be able to model CSF inflow along arteries and outflow along veins.⁴

In this paper we consider the solution of a convection-diffusion-reaction equation with random fields as coefficients for the movement of tracer within the brain. The convection-diffusion-reaction equation reads as: find the tracer concentration $c = c(t, \mathbf{x}, \omega)$ for $\mathbf{x} \in G$, $\omega \in \Omega$ and $t \geq 0$ such that

$$\dot{c}(t, \mathbf{x}, \omega) + \nabla \cdot (\mathbf{v}(\mathbf{x}, \omega) c(t, \mathbf{x}, \omega)) - \nabla \cdot (D^*(\mathbf{x}, \omega) \nabla c(t, \mathbf{x}, \omega)) + rc(t, \mathbf{x}, \omega) = 0. \quad (2)$$

Here, the domain $G \subset \mathbb{R}^3$ represents the brain, the superimposed dot represents the time derivative, D^* is the effective diffusion coefficient of the tracer in the tissue, \mathbf{v} represents a convective fluid velocity and $r \geq 0$ is a drainage coefficient. Further details, including boundary and initial conditions, will be presented in Section 3.

The solution c is random and varies in time and space as well, while the coefficients $\mathbf{v}(\mathbf{x}, \omega)$ and $D^*(\mathbf{x}, \omega)$ are random fields. A typical assumption is that one is interested in computing one (or more) output functional of interest $Q(\omega) = Q(c)$, for example, Q could be the spatial average of c over a region of interest. Typically, solving (2) then means

to compute the expected value of $Q(\omega)$. Note that most other statistics (such as the variance or the cumulative density function [CDF]) can be rewritten as an expectation and thus computed analogously.

A basic method for solving (2) is the standard Monte Carlo (MC) method, for which the expectation is approximated by the sample average

$$\mathbb{E}[Q] \approx \hat{Q} = \frac{1}{N} \sum_{i=1}^N Q(\omega_i), \quad (3)$$

obtained with $0 < N \in \mathbb{N}$ samples of $Q(\omega)$. Note that each sample of Q requires computing samples of the coefficients $\mathbf{v}(\mathbf{x}, \omega)$, $D^*(\mathbf{x}, \omega)$ and the corresponding solution of (2). Alas, the MC method converges slowly in terms of the number of samples N , namely at a rate of $\mathcal{O}(N^{-1/2})$. This makes standard MC quite expensive: assuming an ε^2 tolerance for the mean square error (MSE) $\mathbb{E}[(Q - \hat{Q})^2]$ and a cost per sample of ε^{-q} for some positive problem-specific number q , standard MC has a total cost complexity of $\mathcal{O}(\varepsilon^{-2-q})$, see Ref 26. In practice, more advanced methods, such as Quasi Monte Carlo (QMC), multilevel Monte Carlo (MLMC) and multilevel quasi Monte Carlo (MLQMC), are to be preferred. In what follows we give a brief description of the QMC and MLMC methods.

2.2 | Quasi Monte Carlo

We now give a quick overview of the quasi Monte Carlo method and we refer the reader to the book by Lemieux³⁵ for a more in-depth description. The advantage of QMC with respect to standard MC is that in QMC methods, the convergence rate with respect to the number of samples N is improved from $\mathcal{O}(N^{-1/2})$ to $\mathcal{O}(N^{-1+\varepsilon})$ for any $\varepsilon > 0$. At the heart of QMC is the approximation or reinterpretation of the expectation as a high-dimensional integral over the unit hypercube:

$$\mathbb{E}[Q] = \int_{\Omega} Q(\omega) d\mathbb{P}(\omega) \approx \int_{[0,1]^s} H(\mathbf{x}) d\mathbf{x}, \quad (4)$$

where \mathbb{P} is a suitable probability measure, $H(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^s$ is some suitable function and s is typically the dimensionality of the random input needed to sample Q . The QMC method can then be expressed as a quadrature rule over $[0, 1]^s$ with equal weights, approximating the right-hand side integral with

$$I = \int_{[0,1]^s} H(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{n=1}^N H(\mathbf{x}_n) = I_N, \quad (5)$$

with $\mathbf{x}_n \in \mathbb{R}^s$. Choosing the $\{\mathbf{x}_n\}_{n=1}^N$ uniformly at random results in a convergence rate of $\mathcal{O}(N^{-1/2})$. However, there exist deterministic point sequences, so-called low-discrepancy point sequences that can yield an improved rate of $\mathcal{O}(N^{-1}(\log N)^s)$.³⁶ This is the key idea of QMC methods: estimating the integral I by a low-discrepancy sequence.

Now, let $\mathbb{V}[Q]$ be the variance of Q . In standard MC, the statistical error is $\mathcal{O}(\mathbb{V}[Q]/N)^{1/2}$ (owing to the central limit theorem) and $\mathbb{V}[Q]$ can be estimated by taking the sample variance of Q . However, in standard QMC, the point sequence $\{\mathbf{x}_n\}_{n=1}^N$ is deterministic and therefore there is no notion of statistical error available. For this reason, a practical estimate for the approximation error introduced by QMC is not available. A common solution is to randomize the sequence (while still preserving the hypercube-covering properties) in order to retain a measure of estimator variability (see eg, Chapter 6 of Ref. 35 for an overview and Ref. 37 for a comparison of different randomization strategies). This yields a randomized QMC method: let $\{\hat{\mathbf{x}}_{n,m}(\omega)\}_{n=1, m=1}^{n=N, m=M}$ be M independent randomizations of a low-discrepancy sequence $\{\mathbf{x}_n\}_{n=1}^N$, then randomized QMC estimates I as

$$I \approx \hat{I}_{M,N}(\omega) = \frac{1}{M} \sum_{m=1}^M I_N^m(\omega) = \frac{1}{M} \sum_{m=1}^M \left(\frac{1}{N} \sum_{n=1}^N H(\hat{\mathbf{x}}_{n,m}(\omega)) \right), \quad (6)$$

where each of the $I_N^m(\omega)$ are now random and a confidence interval can therefore be estimated provided M is large enough. In this paper we use $M = 32$. Assuming a fixed M , a given MSE tolerance ε^2 , a $O(\varepsilon^{-q})$ cost per sample and a QMC convergence order of $O(N^{-1+\varepsilon})$ for any $\varepsilon > 0$, the total cost of randomized QMC is $O(\varepsilon^{-q-1/(1-\varepsilon)})$. If we take ε to be small, this is almost ε^{-1} times better than standard MC.

Remark 1. The $(\log N)^s$ can dominate the early convergence behavior of QMC methods, in which case the suboptimal $O(N^{-1/2})$ convergence is observed initially. This behavior is typically exacerbated when the random input dimensionality s is particularly large (eg, in applications with infinite-dimensional random inputs such as random fields), to the extent that the faster rate is never observed in practice. However, this is not always the case: if the QMC integrand $H(\mathbf{x})$ is inherently lower dimensional in the sense that it can be well approximated by a function only depending on $\bar{s} \ll s$ input dimensions, then it is possible to replace s with \bar{s} and the transition to a faster rate will occur earlier. This is the principle of low-effective dimensionality, first introduced by Caflish et al in Ref. 13, and is fundamental when using QMC in high dimensions. Thus, for good QMC convergence it is important to either reduce the integrand dimensionality or to order the integration variables in order of decaying importance such that the improved convergence rates can be attained.

2.3 | Multilevel Monte Carlo

The multilevel Monte Carlo method was first introduced by Heinrich in Ref. 77 for parametric integration and subsequently popularized by Giles for stochastic differential equations in Ref. 32. MLMC works under the assumption that one can compute a hierarchy $\{Q_\ell(\omega)\}_{\ell=1}^L$ of different approximations of the output functional Q of increasing accuracy. For instance, one could consider solving (2) on a hierarchy of computational meshes of size decreasing with ℓ . At the heart of MLMC is the expansion of $\mathbb{E}[Q]$ into the telescoping sum

$$\mathbb{E}[Q] \approx \mathbb{E}[Q_L] = \sum_{\ell=1}^L \mathbb{E}[Q_\ell - Q_{\ell-1}], \quad Q_0 \equiv 0. \quad (7)$$

Approximating each expectation in the sum on the right-hand side with standard MC then yields the MLMC estimator:

$$\mathbb{E}[Q_L] \approx \hat{Q} = \sum_{\ell=1}^L \left[\frac{1}{N_\ell} \sum_{n=1}^{N_\ell} (Q_\ell - Q_{\ell-1})(\omega_\ell^n) \right], \quad (8)$$

in which a key element is that the term $(Q_\ell - Q_{\ell-1})(\omega_\ell^n) = Q_\ell(\omega_\ell^n) - Q_{\ell-1}(\omega_\ell^n)$ is sampled by the same event ω_ℓ^n . This aspect is referred to as the *MLMC coupling* and is essential for the improved performance of MLMC over MC. In fact, while the variance of each Q_ℓ might be large, the variance of the difference $Q_\ell - Q_{\ell-1}$ is typically much smaller due to the strong correlation between the two terms. For this reason, while the estimation of each $\mathbb{E}[Q_\ell - Q_{\ell-1}]$ term still occurs at a $O(N^{-1/2})$ rate, the number of samples needed to achieve a given statistical error tolerance is smaller and decreases with ℓ . More formally, the convergence of MLMC is ensured by the following Theorem³⁹:

Theorem 1. (Ref 26, Theorem 1). Let $Q(\omega)$ be a random variable with finite second moment and let Q_ℓ be its level ℓ approximation. Let Y_ℓ be the (unbiased) MC estimator of $\mathbb{E}[Q_\ell - Q_{\ell-1}]$ on level ℓ , and let C_ℓ and V_ℓ be the expected cost and variance, respectively, of each of the N_ℓ Monte Carlo samples needed to compute Y_ℓ . If the estimators Y_ℓ are independent and there exist positive constants $\alpha, \beta, \gamma, c_1, c_2, c_3$, such that $\alpha \geq \frac{1}{2} \min(\beta, \gamma)$ and

$$|\mathbb{E}[Q_\ell - Q]| \leq c_1 2^{-\alpha\ell}, \quad V_\ell \leq c_2 2^{-\beta\ell}, \quad C_\ell \leq c_3 2^{\gamma\ell}, \quad (9)$$

then there exist a positive constant c_4 such that, for all $\varepsilon < \varepsilon^{-1}$, there is a level number L and number of samples N_ℓ , such that the mean square error (MSE) of the MLMC estimator $\hat{Q} = \sum_{\ell=1}^L Q_\ell$ is bounded:

$$\text{MSE} = \mathbb{E} \left[(\hat{Q} - \mathbb{E}[Q])^2 \right] \leq \varepsilon^2, \quad (10)$$

and its total computational complexity is bounded:

$$\mathbb{E}[C_{\text{tot}}] \leq \begin{cases} c_4 \varepsilon^{-2}, & \beta > \gamma, \\ c_4 \varepsilon^{-2} (\log \varepsilon)^2, & \beta = \gamma, \\ c_4 \varepsilon^{-2 - (\gamma - \beta)/\alpha}, & \beta < \gamma. \end{cases} \quad (11)$$

Remark 2. The MLMC parameters α , β , and γ must be estimated if not known a priori. However, for PDE applications, as it is the case here, a priori error estimates are typically available. The optimal number of samples on each level N_ℓ and the maximum level L needed can similarly be estimated.

Remark 3. Theorem 1 assumes that one can increase the maximum level L at will. In practical applications this might not always be possible, for example, when the computational resources available are limited.

In this study, we modify the original MLMC algorithm by weighting the relative importance of bias and statistical error as introduced by Haji-Ali et al.³⁸ The MSE of the MLMC estimator is given by

$$\text{MSE} = \hat{V} + \mathbb{E}[\hat{Q} - Q]^2, \quad (12)$$

where \hat{Q} is the MLMC estimator with variance \hat{V} . To ensure that $\text{MSE} \leq \varepsilon^2$, we enforce the bounds,

$$\hat{V} \leq (1 - \theta) \varepsilon^2, \quad \mathbb{E}[\hat{Q} - Q]^2 \leq \theta \varepsilon^2, \quad (13)$$

where $\theta \in (0, 1)$ is a weight balancing the two terms so as to obtain comparable error reduction. Small values of θ reduce the number of samples needed and are therefore preferred when the bias is small. Conversely, large values are beneficial when the bias is large as they allow to achieve smaller tolerances without adding finer levels to the hierarchy.

2.4 | Gaussian field sampling techniques

When $Q(\omega)$ depends on a random field (eg, through the coefficients of (2)), an efficient sampling technique is also needed as part of any Monte Carlo method. Unfortunately, sampling generic random fields from a given distribution is normally a prohibitive task. In the specific case in which the field to be sampled is Gaussian, however, the sampling problem becomes tractable, albeit still computationally expensive. Different Gaussian field sampling methods are available in the literature. The most common are: (a) Factorization of the covariance matrix, including Hierarchical matrix approximation⁴⁰⁻⁴³ (b) Karhunen-Loève expansion of the random field (cf. section 11.1 in Ref. 79 and eg, Ref. 78) and (c) the circulant embedding method which uses FFT.⁴⁶⁻⁴⁹

In this paper we use an alternative method known as the stochastic PDE (SPDE) approach for the sake of efficiency.^{29,30,50,51} This sampling strategy is based on approximately drawing realizations of a Matérn field by solving the following elliptic PDE^{50,51}:

$$(I - \kappa^{-2} \Delta)^k u(\mathbf{x}, \omega) = \eta \dot{W}(\cdot, \omega), \quad \mathbf{x} \in \hat{G} \subseteq \mathbb{R}^d, \quad \omega \in \Omega, \quad \nu = 2k - d/2 > 0, \quad (14)$$

where ν is the smoothness parameter of the Matérn covariance to be sampled (cf. Equation (1)), \dot{W} is spatial Gaussian white noise in \mathbb{R}^d , $d \leq 3$, $k > d/4$, and the equality has to hold almost surely and be interpreted in the sense of

distributions. When k is an integer, as we shall assume here, solving (14) is equivalent to solving a second-order PDE k times. The constant $\eta > 0$ is a scaling factor that depends on the Matérn covariance parameters σ , λ , and ν , cf. Ref. 30. The solution to $u(\mathbf{x}, \omega)$ is a Matérn field only if $\hat{G} \equiv \mathbb{R}^d$. Otherwise, for general $\hat{G} \subset \mathbb{R}^d$, u is still a Gaussian field, but its covariance is only approximately Matérn, with the error in the covariance decaying exponentially away from the boundary of \hat{G} .⁵² For this reason in this paper we will always assume that the field sample is needed on a domain $G \subset \hat{G}$ and that we solve (14) on a domain \hat{G} large enough so that this source of error is negligible over G . We impose homogeneous Dirichlet boundary conditions on $\partial\hat{G}$.

Note that the term \dot{W} is random and for each sample of u needed, we must sample \dot{W} from its distribution and solve (14). The sampling of \dot{W} is non-trivial in itself, especially when samples are needed within a MLMC or QMC framework: in the QMC case, a good variable ordering is required to achieve good convergence with respect to the number of samples while in the MLMC case the multilevel coupling must be enforced correctly to obtain a considerable variance reduction. In this paper, we use the sampling techniques developed in Ref. 19 for the MLMC case and in Ref. 18 for the (ML)QMC case to address these requirements. Both these MLMC and (ML)QMC sampling strategies are efficient with optimal cost complexity, making the sampling of u a linear cost and memory operation. This optimal complexity is especially advantageous when dealing with physiologically relevant applications with complex geometries in 3D.

3 | A STOCHASTIC MODEL OF TRACER TRANSPORT IN BRAIN TISSUE

This paper aims to evaluate the numerical and computational performance of different Monte Carlo methods for quantifying uncertainty in stochastic models of tracer transport in brain tissue. In particular, we focus on two comprehensive coefficient models with random diffusion and velocity fields. In addition to reflecting existing hypotheses on ISF tracer transport, these models are designed to offer a suitable challenge to the MLMC and QMC algorithms presented in Refs. 18, 19.

3.1 | The ISF tracer transport equation

We begin by considering the baseline model derived by the authors in Ref. 20, and briefly introduced in Section 2.1, describing the evolution of tracer concentration within the brain parenchyma under uncertainty: find the tracer concentration $c = c(t, \mathbf{x}, \omega)$ for $\mathbf{x} \in G$, $\omega \in \Omega$ and $t \geq 0$ such that

$$\dot{c}(t, \mathbf{x}, \omega) + \nabla \cdot (\mathbf{v}(\mathbf{x}, \omega) c(t, \mathbf{x}, \omega)) - \nabla \cdot (D^*(\mathbf{x}, \omega) \nabla c(t, \mathbf{x}, \omega)) + rc(t, \mathbf{x}, \omega) = 0. \quad (15)$$

As before, $G \subset \mathbb{R}^3$ is the brain parenchyma domain comprised of white and gray matter from the Colin27 human adult brain atlas FEM mesh⁵³ version 2 (Figure 1). The superimposed dot represents the time derivative, D^* is the effective diffusion coefficient of the tracer in the tissue (depending on the tracer free diffusion coefficient and the tissue tortuosity),⁵⁴ \mathbf{v} represents a convective fluid velocity and $r \geq 0$ is a drainage coefficient potentially representing, for example, capillary absorption⁵⁵ or direct outflow to lymph nodes.¹⁰ The parenchymal domain is assumed to not contain any tracer initially: $c(0, \mathbf{x}, \omega) = 0$. This model reflects the MRI-study of Ringstad et al¹⁰ in which 0.5 mL of 1.0 mmol/mL of the radioactive tracer gadobutrol was injected in the spinal canal (intrathecally) of 15 hydrocephalus patients and eight reference subjects.

3.2 | Boundary conditions

Let the brain boundary $\partial G = \partial G_S \cup \partial G_V$, with ∂G_S representing the interface between the brain parenchyma and the subarachnoid space (SAS), and ∂G_V representing the interface between the brain parenchyma and the cerebral ventricles, respectively. We assume the tracer concentration on the SAS interface to be known and impose no ventricular outflux. As boundary conditions for (15), we thus prescribe

$$c = g(t, \mathbf{x}) \text{ on } \partial G_S, \quad (16)$$

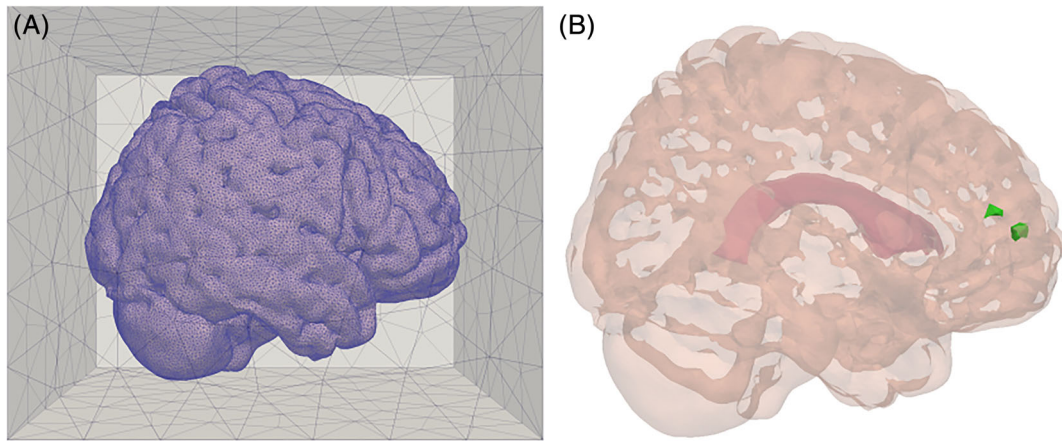


FIGURE 1 The computational domains and points of interest. A, The coarsest computational mesh, embedded in the auxiliary box used to sample the Matérn fields. B, The domain representing the brain parenchyma. The lateral ventricles are shown in red, while the parenchyma is pink. Two smaller regions of interest are marked in green: The leftmost green region, S_w , is within the white matter, while the rightmost green region, S_g , is within the gray matter

$$D^* \nabla c \cdot \mathbf{n} = 0 \text{ on } \partial G_V, \quad (17)$$

where \mathbf{n} is the unit normal vector pointing outward from ∂G . The function $g(t, \mathbf{x})$ models the movement of tracer starting from the lower cranial SAS and traveling upward in the CSF surrounding the brain as observed in the study by Ringstad et al,¹⁰ and the form

$$\begin{aligned} g(t, \mathbf{x}) &= c_{\text{CSF}}(t) h(t, \mathbf{x}), \\ h(t, \mathbf{x}) &= \left(0.5 + \frac{1}{\pi} \arctan(-a(x_3 - z_0 - v_z t)) \right), \end{aligned} \quad (18)$$

where $\mathbf{x} = [x_1, x_2, x_3]$ and $c_{\text{CSF}}(t)$ is the average tracer concentration in the SAS, while $h(t, \mathbf{x})$ represents its spatial distribution. The variable $v_z = 1.5 \times 10^{-5}$ m/sec is the speed of tracer movement upwards in the SAS, while $a = 20 \text{ m}^{-1}$ reflects the gradient of tracer concentration from the lower to the upper cranial SAS. The value $z_0 = -0.2 \text{ m}$ is the initial distance from the lateral ventricles reached by the tracer at $t = 0$.

The average SAS tracer concentration c_{CSF} is modeled as follows. Let $n_0 = 0.5 \text{ mmol}$ be the total amount of tracer initially injected in the CSF and let $V_{\text{CSF}} = 140 \text{ mL}$ be the total CSF volume in the human SAS and ventricles.⁵⁶ Then, the average concentration in the SAS right after injection is given by $c_{\text{CSF}}(0) = 0.5 \text{ mmol}/140 \text{ mL} = 3.57 \text{ mol/m}^3$. Assuming conservation of tracer molecules, the total amount of tracer in the brain and in the SAS plus or minus the tracer otherwise absorbed or produced is constant in time, and is equal to the initial amount n_0 . This observation gives the approximate relation

$$\int_G \bar{c}(t, \mathbf{x}) d\mathbf{x} + c_{\text{CSF}}(t) V_{\text{CSF}} + \int_0^t \int_G r\bar{c}(\tau, \mathbf{x}) d\mathbf{x} d\tau = n_0. \quad (19)$$

where, for simplicity, \bar{c} is given by a deterministic solution of (15) with boundary conditions (18) in which all the random coefficients are replaced by their average. Solving for c_{CSF} , we finally let

$$c_{\text{CSF}}(t) = \frac{1}{V_{\text{CSF}}} \left(n_0 - \int_G \bar{c}(t, \mathbf{x}) d\mathbf{x} - \int_0^t \int_G r\bar{c}(\tau, \mathbf{x}) d\mathbf{x} d\tau \right). \quad (20)$$

3.3 | Quantities of interest

We consider different output quantities (functionals) describing the characteristics of tracer movement within the brain. For each time $\tau = 30k$ min for $k = 1, \dots, 48$ (from half an hour from injection to 1 day after), we consider the total amount of tracer in the gray matter Q_g and in the white matter Q_w :

$$Q_g(\tau, \omega) = \int_{D_g} c(\tau, \mathbf{x}, \omega) d\mathbf{x}, \quad Q_w(\tau, \omega) = \int_{D_w} c(\tau, \mathbf{x}, \omega) d\mathbf{x}. \quad (21)$$

Additionally, we consider two localized concentration measures: the average tracer concentrations q_g and q_w in two smaller regions, one within the gray matter S_g and one within the white matter S_w respectively:

$$q_g(\tau, \omega) = \frac{1}{V_g} \int_{S_g} c(\tau, \mathbf{x}, \omega) d\mathbf{x}, \quad q_w(\tau, \omega) = \frac{1}{V_w} \int_{S_w} c(\tau, \mathbf{x}, \omega) d\mathbf{x}, \quad (22)$$

where V_g and V_w is the volume of the gray and white matter subregions, respectively. The size and location of these subregions are shown in Figure 1B.

4 | COEFFICIENT MODELS

The effective diffusion coefficient of a solute (or tracer) and the velocity field are heterogeneous^{6,57-59} within the parenchyma and also vary from individual to individual. To account for both these types of variation and for the uncertainty in the coefficient magnitude we model them as random variables or fields.

4.1 | Diffusion coefficient

Let $D_{\text{Gad}}^* = 1.2 \times 10^{-10}$ m/s² be the average parenchymal gadobutrol diffusivity.⁹ We model the effective diffusion coefficient as

$$D^*(\mathbf{x}, \omega) = 0.25 \times D_{\text{Gad}}^* + D_\gamma^*(\mathbf{x}, \omega), \quad (23)$$

where D_γ^* is a random field such that for each fixed $\mathbf{x} \in G$, $D_\gamma^*(\mathbf{x}, \omega)$ is a gamma-distributed random variable with shape $k = 3$ and scale $\theta = 0.75 \times D_{\text{Gad}}^*/k$. This choice of parameters ensures the positivity of D^* with probability 1. Furthermore, it reflects the average value and variability reported in the literature, namely we have $\mathbb{E}[D^*] = D_{\text{Gad}}^*$, with values larger than 3 times the average being attained with very low probability.^{9,54} The probability distribution of $D^*(\mathbf{x}, \cdot)$ is shown in Figure 2B.

To sample D_γ^* from its distribution, we first sample a Matérn field $X(\mathbf{x}, \omega)$ using the techniques presented in Refs. 18, 19, and then transform it into a gamma random field by using a copula.⁶⁰ This consists in setting $D_\gamma^*(\mathbf{x}, \omega) = F^{-1}(\Phi(X(\mathbf{x}, \omega)))$, where F^{-1} is the inverse cumulative density function (CDF) of the target (gamma) distribution, Φ is the CDF of the standard normal distribution and $X(\mathbf{x}, \omega)$ is a standard (zero mean, unit variance) Matérn field with smoothness parameter $\nu = 2.5$ and correlation length $\lambda = 0.01$ m, cf. (1). Note that Φ maps any standard normal random variable to a standard uniform random variable and that F^{-1} maps any standard uniform random variable to the target distribution, hence the function $F^{-1}(\Phi(x))$ maps standard random variables to the target gamma distribution. Samples of $D_\gamma^*(\mathbf{x}, \omega)$ obtained this way will preserve the same Spearman correlation and smoothness properties of $X(\mathbf{x}, \omega)$, but will present a different covariance structure.⁶⁰

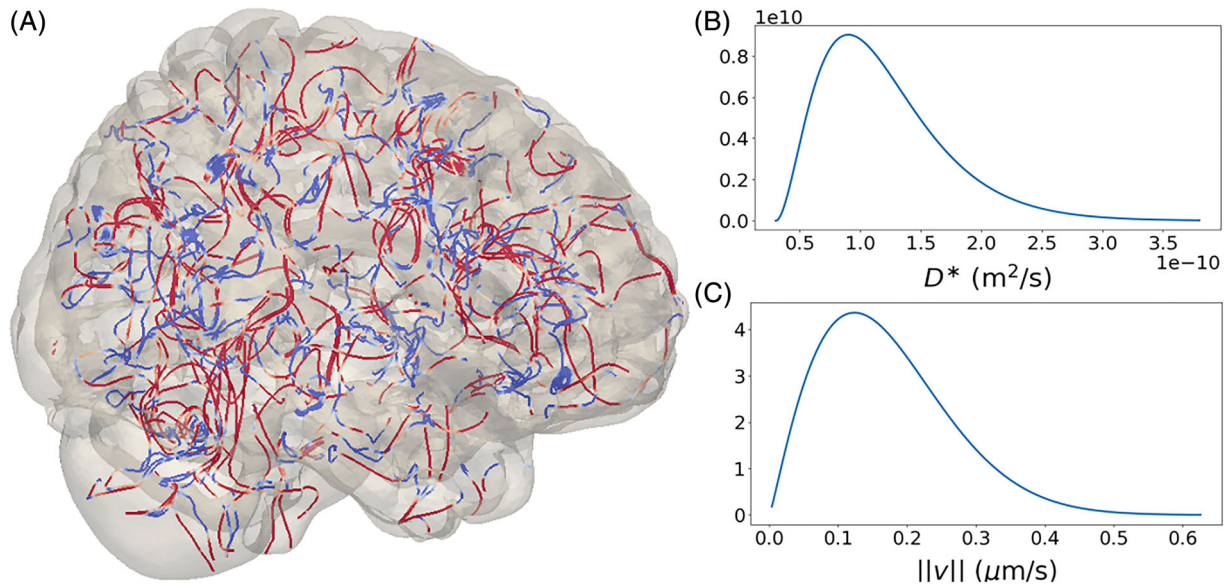


FIGURE 2 Stochastic aspects of diffusion and velocity fields. A, Streamlines of the velocity field \mathbf{v}_{base} , representing a random distribution of blood vessels. Colors indicate velocity magnitude, and an arbitrary small scaling range is chosen for visual purposes. B, Probability density of the diffusion coefficient D^* . C, Probability density of the glymphatic circulation velocity magnitude $\|\mathbf{v}\|_2$ cf. (25)

4.2 | Velocity and drainage coefficients

We now turn to define two models (Model 1, 2) for brain tissue fluid movement and clearance. Both models are combined with the random diffusion field defined by (23). For further physiological considerations, see, for example,¹⁷

4.2.1 | Model 1—Glymphatic velocity model with directionality

For Model 1, we model fluid transport along paravascular spaces in direct communication with the SAS⁶ under the following assumptions. (a) Substantial changes in the velocity field occurs at a distance proportional to a characteristic distance (correlation length λ) between an arteriole and a venule. (b) The velocity field can be represented as a sum of a glymphatic velocity field associated with arterioles and venules and a velocity field represents a directional movement due to larger blood vessel structures. (c) Almost no fluid is filtrated or absorbed by capillaries, thus \mathbf{v}_{base} is divergence-free, while \mathbf{v}_{dir} only has a small net flow of 0.007 mL/min out of the parenchyma. We then let the drainage term $r = 0$ in (15) and define the stochastic velocity field

$$\mathbf{v}(\mathbf{x}, \omega) = \mathbf{v}_{\text{base}}(\mathbf{x}, \omega) + \mathbf{v}_{\text{dir}}(\mathbf{x}). \quad (24)$$

The stochastic glymphatic velocity field \mathbf{v}_{base} is given by

$$\mathbf{v}_{\text{base}}(\mathbf{x}, \omega) = v_{\text{avg}} \cdot \bar{\eta}(\lambda) \sqrt{U(\omega)} \left(\nabla \times \begin{bmatrix} X(\mathbf{x}, \omega) \\ Y(\mathbf{x}, \omega) \\ Z(\mathbf{x}, \omega) \end{bmatrix} \right), \quad (25)$$

where $\bar{\eta}(\lambda) = \lambda/\sqrt{8}$ is a scaling constant chosen such that the magnitude of \mathbf{v} (denoted $\|\mathbf{v}\|$) satisfies $\mathbb{E}[\|\mathbf{v}\|^2]^{1/2} = v_{\text{avg}}$, $U(\omega)$ is an independent standard uniform random variable and $X(\mathbf{x}, \omega)$, $Y(\mathbf{x}, \omega)$ and $Z(\mathbf{x}, \omega)$ are standard i.i.d. Matérn fields with $\nu = 2.5$ and correlation length $\lambda = 1020 \mu\text{m}$. Taking the curl of the random vector field $[X, Y, Z]^T$ ensures that \mathbf{v}_{base} is divergence free. A sample of the streamlines of \mathbf{v}_{base} is shown in Figure 2A while its velocity magnitude distribution is shown in Figure 2C. The deterministic second term \mathbf{v}_{dir} represents a directional velocity field induced by large vascular structures¹⁷ and is given by

$$\mathbf{v}_{\text{dir}}(\mathbf{x}) = -v_f \begin{bmatrix} \arctan(15x_1)(|x_1| - 0.1) \\ \arctan(15x_2)(|x_2| - 0.1) \\ -0.9x_3 + 0.06 - \sqrt{x_1^2 + x_2^2} \end{bmatrix}, \quad (26)$$

where $v_f = 2 \times 10^{-6}$ m/s.

Remark 4. We now briefly show that $\mathbb{E}[\mathbf{v}_{\text{base}}] = \mathbf{0}$ and hence $\mathbb{E}[\mathbf{v}] = \mathbf{v}_{\text{dir}}(\mathbf{x})$. In fact, note that the partial derivative $\partial X / \partial x_i$ of a zero-mean Gaussian field $X(\mathbf{x}, \omega)$ with a twice differentiable covariance $C(\mathbf{x}, \mathbf{y})$ is still a zero-mean Gaussian field with covariance given by $\partial^2 C(\mathbf{x}, \mathbf{y}) / (\partial x_i \partial y_i)$ (cf. section 2.3 in Ref. 76). The curl components of the field within the brackets in (25) are therefore just sums of independent Gaussian fields, and hence Gaussian as well. Similarly, this also shows that the covariance of \mathbf{v} has the same correlation length as the Matérn fields X , Y and Z since the second derivative of a Matérn covariance (cf. (1)) has the same correlation length as the original covariance function, although it is not Matérn anymore.

4.2.2 | Model 2—Capillary filtration model with arterial inflow and sink term

For Model 2, we consider an alternative velocity representation in which CSF enters the brain parenchyma along spaces surrounding penetrating arteries.^{6,14,15,61} In this case, the velocity field is taken to be

$$\mathbf{v}(\mathbf{x}, \omega) = \bar{v}(\omega) \exp\left(-\frac{3(R - \|\mathbf{x} - \mathbf{x}_c\|)^2}{R^2 - (R - \|\mathbf{x} - \mathbf{x}_c\|)^2}\right) (\mathbf{x}_c - \mathbf{x}), \quad (27)$$

representing a net flow of CSF into the brain. The flow field is radially symmetric and directed inwards from the outer SAS to a spherical region of radius $R = 8$ cm around a center point \mathbf{x}_c within the lateral ventricles. Here $\bar{v}(\omega)$ is a gamma-distributed random variable chosen such that the probability distribution of the velocity magnitude is comparable to that of Model 1. The shape parameter is $k = 2$ and the scale parameter is set such that again $\mathbb{E}[|\mathbf{v}|^2]^{1/2} = v_{\text{avg}}$. Note that in this model $\mathbb{E}[\mathbf{v}] \neq \mathbf{0}$ and the main source of uncertainty is the random variable $(\bar{v}(\omega))$ rather than the spatially dependent random field.

Finally, we set a nonzero sink coefficient of $r = 1 \times 10^{-5} \text{ s}^{-1}$, to model the assumption that ISF is drained along some alternative outflux route within the brain parenchyma. The value of r chosen corresponds to a 40% drainage of the injected tracer over 48 hours.

5 | NUMERICAL SOLUTION OF THE STOCHASTIC MODELS

More advanced Monte Carlo methods, such as QMC, MLMC and MLQMC, are known to outperform standard Monte Carlo methods, given an appropriate problem setting. In particular, QMC requires either the output functional to be of low effective dimensionality with respect to the random input, or the input dimensions to be ordered in order of decaying importance.²⁵ Even though our QMC method²⁹ is designed to expose the leading-order dimensions in each random field input, hence providing a suitable variable ordering, it is not known whether this strategy would prove to be effective for the problem at hand. In fact, two main challenges arise: (a) The state Equation (15) must be solved on a complicated 3D geometry, and (b) the random input dimensionality is large. The latter point is especially relevant in connection with Model 1 in which 4 infinite-dimensional random fields appear as coefficients. Both these challenges significantly increase the input dimensionality, possibly affecting QMC performance.

On the other hand, MLMC brings a different set of performance requirements. A good MLMC coupling is ensured by our coupling technique,³⁰ but the technique also hinges on the construction of an appropriate mesh hierarchy on which a good rate of decay of bias and variance can be appreciated. Even though such a hierarchy can always be obtained by increasing the mesh refinement in theory, the availability of computational resources and time may limit the maximal refinement level in practice. As it is a priori unclear whether QMC and/or MLMC actually bring any

significant advantages with respect to standard MC when solving (15), we here evaluate both algorithms to determine which if any method performs the better.

We refer the reader to the book by Lemieux³⁵ and to the review article by Giles²⁶ for further information about QMC and MLMC methods. In what follows, we detail the numerical approach adopted for the solution of (15).

5.1 | Weak form and discretization

We solve (15) numerically using the finite element method (FEM). Let $H_S^1(G)$ be the standard Sobolev space of weakly differentiable functions that are zero on the boundary S .⁶² For $c, s \in H^1(G)$, we define

$$a(c, s) = (\operatorname{div}(\mathbf{v}c), s) + (D^* \nabla c, \nabla s) + (rc, s), \quad (28)$$

whereby (\cdot, \cdot) we indicate the standard $L^2(G)$ inner product:

$$(c, s) = \int_G cs \, d\mathbf{x}. \quad (29)$$

After a (second-order) implicit mid-point time discretization with time step Δt , the continuous weak form of (15) reads: find $c^n \in H^1(G)$ such that for all $s \in H_S^1(G)$ and a.s.,

$$(c^n - c^{n-1}, s) + \frac{\Delta t}{2} (a(c^n, s) + a(c^{n-1}, s)) = 0, \quad (30)$$

$$c^0 \equiv 0, \quad c^n = c_{\text{CSF}}^n h(t^n, \mathbf{x}) \text{ on } \partial G_S, \quad (31)$$

where c^n thus represents an approximation to $c(t^n, \cdot)$ with $t^n = n\Delta t$ for $n = 0, \dots, n_T - 1$ and $n_T - 1 = T/(\Delta t)$ and c_{CSF}^n is an approximation of $c_{\text{CSF}}(t^n)$, defined in (20). We approximate c_{CSF}^n by approximating the time integral in (20) with the trapezoidal rule:

$$c_{\text{CSF}}(t^n) \approx c_{\text{CSF}}^n = \frac{1}{V_{\text{CSF}}} \left(n_0 - \int_G \bar{c}^n \, d\mathbf{x} - \frac{\Delta t}{2} \left(2 \sum_{i=1}^{n-1} \int_G r \bar{c}^i \, d\mathbf{x} + \int_G r \bar{c}^n \, d\mathbf{x} \right) \right). \quad (32)$$

We note that the expression on the right-hand side is known as \bar{c} can be pre-computed once. This decoupling results in a second-order scheme in time for c . To compute \bar{c} we adopt the same discretization, but we approximate c_{CSF}^n explicitly (by replacing n with $n - 1$ in the right-hand side of (32)) thus avoiding the non-local, implicit boundary condition. This approximation gives to a first-order scheme for \bar{c} , for which we compensate by computing the approximation using a very small time step ($\Delta t_{\bar{c}} = 30 \times 2^{-6}$ min) and we solve for \bar{c} on the finest mesh available (see later in this section for a description of the meshes and time step sizes used).

We discretize (30) in space using the FEM. Given a FEM approximation subspace $V_h \subset H_S^1(G)$, the fully discrete weak form of (15) reads: find $c_h^n \in V_h$ such that, for all $s \in V_h$ and a.s.,

$$(c_h^n - c_h^{n-1}, s) + \frac{\Delta t}{2} (a(c_h^n, s) + a(c_h^{n-1}, s)) = 0, \quad (33)$$

$$c_h^0 \equiv 0, \quad c_h^n = c_{\text{CSF}}^{n,h} h(t^n, \mathbf{x}) \text{ on } \partial G_S, \quad (34)$$

where $c_{\text{CSF}}^{n,h}$ is given by (32) in which c^{n-1} and c^i are replaced by c_h^{n-1} and c_h^i respectively. We let V_h be composed of piecewise linear continuous Lagrange elements defined relative to a mesh \mathcal{T}_h of G of mesh resolution h .

TABLE 1 Number of cells and vertices of the Colin27 mesh and of the meshes of the MLMC hierarchy

	Colin27	$\ell = 1$	$\ell = 2$	$\ell = 3$
n cells	249 361	1 994 888	15 959 104	127 672 832
n vertices	55 066	391 559	2 905 985	22 282 705

5.2 | Meshes and time steps

We discretize the domain G by using various refinements of the Colin27 human adult brain atlas simplicial mesh⁵³ (version 2). We construct a multilevel hierarchy in which the coarsest level is given by one uniform refinement of the original brain mesh and the other 2 levels are obtained through uniform refinement. On level ℓ , we fix $(\Delta t)_\ell = 15 \times 2^{-\ell}$ min and we terminate the simulations after $T = 1$ day. The numbers of cells and vertices of the meshes in the hierarchy are given in Table 1. The Matérn fields are sampled on an extended domain, i.e. a mesh of a larger box domain \hat{G} of size sufficiently large to make the domain truncation error negligible (dimensions $0.16 \times 0.21 \times 0.17$ m³),⁵² and then restricted to the brain mesh. The outer box together with the embedded Colin27 mesh is shown in Figure 1A. Each outer box mesh is constructed with the meshing software Gmsh⁶³ (dev version 4.2.3-git-023542a) such that the corresponding brain mesh is nested within. Furthermore, the box meshes are graded such that the cell size gradually gets larger away from the brain boundary (Matérn field values are only needed in the brain domain).

5.3 | Numerical stability considerations

For the parameter regimes considered here, the problem (15) is only mildly convection-dominated, with an upper estimate of the Péclet number of

$$Pe \approx \frac{\hat{L} v_{\text{avg}}}{D_{\text{Gad}}^*} = O(10^2), \quad (35)$$

where $\hat{L} \approx 0.084$ m is half the diameter of the computational domain, $v_{\text{avg}} = 0.17 \mu\text{m/s}$, and $D_{\text{Gad}}^* = 1.2 \times 10^{-10}$ m²/s. Given the fine computational meshes, we obtain low-probability worst-case cell Péclet numbers of $\approx 43 \times 2^{-\ell}$ on level ℓ of the MLMC hierarchy used. In the numerical experiments, convection-related numerical instabilities were not observed.

However, in initial investigations, we observed that the FEM solution undershoots near the boundary, attaining negative concentration values. This is a known phenomenon in the literature and it does not depend on the velocity field, but it is typical of diffusion problems with Dirichlet-type boundary conditions.⁶⁴ We address this problem by a mass-lumping technique, which is known to reduce this effect.⁶⁴ This ill-behavior disappears as the mesh is refined to the extent that no undershoots were observed on the finer levels of the MLMC hierarchy. In our numerical experiments, MLMC convergence was resilient to non-physical behavior on the coarser levels. In fact, we noted that these coarse level solutions would still act as a good control variate for the finer levels.

5.4 | Solver and software

For the computations, we combined the University of Oxford Mathematical Institute computing servers and the University of Oslo Abel supercomputing cluster. The total amount of CPU hours (CPUh) available on the Abel cluster for the project were approximately 400 000 CPUh with the rule that computations with high memory requirements (>4 GB) would cost an amount of CPU hours given by the formula

$$\text{CPU hours cost} = \text{n hours} \times (\text{memory per CPU}) / (4 \text{ GB}). \quad (36)$$

The computation was memory bound, therefore causing the overall computational resources to be limited. We used the Abel cluster exclusively for the samples on the finest MLMC level.

We used the FEniCS FEM software.⁶⁵ The linear systems of equations were solved using the PETSc⁶⁶ implementation of the GMRES algorithm preconditioned with the BoomerAMG algebraic multigrid algorithm from Hypre.⁶⁷ We use the MLMC and QMC routines from the open-source software FEMLMC,³⁴ which contains the implementation of the algorithms presented in Refs. 18, 19. For the FEMLMC Matérn field sampling solver, we declare convergence when the absolute size of the preconditioned residual norm is below a tolerance of 10^{-8} . We use the same stopping criterion for the GMRES solver.

Convergence of the numerical solver was verified with a convergence test comparing different mesh refinements and time steps for a set of deterministic worst-case models (with large velocities and small diffusion coefficients), see, for example, Ref. 20 (Supplementary Material).

5.5 | QMC and MLMC algorithms

We now describe the QMC and MLMC algorithms used more in details. We adopt the following (standard) randomized QMC algorithm³⁵:

QMC algorithm

1. Set the required tolerance ε , $\theta \in (0, 1)$. Set the mesh size h and Δt to ensure that the bias estimate is lower than $\sqrt{\theta}\varepsilon$;
2. Get an initial estimate of $\hat{V}_{\text{QMC}} = \mathbb{V}[I_N(\omega)]$ (cf. (6)) with $N = 1$ samples and $M = 32$ randomizations;
3. While $\hat{V}_{\text{QMC}} > (1 - \theta)\varepsilon^2$, double N .

The total QMC cost is given by $C_{\text{tot}} = \bar{C}NM$, where N is the final number of samples taken and \bar{C} is the (expected) computational cost of computing one QMC sample.

Remark 5. In the random PDE case, typically the bias stems directly from the deterministic PDE solver (the FEM error in our case)^{68,69} and as such it can be estimated empirically, that is, by experimenting with the mesh size and time step, and/or by using theoretical error estimates.⁷⁰ Here, we estimate the bias empirically (cf. description of the MLMC algorithm) and we check that the empirical bias convergence order with h and Δt is second-order as expected given our discretization.

For MLMC, we first need to estimate the optimal number of samples on each level for a given tolerance ε . Let C_ℓ , V_ℓ be the cost and variance of one sample $Q_\ell(\cdot) - Q_{\ell-1}(\cdot)$ respectively. The total MLMC cost and variance are

$$C_{\text{tot}} = \sum_{\ell=1}^L N_\ell C_\ell, \quad \hat{V} = \sum_{\ell=1}^L N_\ell^{-1} V_\ell. \quad (37)$$

It is possible to minimize the estimator variance for a fixed total cost.²⁶ For a fixed MSE tolerance ε^2 , the optimal number of samples for each level and related total cost are,

$$N_\ell = \left((1 - \theta)^{-1} \varepsilon^{-2} \sum_{l=1}^L \sqrt{V_l C_l} \right) \sqrt{V_\ell / C_\ell}, \quad C_{\text{tot}} = (1 - \theta)^{-1} \varepsilon^{-2} \left(\sum_{\ell=1}^L \sqrt{V_\ell C_\ell} \right)^2. \quad (38)$$

We can now describe the MLMC algorithm used in this work:

MLMC algorithm (taken from Ref. 33)

- Set the required tolerance ε , $\theta \in (0, 1)$, the maximum level L_{\max} , the initial number of levels L and the initial number of samples \bar{N}_ℓ to be taken on each.
- **while** extra samples need to be evaluated ($\exists \ell : \bar{N}_\ell > 0$):
 1. for each level, evaluate all the samples scheduled to be taken;
 2. compute/update estimates for the level variance V_ℓ , $\ell = 1, \dots, L$;
 3. compute optimal number of samples N_ℓ by using (38) and update the numbers of extra samples needed \bar{N}_ℓ accordingly;
 4. check whether the weak error (ie, the bias) $|\mathbb{E}[\hat{P} - P]|$ is below the required tolerance $\sqrt{\theta}\varepsilon$; to see how the bias is estimated in practice without knowing the exact solution, see section 3 in Ref. 33.
 5. if not: if $L = L_{\max}$ report failed convergence; otherwise set $L := L + 1$, update N_ℓ and \bar{N}_ℓ and compute $N_L = \bar{N}_L$ (again using (38)).

Remark 6. The standard way^{26,71} to compare the efficiency of these algorithms is to fix the same tolerance for both methods and compare their total costs, given by $C_{\text{tot}} = \bar{C}NM$ for QMC and by $C_{\text{tot}} = \sum_{\ell=1}^L N_\ell C_\ell$ for MLMC (cf. (37)). Furthermore, to ensure that the bias level is the same for both methods, the QMC routine must be run on the finest mesh and time step size of the MLMC hierarchy. This gives $\bar{C} = C_L$. In practice, for the sake of comparing methods, the actual costs C_ℓ , $\ell = 1, \dots, L$ can be replaced by pseudo-costs, i.e. by setting $C_\ell \approx \hat{c}_3 2^{\hat{\gamma}\ell}$, where \hat{c}_3 and $\hat{\gamma}$ are the values, estimated with CPU timings, of the constants c_3 and γ appearing in Theorem 2.3. We use this latter approach.

6 | NUMERICAL RESULTS

We now compare the efficiency of standard MC, QMC and MLMC when employed to solve Models 1 and 2. In what follows, we let $\mathcal{T} = \{30k\text{min}\}_{k=1}^{k=48}$ and define

$$\mathcal{Q} = \{Q_g(t), t \in \mathcal{T}\} \cup \{Q_w(t), t \in \mathcal{T}\} \cup \{q_g(t), t \in \mathcal{T}\} \cup \{q_w(t), t \in \mathcal{T}\}, \quad (39)$$

to be the set of all the output functionals of interest considered (cf (21) and (22)).

6.1 | Estimation of MLMC parameters

We first estimate the MLMC parameters α , β and γ of Theorem 1. Since we are considering the estimation of multiple output functionals, we estimate α and β by monitoring the bias and variance at each level ℓ :

$$\max_{Q \in \mathcal{Q}} |\mathbb{E}[Q_\ell - Q_{\ell-1}]|, \quad \max_{Q \in \mathcal{Q}} \mathbb{V}[Q_\ell - Q_{\ell-1}]. \quad (40)$$

We expect $\alpha = 2$, $\beta = 2\alpha = 4$ and $\gamma = 4$ since for random PDEs we have that the bias convergence is typically the same as the deterministic FEM convergence order, $\beta = 2\alpha$,^{68,69} the numerical method is second-order in both time and space and we are using a multigrid-preconditioned Krylov method (cf. Section 5). The value for γ stems from the fact that the number of time steps on level ℓ are proportional to 2^ℓ and the linear solver used has (essentially) linear complexity in the number of degrees of freedom, which in turn scale proportionally to $2^{3\ell}$. We therefore get a cost per level proportional to $2^{(3+1)\ell}$ and a $\gamma = 4$.

To estimate the bias and variance in (40) we take $N = 4000$ samples on the first two levels and $N = 100$ on the finest level. The choice of the number of samples on the finest level is motivated by the following considerations. The computational resources available were limited, cf. Section 5.4. The number of vertices on the finest level of the MLMC hierarchy is large (cf. Table 1), resulting in a memory burden of ≈ 50 GB to load the mesh, the box mesh in which the brain mesh is embedded (cf. Section 2) and the FEM subspaces. Additionally, solving one instance of (15) on this mesh takes more than 24 hours in serial. The overall sampling procedure took around 2 weeks of combined use of the Oxford Mathematical Institute servers and the University of Oslo Abel cluster. The same computation performed in serial would have taken roughly 2 years. Adding a finer level to the hierarchy would have therefore been prohibitive.

Figures 3 and 4 show the bias and variance versus refinement level for Models 1 and 2, respectively. We observe that the numerical estimates closely match the theoretical expectations. The estimated variance convergence order for both models is $\hat{\beta} \approx 4.2$, which is just above the theoretical value of $\beta = 4$. For Model 2, the estimated bias order is $\hat{\alpha} = 2.09$ which again closely matches the theoretical estimate of $\alpha = 2$. However for Model 1, we observe that the bias decays

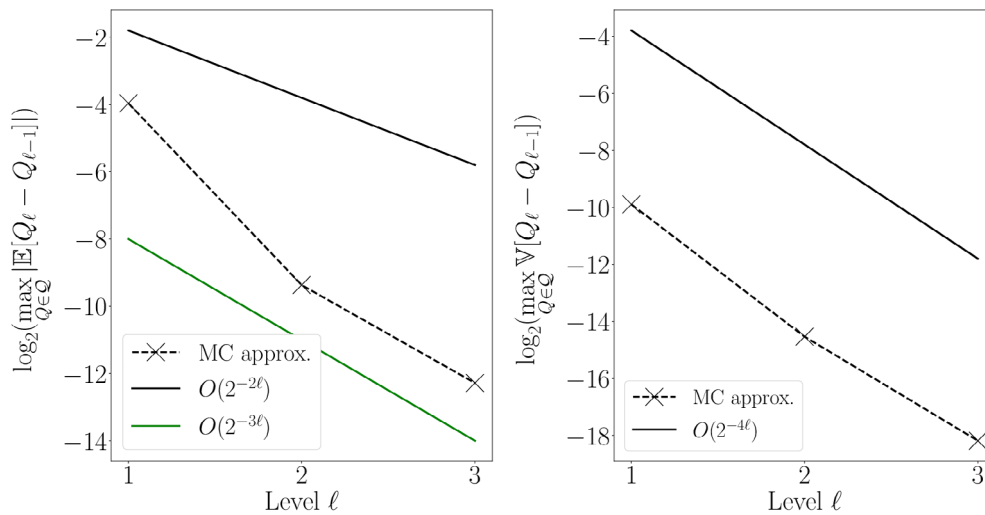


FIGURE 3 Convergence behavior of the FEM approximation to the solution of Model 1. The estimated convergence order for the variance agrees with our predictions and with what expected by the theory in the diffusion-only case.⁶⁹ The bias convergence order observed is instead higher than expected. Estimated parameters via linear regression: $\hat{\alpha} \approx 4.16$, $\hat{\beta} \approx 4.15$

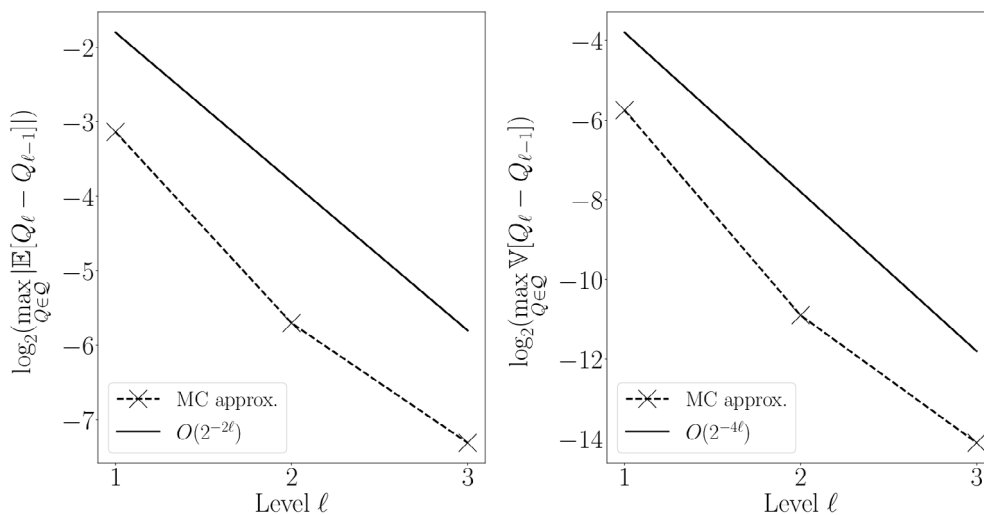


FIGURE 4 Convergence behavior of the FEM approximation to the solution of Model 2. The estimated convergence orders agree with our predictions and with what expected by the theory in the diffusion-only case.⁶⁹ Estimated parameters via linear regression: $\hat{\alpha} \approx 2.09$, $\hat{\beta} \approx 4.18$

more rapidly than expected ($\hat{\alpha} = 4.16$ versus $\alpha = 2$). In this case, we are likely observing a pre-asymptotic regime and the convergence rate seems to be decaying as ℓ increases; this notion would be in agreement with theoretical convergence results.^{69,70} Finally, when estimating γ as taking the average wall-clock time of each sample as a proxy (data not shown), we obtain $\hat{\gamma} \approx 4.09$, which is close to the theoretical prediction of $\gamma = 4$.

6.2 | Mean square error weighting under limited computational resources

Note that we have a finite number of meshes available and consequently the version of MLMC considered here is “weaker” than the true MLMC algorithm, described in Section 5.5, since the maximum level L is bounded, cf. Remark 3. In fact, we are unable to reduce the bias below the threshold given by the finest mesh of the hierarchy without resorting to more advanced techniques (see later Remark 8). However, we can still balance the relative weight of bias and statistical error by choosing two different values of the MLMC weight parameter θ (cf. Ref. 38 and Section 2). Based on the fact that for this problem: (a) the maximum L is restricted; (b) N_L is restricted; (c) we have estimated values for the MLMC parameters α , β , and γ (cf. Theorem 1); we can estimate a priori the largest possible values of θ that we can use without making the number of samples on the finest level exceed 100. This yields the values $\theta = 0.041$ for Model 1 and $\theta = 0.72$ for 2, corresponding to the best achievable MSE tolerances of roughly $\varepsilon = 3.3 \times 10^{-4}$ for Model 1 and of $\varepsilon = 2.5 \times 10^{-3}$ for Model 2. Note that in the Model 1 case, the bias is much smaller (compare Figures 3 and 4), hence why the chosen θ is smaller as well.

Remark 7. Note that the bias (ie, the discretization error) can be estimated by the MLMC algorithm without requiring a reference or exact solution (recall the MLMC algorithm in Section 5.5) and is the same independently from which Monte Carlo method is used. Furthermore, all methods can compute their own estimates of the statistical error. For this reason, no reference or exact solution is needed to compute the MSE or to construct the figures shown in this section.

Remark 8. In practice, it is possible to reduce the MLMC estimator bias by augmenting MLMC with Richardson-Romberg extrapolation.^{28,39} However, we leave this enhancement for future research.

6.3 | Tracer evolution over time

In Figures 5 and 6 we present the expected value and SD of all quantities of interest as a function of time for both models. These results have been obtained with MLMC by setting the lowest achievable MSE tolerances given the restriction on the maximum number of samples on the finest level, cf. Section 6.2. The total number of samples used for these simulations are $N_\ell = [26060, 1372, 100]$ for Model 1 and $N_\ell = [26567, 1171, 100]$ for Model 2 for $\ell = 1, 2, 3$. We remark that the figure errorbars represent one SD away from the mean, and not necessarily the actual variability of these quantities.

In both models, tracer first spreads to the gray matter, reaching a peak amount of tracer in 8–12 hours (Figures 5A and 6A). Over time, the tracer also spreads into the white matter (Figures 5B and 6(B)). The time-to-peak amount of tracer in white matter is estimated at ≈ 24 hours in Model 2 (Figure 6B), while for Model 1 this peak would occur considerably later (Figure 5B). In Figure 5, we observe that for Model 1, the variances of the global amounts of tracer Q_g and Q_w are small. However, the variability in the local tracer concentration q_g and q_w is much larger and, as such, substantially contributes to the computational complexity of the Monte Carlo simulations. We observe that for Model 1, the local variability of the random coefficients “averages out” when computing global output functionals. Thus, if global quantities were the only focus, a MC simulation with small sample size could be sufficient to achieve good accuracy for this model. Conversely, we note that all quantities of interest for Model 2 (amount of tracer in gray and white matter, and tracer concentrations in local regions) have large SDs (Figure 6).

Comparing these results with the MSE tolerances imposed we can estimate the number of correct digits of the computed expectations: the estimates obtained for Model 1 are roughly within 3 digits of accuracy for the gray matter quantities Q_g and q_g , 2 digits for Q_w and 1 digit only for q_w . Similarly, all Model 2 gray matter quantities Q_g and q_g are roughly correct within 2 digits, while the white matter outputs Q_w and q_w are just slightly less accurate.

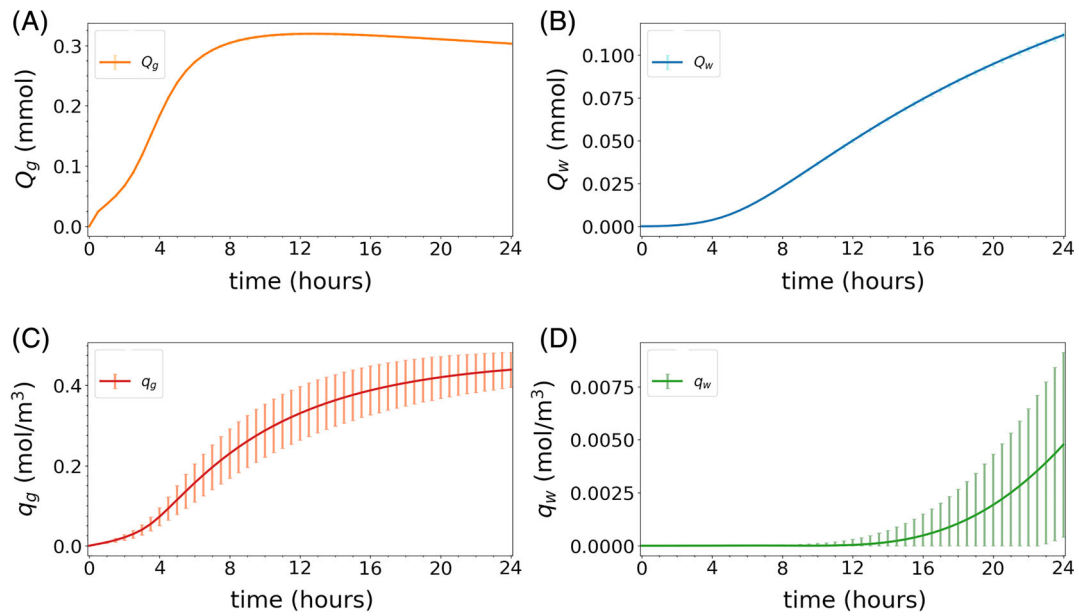


FIGURE 5 Mean and SD vs time for all quantities of interest for Model 1. The continuous lines correspond to expected values and the vertical bars indicate plus or minus one SD away from the mean (negative values are excluded here). Both were computed using MLMC with a MSE tolerance of $\varepsilon = 3.3 \times 10^{-4}$, which is the best achievable tolerance given the discretization error and the restrictions on the number of samples on the finest level, cf. Section 6.2

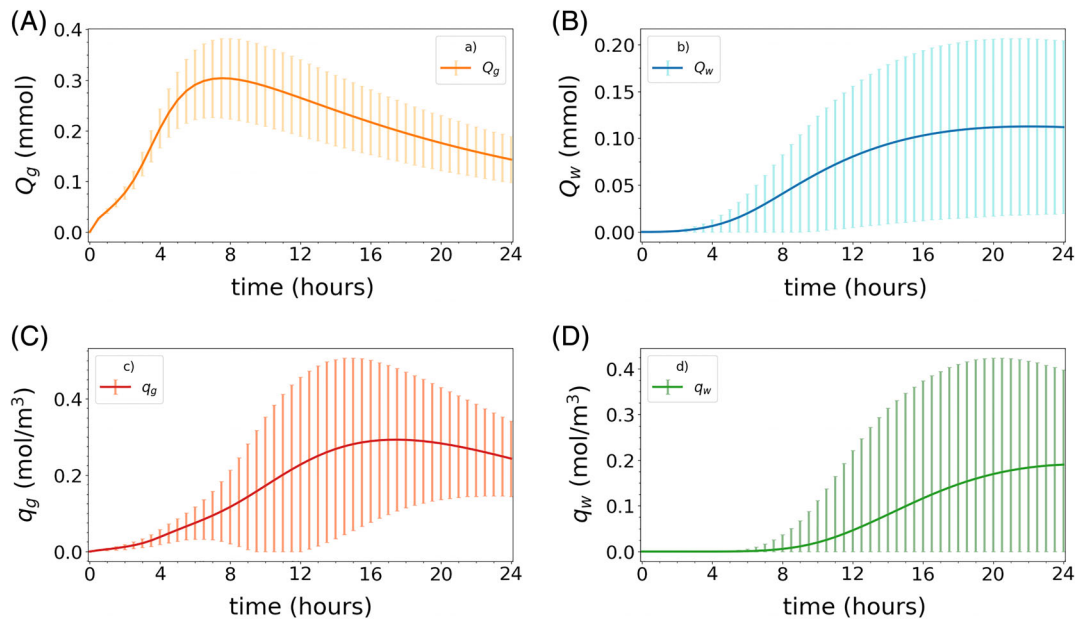


FIGURE 6 Mean and SD vs time for all quantities of interest for Model 2. The continuous lines correspond to expected values and the vertical bars indicate plus or minus one SD away from the mean (negative values are excluded here). Both were computed using MLMC with a MSE tolerance of $\varepsilon = 2.5 \times 10^{-3}$, which is the best achievable tolerance given the discretization error and the restrictions on the number of samples on the finest level, cf. Section 6.2

6.4 | Comparison of MC, QMC, and MLMC performance

In Figures 7 and 8 (left) we show the optimal number of samples chosen automatically by MLMC on each level as the root mean square error tolerance ε is reduced. The maximum level chosen is increased as ε decreases in order to satisfy

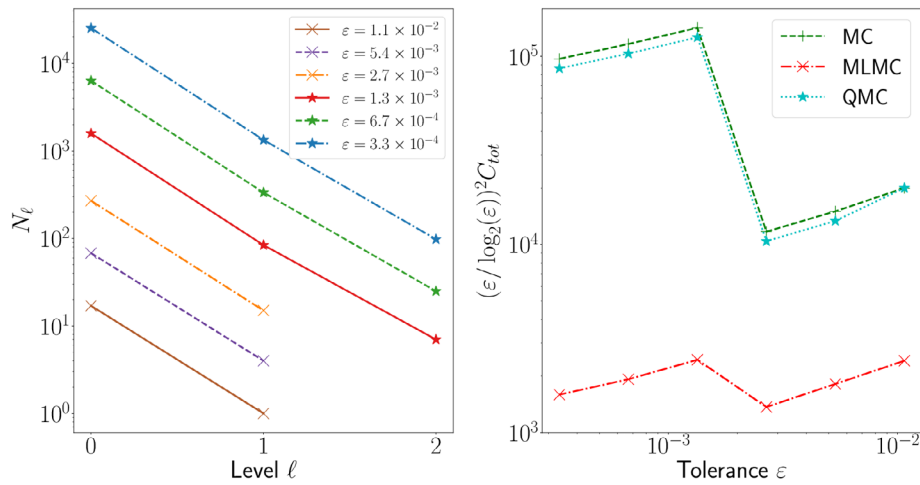


FIGURE 7 Convergence of standard MC, QMC and MLMC for the solution of Model 1 ($\theta = 0.041$). In the plot on the left we show how the MLMC algorithm automatically selects the optimal number of samples N_ℓ on each level to achieve a given tolerance ε . In the plot on the right we compare the efficiency of the methods for different tolerances. The savings of MLMC with respect to standard MC and QMC are considerable, while QMC barely improves on standard MC (see main text)

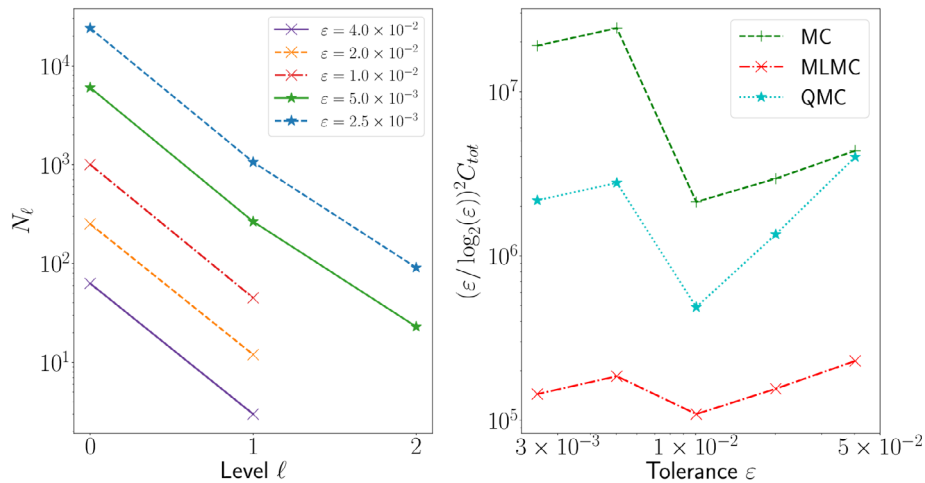


FIGURE 8 Convergence of standard MC, QMC and MLMC for the solution of Model 2 ($\theta = 0.72$). In the plot on the left we show how the MLMC algorithm automatically selects the optimal number of samples N_ℓ on each level to achieve a given tolerance ε . In the plot on the right we compare the efficiency of the methods for different tolerances. MLMC significantly outperforms QMC, which in turn considerably outperforms standard MC

the bias tolerance. Note that the smallest values of ε considered correspond to the lowest bias tolerance that standard MLMC can achieve with an upper limit of 100 samples on the finest level (cf. Remark 8).

In Figures 7 and 8 (right), we compare the total computational cost C_{tot} of standard MC, MLMC, and QMC for the solution of Model 1 and 2, respectively. The cost of running a full standard MC simulation for this problem is prohibitive (on the order of 100 years in serial!) so we estimate the MC cost by using the level variance and cost. This is a standard procedure and essentially derives from the application of the central limit theorem.²⁶ We find that QMC is significantly slower than MLMC. For this reason we only estimate the number of QMC samples needed by running QMC on the second finest mesh of the hierarchy rather than on the finest (cf. Remark 6). The difference should be minimal since the number of samples needed appears to be approximately constant on the finer levels of the hierarchy in numerical experiments.²⁹ Since $\beta = \gamma$, Theorem 1 predicts an overall MLMC complexity of $\varepsilon^{-2}(\log \varepsilon)^2$ for a root mean square error tolerance ε . We therefore expect a near constant cost curve for $\varepsilon^2(\log \varepsilon)^{-2} C_{tot}$ versus ε in the MLMC case. The numerical results corroborate the theoretical expectations: while the MLMC cost lines oscillate some, they are

well-fitted by a horizontal line (estimated slope ≈ 0.05 for Model 1 and ≈ 0.1 for Model 2). Overall, for both models MLMC significantly outperforms both QMC and standard MC, with a $O(100)$ factor of improvement with respect to standard MC.

While the qualitative behavior of standard MC and MLMC is consistent between the two models, QMC behaves differently. For Model 2 (Figure 8), the performance of QMC considerably improves on that of standard MC, especially for smaller MSE tolerances. On the other hand, for Model 1 (Figure 7), the improvement is negligible and QMC performs essentially the same as standard MC.

This behavior could be interpreted in the context of the formulations of Models 1 and 2 (cf. Section 4). While the stochastic input in Model 2 includes 1 random field and 1 random variable, Model 1 depends on 4 random fields and 1 random variable. Given the lack of performance gain for QMC applied to Model 1, we hypothesize that the higher input dimensionality affects the QMC convergence, causing the rate to decay to a standard $O(N^{-1/2})$ MC rate. Indeed, the fact that QMC performance degrades with high input dimensions is well-known.²⁵ It therefore appears that the (ML)QMC method presented in Ref. 18 is not robust with respect to the number of random field inputs, at least in 3D where the dimensionality is larger. We did not observe this ill-behavior in analogous numerical tests performed on a convection-diffusion PDE with random coefficients on a square domain.

Remark 9. Adding a coarser level to the mesh hierarchy, given by the original Colin27 human adult brain atlas mesh⁵³ (version 2) did not improve the performance of MLMC. This was determined by first estimating the cost and variance of the coarsest MLMC levels in the extended hierarchy using 4000 MC pilot samples and then determining which hierarchy would give the best savings in terms of cost (see eg, section 2.6 in Ref. 33).

7 | DISCUSSION

Most theoretical UQ research on equations with random coefficients focuses on the same model problem,^{45,69,72,73} with numerical results often performed on simple boxes and in low dimensions. For this reason, and because of their different requirements and convergence properties, it is a priori unknown which method performs best for a given problem. While asymptotic complexity results might be available,^{26,69,73} it might not be possible to observe the asymptotic behavior due to practical limitations that arise in large-scale problems (eg, limited computational resources, small hierarchy size, poor low-fidelity models). In fact, when considering large-scale applications with physics-inspired problems on MRI-derived geometries, currently, the only way to compare the performance of different methods is via numerical experiments. Nevertheless, the application of UQ to realistic large-scale problems is scarce in the literature, to the extent that (that we know of) this is the first study employing advanced UQ techniques to study the human brain.

The only related study that we are aware of comes from our previous work on the subject.¹⁷ In this study we found that diffusion alone could not explain MRI data, suggesting flow velocities may impact the distribution of tracer seen in MRI data, possibly also affecting both brain clearance and drug delivery. However, in Ref. 20 standard MC was used, which forced us to resort to coarser meshes and shorter simulation time. The need for a more efficient UQ methodology motivated the current study.

In terms of limitations, in this study we considered a single brain geometry deriving from the Colin27 mesh.⁷⁴ A set of different geometries should be considered to ensure that the results obtained are patient-independent. We only modeled a single fluid flow compartment, while future research could consider more elaborate models accounting for all the different fluids flowing in the parenchyma (venous, arterial and capillary blood, interstitial fluid, cerebrospinal fluid). The computational resources available for this study were limited and with more computational power we could have obtained more accurate results by adding a finer level to the hierarchy and/or increasing the number of samples. However, we did not encounter any significant issues with setting up a hierarchy on which good MLMC convergence could be observed, a problem mentioned in recent work by Quaglino et al,²⁴ in which they apply MFMC for a cardiac electrophysiology application.

We also point at a limitation related to the choice of covariance structure for the diffusion and velocity fields. Here, these fields have been assigned an isotropic covariance. As such, the correlation between any two points only depends on their Euclidean distance, and not on the (domain) interior path length between the two points. Consequently, brain surface regions such as the different gyri that are closer to each other than the field correlation length, become

correlated even if separated by a sulcus. Since the average sulcal width is estimated to range between 1.5 and 3 mm in humans,^{44,74} this aspect does not play a role for the velocity field ($\lambda \approx 1$ mm), but for the diffusion coefficient it may ($\lambda = 1$ cm). We expect global quantities (eg, Q_g and Q_w) and local quantities depending on white matter regions which are far from the brain surface (eg, q_w) to be largely unaffected by this modeling choice. Conversely, local quantities that depend on gray matter local regions (eg, q_g) might be affected and further investigation would be needed. We remark, however, that isotropic covariances make for very efficient sampling methods, and designing and efficiently sampling a random field which is topology-aware at the high resolution scale of the sulci is highly challenging.

Overall, the random fields parameters were chosen to represent the current understanding of the relevant physiology to the extent possible, with physical considerations as the primary factor in the choice of fields and parameters. While most of the model parameters do not affect the performance of the numerical solver, QMC or MLMC in any way, the following aspects should be considered for alternative models:

- **Velocity magnitude and diffusion coefficient:** a much larger velocity magnitude (or a much smaller diffusion coefficient) would increase the Péclet number and harm the stability of the finite element solver used (cf. Section 5.3). A more advanced numerical solver would be needed in this case. We remark that for random PDEs the convergence of MLMC directly depends on the convergence properties of the deterministic solver used.^{68,69}

- **Random field variance:** larger variability implies that more samples would be needed for an accurate result. This does not affect the efficacy of the methods, but it does increase the CPU hours needed for the computation, which might be a problem when the computational resources are limited.

- **Smoothness parameter and correlation length:** small values of these parameters are known to affect most random field sampling techniques and extra care must be taken. At the same time small values would also introduce sharp and abrupt spatial changes in the sampled field that are acceptable, for example, oil reservoir modeling,⁷⁵ but would be nonphysical in a healthy brain. We remark that the velocity correlation length chosen in this study is quite small, showing that the sampling techniques in Refs. 18, 19 still work well in this case.

8 | CONCLUSIONS

We have compared and evaluated the MC, QMC, and MLMC uncertainty quantification and sampling methods for two physiologically realistic brain transport models. Even under restriction on the maximum number of samples on the finest mesh level and on the finest mesh available, MLMC significantly outperforms all other methods, yielding an improvement factor of roughly 100 with respect to standard MC. QMC outperforms standard MC by a factor of approximately 10 for one of the models (Model 2), but yields no performance gains for the other (Model 1). Overall, for this application, MLMC achieves the best performance and should be preferred. We remark that a standard MC simulation would have taken on the order of 10 years to complete in parallel ($O(100)$ years in serial!), while only 2 weeks are needed for MLMC.

MLQMC methods were not investigated directly in this study. However, our numerical findings indicate that no additional improvement can be achieved in the Model 1 case. On the other hand, Model 2 seems amenable to MLQMC. Thus, MLQMC could also bring additional computational gains for similar models with a small number of input random fields.

In conclusion, even large-scale problems on MRI-derived geometries are amenable to advanced Monte Carlo methods and we hope that our work can inspire further research on uncertainty quantification for biomedical engineering problems.

ACKNOWLEDGEMENTS

This research is supported by the European Research Council via Grant #714892 (Waterscales), the Research Council of Norway via Grant #250731 (Waterscape), by the NOTUR grant NN9316K, and by the EPSRC Centre For Doctoral Training in Industrially Focused Mathematical Modeling (EP/L015803/1). Part of this work was performed on the Abel Cluster, owned by the University of Oslo and Uninett/Sigma2, and operated by the Department for Research Computing at USIT, the University of Oslo IT-department <http://www.hpc.uio.no/>.

ORCID

Matteo Croci  <https://orcid.org/0000-0003-1669-9445>

Vegard Vinje  <https://orcid.org/0000-0003-2928-3385>

Marie E. Rognes  <https://orcid.org/0000-0002-6872-3710>

REFERENCES

1. Sokoloff L. Circulation and energy metabolism of the brain. *Basic Neurochem.* 1989;2:338-413.
2. Aldea R, Weller RO, Wilcock DM, Carare RO, Richardson G. Cerebrovascular smooth muscle cells as the drivers of intramural per-arterial drainage of the brain. *Front Aging Neurosci.* 2019;11:1-17.
3. Carare RO, Bernardes-Silva M, Newman TA, et al. Solutes, but not cells, drain from the brain parenchyma along basement membranes of capillaries and arteries: significance for cerebral amyloid angiopathy and neuroimmunology. *Neuropathol Appl Neurobiol.* 2008;34(2):131-144.
4. Iliff J. J., M. Wang, Y. Liao, B. A. Plogg, W. Peng, G. A. Gundersen, and Others, 2012: A paravascular pathway facilitates CSF flow through the brain parenchyma and the clearance of interstitial solutes, including amyloid beta. *Sci Transl Med*, 4, no. 147, 147ra111.
5. Iliff JJ, Wang M, Zeppenfeld DM, et al. Cerebral arterial pulsation drives paravascular CSF—interstitial fluid exchange in the murine brain. *J Neurosci.* 2013;33(46):18190-18199.
6. Jessen NA, Munk ASF, Lundgaard I, Nedergaard M. The glymphatic system: a beginner's guide. *Neurochem Res.* 2015;40(12):2583-2599.
7. Smith AJ, Yao X, Dix JA, Jin B-J, Verkman AS. Test of the 'glymphatic' hypothesis demonstrates diffusive and aquaporin-4-independent solute transport in rodent brain parenchyma. *eLife.* 2017;6(e27679):16.
8. Wolak DJ, Thorne RG. Diffusion of macromolecules in the brain: implications for drug delivery. *Mol Pharm.* 2013;10(5):1492-1504.
9. Ringstad G, Valnes LM, Dale AM, et al. Brain-wide glymphatic enhancement and clearance in humans assessed with MRI. *JCI Insight.* 2018;3(13):16.
10. Ringstad G, Vatnehol SAS, Eide PK. Glymphatic MRI in idiopathic normal pressure hydrocephalus. *Brain.* 2017;140(10):2691-2705.
11. Asgari M, De Zélicourt D, Kurtcuoglu V. Glymphatic solute transport does not require bulk flow. *Sci Rep.* 2016;6:38635.
12. Daversin-Catty C, Vinje V, Mardal K-A, Rognes ME. The mechanisms behind perivascular fluid flow. *bioRxiv.* 2020:1-11.
13. Sharp MK, Carare RO, Martin BA. Dispersion in porous media in oscillatory flow between flat plates: applications to intrathecal, per-arterial and paraarterial solute transport in the central nervous system. *Fluids Barriers CNS.* 2019;16(1):13.
14. Bedussi B, Almasian M, de Vos J, van Bavel E, Bakker ENTP. Paravascular spaces at the brain surface: low resistance pathways for cerebrospinal fluid flow. *J Cerebral Blood Flow Metab.* 2017;38(4):719-726.
15. Mestre H, Tithof J, Du T, et al. Flow of cerebrospinal fluid is driven by arterial pulsations and is reduced in hypertension. *Nat Commun.* 2018;9(1):4878.
16. Holter KE, Kehlet B, Devor A, et al. Interstitial solute transport in 3D reconstructed neuropil occurs by diffusion rather than bulk flow. *Proc Natl Acad Sci.* 2017;114(37):9894-9899.
17. Croci M, Vinje V, Rognes ME. Uncertainty quantification of parenchymal tracer distribution using random diffusion and convective velocity fields. *Fluids Barriers CNS.* 2019b;16(1):32.
18. Smith JH, Humphrey JA. Interstitial transport and transvascular fluid exchange during infusion into brain and tumor tissue. *Microvasc Res.* 2007;73(1):58-73.
19. Fultz NE, Bonmassar G, Setsompop K, et al. Coupled electrophysiological, hemodynamic, and cerebrospinal fluid oscillations in human sleep. *Science.* 2019;366(6465):628-631.
20. Asgari M, De Zélicourt D, Kurtcuoglu V. How astrocyte networks may contribute to cerebral metabolite clearance. *Sci Rep.* 2015;5:15024.
21. Bilston LE, Fletcher DF, Brodbelt AR, Stoodley MA. Arterial pulsation-driven cerebrospinal fluid flow in the perivascular space: a computational model. *Comput Methods Biomech Biomed Eng.* 2003;6(4):235-241.
22. Biehler J, Gee MW, Wall WA. Towards efficient uncertainty quantification in complex and large-scale biomechanical problems based on a Bayesian multi-fidelity scheme. *Biomech Model Mechanobiol.* 2015;14:489-513.
23. Quaglino A, Pezzuto S, Koutsourelakis PS, Auricchio A, Krause R. Fast uncertainty quantification of activation sequences in patient-specific cardiac electrophysiology meeting clinical time constraints. *Int J Numer Methods Biomed Eng.* 2018a;34(7):e2985. <https://doi.org/10.1002/cnm.2985>.
24. Quaglino, A., S. Pezzuto, and R. Krause, 2018b: Generalized multifidelity Monte Carlo estimators.
25. Caflish RE, Morokoff W, Owen A. Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension. *Journal of Computational Finance.* 1997;1:27-46.
26. M. B. Giles. Multilevel Monte Carlo methods. *Acta Numer.* 2015;24:259-328.
27. Peherstorfer B, Willcox K, Gunzburger M. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Rev.* 2018;60(3):550-591.
28. Lemaire V, Pagès G. Multilevel Richardson-Romberg extrapolation. *Ther Ber.* 2017;23(4A):2643-2692.
29. Croci, M., M. B. Giles, and P. E. Farrell, 2019a: Multilevel quasi monte carlo methods for elliptic pdes with random field coefficients via fast white noise sampling. arXiv preprint arXiv:1911.12099.

30. Croci M, Giles MB, Rognes ME, Farrell PE. Efficient white noise sampling and coupling for multilevel Monte Carlo with non-nested meshes. *SIAM/ASA J Uncertain Quantif*. 2018;6(4):1630-1655.
31. Alnæs MS, Blechta J, Hake J, et al. The FEniCS project version 1.5. *Archive of Numerical Software*. 2015;3(100):9-23.
32. Maddison, J. R., P. E. Farrell, and I. Panourgias, 2013: Parallel supermeshing for multimesh modelling. Archer, eCSE03-08.
33. Maddison, J. R., I. Panourgias, and P. E. Farrell, 2017: *Libsupermesh*. Available at <https://bitbucket.org/libsupermesh/libsupermesh>
34. Croci, M., 2019: *FEMLMC—A Library of Finite Element (Quasi) Monte Carlo Methods for FEniCS*. Available at <https://bitbucket.org/croci/femlmc/>
35. Lemieux, C., 2009: *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer Series in Statistics, 1st ed., Springer-Verlag New York.
36. Morokoff WJ, Caflisch RE. Quasi-Monte Carlo integration. *J Comput Phys*. 1995;122(2):218-230.
37. Owen AB. Variance and discrepancy with alternative scramblings. *ACM Trans Model Comput Simul*. 2003;13(4):363-378.
38. Haji-Ali A-L, Nobile F, von Schwerin E, Tempone R. Optimization of mesh hierarchies in multilevel Monte Carlo samplers. *Stoch Partial Differ Equ Anal Comput*. 2016;4(1):76-112.
39. Giles MB. Multilevel Monte Carlo path simulation. *Oper Res*. 2008;56(3):607-617.
40. Dölz J, Harbrecht H, Schwab C. Covariance regularity and H-matrix approximation for rough random fields. *Numer Mathem*. 2017;135(4):1045-1071.
41. Feischl M, Kuo FY, Sloan IH. Fast random field generation with H-matrices. *Numer Mathem*. 2018;140:639-676.
42. Hackbusch W. *Hierarchical Matrices: Algorithms and Analysis*. Berlin-Heidelberg: Springer; 2015:511.
43. Khoromskij BN, Litvinenko A, Matthies HG, Litvinenko A, Matthies HG. Application of hierarchical matrices for computing the Karhunen-Loève expansion. *Comput Secur*. 2009;84(1-2):49-67.
44. Madan CR. Robust estimation of sulcal morphology. *Brain Inform*. 2019;6(1):1-11. <https://doi.org/10.1186/s40708-019-0098-1>.
45. Charrier J, Scheichl R, Teckentrup AL. Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods. *SIAM J Numer Anal*. 2013;51(1):322-352.
46. Bachmayr, M., I. G. Graham, V. K. Nguyen, and R. Scheichl, 2019: Unified analysis of periodization-based sampling methods for matern covariances.
47. Dietrich CR, Newsam GN. Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM J Sci Comput*. 1997;18(4):1088-1107.
48. Graham IG, Kuo FY, Nuyens D, Scheichl R, Sloan IH. Analysis of circulant embedding methods for sampling stationary random fields. *SIAM J Numer Anal*. 2018;56(3):1871-1895.
49. Wood ATA, Chan G. Simulation of stationary Gaussian processes in $[0, 1]^d$. *J Comput Graph Stat*. 1994;3(4):409-432.
50. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J R Stat Soc Ser B Stat Methodol*. 2009;73(4):423-498.
51. Whittle P. On stationary processes in the plane. *Biometrika*. 1954;41(3-4):434-449.
52. Khristenko, U., L. Scarabosio, P. Swierczynski, E. Ullmann, and B. Wohlmuth, 2018: Analysis of boundary effects on PDE-based sampling of Whittle-Matérn random fields.
53. Fang Q. Mesh-based Monte Carlo method using fast ray-tracing in Plücker coordinates. *Biomed Opt Express*. 2010;1(1):165-175.
54. Nicholson C. Diffusion and related transport mechanisms in brain tissue. *Rep Prog Phys*. 2001;64(7):815-884.
55. Orešković D, Radoš M, Klarica M. New concepts of cerebrospinal fluid physiology and development of hydrocephalus. *Pediatr Neurosurg*. 2017;52(6):417-425.
56. Wood JH. *Neurobiology of Cerebrospinal Fluid*. 2nd ed. Boston, MA: Springer Science & Business Media; 2013.
57. Kiviniemi, V., X. Wang, V. Korhonen, T. Keinänen, T. Tuovinen, J. Autio, P. LeVan, S. Keilholz, Y.-F. Zang, J. Hennig, and Others, 2016: Ultra-fast magnetic resonance encephalography of physiological brain activity—glymphatic pulsation mechanisms? *J Cerebral Blood Flow Metab*, 36, no. 6, 1033-1045.
58. Rajna Z, Raitamaa L, Tuovinen T, Heikkilä J, Kiviniemi V, Seppänen T. 3D multi-resolution optical flow analysis of cardiovascular pulse propagation in human brain. *IEEE Trans Med Imaging*. 2019;38(9):9.
59. Tuch DS, Reese TG, Wiegell MR, Makris N, Belliveau JW, Wedeen VJ. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn Reson Med*. 2002;48(4):577-582.
60. Nelsen RB. *An Introduction to Copulas*. New York: Springer Science & Business Media; 2007.
61. Albargothy NJ, Johnston DA, MacGregor-Sharp M, et al. Convective influx/glymphatic system: tracers injected into the CSF enter and leave the brain along separate periarterial basement membrane pathways. *Acta Neuropathol*. 2018;136(1):139-152.
62. Evans LC. *Partial Differential Equations*. Providence, RI: American Mathematical Society; 2010.
63. Geuzaine C, Remacle J-F. Gmsh: a 3-D finite element mesh generator with built-in pre-and post-processing facilities. *Int J Numer Methods Eng*. 2009;79(11):1309-1331.
64. Thomée V. On positivity preservation in some finite element methods for the heat equation. *International Conference on Numerical Methods and Applications*. Cham: Springer; 2014:13-24.
65. Logg A, Mardal K-A, Wells G. *Automated Solution of Differential Equations by the Finite Element Method: the FEniCS Book, Volume 84*. Berlin-Heidelberg: Springer Science & Business Media; 2012.
66. Balay S, Abhyankar S, Adams M, et al. *PETSc Users Manual Revision 3.8*. Lemont, IL: Argonne National Laboratory (ANL); 2017.
67. Falgout RD, Yang UM. hypre: A Library of High Performance Preconditioners. *International Conference on Computational Science*. Berlin-Heidelberg: Springer; 2002:632-641.

68. Cliffe KA, Giles MB, Scheichl R, Teckentrup AL. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Comput Visual Sci.* 2011;14(1):3-15.
69. Teckentrup AL, Scheichl R, Giles MB, Ullmann E. Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. *Numer Mathem.* 2013;125(3):569-600.
70. Brenner S, Scott RL. *The Mathematical Theory of Finite Element Methods, Volume 15.* New York: Springer Science and Business Media; 1994.
71. Giles MB, Waterhouse BJ. Multilevel quasi-Monte Carlo path simulation. *Advanced Financial Modelling (Radon Series on Computational and Applied Mathematics). De Gruyter, Number.* 2009;8:165-181.
72. Kuo FY, Scheichl R, Schwab C, Sloan IH, Ullmann E. Multilevel quasi-Monte Carlo methods for lognormal diffusion problems. *Mathem Comput.* 2017;86(308):2827-2860.
73. Kuo FY, Schwab C, Sloan IH. Multi-level quasi-Monte Carlo finite element methods for a class of elliptic PDEs with random coefficients. *Found Comput Mathem.* 2015;15(2):411-449.
74. Jin K, Zhang T, Shaw M, Sachdev P, Cherbuin N. Relationship between sulcal characteristics and brain aging. *Front Aging Neurosci.* 2018;10:339. <https://doi.org/10.3389/fnagi.2018.00339>.
75. Potsepaev R, Farmer CL. Application of stochastic partial differential equations to reservoir property modelling. *ECMOR XII-12th European Conference on the Mathematics of Oil Recovery, Vol. 2.* Houten: The European Association of Geoscientists and Engineers; 2014.
76. Abrahamsen P. *A Review of Gaussian Random Fields and Correlation Functions.* 2nd ed. Oslo, Norway: Norwegian Computing Center; 1997.
77. Heinrich S. Multilevel Monte Carlo Methods. *International Conference on Large-Scale Scientific Computing.* Berlin-Heidelberg: Springer Berlin Heidelberg; 2001:58-67.
78. Rodriguez-Cantano R, Sundnes J, Rognes ME. Uncertainty in cardiac myofiber orientation and stiffnesses dominate the variability of left ventricle deformation response. *Int J Numer Methods Biomed Eng.* 2019;35(5):e3178.
79. Sullivan TJ. *Introduction to Uncertainty Quantification, Vol. 63.* Switzerland: Springer International Publishing; 2015.

How to cite this article: Croci M, Vinje V, Rognes ME. Fast uncertainty quantification of tracer distribution in the brain interstitial fluid with multilevel and quasi Monte Carlo. *Int J Numer Meth Biomed Engng.* 2021;37: e3412. <https://doi.org/10.1002/cnm.3412>