UNIVERSITY OF OXFORD
Mathematical Institute

Non-nested multilevel Monte Carlo methods with applications to brain simulation

Matteo Croci[†], M. B. Giles[†], P. E. Farrell[†], M. E. Rognes[*], V. Vinje[*]

[†] *Mathematical Institute, University of Oxford, UK*
[*] *Simula Research Laboratory, Norway*

Oden Institute, 10 February 2021

Oxford Mathematics

Overview

Motivation and physiological background

Monte Carlo methods for PDEs with random coefficients

Application: uncertainty quantification in brain tracer simulation

Conclusions

Overview

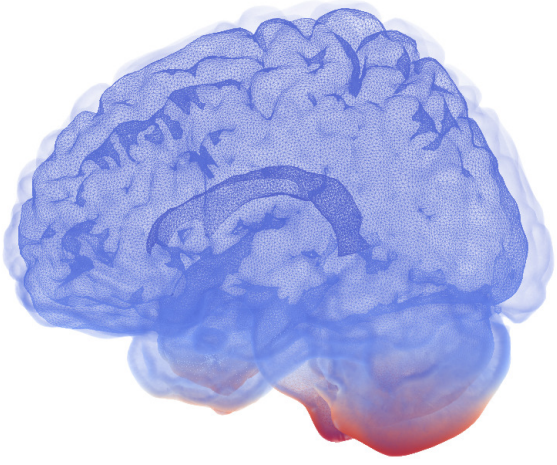
Motivation and physiological background

Monte Carlo methods for PDEs with random coefficients

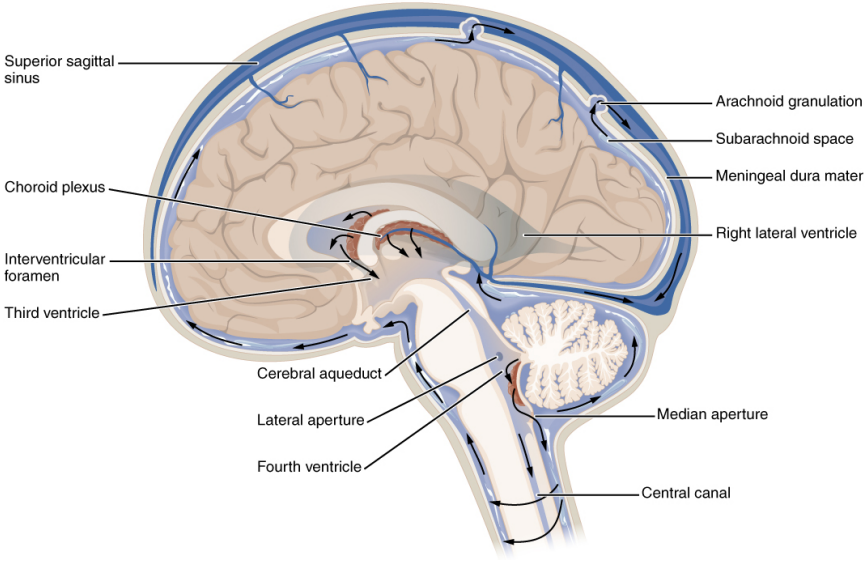
Application: uncertainty quantification in brain tracer simulation

Conclusions

Brain simulation. Why?



Physiological background: CSF and ISF flow



[image source: Wikipedia]

Here we only model **ISF tracer concentration**. Let $G \subset \mathbb{R}^3$ represent the brain,

$$\dot{c}(t,x,\omega) + \nabla \cdot (\boldsymbol{v}(x,\omega)c(t,x,\omega)) - \nabla \cdot (D^*(x,\omega)\nabla c(t,x,\omega)) = 0,$$
$$+ \text{ BCs on } \partial G \text{ (defined later)} \quad | \quad c(0,x,\omega) = 0.$$

Here we model the ISF velocity $\boldsymbol{v}$ and tracer diffusivity D^* to be **random fields**.

Solution via the FEM and advanced Monte Carlo methods.

Overview

Motivation and physiological background

Monte Carlo methods for PDEs with random coefficients

Application: uncertainty quantification in brain tracer simulation

Conclusions

In this part of the talk, we present some new advanced Monte Carlo algorithms for the solution of PDEs with random coefficients. For simplicity, consider the model problem,

$$\begin{aligned} -\nabla \cdot (D^*(x,\omega)\nabla p(x,\omega)) &= 1, \qquad x \in G \subset \mathbb{R}^d, \qquad \omega \in \Omega, \\ p(x,\omega) &= 0, \qquad x \in \partial G, \qquad \omega \in \Omega. \end{aligned}$$

Say that we are interested in computing $\mathbb{E}[P]$, where $P(\omega) = \int_G p^2 \mathrm{d}x = ||p||^2_{L^2(G)}(\omega)$. Set $D^* = e^{u(x,\omega)}$, where $u \sim \mathcal{N}(0, \mathcal{C}(x,y))$ is a (Matérn) Gaussian random field.

In this part of the talk, we present some new advanced Monte Carlo algorithms for the solution of PDEs with random coefficients. For simplicity, consider the model problem,

$$\begin{aligned} -\nabla \cdot (D^*(x,\omega)\nabla p(x,\omega)) &= 1, \qquad x \in G \subset \mathbb{R}^d, \qquad \omega \in \Omega, \\ p(x,\omega) &= 0, \qquad x \in \partial G, \qquad \omega \in \Omega. \end{aligned}$$

Say that we are interested in computing $\mathbb{E}[P]$, where $P(\omega) = \int_G p^2 \mathrm{d}x = ||p||^2_{L^2(G)}(\omega)$. Set $D^* = e^{u(x,\omega)}$, where $u \sim \mathcal{N}(0, \mathcal{C}(x,y))$ is a (Matérn) Gaussian random field.

Simple approach: standard Monte Carlo (MC) method: $\mathbb{E}[P] \approx \frac{1}{N}\sum_{n=1}^{N} P_L(\omega^n)$.

$$\text{root-mean-square error} = \text{bias} + \text{statistical error} < \varepsilon$$

Bias: Assume cost of sampling one realisation of P_L that is accurate enough is $O(\varepsilon^{-q})$.
Statistical error: MC convergence rate $O(N^{-1/2})$ means we need $O(\varepsilon^{-2})$ samples.
Total MC complexity: $O(\varepsilon^{-2-q})$.

Solve for P using a hierarchy of $L+1$ (possibly non-nested) meshes to obtain the approximations P_ℓ of different accuracy for $\ell = 0, \dots, L$, then

$$\mathbb{E}[P] \approx \mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{\ell=0}^{L} \mathbb{E}[P_\ell - P_{\ell-1}].$$

Solve for P using a hierarchy of $L+1$ (possibly non-nested) meshes to obtain the approximations P_ℓ of different accuracy for $\ell = 0, \dots, L$, then

$$\mathbb{E}[P] \approx \mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{\ell=0}^{L} \mathbb{E}[P_\ell - P_{\ell-1}].$$

Apply standard MC to each term on the RHS:

$$\mathbb{E}[P] \approx \frac{1}{N_0} \sum_{n=1}^{N_0} P_0(\omega_0^n) + \sum_{\ell=0}^{L} \frac{1}{N_\ell} \sum_{n=1}^{N_\ell} [P_\ell(\omega_\ell^n) - P_{\ell-1}(\omega_\ell^n)].$$

Under suitable conditions $\Longrightarrow$ optimal N_ℓ known and $O(\varepsilon^{-2})$ complexity, $O(\varepsilon^{-q})$ better than MC.

Other options: quasi Monte Carlo $\rightsquigarrow$ $O(\varepsilon^{-1-q})$ complexity.
multilevel Quasi Monte Carlo $\rightsquigarrow$ $O(\varepsilon^{-1})$ complexity.

Sampling the Gaussian field $u(x, \omega) \sim \mathcal{N}(0, \mathcal{C})$ is hard!

Gaussian field sampling

Sampling the Gaussian field $u(x, \omega) \sim \mathcal{N}(0, \mathcal{C})$ is hard!

Naïve approach

- Discretise u and compute a Cholesky factorization of the **dense** covariance matrix.

Better approaches

- Karhunen-Loève.
- Hierarchical matrices [Dölz et al. 2017, Feischl et al. 2018].
- FFT + circulant embeddings [Wood and Chan 1994, Dietrich and Newsam 1997].
- **SPDE approach** [Lindgren et al. 2009] (see next).

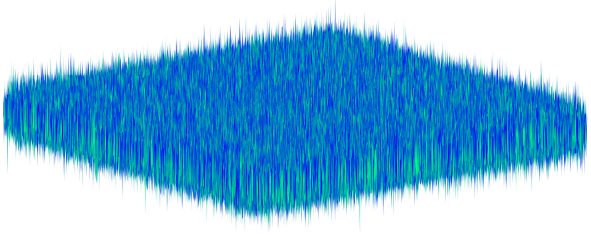
SPDE approach to Matérn field sampling [Lindgren et al. 2009]

UNIVERSITY OF OXFORD

The SPDE approach re-casts the sampling problem as the solution of a SPDE, e.g.

$$\mathcal{L}u = u - \Delta u = \dot{W}.$$

The (deterministic!) operator $\mathcal{L}$ determines the covariance of u.
$\dot{W}$ is **spatial white noise** and is defined through its action against test functions.



Spatial white noise in $[0,1]^2$.

Refs: [Abrahamsen 1997, Scheuerer 2010, Lindgren et al. 2009, Bolin et al. 2017, Khristenko et al. 2018]

The SPDE approach re-casts the sampling problem as the solution of a SPDE, e.g.

$$\mathcal{L}u = u - \Delta u = \dot{W}.$$

The (deterministic!) operator $\mathcal{L}$ determines the covariance of u.
$\dot{W}$ is **spatial white noise** and is defined through its action against test functions.

Let $V_h = \text{span}(\{\phi_i\}_{i=1}^{n_{\text{dofs}}})$ be the FEM approximation subspace used to solve the SPDE.
After discretisation, we obtain the linear system

$$\boxed{A\boldsymbol{u} = \boldsymbol{b}}, \quad \text{where} \quad \boldsymbol{b} \sim \mathcal{N}(0, M), \quad M_{ij} = \int_D \phi_i \phi_j \, \mathrm{d}x,$$

$$A_{ij} = \int_D \phi_i \phi_j \, \mathrm{d}x + \int_D \nabla\phi_i \cdot \nabla\phi_j \, \mathrm{d}x$$

Refs: [Abrahamsen 1997, Scheuerer 2010, Lindgren et al. 2009, Bolin et al. 2017, Khristenko et al. 2018]

Let $\{V_\ell\}_{\ell=0}^{L}$ be a hierarchy of (possibly non-nested) FEM approximation subspaces with $V_\ell = \text{span}(\{\phi_i^\ell\}_{i=1}^{n_{\text{dofs}}^\ell})$. On each MLMC level, we need to solve for u_ℓ and $u_{\ell-1}$,

$$\mathcal{L}u_\ell = \dot{W}, \quad \text{and} \quad \mathcal{L}u_{\ell-1} = \dot{W}, \quad \text{if } \ell > 0,$$

where we use the same white noise sample on both levels to enforce the MLMC coupling.

Let $\{V_\ell\}_{\ell=0}^{L}$ be a hierarchy of (possibly non-nested) FEM approximation subspaces with $V_\ell = \text{span}(\{\phi_i^\ell\}_{i=1}^{n_{\text{dofs}}^\ell})$. On each MLMC level, we need to solve for u_ℓ and $u_{\ell-1}$,

$$\mathcal{L}u_\ell = \dot{W}, \quad \text{and} \quad \mathcal{L}u_{\ell-1} = \dot{W}, \quad \text{if } \ell > 0,$$

where we use the same white noise sample on both levels to enforce the MLMC coupling.

After discretisation, the above yields the linear system,

$$\left[\begin{array}{c|c} A^\ell & 0 \\ \hline 0 & A^{\ell-1} \end{array}\right] \left[\begin{array}{c} \mathbf{u}^\ell \\ \hline \mathbf{u}^{\ell-1} \end{array}\right] = \left[\begin{array}{c} \mathbf{b}^\ell \\ \hline \mathbf{b}^{\ell-1} \end{array}\right] = \mathbf{b}, \quad \text{where} \quad \boldsymbol{b} \sim \mathcal{N}(0, M).$$

Here M is the mass matrix over $V_\ell + V_{\ell-1}$, (set $V_{-1} = \emptyset$), i.e.

$$M = \left[\begin{array}{c|c} M^\ell & M^{\ell,\ell-1} \\ \hline (M^{\ell,\ell-1})^T & M^{\ell-1} \end{array}\right], \quad M_{ij}^{\ell,\ell-1} = \int \phi_i^\ell \phi_j^{\ell-1} \, \mathrm{dx}, \quad \text{if } \ell > 0.$$

NOTE: we do not require the FEM approximation subspaces to be nested!

How to sample b?

Sampling b is hard!

Sampling b is hard!

Naïve approach

- Factorise the covariance M using Cholesky (**cubic complexity!**)
- Works well if M **diagonal**. Previous work under this assumption [Lindgren et al. 2009, Osborn et al. 2017, Drzisga, et al. 2017, Du and Zhang 2002].

How to sample b?

Sampling b is hard!

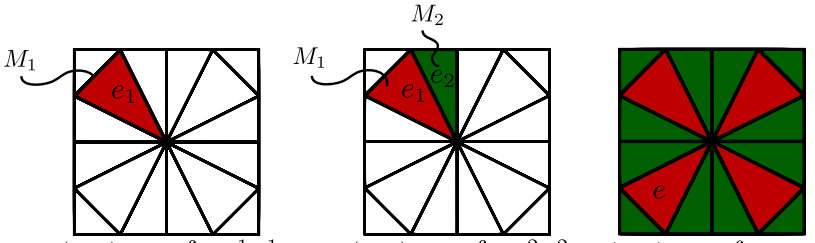
Naïve approach

- Factorise the covariance M using Cholesky (**cubic complexity!**)
- Works well if M **diagonal**. Previous work under this assumption [Lindgren et al. 2009, Osborn et al. 2017, Drzisga, et al. 2017, Du and Zhang 2002].

Our Work [C. et al. 2018]

- We do not require M to be diagonal.
- We can sample **b** with **linear complexity**.

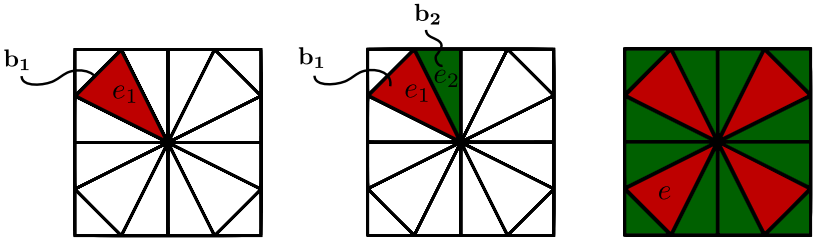
The idea is to exploit the FEM assembly of the mass matrix [Wathen 1987]:



$(M_1)_{ij} = \int_{e_1} \phi_i^1 \phi_j^1$ $(M_2)_{ij} = \int_{e_2} \phi_i^2 \phi_j^2$ $(M_e)_{ij} = \int_e \phi_i^e \phi_j^e$

$$M = L^T \begin{bmatrix} M_1 & 0 & \cdots \\ 0 & M_2 & \ddots \\ \vdots & \ddots & \ddots \end{bmatrix} L = L^T \mathrm{diag}_e(M_e) L. \tag{1}$$

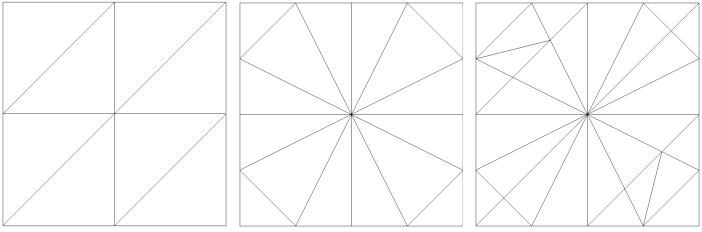
Disjoint pieces of white noise are independent. Sample small independent **local white noise vectors** on each cell and assemble the contributions together.



$$\mathcal{N}(0, M) \sim \mathbf{b} = L^T \begin{bmatrix} \mathbf{b_1} \\ \mathbf{b_2} \\ \vdots \end{bmatrix} = L^T \mathsf{vstack}_e(\mathbf{b}_e) \tag{2}$$

NOTE: only local Cholesky factorisations are needed, and trivially parallelisable!

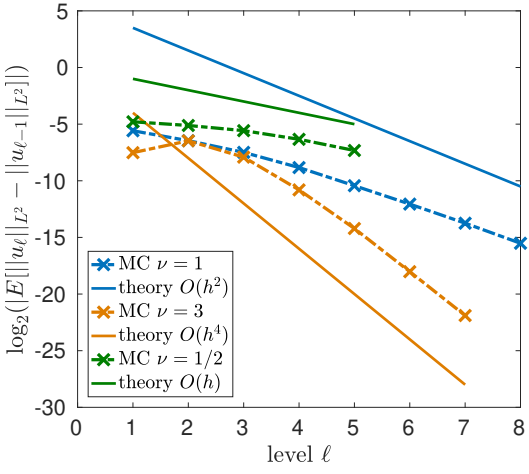
Construct a FEM subspace S_ℓ such that V_ℓ and $V_{\ell-1}$ are both nested within S_ℓ. This requires a **supermesh construction** [Farrell 2009]:

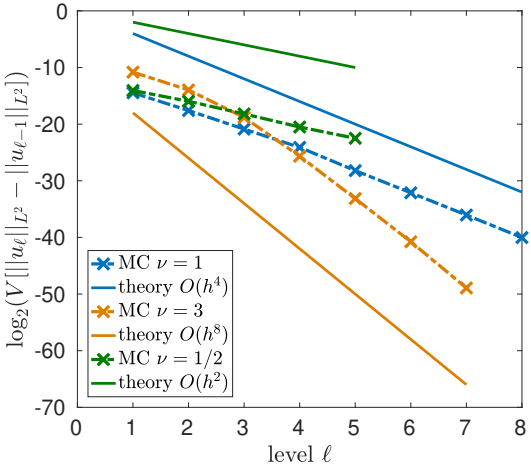


Sample $\boldsymbol{b}_S^\ell \sim \mathcal{N}(0, M_S^\ell)$ where M_S^ℓ is the mass matrix over S_ℓ and get $\boldsymbol{b}_\ell$ and $\boldsymbol{b}_{\ell-1}$ by transferring $\boldsymbol{b}_S^\ell$ onto V_ℓ and $V_{\ell-1}$ using **nested interpolation**.

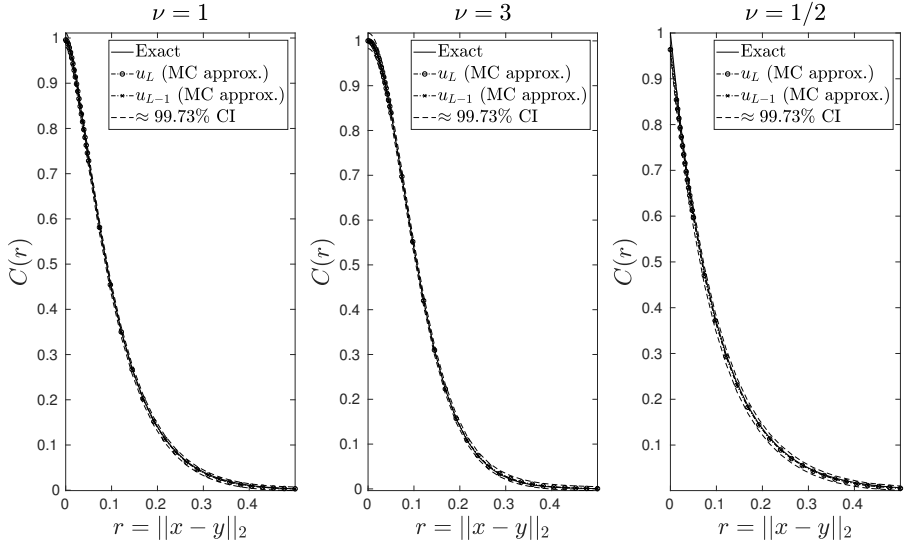
Linear cost complexity [C. and Farrell 2020], and trivially parallelisable!
Can also sample (ML)QMC-ready white noise samples in linear cost [C. et al. 2020]!

Numerical results: FEM convergence (2D and 3D)

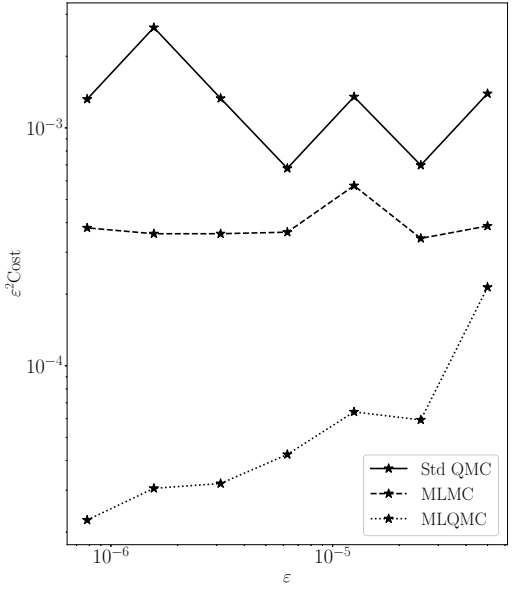




Numerical results: covariance convergence (2D and 3D)



QMC vs MLMC vs MLQMC cost comparison (2D)



Overview

Motivation and physiological background

Monte Carlo methods for PDEs with random coefficients

Application: uncertainty quantification in brain tracer simulation

Conclusions

UNIVERSITY OF OXFORD

doi:10.1093/brain/awx191 BRAIN 2017: 140; 2691–2705 | 2691

BRAIN
A JOURNAL OF NEUROLOGY

Glymphatic MRI in idiopathic normal pressure hydrocephalus

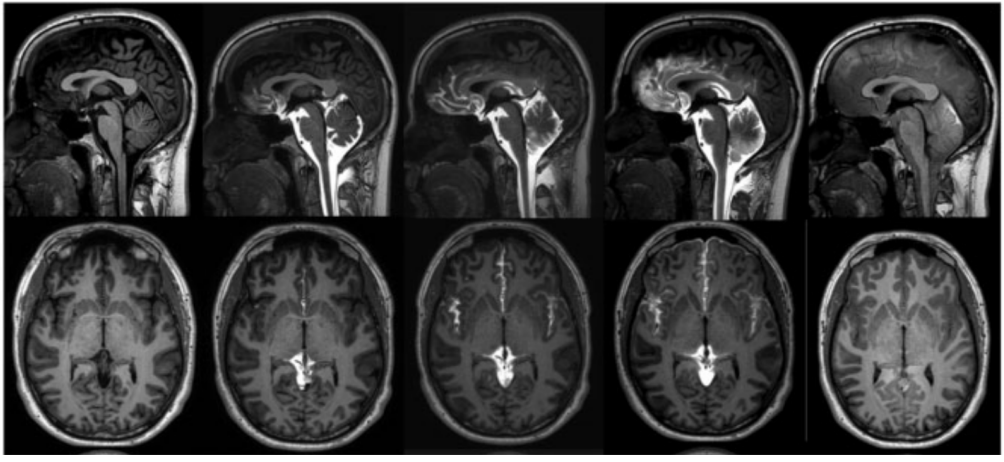
Geir Ringstad,[1,2] Svein Are Sirirud Vatnehol[3] and Per Kristian Eide[2,4]

Recall: $$\dot{c}(t,x,\omega) + \nabla \cdot (\boldsymbol{v}(x,\omega)c(t,x,\omega)) - \nabla \cdot (D^*(x,\omega)\nabla c(t,x,\omega)) = 0,$$

We consider different models for $\boldsymbol{v}$ and D^*, each corresponding to a different hypothesis on solute movement available in the medical literature. See [C. et al. 2019, C. et al. 2020].

Objective: performing UQ to find which models are more likely to match with experimental data and thus hold true in practice, and which MC method works best.

Glymphatic MRI in humans: clinical results

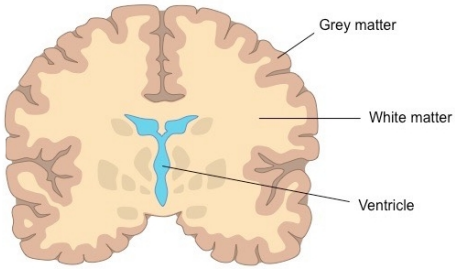


Ringstad et al. inject a contrast agent (gadobutrol) directly into the spinal canal at the lumbar level and they use MRI to measure how it spreads in the brain.

We impose no tracer flux across the ventricles and we model the upward movement of the tracer with the external interface BC

$$c(t,x) = c_{\text{CSF}}(t,c)h(t,x),$$

$c_{\text{CSF}}(t,c) =$ CSF tracer concentration,
$h(t,x) =$ spatial distribution of tracer.



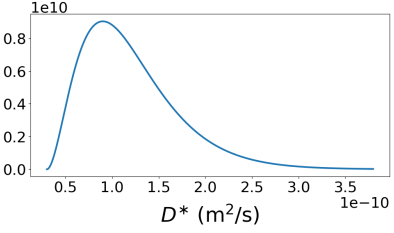
We compute the probability density function of the following **quantities of interest**:

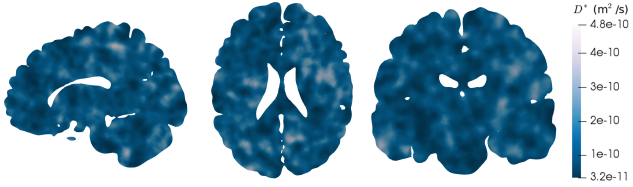
- $Q_g(t,\omega)$ = total amount of tracer in the **grey matter** at time t.
- $Q_w(t,\omega)$ = total amount of tracer in the **white matter** at time t.

Physical and physiological restrictions:

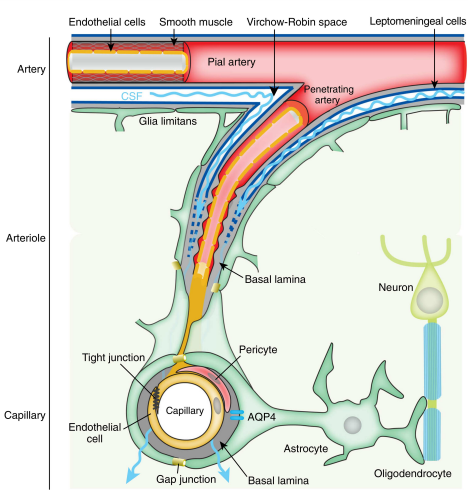
- Average tracer diffusivity and patient-variability known.
- $D^* > 0$ (diffusion coefficient must be positive).
- $D^*(x, \omega)$ varies in space at a given length scale $\lambda_D \approx 0.01$m.

Modelling solution:



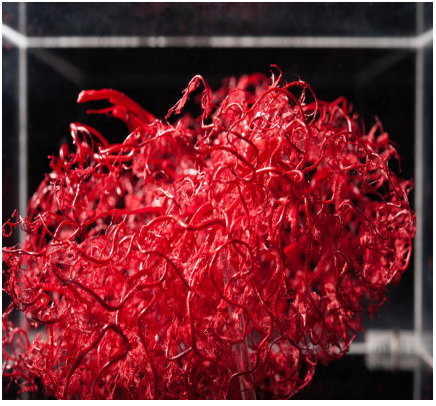


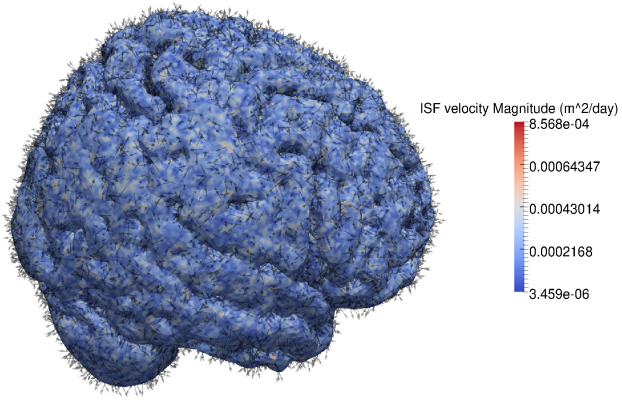
According to the glymphatic hypothesis CSF enters the brain from para-arterial spaces (arteries) and exits from para-venous spaces (veins).



[Source: *The Glymphatic System: A Beginner's Guide* by Jessen et al., 2015]

A complex vascular structure



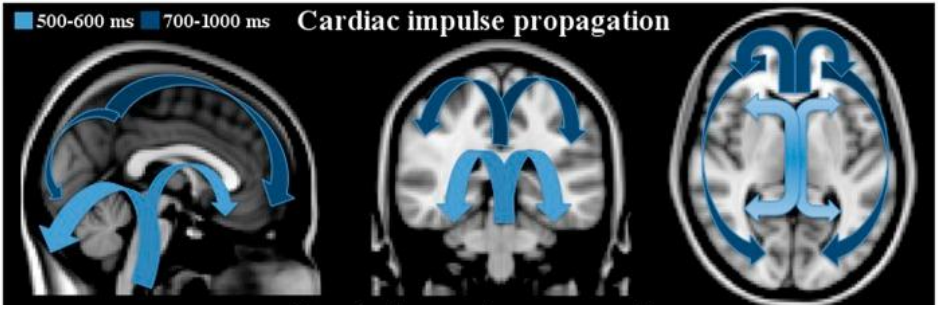


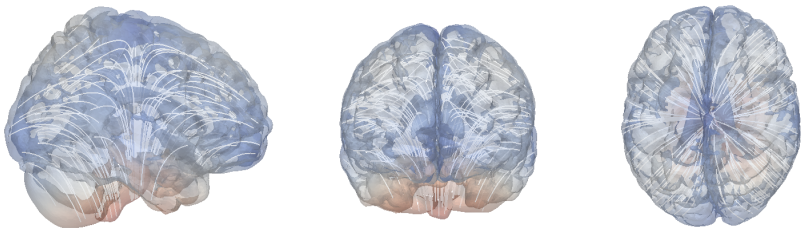
We assume that significant changes within the velocity field happen after a distance proportional to the mean distance between arterioles and venules. Set $\lambda_v = 560\mu$m.

Other requirements: Prescribed average velocity magnitude. No ISF enters/leaves the system ($\nabla \cdot \boldsymbol{v} = 0$). Arteries and veins equally likely at any point in space.

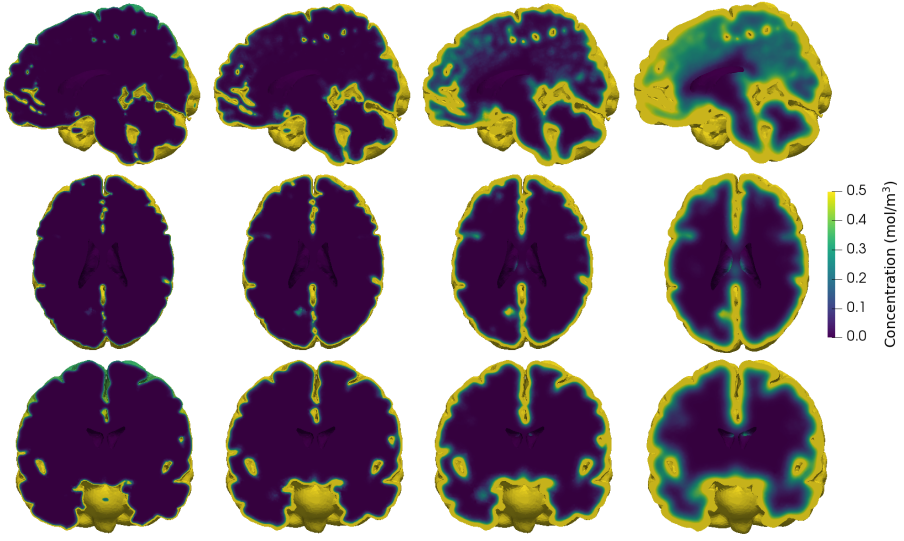
Glymphatic velocity model with directionality

Top picture from [Kiviniemi et al. 2016]

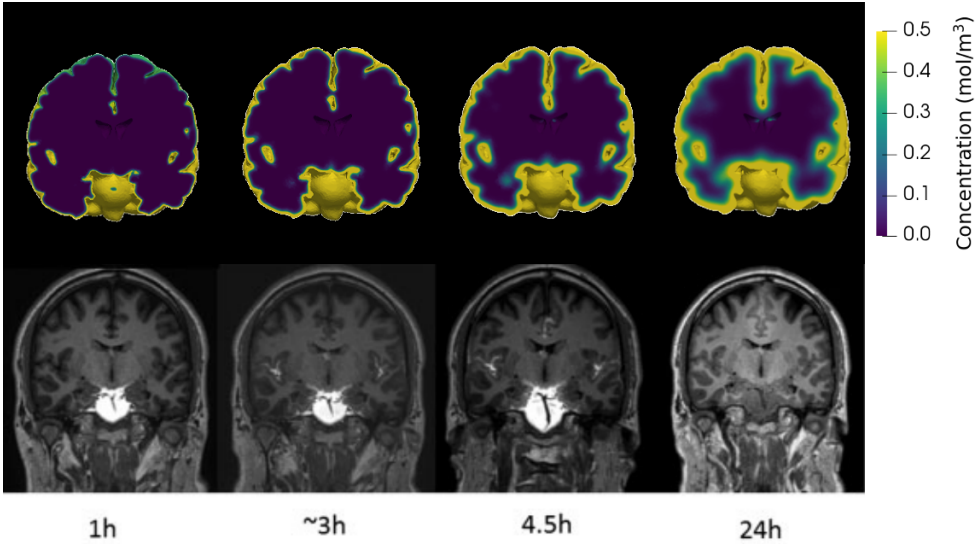




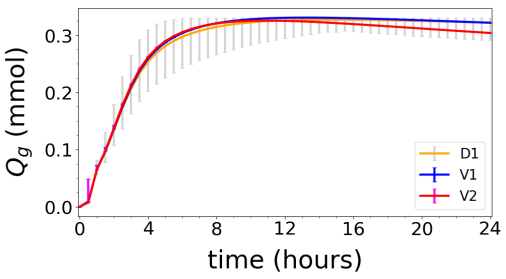
Simulation results (pure homogeneous diffusion)

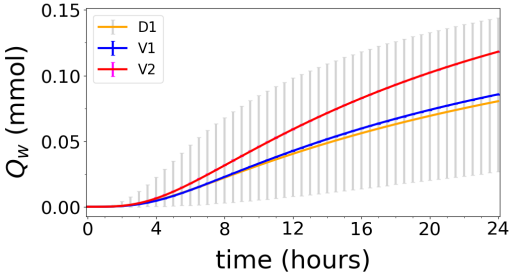


Simulation results (pure homogeneous diffusion)



Key observation: tracer penetrates into gray matter, but not into deep central regions.





$D1$ = pure hom. diffusion, $V1$ = glymphatic model, $V2 = V1$ + directionality.

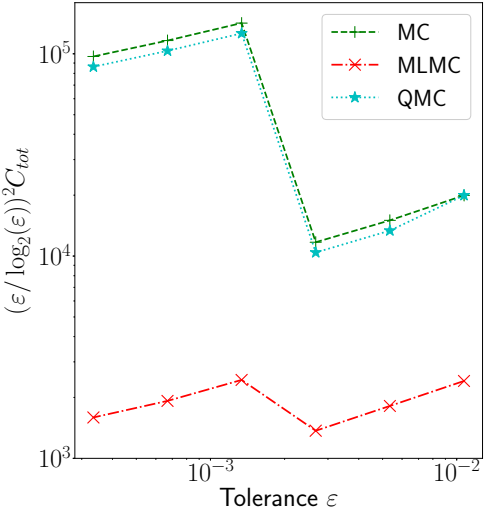
Key observations:

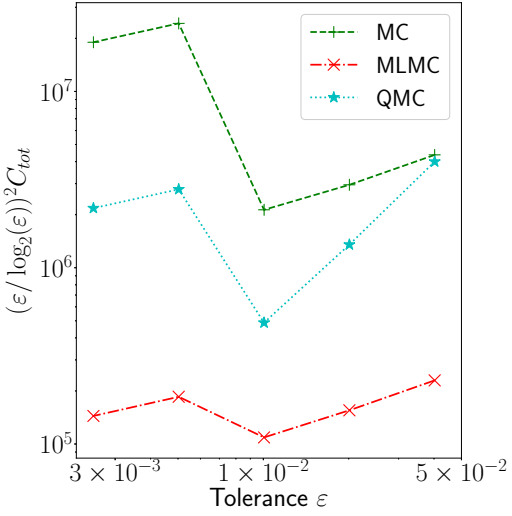
1) Glymphatic alone and pure diffusion are not sufficient to match experimental data and a bulk flow with directionality is required.

2) Models using only a homogeneous diffusion coefficient are subject to large patient-to-patient variability.

Algorithm comparison: MC vs MLMC vs QMC

Solve on Abel, the Norwegian supercomputing cluster (TOP500 in 2012, now retired).





Conclusion: MLMC superior ($O(1)$ wks time), vs standard MC $O(10)$ yrs (in parallel)!

Overview

Motivation and physiological background

Monte Carlo methods for PDEs with random coefficients

Application: uncertainty quantification in brain tracer simulation

Conclusions

Conclusions

Gaussian field sampling and non-nested ML(Q)MC:

- When dealing with uncertainty in spatially varying quantities random fields are required, but their sampling can be expensive. We used the SPDE approach.
- Thanks to our research, we can sample Matérn-Gaussian fields in linear cost complexity, even in a non-nested ML(Q)MC setting via a supermesh construction.
- This is because we proved that supermeshes between quasi-uniform meshes have numbers of cells which are linear in the number of cells of the parent meshes.

UQ in brain tracer simulation:

- The human brain is a challenging subject to study and model parameters (and sometimes models themselves) are affected by a significant degree of uncertainty.
- The uncertainty in model predictions can be quantified with advanced Monte Carlo methods. ML(Q)MC seems to be the most efficient approach.
- Thanks to UQ we discovered the only two medical hypotheses leading to mathematical models that match experimental data.

Thank you for listening! If you want to know more:

UNIVERSITY OF OXFORD

Papers, slides, and more info at: `https://croci.github.io`

Shameless advert: My SIAM CSE talk on *Reduced precision solvers for parabolic PDEs*, Tue 2 March at 10:25 (minisymposium code: MS112).

---

[1] M. Croci. *Multilevel Monte Carlo methods for uncertainty quantification in brain simulations*. PhD thesis, University of Oxford, 2020.

[2] M. Croci, M. B. Giles, and P. E. Farrell. Multilevel quasi Monte Carlo methods for elliptic PDEs with random field coefficients via fast white noise sampling. *submitted to SIAM Journal on Scientific Computing*, 2020. URL `https://arxiv.org/abs/1911.12099`.

[3] M. Croci and P. E. Farrell. Complexity bounds on supermesh construction for quasi-uniform meshes. *Journal of Computational Physics*, 414:1–7, 2020. doi: 10.1016/j.jcp.2020.109459.

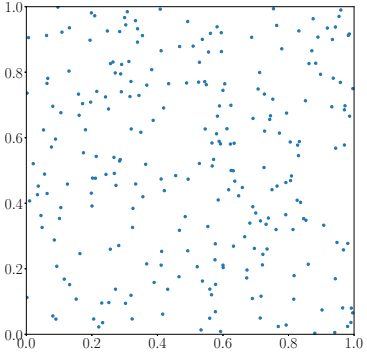
[4] M. Croci, V. Vinje, and M. E. Rognes. Fast uncertainty quantification of tracer distribution in the brain interstitial fluid with multilevel and quasi Monte Carlo. *International Journal for Numerical Methods in Biomedical Engineering*, pages 1–24, 2020. doi: 10.1002/cnm.3412.

[5] M. Croci, V. Vinje, and M. E. Rognes. Uncertainty quantification of parenchymal tracer distribution using random diffusion and convective velocity fields. *Fluids and Barriers of the CNS*, 16(32):1–21, 2019. doi: 10.1101/665109.

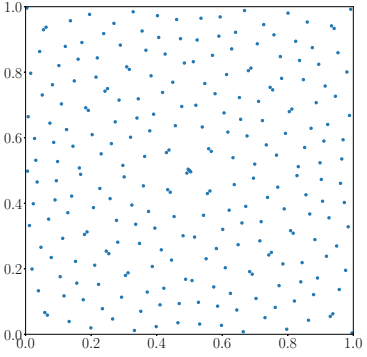
[6] M. Croci, M. B. Giles, M. E. Rognes, and P. E. Farrell. Efficient white noise sampling and coupling for multilevel Monte Carlo with non-nested meshes. *SIAM/ASA Journal on Uncertainty Quantification*, 6(4):1630–1655, 2018. doi: 10.1137/18M1175239.

Approximate $\mathbb{E}[P]$ with an s-dimensional integral over $[0,1]^s$:

$$\mathbb{E}[P] \approx \int_{[0,1]^s} Y(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \approx \frac{1}{N} \sum_{n=1}^{N} Y(\boldsymbol{x}_n),$$



Pseudo-random points.



Low-discrepancy point sequence (Sobol').

QMC convergence rate up to $O(N^{-1}) \Rightarrow$ up to $O(\varepsilon^{-1-q})$ complexity.
Combine with MLMC to get $O(\varepsilon^{-1})$ complexity and MLQMC.

Extending work to QMC and MLQMC: the challenge

UNIVERSITY OF OXFORD

- Low-discrepancy sequences are extremely uniform in the first few dimensions and in low-dimensional projections, but less so across the whole hypercube. Therefore **QMC works best when the integrand has low effective dimension** (adapted from [Caflish et al. 1997 and Joe and Kuo 2008]).
- For good QMC convergence we need to **order the dimensions** in our QMC integrands in order of decaying importance so that the largest error components are on the first dimensions.

Extending work to QMC and MLQMC: the challenge

UNIVERSITY OF OXFORD

- Low-discrepancy sequences are extremely uniform in the first few dimensions and in low-dimensional projections, but less so across the whole hypercube. Therefore **QMC works best when the integrand has low effective dimension** (adapted from [Caflish et al. 1997 and Joe and Kuo 2008]).
- For good QMC convergence we need to **order the dimensions** in our QMC integrands in order of decaying importance so that the largest error components are on the first dimensions.
- **Solution**: rewrite $\dot{W}$ in terms of basis functions that naturally expose the leading order dimensions in the QMC integrand. Wavelets work well [Kuo et al. 2015, Hermann and Schwab 2017].
- **Result:** we can sample (ML)QMC-ready white noise in linear complexity, even with non-nested hierarchies [C. et al. 2020]!

Definition [Spatial White Noise $\dot{W}$ (Hida et al. 1993)]

For any $\phi \in L^2(D)$, define $\langle \dot{W}, \phi \rangle := \int_D \phi \, \mathrm{d}\dot{W}$. For any $\phi_i,\ \phi_j \in L^2(D)$, $b_i = \langle \dot{W}, \phi_i \rangle$, $b_j = \langle \dot{W}, \phi_j \rangle$ are zero-mean Gaussian random variables, with,

$$\mathbb{E}[b_i b_j] = \int_D \phi_i \phi_j \, \mathrm{d}x =: M_{ij}, \quad \mathbf{b} \sim \mathcal{N}(0, M).$$

Let $V_h = \mathrm{span}(\{\phi_i\}_{i=1}^{n_{\mathrm{dofs}}}) \subseteq H_0^1$ be the FEM approximation subspace used to solve the SPDE $\mathcal{L}^k u = \dot{W}$. After discretisation, we obtain the linear system

$$\boxed{A^k \boldsymbol{u} = \boldsymbol{b}}, \quad \text{where} \quad \boldsymbol{b} \sim \mathcal{N}(0, M), \quad M_{ij} = \int_D \phi_i \phi_j \, \mathrm{d}x,$$

$$A_{ij} = \int_D \phi_i \phi_j \, \mathrm{d}x + \int_D \kappa^{-2} \nabla \phi_i \cdot \nabla \phi_j \, \mathrm{d}x$$