Calculating the Malliavin derivative of some stochastic mechanics problems

Paul Hauseux[1], Jack S. Hale[1], Stéphane P. A. Bordas[1,2]*

1 Institute of Computational Engineering, University of Luxembourg, 6 Avenue de la Fonte, 4362 Esch-sur-Alzette, Luxembourg, **2** Cardiff School of Engineering, Cardiff University, The Queen's Building, The Parade, Cardiff, Wales, CF24 3AA, United Kingdom

* stephane.bordas@uni.lu

Abstract

The Malliavin calculus is an extension of the classical calculus of variations from deterministic functions to stochastic processes. In this paper we aim to show in a practical and didactic way how to calculate the Malliavin derivative, the derivative of the expectation of a quantity of interest of a model with respect to its underlying stochastic parameters, for four problems found in mechanics. The non-intrusive approach uses the Malliavin Weight Sampling (MWS) method in conjunction with a standard Monte Carlo method. The models are expressed as ODEs or PDEs and discretised using the finite difference or finite element methods. Specifically, we consider stochastic extensions of; a 1D Kelvin-Voigt viscoelastic model discretised with finite differences, a 1D linear elastic bar, a hyperelastic bar undergoing buckling, and incompressible Navier-Stokes flow around a cylinder, all discretised with finite elements. A further contribution of this paper is an extension of the MWS method to the more difficult case of non-Gaussian random variables and the calculation of second-order derivatives. We provide open-source code for the numerical examples in this paper.

Introduction

The classical derivative is a fundamental tool of calculus that is widely used across every field of mathematics, science and engineering. Various generalisations and extensions of the classical derivative, e.g. local and/or partial Frechét and Gâteaux derivatives [1], are now common tools in the repertoire of practitioners working in many fields. Modern extensions such as fractional and non-local derivatives are finding increasing use in several fields of science and technology, see e.g. [2–6]. The semi-inverse method of [7] is a powerful tool for the establishment of variational principles (Euler-Lagrange) from governing equations for physical problems.

By contrast, the Malliavin calculus [8], an extension of the notions of classical calculus of variations to stochastic processes, is certainly less widely known. In our view, this is probably because the vast majority of papers written on the subject require study of mathematics and stochastics to an advanced level. However, we think that Malliavin calculus deserves a wider audience. The objective of this paper then is introduce the Malliavin derivative as a useful numerical tool for practitioners to understand the behaviour of stochastic PDEs in mechanics,

Competing interests: The authors have declared that no competing interests exist.

rather than to fully explain the technicalities of Malliavin calculus. Interested readers are referred to e.g. [8–10] for a full mathematical treatment.

We are not the first to apply Malliavin calculus as a useful tool for practical computation. The Malliavin calculus can be used to efficiently calculate the Greeks, the sensitivity of financial instruments to their underlying parameters e.g. [11–14]. In the physical sciences we are aware of only a handful of recent papers that use techniques inspired by the Malliavin calculus to understand the behaviour of systems with stochastic behaviour. We are not aware of any papers in the engineering mechanics community on the topic. [15] introduced the methodology of Malliavin Weight Sampling (MWS), the method we adopt in this paper, and applied it to the simulation of particles undergoing Brownian motion. [16] presented a more general framework for deriving the MWS update rules and its practical implementation. [17] used the MWS to evaluated linear response functions of particle systems forced by coloured noise. When the coefficients of the models are assumed to follow known statistical distributions, then the likelihood ratio method can be seen as a Malliavin weighting function [11, 12, 18]. The Malliavin theory is however more general and allows the determination of the optimal weight with minimum variance even if the specification of the stochastic parameters involved in the model are not known explicitly.

The contribution of this paper is as follows; we show the application of the Malliavin Weight Sampling method [15] to four archetypal problems in mechanics. Unlike the examples in [15], we consider some models defined by partial differential equations (PDEs) that are discretised using the finite element method. We make a new extension of the MWS method to parameters defined by non-Gaussian distributions. This has important practical value because it is often important to model parameters with distributions that preclude realisations with non-physical values, e.g. positive viscosity in a fluid mechanics problem. Finally we extend the MWS method in [15] to the calculation of second-order derivatives.

An outline of this paper is as follows; we give an outline of the MWS method and use the MWS method to study the behaviour of a simple Kelvin-Voigt visco-elastic system with Gaussian and non-Gaussian stochastic variables respectively. We extend the analysis of the Kelvin-Voigt system to the second derivative. We then study; a 1D elastic bar, a hyperelastic bar prone to buckling, and Navier-Stokes flow around a cylinder, all discretised in space using the finite element method. Finally we summarise and suggest some interesting avenues for future research.

The Malliavin Weight Sampling (MWS) method

Problem setting

Consider a non-linear, possibly time-dependent stochastic partial differential equation $F(u, m) = 0$ with random parameter m. For each possible value of m, u is the solution of the PDE and therefore u depends explicitly on m ($m \mapsto u(m)$). To simplify the notation, the spatial position x and time t are omitted but it is understood that u can also depend on $x \in \Omega \subset \mathbb{R}^d$ where $d = \{1, 2, 3\}$ is the spatial dimension of the domain and/or $t \in \mathbb{R}^+$.

Let $(\Omega_p, \mathcal{F}, P)$ a probability space where Ω_p is the sample space, $\mathcal{F}$ is a σ-algebra of subsets of Ω_p and P is a probability measure. We are interested to evaluate the expected value of a quantity of interest $J(m) = J(u(m))$ denoted by $\mathbb{E}[J]$ [19]:

$$\mathbb{E}[J] := \int_{\Omega_p} J(u(\omega)) \cdot dP(\omega). \tag{1}$$

In a practical way, if m is a random variable with probability density function f_m, Eq (1) writes:

$$\mathbb{E}[J] := \int_{\mathbb{R}} J(u(x)) \cdot f_m(x) \mathrm{d}x. \tag{2}$$

As we will see, the Malliavin Weight Sampling method (MWS) [16] allows the evaluation of the sensitivity of the expected value of the quantity of interest with respect to the mean value of the stochastic parameter m as [9, 11, 16, 18]:

$$\frac{\partial \mathbb{E}[J]}{\partial \bar{m}} = \mathbb{E}[Jq_m], \tag{3}$$

where q_m is the Malliavin weight for the parameter m and $\bar{m}$ is the mean of m. Under certain condition of regularity [11, 20, 21] when the probability density function (PDF) of the parameter m is known, the Malliavin weight q_m associated can be computed directly from the PDF of m. This approach can be viewed as an integration by parts, and is a direct result of Malliavin calculus where we take the derivative of random functions rather than the classical derivative. We emphasise again the quite different nature of the above derivative Eq (3) to the classical notion of a derivative from elementary calculus.

In Eq (3), we suppose that the quantity of interest J does not depend explicitly on the parameter m. Later we introduce a more general equation Eq (35) that must be considered if in fact J does depend on m.

The simplest approach to calculate $\mathbb{E}[Jq_m]$, and the one we use exclusively in this paper, is to use the standard Monte Carlo estimator; that is, take Z independent and identically distributed (iid) realisations m_z of m, solve for $J_z := J(u(m_z))$ before taking the sample mean of the set of realisations $\{J_1, \ldots, J_Z\}$:

$$\frac{\partial \mathbb{E}[J]}{\partial \bar{m}} = \mathbb{E}[Jq_m] \approx \frac{1}{Z} \sum_{z=1}^{Z} J(m_z) \cdot q_m(m_z). \tag{4}$$

From the central limit theorem, the error in Eq (4) is normally distributed with variance $Z^{-1} V$ where V is the variance of Jq_m.

What will not be clear to the reader at this stage is how to determine the Malliavin weights. Through a simple practical examples in the next section, we will explain how to use the MWS method, determine the specification of the weights for both Gaussian and non-Gaussian distributions on the parameter m, and thus calculate the Malliavin derivative Eq (3).

Kelvin-Voigt model

The Kelvin-Voigt constitutive model with Young's modulus E, viscosity η and loading stress σ can be written as the following linear ordinary differential equation:

$$E\epsilon(t) + \eta \frac{\mathrm{d}\epsilon(t)}{\mathrm{dt}} = \sigma. \tag{5}$$

A schematic of this model is shown in Fig 1.

The initial condition on the strain is $\epsilon(t = 0) = 0$ and we study the response of the system for time $t \in [0, T]$. Our quantity of interest functional is the value of the strain at time t, i.e.:

$$J := \epsilon(t) \tag{6}$$

and we are interested in its expected value (mean) $\mathbb{E}[\epsilon(t)]$.
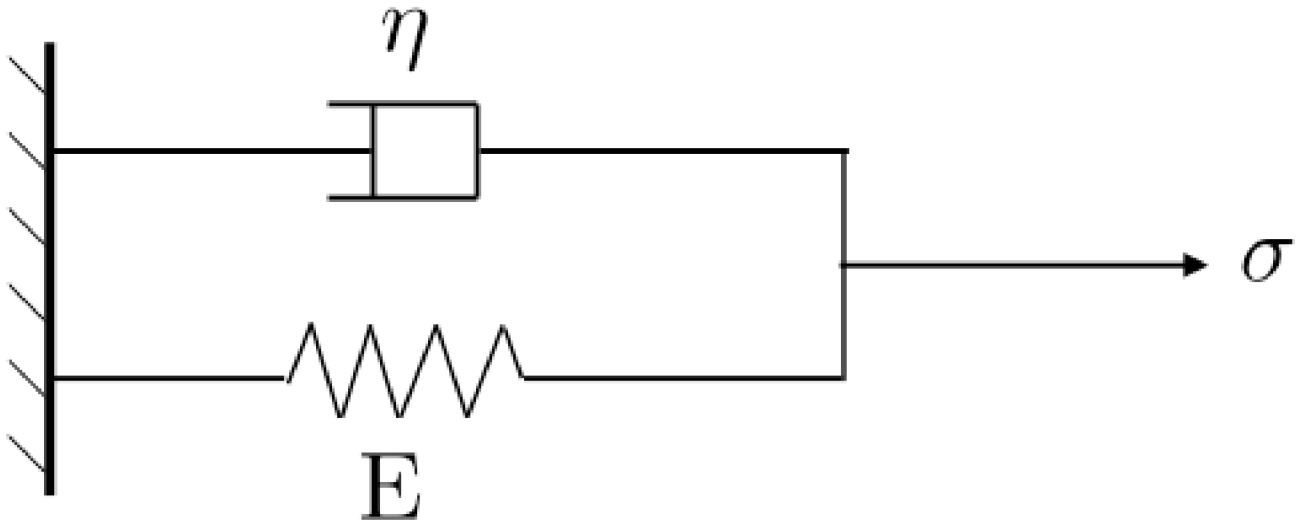
Fig 1. Schematic of a Kelvin-Voigt model with Young's modulus *E*, viscosity *η* and loading stress *σ*. We model the loading stress *σ* as a random noise (random variable), inducing a random strain ϵ as the output of the model.

Gaussian case

We first consider the case that the randomness can be modelled as a Gaussian random variable. A similar model is shown in [16].

We choose choose to model the stress as a random noise:

$$\sigma(t) = \sigma_0 + \alpha\xi, \tag{7}$$

where σ_0 and α are constant and ξ is a a Gaussian random variable with zero mean and unit variance. ξ represents the uncertainty related to the value of the stress σ.

From Eq (7), the mean value of σ is σ_0 and the variance of σ is equal to α^2. We assume throughout that the Young's modulus E and the viscosity η are perfectly known. Given that the forcing stress σ for the system is random, the strain ϵ is also random. The goal then is to evaluate the derivative of the expected value of the strain with respect to the mean value σ_0:

$$\frac{\partial \mathbb{E}[\epsilon(t)]}{\partial \sigma_0} \tag{8}$$

using the method of MWS.

We choose to solve the ODE Eq (5) using an Euler explicit finite difference method with time step δt:

$$\epsilon(t + \delta t) = \epsilon(t) + \frac{\delta t}{\eta}\left[\sigma_0 - E\epsilon(t) + \frac{\alpha\xi}{\sqrt{\delta t}}\right]. \tag{9}$$

Remark. Note that the multiplying term before ξ contains $\sqrt{\delta t}$ and not δt. This is a 'conforming' discretisation of the stochastic noise term, resulting in a dependence of the variance of the random parameter on the discretisation size. Informally, taking the limit, we can recover the original ODE Eq (5) as:

$$\mathbb{E}\left[\frac{\mathrm{d}\epsilon}{\mathrm{d}t}\right] = (\sigma_0 - E\epsilon(t))/\eta, \tag{10}$$

and:

$$\mathbb{V}\left[\eta\frac{\mathrm{d}\epsilon}{\mathrm{d}t} - (\sigma_0 - E\epsilon(t))\right] = \alpha^2. \tag{11}$$

Where $\mathbb{V}[\cdot]$ is the the variance. Given that $\mathbb{E}[\sqrt{\delta t}\xi] = 0$ and $\mathbb{E}[(\sqrt{\delta t}\alpha\xi)^2] = \alpha^2\delta t$, the numerical method in Eq (9) is consistent in the following sense:

$$\lim_{\delta t \to 0}\frac{1}{\delta t}\mathbb{E}[\epsilon(t + \delta t) - \epsilon(t)] = (\sigma_0 - E\epsilon(t))/\eta, \tag{12}$$

$$\lim_{\delta t \to 0}\frac{1}{\delta t}\mathbb{E}[[\eta(\epsilon(t + \delta t) - \epsilon(t)) - \delta t(\sigma_0 - E\epsilon(t))]^2] = \alpha^2. \tag{13}$$

For this next part, we adopt the same notation as [16]. We denote ϵ the strain of the the system at time t and we denote ϵ' the strain of the the system at time $t + \delta t$. Furthermore, we let $P(\epsilon)$ and $P(\epsilon')$ be the probability that the strain of the system is ϵ and ϵ' respectively. The propagator $W(\epsilon \to \epsilon')$ must satisfy:

$$P(\epsilon') = \int_\epsilon W(\epsilon \to \epsilon')P(\epsilon)\mathrm{d}\epsilon, \tag{14}$$

$$\int_{\epsilon'} W(\epsilon \to \epsilon')\mathrm{d}\epsilon' = 1. \tag{15}$$

Eq (14) means that the probability that the strain of the system is ϵ' is the sum (integral) of all the probabilities to be at ϵ multiplied by the probability that the system passes from state ϵ to ϵ' during δt. Condition Eq (15) comes from the integration of the first condition over ϵ'.

To derive the analytical expression of the propagator in Eq (18) we start with the fact that $\xi^\star = \alpha\sqrt{\delta t}\xi$ is Gaussian $N \sim (0, \delta t\alpha^2)$, hence the probability density function is known and must satisfy the following condition:

$$\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\delta t\alpha^2}}\exp\left(-\frac{\xi^{\star 2}}{2\delta t\alpha^2}\right)\mathrm{d}\xi^\star = 1. \tag{16}$$

With an integration by substitution from Eq (9), with:

$$\xi^\star = ((\epsilon' - \epsilon)\eta - \delta t\sigma_0 + \delta tE\epsilon), \tag{17}$$

we can then show the expression of the propagator, the probability that the system passes from state ϵ to ϵ' during δt is given by [16]:

$$W(\epsilon \to \epsilon') = \frac{\eta}{\sqrt{2\pi\delta t\alpha^2}}\exp\left(-\frac{((\epsilon' - \epsilon)\eta - \delta t\sigma_0 + \delta tE\epsilon)^2}{2\delta t\alpha^2}\right). \tag{18}$$

With the propagator in hand we will now see how it is possible to evaluate the Malliavin derivative with the MWS method. To recap, we denote by $J(\epsilon)$ a quantity of interest of our system and we want to compute the derivative of the mean value of this quantity of interest $\mathbb{E}[J]$ with respect to a parameter $\bar{m}$, in this case σ_0 when $\sigma = \sigma_0 + \alpha\xi$.

The form of the Malliavin weights q_m can be obtained using the following procedure. First, we know that with $dP(\epsilon) = P(\epsilon)d\epsilon$, Eq (1) we can write:

$$\mathbb{E}[J] = \int_\epsilon JP(\epsilon)\mathrm{d}\epsilon, \tag{19}$$

and by taking the derivative of Eq (19) [11, 16, 18]:

$$\frac{\partial \mathbb{E}[J]}{\partial \bar{m}} = \int_\epsilon JP(\epsilon) \frac{\partial \ln P}{\partial m} \, d\epsilon. \tag{20}$$

To define a set of rules for updating q_m, we differentiate Eq (14) with respect to m:

$$P(\epsilon') \frac{\partial \ln P'}{\partial \bar{m}} = \int_\epsilon W(\epsilon \rightarrow \epsilon') P(\epsilon) \frac{\partial \ln P + \partial \ln W}{\partial m} \, d\epsilon, \tag{21}$$

and we obtain the following rule for updating the Malliavin weight:

$$q_m(t + \delta t) = q_m(t) + \frac{\partial \ln W}{\partial m}. \tag{22}$$

In the example of random stress with $\sigma = \sigma_0 + \xi$, we obtain:

$$\frac{\partial \ln W}{\partial \sigma} = \sqrt{\delta t} \xi / \alpha. \tag{23}$$

For random Young's modulus $E = E_0 + \xi$ we would have the same expression. In the case of random viscosity $\eta = \eta_0 + \xi$ we would obtain following the same logic:

$$\frac{\partial \ln W}{\partial \eta} = \xi / \alpha. \tag{24}$$

With the expression for the Malliavin weight Eq (23) in hand we can now implement an algorithm to calculate the derivative. The procedure is very simple; we take Z samples of the evolutions of the stochastic ODE using the explicit Euler scheme whilst simultaneously evolving the Malliavin weight q_m. At teach time step Algorithm 1 describes this procedure in more detail.

The deterministic constants are given to be $\eta = 1$, $E = 1$ and we take a time step of $\delta t = 0.01$ for the finite difference scheme. We evaluate by the MWS method the derivative of the expected value of ϵ with respect to σ_0 for a loading time $t \in [0, T]$ with $T = 30s$. In this example the number of realisations is fixed at $Z = 20000$. We compare the results with the analytical solution which is:

$$\frac{\partial \mathbb{E}[\epsilon]}{\partial \sigma_0} = \frac{1}{E} \left(1 - \exp \left\{ -\frac{Et}{\eta} \right\} \right). \tag{25}$$

We briefly remark that for all numerical results presented in this paper there are two sources of errors committed with respect to the undiscretised problem. The first error is due to the deterministic approximation of the PDE (finite difference or finite element method), and the second due to the stochastic approximation (Monte Carlo estimator). In all cases we drive the error in the deterministic approximation of the PDE far lower than that in the stochastic approximation, such that the error is dominated by the number of realisations Z used in the Monte Carlo estimator.

In Fig 2 we can see that the MWS method gives a good estimation of the Malliavin derivative, particularly in the non-steady state regime $t \leq 5s$. The relative error and so the global statistical error can become very high when the system reaches a steady state because the value of the sensitivity derivative is constant but the statistical error is compounded after each time step. To address this issue a technique can be employed based on the correlation function [16]. The reader is referred to [16] for further details. We have implemented this correlation correction, which we denote MWS-steady-state, and we can see in Fig 2 that the error is greatly
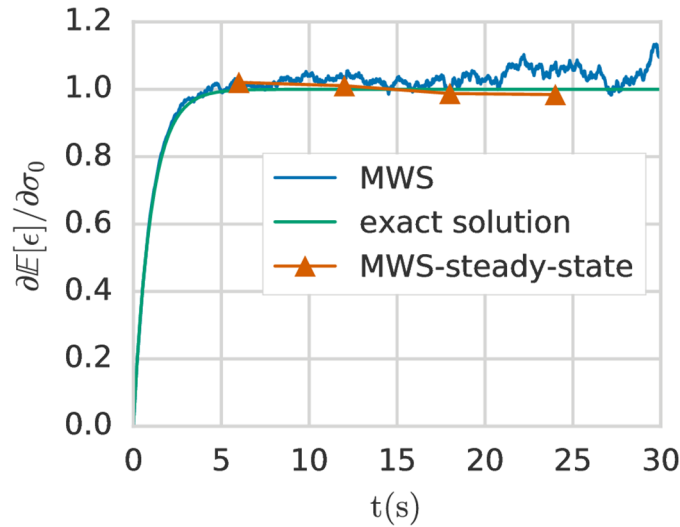
Fig 2. Malliavin derivative of the expected value of the strain with respect to the loading σ_0 for the Kelvin-Voigt model with uncertain stress modelled as a Gaussian random variable. Comparison between the exact solution, the MWS method and the the MWS-steady-state method with a correction using the correlation function to improve the convergence of the MWS method when the system transitions into the steady state. For the MWS method we use $Z = 20000$ realisations at each time step.

reduced in the steady-state regime. For the numerical examples presented in the following sections, we will consider only the systems undergoing transition or purely steady state systems. Therefore we will not use the MWS-steady-state method again in this paper.

Non-Gaussian case

In this section we explain how to calculate the derivative for non-Gaussian stochastic parameters using the MWS method. The procedure is similar to that shown in the previous part but the rule for updating the Malliavin weight must be modified.

We begin as before by considering uncertainty in the stress σ:

$$
\begin{aligned}
\sigma &= \sigma_1 + c\xi \\
&= \underbrace{\sigma_1 + c\mathbb{E}[\xi]}_{\sigma_0} + c(\xi - \mathbb{E}[\xi]).
\end{aligned}
\tag{26}
$$

where ξ a random variable with probability density function $f(x)$ and c and σ_1 are two constants. We have written Eq (26) in the form of a constant $\sigma_0 = \sigma_1 + c\mathbb{E}[\zeta]$ plus a random variable $c(\zeta - \mathbb{E}[\zeta])$ with zero mean. Then it follows that σ_0 is the mean of the uncertain stress σ. We will use the MWS method to evaluate the derivative of the expected value of the quantity of interest with respect to σ_0.

Algorithm 1: Malliavin Weight Sampling algorithm for time dependent problem. The notation used is that of the Kelvin-Voigt example in but the same basic algorithm is used throughout the paper. Note that a correction term is needed for systems in steady state, see [16].

```
Data: σ₀, E, η and the random variable ξ ∈ N(0, 1).
Result: ∂E[ϵ(t)]/∂σ₀, the derivative of the mean of ϵ with respect to the
mean stress σ₀ at time t.
```

```
∂𝔼[ε(t)]/∂σ₀ = 0 for all t.                                    ▷initialisation
for z = 0 to Z - 1 do
  t = 0;                                                              ▷time
  ε(t) = 0;                                                ▷initial condition
  q_σ = 0;                                                       ▷MWS weight
  for i = 1 to n do
    Draw realisation of random variable ξᵢ;
    ε(t + δt) = ε(t) + σ₀δt/η − Eε(t)δt/η + √δt ξᵢ/η;
    q_σ(t + δt) = q_σ(t) + ∂lnW/∂σ = q_σ + √δt ξᵢ;
    ∂𝔼[ε(t + δt)]/∂σ₀+ = ε(t + δt)q_σ/Z;
    t+ = δt;
  end
end
```

To be able to use the MWS method the probability density function on the parameter must satisfy some regularity conditions, see [11, 20, 21]. Intuitively, the probability density function must be sufficiently "regular" on $\mathbb{R}$ which holds for the Gaussian, log-normal, Beta($\alpha > 1, \beta > 1$) and Gamma($k > 1, \theta$) distributions. However, a uniform distribution between two values a and b can not be considered "regular" because the probability density function is not differentiable at a and b. Instead, we choose to regularised approximation of a uniform distribution using a Beta($1 + e, 1 + e$) random variable with $e \ll 1$.

Continuing, we again discretise Eq (5) using an explicit Euler method with time step δt:

$$\epsilon(t + \delta t) = \epsilon(t) + \frac{\delta t}{\eta}\left(\sigma_0 - E_0\epsilon(t) + \frac{c}{\sqrt{\delta t}}(\xi - \mathbb{E}[\xi])\right). \tag{27}$$

Alternatively, in the case of uncertainty in the Young's modulus, the discretisation can be written:

$$\epsilon(t + \delta t) = \epsilon(t) + \frac{\delta t}{\eta}\left(\sigma_0 - E_0\epsilon(t) + \frac{c\ \epsilon(t)}{\sqrt{\delta t}}(\mathbb{E}[\xi] - \xi)\right), \tag{28}$$

or, for uncertainty related to the viscosity:

$$\epsilon(t + \delta t) = \epsilon(t) + \frac{\sigma_0 - E\epsilon(t)}{\xi + \eta}\delta t. \tag{29}$$

The probability density function of the beta distribution, for $0 \leq x \leq 1$, and shape parameters $\alpha, \beta > 0$, is:

$$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)}x^{\alpha-1}(1 - x)^{\beta-1}. \tag{30}$$

The beta function B is a normalisation constant to ensure that the total probability integrates to 1. In general we will evaluate and update the Malliavin weight for the parameter m as:

$$q_m(t + \delta t) = q_m(t) + \frac{\partial \ln W}{\partial m} = q_m(t) + \frac{\partial \ln f(\xi)}{\partial \xi}\frac{\partial \xi}{\partial m}. \tag{31}$$

Note that in Eq (31), it is important to check that the condition $\mathbb{E}[q_m(t)] = 0$ is verified. If $\mathbb{E}[q_m(t)] \neq 0$, the updated rule must be corrected. An example of performing this correction is given in the next section entitled extension to second derivative. Finally, we note that for the initial condition we always impose $q_m(t = 0) = 0$.

For the uncertain Young's modulus modelled with a beta distribution, we have:

$$\frac{\partial \ln W}{\partial m} = \frac{(\beta - 1)\sqrt{\delta t}}{c(1 - m)} - \frac{(\alpha - 1)\sqrt{\delta t}}{c \, m}. \tag{32}$$

For the uncertain stress with beta distribution, we have:

$$\frac{\partial \ln W}{\partial m} = \frac{(\beta - 1)\sqrt{\delta t}}{c(1 - m)} - \frac{(\alpha - 1)\sqrt{\delta t}}{c \, m}. \tag{33}$$

For the uncertain viscosity with beta distribution, we have:

$$\frac{\partial \ln W}{\partial m} = \frac{\beta - 1}{c(1 - m)} - \frac{\alpha - 1}{c \, m}. \tag{34}$$

These results and further calculations are summarised in Table 1.

The Malliavin derivatives of the Kelvin-Voigt model with respect to the mean of the three parameters $\{\sigma_0, \eta_0, E_0\}$ modelled as beta(2, 2) distributions are shown in Fig 3. The exact solution is computed semi-analytically using standard integration rules. Good agreement between the MWS and semi-analytical solution is observed for E_0 and σ_0. For the viscosity η_0 the number of Monte Carlo samples is not sufficient to achieve negligible error, but the overall trend is followed.

Extension to second derivative

The MWS method can also be used to compute the second Malliavin derivative of the expected value of a quantity of interest J. If the quantity of interest does not depend explicitly of the random parameter, the expression given in Eq (3) is valid, but the more general form is the following:

$$\frac{\partial \mathbb{E}[J]}{\partial \bar{m}} = \mathbb{E}\left[\frac{\partial J}{\partial \bar{m}}\right] + \mathbb{E}[J q_m]. \tag{35}$$

In Eq (35), when we want to compute the second derivative the term $\left[\frac{\partial J}{\partial \bar{m}}\right]$ does not vanish because in this case J is the first derivative with respect to $\bar{m}$ and therefore depends on the parameter $\bar{m}$ in general. By applying Eq (35), we can show for example in the case of uncertain Young's modulus that:

$$\frac{\partial^2 \mathbb{E}[\epsilon(t)]}{\partial E_0^2} = \mathbb{E}[\epsilon(t)(q_{EE}(t) + q_E(t)^2 - C_{EE} - C_E^2)], \tag{36}$$

Table 1. Summary of main results for Kelvin-Voigt model with three distributions on three different model parameters.

distribution	$\frac{\partial \ln W}{\partial \sigma}$ or $\frac{\partial \ln W}{\partial E}$	$\frac{\partial \ln W}{\partial \eta}$
Beta(α, β)	$\frac{(\beta-1)\sqrt{\delta t}}{c(1-m)} - \frac{(\alpha-1)\sqrt{\delta t}}{mc}$	$\frac{(\beta-1)}{c(1-m)} - \frac{(\alpha-1)}{mc}$
Gamma(κ, θ)	$\frac{-(\kappa-1)\sqrt{\delta t}}{mc} + \frac{\sqrt{\delta t}}{\theta c}$	$\frac{-(\kappa-1)}{mc} + \frac{1}{\theta c}$
log-Normal(μ, σ)	$\frac{\sqrt{\delta t}}{mc}\left(1 + \frac{\ln m - \mu}{\sigma^2}\right)$	$\frac{1}{mc}\left(1 + \frac{\ln m - \mu}{\sigma^2}\right)$
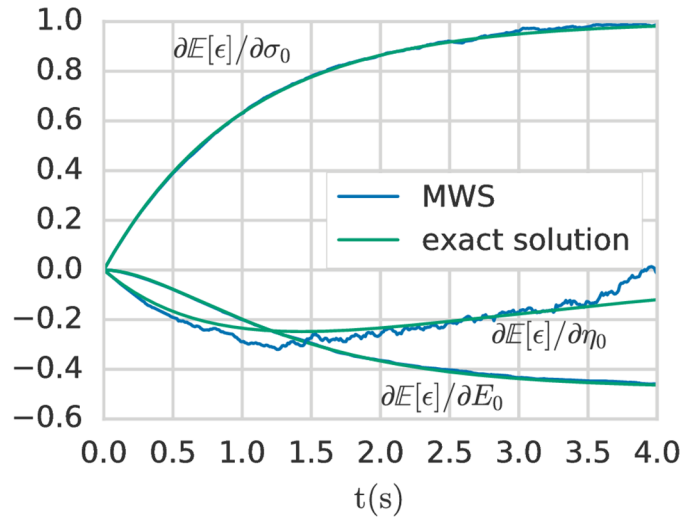
Fig 3. Malliavin derivatives of the expected value of the strain with respect to the mean of the stochastic parameters (Young's modulus E_0, viscosity η_0 and stress loading σ_0). Comparison between the exact solution and the MWS method. All uncertain parameters are modeled with a beta(2, 2) distribution. $Z = 10^5$ realisations are performed for each estimator and the mean value of 10 estimators is plotted for each parameter. Note that the value of Z is not large enough for the viscosity to converge with an negligible error compared to the two other parameters. By increasing Z, this error could be reduced.

with the following updating rule:

$$q_{EE}(t + \delta t) = q_{EE}(t) + \frac{\partial^2 \ln W}{\partial E^2}, \tag{37}$$

and:

$$q_E(t + \delta t) = q_E(t) + \frac{\partial \ln W}{\partial E}. \tag{38}$$

The constant C_E^2 and C_{EE} allow to ensure that the expected value of the global Malliavin weight $\left(q_{EE}(t) + q_E(t)^2 - C_{EE} - C_E^2\right)$ has an expected value equal to zero. In this specific case we have:

$$C_E^2 = \mathbb{E}\left[\left(\frac{\partial \ln W}{\partial E}\right)^2\right], \tag{39}$$

$$C_{EE} = \mathbb{E}\left[\frac{\partial^2 \ln W}{\partial E^2}\right]. \tag{40}$$

The precise specification of the constants depends on the distribution. We compute them analytically or by using standard numerical integration techniques found in e.g. Scipy or Maple.

In Fig 4, a comparison between the analytical solution and the MWS method is given for the value of the second derivative depending on time of the expected value of ϵ with respect to the Young's modulus. For the sake of example, the problem specification is the same as in previous sections, with the exception that the random variable follows a log-normal(μ, σ) distribution with mean equal to 0.5 and standard deviation equal to 0.25 which corresponds to $\mu =$
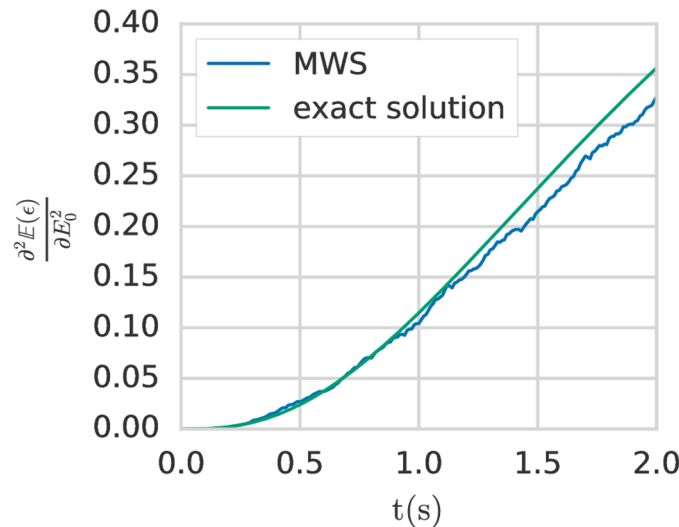
Fig 4. Second sensitivity derivative of the expected value of the strain with respect to the Young's modulus for the Kelvin-Voigt model. Comparison between the exact solution solution and the MWS method. The Young's modulus is modelled with a log-normal distribution. For the MWS method, $Z = 10^7$ realisations are performed. Note that the value of Z for the same order of magnitude for the error is higher for the second derivative compared to the first derivative because the variance V in the Malliavin estimator is bigger and we know from the central limit theorem that the error is in $\mathcal{O}(V^{1/2}Z^{-1/2})$.
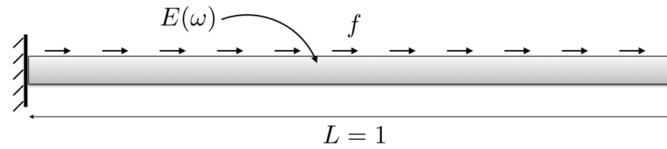
-0.804 and $\sigma = 0.473$. The analytical solutions for the two constants C_{EE} and C_E^2 are in this case:

$$C_{EE} = C_E^2 = \left(1 + \tfrac{1}{\sigma^2}\right)\exp\left(2\sigma^2 - 2\mu\right). \qquad (41)$$

As we can see in Fig 4, the MWS method gives a good approximation for the evaluation of the second derivative with $Z = 10^7$ realisations.

Extension to random process

In this paper we deal with random noise and in the next section we show numerical results of stochastic mechanics problems where models are defined as PDEs. The probability density function of the random variables used in these examples does not depend on time. Similarly to the Kelvin-Voigt model presented before, we study a time dependent problem in a finite dimensional space by splitting the time interval $[0, T]$ into a finite number of increments. Note that it is also possible to take into account the random noise only at the initial time instead of generating random variables at each time step. It would be possible to extend this work to random process, e.g. by using a Wiener process $W(t)$ which verifies in particular $(W(t + \delta t) - W(t)) \sim N(0, \delta t)$. In this case, even for simple ODEs, it is very difficult to obtain analytical solutions because the probability density function of a random process evolves in time. The Malliavin calculus is very well adapted to address these stochastic problems but requires much more advanced mathematical tools as those presented in this paper. In addition, the Malliavin calculus has the advantage and the specificity that it is possible to directly work in the continuum (infinite dimensional space) to evaluate the sensitivity derivatives. We hope that the first and simple approach restricted to random variables presented in this paper may be of interest to the engineering community and encourage them to investigate the benefits that the Malliavin calculus could provide in the context of stochastic PDEs.

Fig 5. Elastic bar with stochastic Young's modulus.

https://doi.org/10.1371/journal.pone.0189994.g005

PDE examples

We now turn our attention to models that are defined as PDEs. To solve the deterministic evaluations of the PDEs we use the finite element method. We have chosen to use DOLFIN, part of the FEniCS Project to implement the finite element method solvers [22].

Elastic bar with stochastic Young's modulus

The strong form PDE and boundary conditions of the behaviour of a 1-dimensional elastic bar (see Fig 5) are:

$$E \; \frac{\mathrm{d}^2 u(x)}{\mathrm{d}x^2} + f = 0; \; u(0) = 0 \text{ and } \frac{\mathrm{d}u(L)}{\mathrm{d}x} = 0. \tag{42}$$

We take $f = 1$, $L = 1$ and a stochastic Young's modulus:

$$E = 2(1 + \xi), \tag{43}$$

with ξ a random variable with beta(2, 2) distribution.

The forward model is described by the following weak residual formulation, find $u \in H_D^1(\Omega_s)$ such that:

$$F(u; \tilde{u}) := - \int_{\Omega_s} E \; \nabla u \cdot \nabla \tilde{u} \; \mathrm{d}x + \int_{\Omega_s} f \tilde{u} \; \mathrm{d}x = 0 \quad \forall \tilde{u} \in H_0^1(\Omega_s), \tag{44}$$

where the space $H_D^1(\Omega_s)$ is the usual Sobolev space of square-integrable functions with square-integrable weak derivatives on the domain $\Omega_s := [0, 1]$ that satisfy the Dirichlet boundary condition $u(0) = 0$ and $H_0^1(\Omega_s)$ vanish on the whole boundary. We solve the forward model using a piecewise linear finite element method with 1024 cells in the mesh.

The quantity of interest is:

$$J = \int_0^1 u(x)\mathrm{d}x. \tag{45}$$

The derivative of the expected value of J with respect to the mean value of the Young's modulus $\bar{E}$ can be computed analytically in this case:

$$\frac{\partial \mathbb{E}[J]}{\partial \bar{E}} = - \int_0^1 \frac{1}{3(2 + 2x)^2} \frac{x(1 - x)}{\mathrm{B}(2, 2)} \; \mathrm{d}x = 1 - \frac{3}{2}\ln(2). \tag{46}$$

This problem is a stationary (not time-dependent), in contrast to the Kelvin-Voigt model considered previously. However, this stationary problem can be solved using the same techniques. We introduce the concept of pseudo-time, where the system evolves from its initial state at $t = 0$ to the final solution at time $t = T$ through the a single solution of the PDE Eq (44).
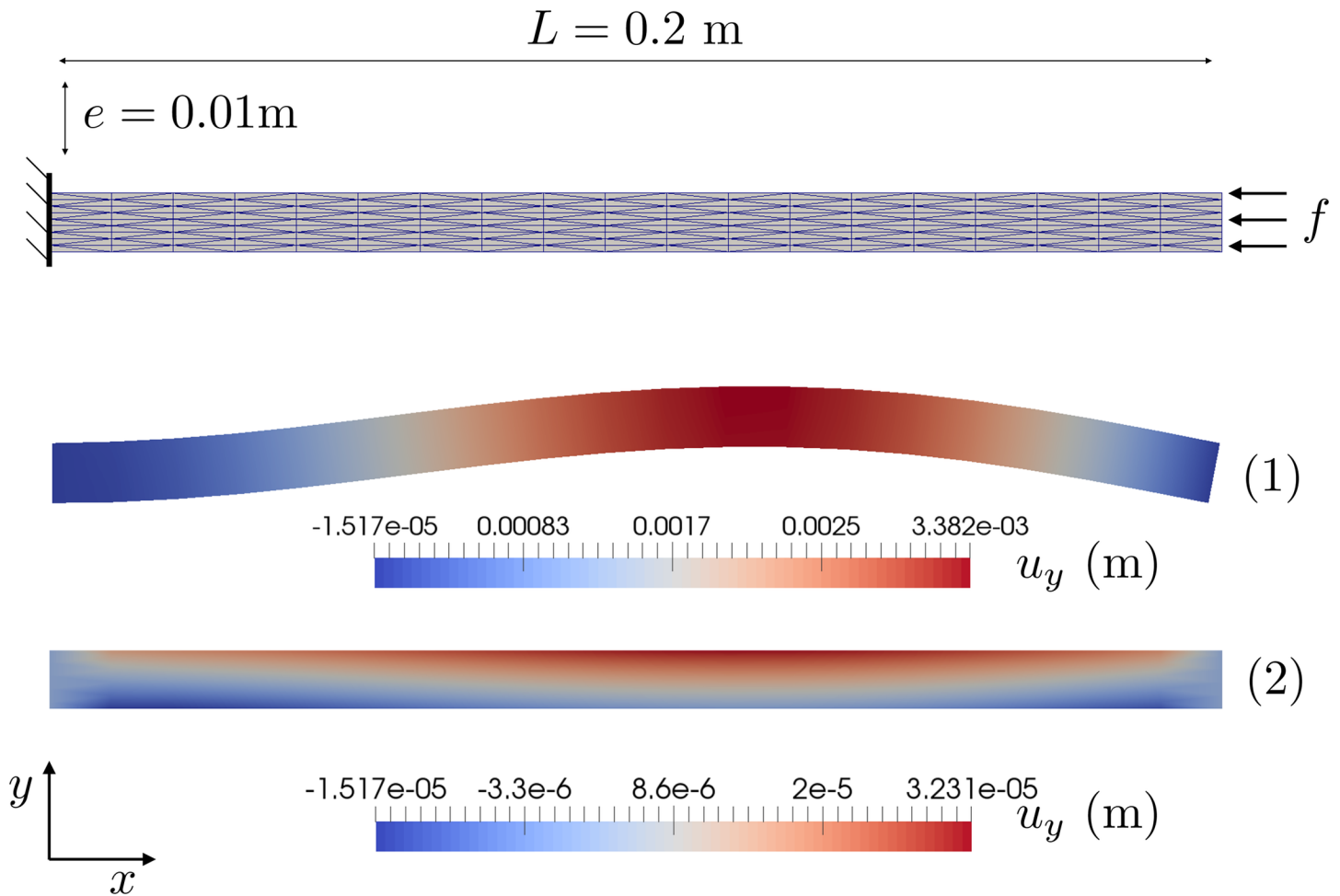
Fig 6. Hyperelastic beam: Mesh and schematic of boundary conditions. (1) a realisation of the problem where there is a geometric instability (buckling) and (2) another without.

Therefore in algorithm 1 we take the pseudo-time step as $\delta t = T$ and hence $n = 1$. As before, the Malliavin weight still has initial condition $q_m(0) = 0$.

Finally, the relative error between the MWS method with $Z = 5 \times 10^5$ realisations and the analytical solution is 3.0×10^{-3}.

Buckling of a hyperelastic beam with stochastic Young's modulus

We study the deformation of a 2D geometrically non-linear hyperelastic beam with stochastic Young's modulus E. We have deliberately designed this problem so that for some values of E the beam undergoes buckling, and for others not.

Consider a hyperelastic body that in its undeformed state occupies the domain $\Omega_0 = [0, L] \times [0, e] \subset \mathbb{R}^2$ with $L = 0.2m$ and $e = 0.01m$ (see Fig 6), and in its deformed state occupies some (unknown) domain $\Omega \subset \mathbb{R}^2$. $\boldsymbol{\varphi}$ is the map between the material points $\mathbf{X}$ in the undeformed domain Ω_0 and points $\mathbf{x}$ in the deformed domain Ω:

$$\boldsymbol{\varphi} : \Omega_0 \ni \mathbf{X} \to \mathbf{x} \in \Omega, \tag{47}$$

The deformation gradient can be written $\mathsf{F}(\mathbf{X}) := \frac{\partial \varphi}{\partial \mathbf{X}}$. The right Cauchy-Green tensor is then defined as $\mathsf{C} := \mathsf{F}^T \mathsf{F}$.

The Neo-Hookean stored energy density of the body is then:

$$\mathcal{W}(\mathsf{F}) := \mu(I_c - 2)/2 - \mu \log I_3 + \lambda (\log I_3)^2/2. \tag{48}$$

where $I_3 := \det(\mathsf{F})$ and and $I_C = \operatorname{tr}(\mathsf{C})$. λ and μ are the Lame parameters and can be expressed as a function of the Young's modulus E and Poisson's ratio v as:

$$\lambda = \frac{Ev}{(1+v)(1-2v)} \quad \text{and} \quad \mu = \frac{E}{2(1+v)}. \tag{49}$$

We choose to model the Young's modulus as a log-normal random variable with mean value 11 MPa and standard deviation 2 MPa. We take Poisson's ratio as a fixed constant $v = 0.3$.

Defining the displacement field as $\mathbf{u} := \boldsymbol{\varphi} - \mathbf{X}$ and a linear functional $\mathbf{f}$ that encodes the external tractions we can characterise the elastic equilibrium displacement field $\mathbf{u}^*$ as the solution to the following minimisation problem:

$$\begin{aligned}
\mathbf{u}^* &= \underset{\mathbf{u} \in [H_D^1(\Omega_0)]^2}{\arg\min} \; L(\mathbf{u}) \\
&= \underset{\mathbf{u} \in [H_D^1(\Omega_0)]^2}{\arg\min} \left\{ \int_{\Omega_0} \mathcal{W}(\mathsf{F}) \; dx_0 - \langle \mathbf{f}, \mathbf{u} \rangle \right\},
\end{aligned} \tag{50}$$

where $[H_D^1(\Omega_0)]^2$ is the usual vector-valued Sobolev space of square integrable functions with square integrable derivatives that satisfies the given Dirichlet boundary conditions and dx_0 is a measure on Ω_0. We fix the left hand of the beam, $u(0, y) = 0$ and apply a surface traction in the negative x direction on the right hand of the beam of magnitude f.

For one Monte Carlo realisation we solve the non-linear problem using a Newton method from SNES [23] with continuation in the loading parameter f and a third-order backtracking line search. We let the symbolic differentiation capabilities of UFL derive the residual and Jacobian of the forward model for use in the Newton method. We solve the linear systems arising from the Newton iterations using a conjugate gradient method preconditioned using algebraic multigrid (Hypre BoomerAMG [24]) interfaced from PETSc [23].

The quantity of interest is defined as:

$$J = \int_{\Omega} |u_y| \; dx. \tag{51}$$

The Malliavin derivative of $\mathbb{E}[J]$ with respect to the mean Young's modulus obtained with the MWS method with $Z = 3 \times 10^3$ realisations is:

$$\frac{\partial \mathbb{E}[J]}{\partial E_0} \approx -3.1 \times 10^{-6} \; \text{m}^3/\text{MPa}. \tag{52}$$

No analytical solution exists for comparison. If we use dolfin-adjoint [25], we can compute the classical derivative of J with respect to the Young's modulus around the mean parameter:

$$\left. \frac{\partial J}{\partial E} \right|_{E=E_0} \approx -3.5 \times 10^{-8} \; \text{m}^3/\text{MPa}. \tag{53}$$

In this example the difference between the classical derivative and the Malliavin derivative is quite pronounced. This difference is caused by the presence of an instability (buckling). This instability is not activated when $E = E_0$, hence, the classical derivative tells us that J is relatively
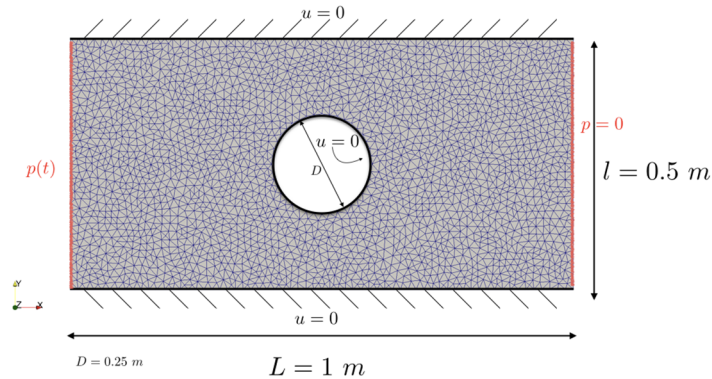
Fig 7. Mesh, geometry and boundary conditions for the incompressible Navier-Stokes problem.

insensitive to perturbations in the Young's modulus about E_0. However, the Malliavin derivative tells us that $\mathbb{E}[J]$ *is* in fact quite sensitive to changes in the mean of the Young's modulus E_0. The Malliavin derivative gives us quite a different perspective on the sensitivity of this problem than the classical one.

Incompressible Navier-Stokes equations with stochastic viscosity

We consider the incompressible Navier-Stokes equations on a domain Ω in $\mathbb{R}^2$ consisting of a pair of momentum and continuity equations:

$$
\begin{aligned}
\dot{\mathbf{u}} + \nabla \mathbf{u} \cdot \mathbf{u} - v\Delta \mathbf{u} + \nabla p &= \mathbf{f}, \\
\nabla \cdot \mathbf{u} &= 0.
\end{aligned}
\tag{54}
$$

In Eq (54), $\mathbf{u}$ refers to the unknown velocity of the fluid, v is the viscosity of the fluid, p the unknown pressure and $\mathbf{f}$ is a given source. The mesh, geometry and boundary conditions for the incompressible Navier-Stokes problem are shown in Fig 7. The viscosity is modelled as a random variable:

$$
v = 0.015 + 0.01(\xi - 0.005),
\tag{55}
$$

with ξ a log-normal distribution with mean equal to 0.5 and standard deviation equal to 0.25.

We solve the PDE for a given parameter v with FEniCS [26] (FE approximation) using Chorin's method with time step $\delta t = 0.01$ for $t \in [0, 1]$, see [27] for more details on the implementation. For one realisation of the viscosity the velocity at time $t = 1$ s is show in Fig 8.
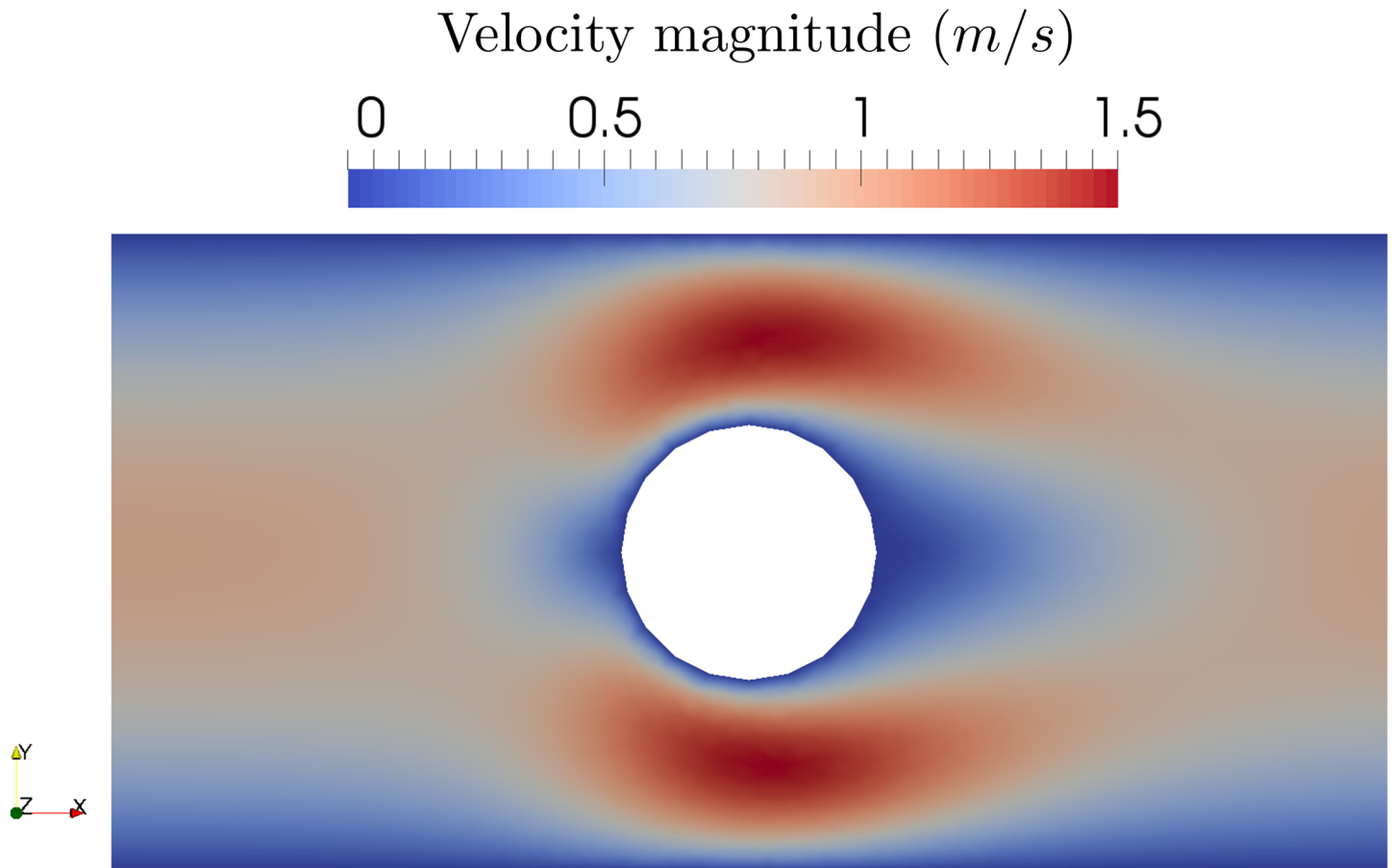
The quantity of interest is the total volume of fluid that exits the right end of the domain:

$$
J = \int_{t=0}^{t=1} \int_{S_{p=0}} \mathbf{u} \cdot \mathbf{n} \, \mathrm{d}s\mathrm{d}t,
\tag{56}
$$

where $S_{p=0}$ is the surface with normal vector $\mathbf{n}$ on the right side where the pressure is imposed to zero.

The derivative of $\mathbb{E}[J]$ with respect to η_0 obtained with the MWS method for $Z = 4 \times 10^5$ realisations is:

$$
\frac{\partial \mathbb{E}[J]}{\partial \eta_0} \approx -7.9 \ \mathrm{s/m^2}.
\tag{57}
$$

Fig 8. Velocity magnitude at time *t* = 1 s for one realisation of the viscosity.

No analytical solution exists for comparison. If we use dolfin-adjoint [25], we can compute the derivative of *J* with respect to the viscosity around the mean parameter:

$$\left.\frac{\partial J}{\partial \eta}\right|_{\eta=\eta_0} \approx -7.85 \ \text{s/m}^2. \tag{58}$$

The two sensitivity derivatives are close. In this example, contrary to the hyperelastic example, the Malliavin approach does not give us a particularly different interpretation of the sensitivity.

Conclusion

In this paper we have shown how to calculate the Malliavin derivative using the method of Malliavin Weight Sampling. We have applied the method to some typical mechanics models that can be described by ODEs and PDEs, and solved those models using finite difference and finite element methods. In addition, we have extended the existing practical literature on MWS to non-Gaussian random variables and the calculation of second-order derivatives. We are currently investigating the extension of this work from random parameters to problems with variables modelled as random fields. We are also exploring the use of the Malliavin

derivative in derivative-driven variance reduction methods e.g. [28]. Code showing the calculation of the Malliavin derivative for the examples in this paper is freely available at: https://dx.doi.org/10.6084/m9.figshare.5432722 [29].

Acknowledgments

Author Contributions

Conceptualization: Paul Hauseux.

Formal analysis: Paul Hauseux, Jack S. Hale.

Funding acquisition: Stéphane P. A. Bordas.

Methodology: Paul Hauseux, Jack S. Hale.

Project administration: Stéphane P. A. Bordas.

Software: Paul Hauseux, Jack S. Hale.

Supervision: Jack S. Hale, Stéphane P. A. Bordas.

Validation: Paul Hauseux, Jack S. Hale.

Visualization: Paul Hauseux, Jack S. Hale.

Writing – original draft: Paul Hauseux.

Writing – review & editing: Jack S. Hale, Stéphane P. A. Bordas.

References

1. Ambrosetti A, Prodi G. A Primer of Nonlinear Analysis. Cambridge University Press; 1995.

2. Yang XJ, Machado JAT. A new fractional operator of variable order: Application in the description of anomalous diffusion. Physica A: Statistical Mechanics and its Applications. 2017; 481(Supplement C):276–283. https://doi.org/10.1016/j.physa.2017.04.054

3. Atangana A, Koca I. In: Ruzhansky M, Cho YJ, Agarwal P, Area I, editors. On Uncertain-Fractional Modeling: The Future Way of Modeling Real-World Problems. Singapore: Springer Singapore; 2017. p. 121–143.

4. Atangana A, GÃmez-Aguilar JF. A new derivative with normal distribution kernel: Theory, methods and applications. Physica A: Statistical Mechanics and its Applications. 2017; 476(Supplement C):1–14. https://doi.org/10.1016/j.physa.2017.02.016

5. Atangana A, Baleanu D. New Fractional Derivatives with Nonlocal and Non-Singular Kernel: Theory and Application to Heat Transfer Model. Thermal Science. 2016; 20:763–769. https://doi.org/10.2298/TSCI160111018A

6. Yang XJ, Gao F, Machado JAT, Baleanu D. A new fractional derivative involving the normalized sinc function without singular kernel;. Available from: https://arxiv.org/abs/1701.05590

7. He JH. Variational principles for some nonlinear partial differential equations with variable coefficients. Chaos, Solitons & Fractals. 2004; 19(4):847–851. https://doi.org/10.1016/S0960-0779(03)00265-0

8. Stochastic Calculus of Variations in Mathematical Finance | Paul Malliavin | Springer;. Available from: http://www.springer.com/gp/book/9783540434313

9. Malliavin P. Stochastic analysis. Grundlehren der mathematischen Wissenschaften. Berlin, New York: Springer; 1997.

10. Nourdin I, Peccati G. Normal Approximations with Malliavin Calculus: From Stein's Method to Universality. Cambridge Tracts in Mathematics. Cambridge University Press; 2012.

11. Broadie M, Glasserman P. Estimating security price derivatives using simulation. Management science. 1996; 42(2):269–285. https://doi.org/10.1287/mnsc.42.2.269

12. Benhamou E. Optimal Malliavin Weighting Function for the Computation of the Greeks. Mathematical Finance. 2003; 13(1):37–53. https://doi.org/10.1111/1467-9965.t01-1-00004

13. Fournié E, Lasry JM, Lebuchoux J, Lions PL, Touzi N. Applications of Malliavin calculus to Monte Carlo methods in finance. Finance and Stochastics. 1999; 3(4):391–412. https://doi.org/10.1007/s007800050068

14. Chen N, Glasserman P. Malliavin Greeks without Malliavin calculus. Stochastic Processes and their Applications. 2007; 117(11):1689–1723. https://doi.org/10.1016/j.spa.2007.03.012

15. Warren PB, Allen RJ. Malliavin Weight Sampling for Computing Sensitivity Coefficients in Brownian Dynamics Simulations. Physical Review Letters. 2012; 109(25):250601. https://doi.org/10.1103/PhysRevLett.109.250601 PMID: 23368441

16. Warren PB, Allen RJ. Malliavin Weight Sampling: A Practical Guide. Entropy. 2014; 16(1):221–232. https://doi.org/10.3390/e16010221

17. Szamel G. Evaluating linear response in active systems with no perturbing field. EPL (Europhysics Letters). 2017; 117(5):50010. https://doi.org/10.1209/0295-5075/117/50010

18. Nualart D. The Malliavin calculus and related topics. vol. 1995. Springer; 2006. Available from: http://dx.doi.org/10.1007/3-540-28329-3

19. Matthies GH. Stochastic finite elements: Computational approaches to stochastic partial differential equations. Journal of Applied Mathematics and Mechanics. 2008; 88:849–873.

20. L'Ecuyer P. A Unified View of the IPA, SF, and LR Gradient Estimation Techniques. Manage Sci. 1990; 36(11):1364–1383. https://doi.org/10.1287/mnsc.36.11.1364

21. Capriotti L. Reducing the variance of likelihood ratio greeks in Monte Carlo. In: 2008 Winter Simulation Conference; 2008. p. 587–593.

22. Alnaes M, Blechta J, Hake J, A J, Kehlet B, Logg A, et al. The FEniCS Project Version 1.5. Archive of Numerical Software. 2015; 3(100).

23. Balay S, Abhyankar S, Adams MF, Brown J, Brune P, Buschelman K, et al. PETSc Users Manual. Argonne National Laboratory; 2016. ANL-95/11—Revision 3.7. Available from: http://www.mcs.anl.gov/petsc

24. Falgout RD, Yang UM. hypre: A Library of High Performance Preconditioners. In: Sloot PMA, Hoekstra AG, Tan CJK, Dongarra JJ, editors. Computational Science—ICCS 2002. No. 2331 in Lecture Notes in Computer Science. Springer Berlin Heidelberg; 2002. p. 632–641.

25. Farrell PE, Ham DA, Funke SW, Rognes ME. Automated Derivation of the Adjoint of High-Level Transient Finite Element Programs. SIAM Journal on Scientific Computing. 2013; 35(4):C369–C393. https://doi.org/10.1137/120873558

26. Logg A, Wells GN. DOLFIN: Automated Finite Element Computing. ACM Trans Math Softw. 2010; 37(2):20:1–20:28. https://doi.org/10.1145/1731022.1731030

27. Langtangen HP, Logg A. Solving PDEs in Python—The FEniCS Tutorial I. No. 3 in Simula SpringerBriefs on Computing. Springer International Publishing; 2016. Available from: http://dx.doi.org/10.1007/978-3-319-52462-7

28. Hauseux P, Hale JS, Bordas SPA. Accelerating Monte Carlo estimation with derivatives of high-level finite element models. Computer Methods in Applied Mechanics and Engineering. 2017; 318(Supplement C):917–936. https://doi.org/10.1016/j.cma.2017.01.041

29. Hauseux P, Hale JS, Bordas SPA. Calculating the Malliavin Derivative of some numerical models using the Malliavin Weight Sampling method, 2017. https://dx.doi.org/10.6084/m9.figshare.5432722