Uncertainty quantification for waves: introduction and forward UQ

Laura Scarabosio (l.scarabosio@math.ru.nl)

Summer school "Waves: Modeling, Analysis, and Numerics"
Nijmegen, 25–29 August 2025

Radboud University

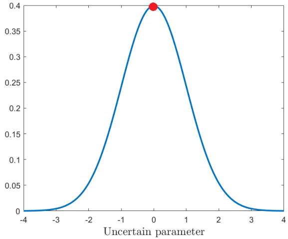
Introduction

Uncertainty quantification in a nutshell

Uncertainty quantification (UQ) takes into account **uncertainties** in the description of a physical system and **quantifies** their effect on the outcome.

Uncertain quantities: **random variables** or **random fields**.

Quantifying: computing **probabilities** or **statistics** of **quantities of interest**.

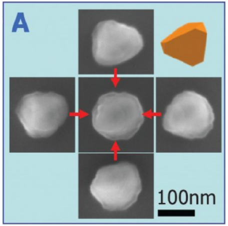


Types of uncertainties

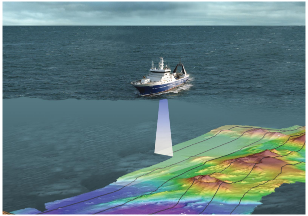
Uncertainties can be **aleatoric** or **epistemic**.

Examples:

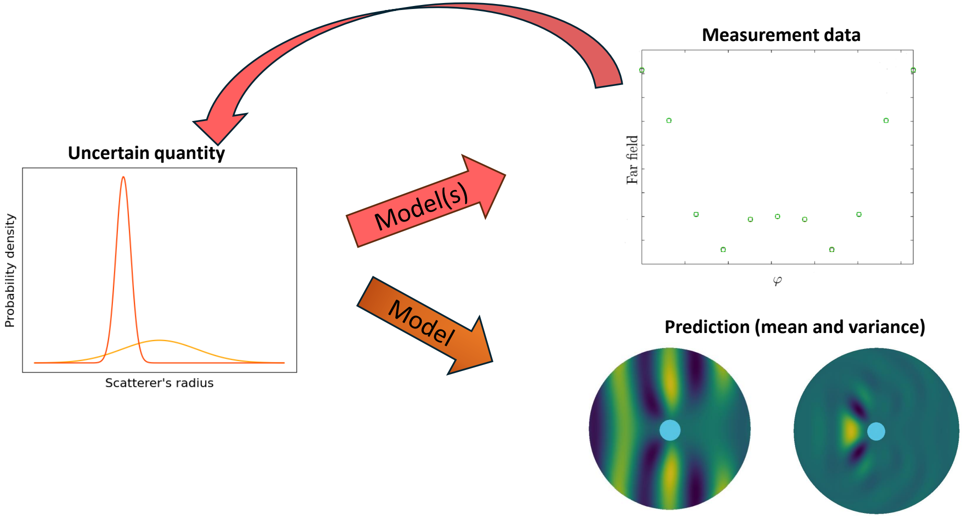
- measurement error
- material properties
- geometry
- forcing terms
- initial conditions
- boundary conditions
- mathematical model
- numerics



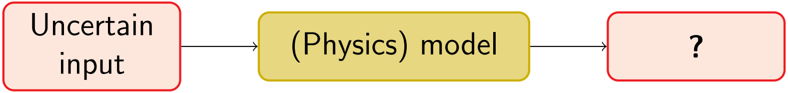
Sannomiya, Diss. ETHZ 18747


Walsh, Proc. US Naval Institute (2020)

The predictive and interpretable estimation process



The forward problem



Given a **probability distribution** of the input, compute:

- **statistics** ($\rightsquigarrow$ quadrature, sampling)
- an **approximant** ($\rightsquigarrow$ interpolation, regression)

for some **quantity of interest**. Usually “easy” distributions are used.

The inverse problem

? → (Physics) model → Measurements

Given educated **prior knowledge** on the input, compute

- the **posterior** distribution of the input **given the data**
- statistics of **quantity of interest** w.r.t. **posterior** distribution.

Why uncertainty quantification?

"It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so." – Mark Twain

- **Predict** system responses under input variability
- **Quantify** reliability of predictions
- **Reduce** development time and prototyping costs
- Analyze **risk** of undesirable events
- Find robust **optimized solutions**

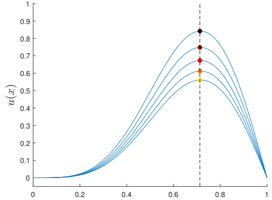
A one-dimensional example

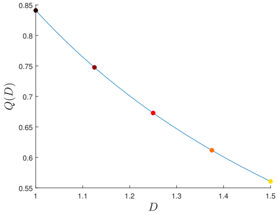
$$\boxed{\begin{gathered}-\frac{\mathrm{d}}{\mathrm{d}x}\left(D\frac{\mathrm{d}u}{\mathrm{d}x}\right) = f \text{ in } (0,1)\\ + \text{ boundary conditions}\end{gathered}}$$

with $D \sim U(a, b)$.

$Q = Q(D)$ quantity of interest.

- $\mathbb{E}[Q]$, $D \mapsto Q(D)$ (robust predictions)
- $\mathrm{Var}[Q]$, $\dfrac{\mathrm{d}Q}{\mathrm{d}\vartheta}$ (quantify reliability)
- $\mathbb{P}(Q \geq q_{\text{thresh}})$ (risk assessment)
- $\mathrm{argmin}_{f\in X}\mathbb{E}[\mathcal{J}(u)]$ (robust optimization)





The high-dimensionality challenge

$$\boxed{\begin{gathered} -\frac{\mathrm{d}}{\mathrm{d}x}\left(D\frac{\mathrm{d}u}{\mathrm{d}x}\right) = f \text{ in } (0,1) \\ u(0) = u_0, \quad u(1) = u_1 \end{gathered}}$$

with $D \sim U(a,b)$, $u_0 \sim U(a_0,b_0)$, $u_1 \sim U(a_1,b_1)$, $\cdots$

⚠ Tensorization of 1d rules is not feasible!

Plan for forward UQ

Refresher of elementary probability

The Monte Carlo estimator

The Multilevel Monte Carlo estimator

- application to time-harmonic scattering

More advanced topics

The (vanilla) Monte Carlo estimator

Monte Carlo sampling

ϑ uncertain parameter with pdf f, Q quantitity of interest.

$$\mathbb{E}[Q(\vartheta)] = \int_{\mathbb{R}^d} Q(\theta) f(\theta)\,\mathrm{d}\theta$$

$$\approx E_M(Q(\vartheta)) = \frac{1}{M} \sum_{i=1}^{M} Q(\vartheta^{(i)})$$ with i.i.d. samples

Convergence of the vanilla Monte Carlo estimator

Note: $E_M(Q(\vartheta))$ is a *random variable*, $\mathbb{E}[Q(\vartheta)]$ is a *number* (or function).

Unbiasedness

$$\mathbb{E}[E_M(Q(\vartheta))] = \mathbb{E}[Q(\vartheta)]$$

Convergence

$$E_M(Q(\vartheta)) \xrightarrow{a.s.} \mathbb{E}[Q(\vartheta)]$$

by **strong law of large numbers**.

Convergence rate

If $Q(\vartheta)$ has finite **variance**, i.e. $Q \in L^2(\Omega, \mathcal{H})$,

$$\mathrm{Var}(E_M(Q(\vartheta))) = \underbrace{\mathbb{E}\left[\|E_M(Q(\vartheta)) - \mathbb{E}\left[Q(\vartheta)\right]\|_{\mathcal{H}}^2\right]}_{\text{Mean Square Error}} = \frac{\mathrm{Var}(Q(\vartheta))}{M}$$

Remark: the rate is *independent* of the dimension d.

The bias-variance decomposition

Often the **exact** $Q(\vartheta)$ **is not accessible** or too expensive to compute, and is replaced by an **approximation** $Q_L(\vartheta)$.

Example: $Q_L(\vartheta)$ comes from a numerical discretization with mesh size h_L.

Approximating Q introduces a **bias** in the Monte Carlo estimator:

$$\underbrace{\mathbb{E}\left[\|E_M(Q_L(\vartheta)) - \mathbb{E}\left[Q(\vartheta)\right]\|_{\mathcal{H}}^2\right]}_{\text{Mean Square Error}} = \underbrace{\|\mathbb{E}\left[Q_L(\vartheta) - Q(\vartheta)\right]\|_{\mathcal{H}}^2}_{\text{bias error}} + \underbrace{\frac{1}{M}\mathrm{Var}\left[Q(\vartheta)\right]}_{\text{statistical error}}$$

$\Rightarrow$ approximation error and number of samples need to be **balanced**.

Multilevel Monte Carlo

Multilevel Monte Carlo estimator

Consider $(Q_l)_{l=0}^{L}$ and set $Q_{-1} := 0$, with *increasing accuracy and cost*.

Example: obtained from nested PDE discretizations with mesh sizes $(h_l)_{l=0}^{L}$.

$$\mathbb{E}\left[Q_L\right] = \sum_{l=0}^{L} \mathbb{E}\left[Q_l - Q_{l-1}\right] \quad \rightsquigarrow \quad \boxed{E^L\left[Q\right] := \sum_{l=0}^{L} E_{M_l}\left[Q_l - Q_{l-1}\right]}$$

$$\underbrace{\mathbb{E}\left[\|E^L\left[Q\right] - \mathbb{E}\left[Q\right]\|_{\mathcal{H}}^2\right]}_{\text{Mean Square Error}} \leq \underbrace{\|\mathbb{E}\left[Q_L - Q\right]\|_{\mathcal{H}}^2}_{\text{bias error}} + \underbrace{\sum_{l=0}^{L} \frac{1}{N_l} \mathrm{Var}\left[Q_l - Q_{l-1}\right]}_{\text{statistical error}}$$

Multilevel Monte Carlo algorithm

Given a tolerance ε^2 and a sequence of models (e.g., discretizations) $(\mathcal{M}_l)_{l\geq 0}$:

1. Split $\varepsilon^2 = \varepsilon^2_{bias} + \varepsilon^2_{stat}$
2. Select level L such that $\|\mathbb{E}\left[Q_L - Q\right]\| < \varepsilon_{bias}$
3. Choose $(M_l)_{l=0}^L$ s.t. stat error $< \varepsilon^2_{stat}$ at minimum cost

Constrained minimization problem

Find $(M_l)_{l=0}^L$: $W_{tot} = \sum_{l=0}^L M_l W_l \downarrow$ and $\mathrm{Var}(E^L(Q)) = \sum_{l=0}^L \frac{V_l}{M_l} = \varepsilon^2_{stat}$

$$\Rightarrow M_l = \mu\sqrt{\frac{V_l}{W_l}}, \quad \mu = \varepsilon^{-2}_{stat}\sum_{l=0}^{L}\sqrt{V_l W_l}$$

.

Cost of MLMC estimator: considerations

$$W_{tot} = \varepsilon_{stat}^{-2} \left(\sum_{l=0}^{L} \sqrt{V_l W_l} \right)^2$$

Observations:

Efficiency relies on **delicate balance** between variances and costs

$V_l = \mathrm{Var}(Q_l - Q_{l-1}) = \mathrm{Var}(Q_l) + \mathrm{Var}(Q_{l-1}) - 2\mathrm{Cov}(Q_l, Q_{l-1})$ (variance reduction)

Cost of vanilla Monte Carlo: $W_{tot}^{MC} = \varepsilon_{stat}^{-2} V_0 C_L$

Multilevel Monte Carlo: complexity theorem

Theorem (Cliffe et al. 2011, Giles 2015)

Suppose there exist $\alpha, \beta, \gamma > 0$, $\alpha \geq \frac{1}{2}\min\{\beta, \gamma\}$ *and* $C_1, C_2, C_3 > 0$ *s.t.*

(i) $\|\mathbb{E}[Q_l - Q]\|_{\mathcal{H}} \leq C_1 2^{-\alpha l}$ *(bias bound)*

(ii) $Var[Q_l - Q_{l-1}] \leq C_2 2^{-\beta l}$ *(variance bound)*

(iii) $W_l \leq C_3 2^{\gamma l}$ *(cost bound).*

Then, for every $\varepsilon < e^{-1}$, *there exist* $L \in \mathbb{N}$ *and* $(M_l)_{l=0}^{L}$ *s.t.*

$$\|E^L[Q] - \mathbb{E}[Q]\|_{L^2(\Omega,\mathcal{H})} < \varepsilon \quad W_{tot}(E^L) \leq \begin{cases} C_4 \varepsilon^{-2} & \text{if } \beta > \gamma, \\ C_4 \varepsilon^{-2} (\log \varepsilon)^2 & \text{if } \beta = \gamma, \\ C_4 \varepsilon^{-2-\frac{\gamma-\beta}{\alpha}} & \text{if } \beta < \gamma. \end{cases}$$

Multilevel Monte Carlo for elliptic PDEs

Levels associated to nested meshes with mesh sizes $(h_l)_{l\geq 0}$.

Theorem (Cliffe et al. 2011, Giles 2015)

Suppose there exist $\alpha, \beta, \gamma > 0$ and positive constants $C_1, C_2, C_3 > 0$ s.t.

(i) $\|\mathbb{E}[Q_l - Q]\|_{\mathcal{Y}} \leq C_1 h_l^{\alpha}$ *(bias bound)*

(ii) $Var[Q_l - Q_{l-1}] \leq C_2 h_l^{\beta}$ *(variance bound)*

(iii) $W_l \leq C_3 h_l^{-\gamma}$ *(cost bound).*

Then, for every $\varepsilon < e^{-1}$, there exist $L \in \mathbb{N}$ and $(M_l)_{l=0}^{L}$ s.t.

$$\|E^L[Q] - \mathbb{E}[Q]\|_{L^2(\Omega,\mathcal{H})} < \varepsilon \quad W_{tot}(E^L) \leq \begin{cases} C_4 \varepsilon^{-2} & \text{if } \beta > \gamma, \\ C_4 \varepsilon^{-2} (\log \varepsilon)^2 & \text{if } \beta = \gamma, \\ C_4 \varepsilon^{-2-\frac{\gamma-\beta}{\alpha}} & \text{if } \beta < \gamma. \end{cases}$$

Multilevel Monte Carlo for elliptic PDEs: convergence rates

If there exist $C > 0$ and $s > 0$ s.t.

$$\|Q_l(u_l) - Q(u)\|_{L^2(\Omega,\mathcal{H})} \leq C h_l^s \|u\|_{L^2(\Omega,\mathcal{W})},$$

we have $\alpha = s$, $\beta = 2s$.

Note: $\mathcal{W}$ is usually a stronger space than what we need for well-posedness.

Notable example [Barth, Schwab, Zollinger 2011], [Charrier, Scheichl, Teckentrup 2013]

$$\begin{aligned} -\nabla \cdot (a\nabla u) &= f \quad \text{on } D, \\ u &= 0 \quad \text{on } \partial D, \end{aligned}$$

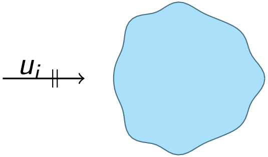
where $a \sim \mathcal{U}([a_-, a_+])$.

Multilevel Monte Carlo for elliptic PDEs: convergence rates

Multilevel Monte Carlo for waves: setting

Helmholtz model scattering problem (sound-soft):

$$
\begin{aligned}
-\nabla\cdot(A\nabla u) - \kappa^2 n u &= 0 && \text{in } \mathbb{R}^d \setminus D \\
u &= 0 && \text{on } \partial D \\
&+ \text{Sommerfeld r.c. for } u - u_i &&
\end{aligned}
$$



Assumptions

$A \sim \mathcal{U}([A_-, A_+])$, $n \sim \mathcal{U}([n_-, n_+])$, with $A_-, n_- > 0$

Non-trapping regime

Discretization with h-finite elements

Domain truncation with exact DtN map

Multilevel Monte Carlo for waves: convergence

Low frequency: essentially same treatment as elliptic case [Scarabosio 2019].
In general: constraint on first level, $h_0 = C\kappa^{-a}$ [Pembery, 2020].

$Q(u)$	a	Monte-Carlo	Multi-Level Monte-Carlo
$\lVert u\rVert_{H^1_k(D)}$	$\frac{2p+1}{2p}$	$k^{d\frac{2p+1}{2p}}\varepsilon^{-2-\frac{d}{2p}}$	$k^{d\frac{2p+1}{2p}}\varepsilon^{-2}$
$\lVert u\rVert_{L^2(D)}$	$\frac{2p+1}{2p}$	$k^{d\frac{2p+1}{2p}}\varepsilon^{-2-\frac{d}{2p}}$	$k^{d}\varepsilon^{-2}$ if $k\varepsilon$ small, otherwise $k^{d\frac{2p+1}{2p}}\varepsilon^{-2-\frac{d}{2p}}$
$\lVert u\rVert_{H^1_k(D)}$	1	$k^{d\frac{2p+1}{2p}}\varepsilon^{-2-\frac{d}{2p}}$	$k^{d+2}\varepsilon^{-2}$
$\lVert u\rVert_{L^2(D)}$	1	$k^{d}\varepsilon^{-2-\frac{d}{2p}}$	$k^{d}\varepsilon^{-2}$

Source: PhD thesis of O. Pembery, 2020, University of Bath.

More advanced topics

Surrogate models: what are they and why we need them?

Computational bottleneck: Often each evaluation of $Q(\vartheta)$ is expensive (e.g., one PDE solve).

> A **surrogate model** is a model that, for every value of ϑ, is an approximation to $Q(\vartheta)$ and is cheap to evaluate.

Offline/online paradigm: Surrogate is built in offline training phase (expensive), and used online (cheap), e.g. in Monte Carlo or optimization.

Types of surrogate models

Goal: Build $\tilde{Q}(\vartheta)$ s.t. $\tilde{Q}(\vartheta) \approx Q(\vartheta)$.

Intrusive methods: the original model is subject to modification.
Non-intrusive methods: original (high-fidelity) model as black box.
Built from training pairs $\{(\vartheta_i, Q(\vartheta_i))\}_{i=1}^N$.

(At least) three possible routes:

1. Acting on ϑ variable, **reduce dimensionality in parameter space**
2. Acting on (x, t) variables, **reduce dimensionality in spatio-temporal space**
3. **Simplifying the physics**, possibly reduce dimensionality in both spaces

Multilevel Monte Carlo with surrogate models

It falls within **multifidelity** approaches.

Note: *multifidelity estimator* as from [Peherstorfer, Willcox, Gunzburger, 2016] is slightly different but based on the same principles.

Reminder: Efficiency relies on **delicate balance** between variances and costs

1. Split $\varepsilon^2 = \varepsilon^2_{bias} + \varepsilon^2_{stat}$
2. Select $\mathcal{M}_L$ such that $\|\mathbb{E}\left[Q_L - Q\right]\| < \varepsilon_{bias}$
3. Select ordered subset $(\mathcal{M}_0, \ldots, \mathcal{M}_{L-1})$ minimizing estimated total cost
4. Choose $(M_l)_{l=0}^{L}$ s.t. stat error $< \varepsilon^2_{stat}$ at minimum cost

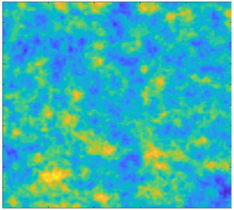
[Scarabosio et al. 2019], [Schaden Ullmann 2020].

See also important review paper [Peherstorfer, Willcox, Gunzburger 2018].

About things that we did not address

Random fields

Beyond scalar-valued random variables, we can use "random functions" to model distributed uncertainties.



Higher order methods

Under stricter smoothness (and dimension-anisotropy) assumptions, higher order methods can be used instead of Monte Carlo.

Catch for Helmholtz: as the frequency increases, their performance deteriorates *unless the variance decreases accordingly*.

[Ganesh et al. 2021], [Spence, Wunsch 2023], [Hiptmair et al. 2024]

MLMC in time domain: see e.g. [Mishra, Schwab, Sukys 2012].