Study of Uncertainty Quantification Techniques for the NEPTUNE Project

Peter V. Coveney, Maxime Vassaux, Wouter Edeling

Centre for Computational Science
University College London
UK

Institute for Informatics
University of Amsterdam
Netherlands

Recap of proposed and completed activities

- **Activity 1: Knowledge report on high-throughput VVUQ capabilities for the NEPTUNE software**
  - **Objective 1:** write a concise set of recommendations as to which UQ methodologies to develop (**March**)
    - **Completed**, submitted on 02/02/2021; review on UQ for fusion submitted on 18/02/2021
  - **Objective 2:** explicitly describe the architecture of UQ workflows for co-design purposes toward exascale (**June**)
    - M33 VECMA Toolkit code release is 29 March; starting initial work on active subspaces and semi-intrusive UQ for a chaotic code (molecular dynamics)

- **Activity 2: Workshops on UQ for ExCALIBUR partners**
  - **Objective 3:** hold a first meeting to introduce EasyVVUQ capabilities, provide hands-on tutorials and partners to present the structure of their codes and UQ requirements (**mid-January**)
    - **Completed**, UQ workshop held on 18/01 and Hackathon organised on 19-22/01 with participation of BOUT++ team
  - **Objective 4:** hold a second meeting to present our shortlist of recommended UQ methods (**early July**)
    - next Hackathon will take place 21-23 April.

Outline of report on high-throughput UQ

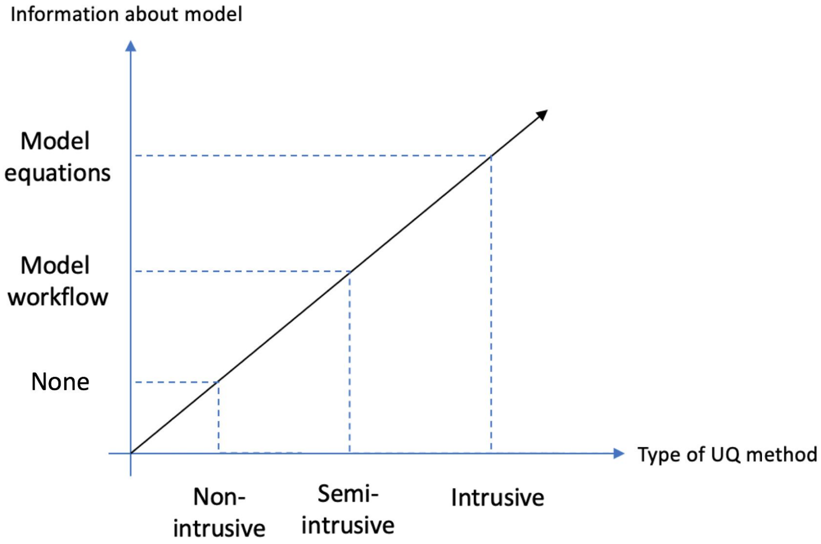
- **Concise set of recommendations as to which UQ methodologies to develop**
  - Establish standard use cases of Bout++ and Nektar++ simulations jointly with NEPTUNE partners
  - Perform UQ on the defined use cases using EasyVVUQ (VECMA) and MOGP (ATI) capabilities

- **Draw on VECMA expertise**
  - open source and open development software
  - D2.2: Report on advanced multiscale UQ algorithms, including intrusive approaches, and mapping thereof in UQPs and first results on V&V.
  - D2.1: Report on multiscale UQ algorithms based on non-intrusive MC and semi-intrusive MC and mapping thereof in UQPs

- **Involvement of David Coster and Jalal Lakhlili (Fusion, MPG-IPP), Wouter Edeling (UQ, CWI) and Eric Daub (MOR, ATI)**
  - in particular during NEPTUNE UQ workshop (18/01) and VECMA hackathon (19-21/01)

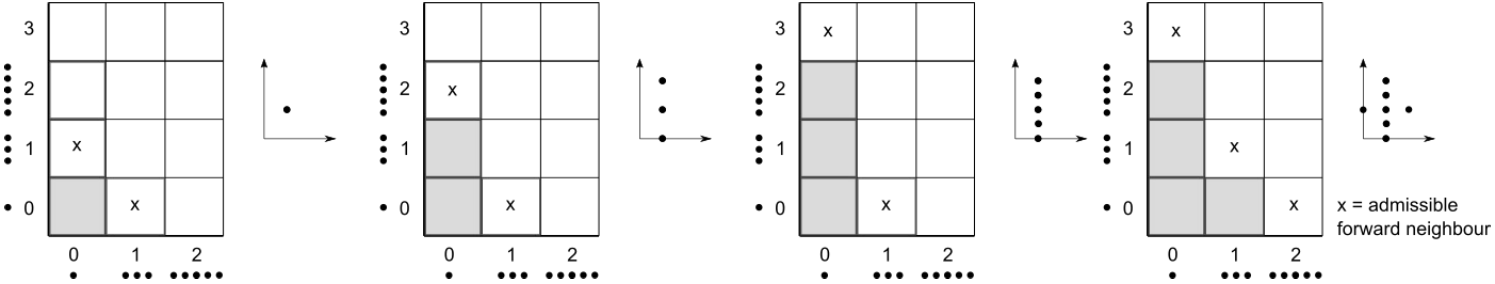
High-throughput UQ reporting

- short overview of approaches to uncertainty quantification most relevant to the Neptune project
- draws upon the experience gathered during the past three years in the VECMA project (www.vecma.eu)
    - facilitate the use of VVUQ techniques in multiscale, multi-physics applications
- classification of UQ methods based on their level of intrusiveness
    - focus on **non-intrusive** and **semi-intrusive** methods
    - intrusive methods are much cheaper but hardly automated
        - no generic solver
        - solver must be reimplemented for every different UQ
        - comparison with non-intrusive PCE on the basis of 2^{nd}-order exp. and 10 parameters:
            - non-intrusive PCE: 59049 samples
            - intrusive (stoch. Galerkin): 66 couples PDEs



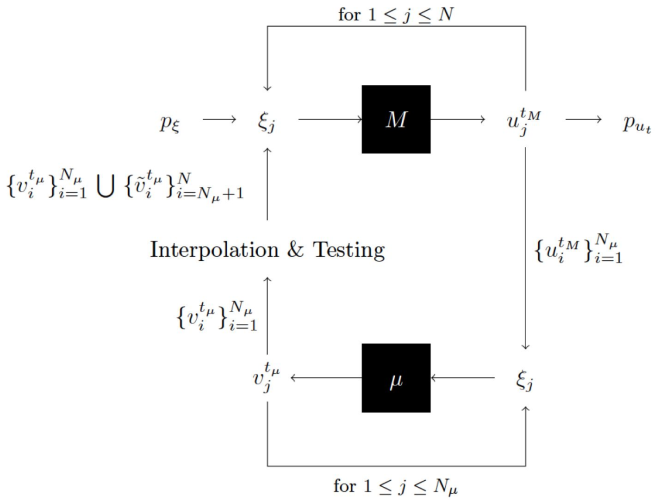
Non-intrusive methods: enhanced sampling

- Monte Carlo sampling: systematic convergence, but slow convergence rate
- Stochastic or point collocation (Spectral UQ): faster convergence (exponentially) in the most favourable cases
- Computational costs increases exponentially with the number of considered input parameters
- Fighting off the curse of dimensionality:
  - **delaying the curse**: adaptive sampling of the stochastic dimensions
  - **circumventing the curse**:
    - High-Dimensional Model Representation: never sample a high-dimensional space, break problem up into a series of low-dimensional subproblems.
    - Active subspaces: project inputs to a rotated (low-dimensional) coordinate system aligned with direction of strongest variability
- Dealing with irregular outputs (e.g. discontinuities or high gradients in stochastic space)
  - Use localised basis in stochastic space (Adaptive sparse grid / Simplex Stochastic Collocation methods)

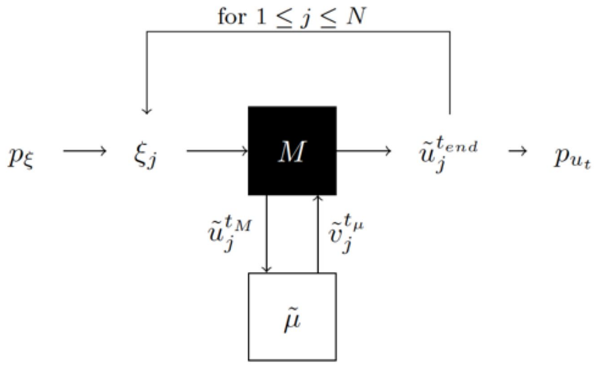


Semi-intrusive methods

- suited for coupled models applications
- individual models are treated as black-boxes, but the workflow is modified to accelerate UQ
- maintains “separation of concerns” for complex workflows



A smaller number of samples of the expensive submodel are simulated using advanced sampling.



The expensive submodel is replace by a cheaper surrogate model when computing ensembles of simulation to perform UQ.

Nikishova, Hoekstra. "Semi-intrusive uncertainty propagation for multiscale models." *Journal of Computational Science* 35 (2019): 80-90.

Surrogate modelling for multiscale models: stochastic surrogates

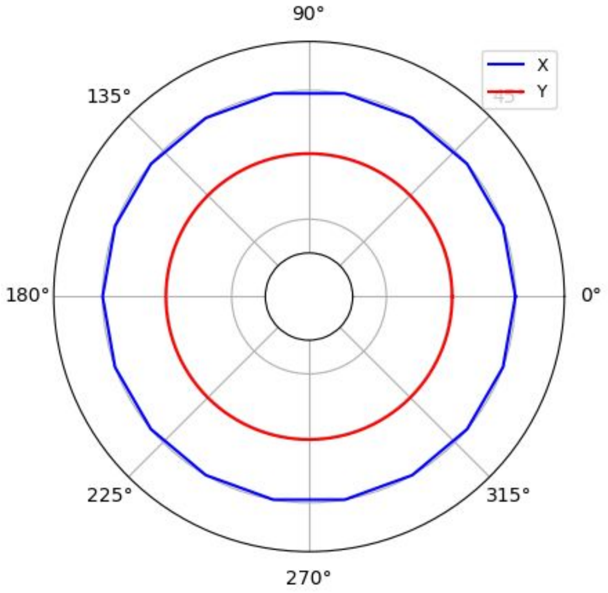
Example L96 multiscale problem:

- Toy model for the atmosphere at constant latitude.
- Full system: N x L couples ODEs (e.g. N=18, L = 20)

$$\frac{dX_n}{dt} = X_{n-1}\left(X_{n+1} - X_{n-2}\right) - X_n - F + r_n$$

$$r_n := \frac{h_x}{L}\sum_{l=1}^{L} Y_{l,n}$$

$$\frac{dY_{l,n}}{dt} = \frac{1}{\epsilon}\left[Y_{l+1,n}\left(Y_{l-1,n} - Y_{l+2,n}\right) - Y_{l,n} + h_y X_n\right].$$



Surrogate modelling for multiscale models: stochastic surrogates

- Works well for this problem, more complex problems might require **online** training. Ongoing research.
- More complex codes will also require coupling library.
  - VECMA option: Multiscale Coupling Library and Environment (MUSCLE 3, https://muscle3.readthedocs.io/)
- Thus far we only tested this on 2D, spectral codes

Edeling, Wouter, and Daan Crommelin. "Reducing data-driven dynamical subgrid scale models by physical constraints." *Computers & Fluids* 201 (2020): 104470.

January UQ workshop and hackathon

- **UQ NEPTUNE workshop held on January 18**
- **VECMA Hackathon held from January 19 to 21**
    - Participation of teams using BOUT++
    - Teams managed to perform basic UQ using either PCE or SC sampling methods
    - Respondents are so far satisfied with the UQ capabilities of VECMAtk and its ease of use
        - Use of EasyVVUQ and Dask
        - Interest in the PilotJob mechanism and FabSim3 to coordinate large ensembles
        - Interest in **adaptive sampling methods** available
        - More sophisticated model may be required for Neptune applications, questions remain:
            - will ever feature such large number of parameters (in the final coupled model)?
            - will there ever be a need to investigate all uncertainties at the same time?
            - each individual application developer might be able to significantly narrow down the number of uncertainties.
        - Both respondents indicate that they foresee **surrogates as being useful** but highlight the challenges in creating them. The main question remains **training**, as usual… That's where recent developments from CWI, when it comes to **online learning/training** could become interesting.
    - Hackathon participants **gained insight in what entails performing UQ** of their application
        - more precise idea of the inputs and outputs of their application (needed to couple single models, could be done with MUSCLE3)
        - distribution of inputs stochasticity

January UQ workshop and hackathon

- **Questionnaires to applications developers on UQ needs**
  - Applications type
    - Single Models (6 respondents)
    - Coupled Models (1 respondent)
  - Typical simulations will have between 2 and 100 uncertainty parameters
    - recommend **expert or sensitivity analysis based dimension reduction**
  - Interest in both **intrusive and non-intrusive** UQ methods
    - the first not compatible at first with "separation of concerns" paradigm
  - A desire for UQ codes which integrate easily with theirs and the availability of suitable surrogates.
    - **stochastic surrogates for turbulent flow** model simulation (see MUSCLE3/EasySurrogate tutorial for coupled stochastic surrogate)
    - concern about challenges in creating surrogate, in particular **training**
      - recent developments from CWI when it comes to **online learning/training** could become interesting