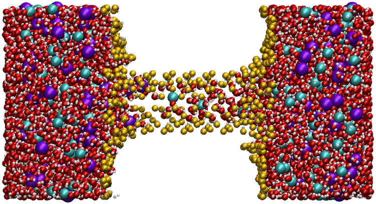
Uncertainty Quantification for predictive modeling and simulation

Francesco Rizzi

Senior Computational Scientist - NexGen Analytics (USA)

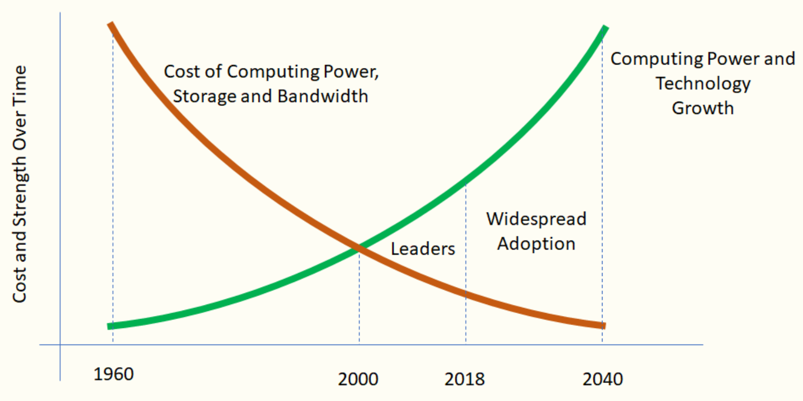
(worked performed while at The Johns Hopkins University, MD, USA)

UNIUD
11 October 2018



Modeling and Simulation: trend

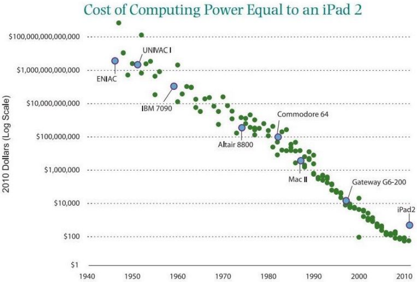
- Every field of science and engineering has seen an increasing demand for and use of modeling & simulation over the last two decades or so. Why?

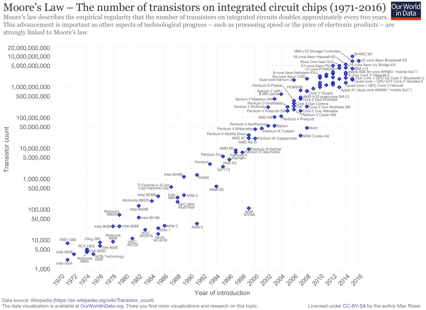


Look for more demanding problems and, in turn, create capabilities to solve them.

Modeling and Simulation: trend

- Every field of science and engineering has seen an increasing demand for and use of modeling & simulation over the last two decades or so. Why?





Cost vs year (left), # transistors vs year (right); (web).

Look for more demanding problems and, in turn, create capabilities to solve them.

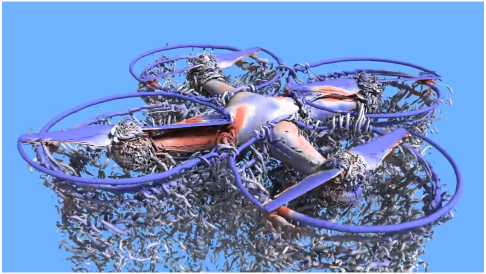
Modeling and Simulation: Why?

- Cost: generally much cheaper than experiments.
- Feasibility: often more simple to setup than a real experiment.
- Safety: they can be safer than conducting real-world experiments. For example, simulating nuclear devices, extreme climate or natural events.
- Flexibility: they allow to freely change the configuration of target parameters.
- Speed: simulations are (generally) conducted faster than experiments. Also accounting for preparation and setup time.

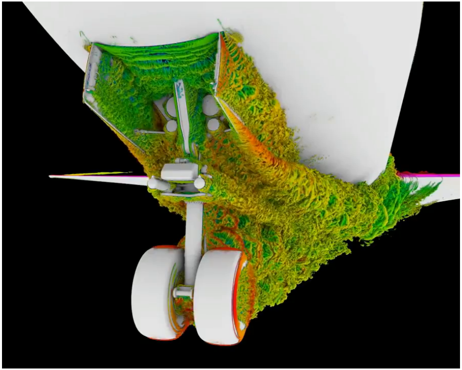
Towards Exascale

- 1 exaFlops (1e18) calculations per second, supposedly arriving by 2023-2024.
- As of June 2018:
  - Summit (USA): 187 PFlops
  - Sunway TaihuLight (China): 125 PFlops
  - Sierra (USA): 120 PFlops
  - Tianhe-2 (MilkyWay-2) (China): 54 PFlops
- Opportunities: simulations at an unprecedented length and time scales.

State-of-the-art simulations



Simulation of quadcopter (NASA)[SC17]



Landing of Boeing 777 (NASA)[SC17]

Meteoroid Airburst (NASA)[SC17]

Background and Motivation

- Predictive modeling & simulation is becoming crucial for science but...

George E. P. Box, 1919 - 2013 (statistician)

"Remember that all models are essentially wrong; the practical question is how wrong do they have to be to not be useful."

- Can we trust a simulation?
  1. Verification: predictions are consistent with the underlying mathematical model.
  2. Validation: are we building the right "tool"? i.e. agreement with experiments.
  3. Prediction: how reliable are the predictions?

- Models with complex physics + many parameters: small uncertainties/errors in the model/parameters can strongly affect the predictions.

- Key role when high-fidelity/risk assessment is of central importance.

- Uncertainty Quantification (UQ): **quantifying/characterizing** uncertainty.

Background and Motivation

- Predictive modeling & simulation is becoming crucial for science but...

George E. P. Box, 1919 - 2013 (statistician)

"Remember that all models are essentially wrong; the practical question is how wrong do they have to be to not be useful."

- Can we trust a simulation?
  1. Verification: predictions are consistent with the underlying mathematical model.
  2. Validation: are we building the right "tool"? i.e. agreement with experiments.
  3. Prediction: how reliable are the predictions?

- Models with complex physics + many parameters: small uncertainties/errors in the model/parameters can strongly affect the predictions.
- Key role when high-fidelity/risk assessment is of central importance.
- Uncertainty Quantification (UQ): **quantifying/characterizing** uncertainty.

Background and Motivation

- Predictive modeling & simulation is becoming crucial for science but...

George E. P. Box, 1919 - 2013 (statistician)

"Remember that all models are essentially wrong; the practical question is how wrong do they have to be to not be useful."

- Can we trust a simulation?
  1. Verification: predictions are consistent with the underlying mathematical model.
  2. Validation: are we building the right "tool"? i.e. agreement with experiments.
  3. Prediction: how reliable are the predictions?

- Models with complex physics + many parameters: small uncertainties/errors in the model/parameters can strongly affect the predictions.

- Key role when high-fidelity/risk assessment is of central importance.

- Uncertainty Quantification (UQ): **quantifying/characterizing** uncertainty.

Background and Motivation

- Predictive modeling & simulation is becoming crucial for science but...

George E. P. Box, 1919 - 2013 (statistician)

"Remember that all models are essentially wrong; the practical question is how wrong do they have to be to not be useful."

- Can we trust a simulation?
  1. Verification: predictions are consistent with the underlying mathematical model.
  2. Validation: are we building the right "tool"? i.e. agreement with experiments.
  3. Prediction: how reliable are the predictions?

- Models with complex physics + many parameters: small uncertainties/errors in the model/parameters can strongly affect the predictions.

- Key role when high-fidelity/risk assessment is of central importance.

- Uncertainty Quantification (UQ): **quantifying/characterizing** uncertainty.

UQ is important for...

- Chaotic systems : sensitivity to initial conditions
- Boundary conditions/initial conditions
  - can be very complex to set for non trivial systems.
- Large Eddy Simulation (LES): subgrid scale models.
- Combustion
  - not much confidence on reaction rates (κ_i).
- Materials
  - e.g. physical parameters, microstructure, isotropy, etc.
- Molecular dynamics: interatomic potentials
- Plasma physics
  - e.g. physical parameters, poorly understood high-temperature kinematics
- Multi-physics/multi-scale simulations
  - subsystems interactions, information propagation, a mix of all of the above.
  - climate models are *the* representative example.
- ...

Example of why correlations matter

- Consider $\phi(x, y, t; a, b)$:

$$\dot{x}(t) = a^2 - b^2$$
$$\dot{y}(t) = ab + 0.01\sin(x)$$

- a, b are **model parameters**:

$$(a, b) \sim \mathcal{N}([2\ 1], Cov)$$

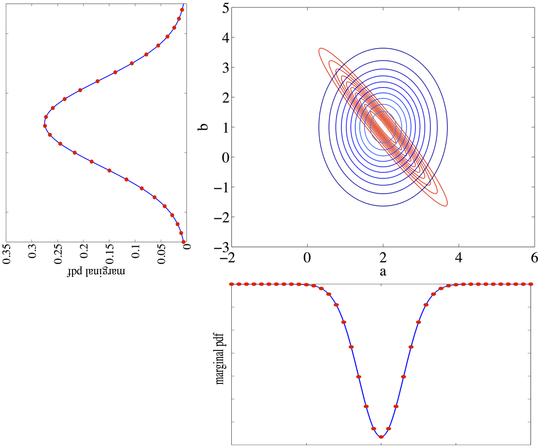
- Two cases:

Uncorrelated parameters:

$$Cov = \begin{bmatrix} 0.6 & 0.0 \\ 0.0 & 1.45 \end{bmatrix}$$

Correlated parameters:

$$Cov = \begin{bmatrix} 0.6 & -0.9 \\ -0.9 & 1.45 \end{bmatrix}$$



- Same marginal densities.
- What is the impact of the correlation?

Example of why correlations matter

- Consider $\phi(x, y, t; a, b)$:

$$\begin{aligned} \dot{x}(t) &= a^2 - b^2 \\ \dot{y}(t) &= ab + 0.01\sin(x) \end{aligned}$$

- a, b are **model parameters**:

$$(a, b) \sim \mathcal{N}([2\ 1], Cov)$$

- Two cases:

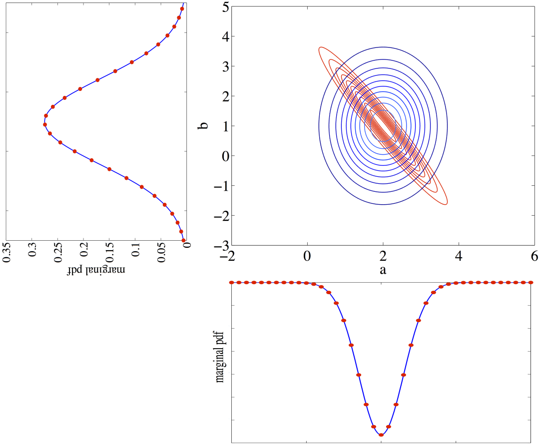
  Uncorrelated parameters:

$$Cov = \begin{bmatrix} 0.6 & 0.0 \\ 0.0 & 1.45 \end{bmatrix}$$

  Correlated parameters:

$$Cov = \begin{bmatrix} 0.6 & -0.9 \\ -0.9 & 1.45 \end{bmatrix}$$

- Same marginal densities.
- What is the impact of the correlation?



Example of why correlations matter

- Consider $\phi(x, y, t; a, b)$:

$$\dot{x}(t) = a^2 - b^2$$
$$\dot{y}(t) = ab + 0.01\sin(x)$$

- $\underline{a, b}$ are **model parameters**:

$$(a, b) \sim \mathcal{N}([2\ 1], Cov)$$

- Two cases:

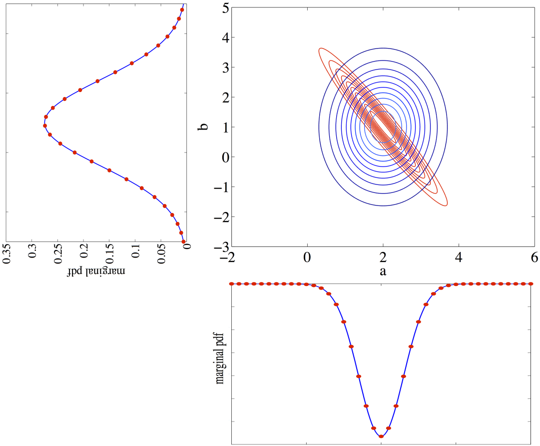
Uncorrelated parameters:

$$Cov = \begin{bmatrix} 0.6 & 0.0 \\ 0.0 & 1.45 \end{bmatrix}$$

Correlated parameters:

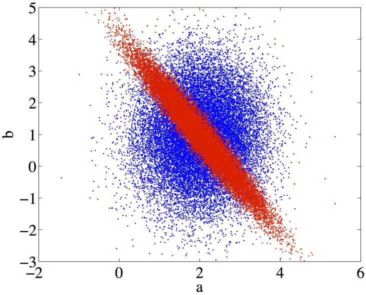
$$Cov = \begin{bmatrix} 0.6 & -0.9 \\ -0.9 & 1.45 \end{bmatrix}$$

- Same marginal densities.
- What is the impact of the correlation?



Example: results

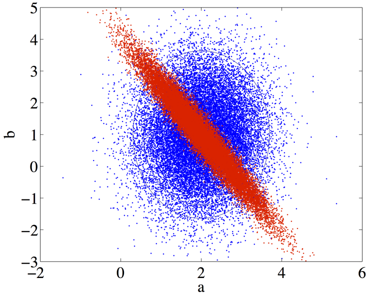
- Sample the joint PDFs: $\{(a_i, b_i)^{U,C}\}_{i=1}^{n}$
- Compute trajectories from $(x_0 = 1, y_0 = 0.5)$.
- Two sets of predictions: $\{(x_j, y_j)^{U,C}\big|_T\}_{j=1}^{n}$
- Estimate the joint PDFs.



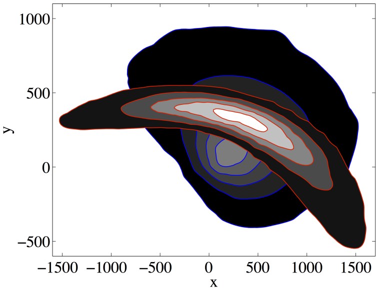
- Model predictions are substantially different.
- Correlation has large impact.
- Especially important for more complicated and non-linear systems.

Example: results

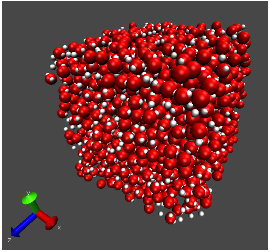
- Sample the joint PDFs: $\{(a_i, b_i)^{U,C}\}_{i=1}^{n}$
- Compute trajectories from $(x_0 = 1, y_0 = 0.5)$.
- Two sets of predictions: $\{(x_j, y_j)^{U,C}|_T\}_{j=1}^{n}$
- Estimate the joint PDFs.



- Model predictions are substantially different.
- Correlation has large impact.
- Especially important for more complicated and non-linear systems.



Forward Propagation

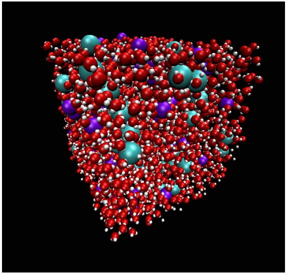


- ⋆ F. Rizzi, H. Najm, B. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson and O. Knio - Part I – *SIAM Multiscale Modeling & Simulation*, 10(4), 1428-1459.
- ⋆ F. Rizzi, H. Najm, B. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson and O. Knio - Part II – *SIAM Multiscale Modeling & Simulation*, 10(4), 1460-1492.
- ⋆ F. Rizzi, Ph.D. thesis, The Johns Hopkins University, Baltimore, MD.

Background and Motivation: MD overview

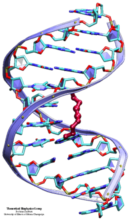
- 1957: seminal work in molecular dynamics (MD). (Alder and Wainwright)
- 1964: first MD simulation based on a realistic potential (Lennard-Jones): liquid Ar (Rahman).
- 1974: first MD simulation of liquid water. (Stillinger and Rahman)

  . . .



MD simulation of **Na** **Cl** in water.

- MD is useful and cheap (vs. experiments).
- Industrial/academic applications: liquids, solids, proteins and nucleic acids (DNA, RNA).
- As every simulation technique, MD is an approximation method with a few weaknesses...



MD snapshot of DNA (Biophys. group, UIUC)

Background and Motivation: MD overview

- Classical MD simulation (Frenkel,2001; Allen & Tildesley,1987):

$$\frac{d^2\mathbf{r}_{(i,t)}}{dt^2} = \frac{\mathbf{f}_{(i,t)}}{m_i} \qquad \mathbf{f}_{(i,t)} = -\nabla_{\mathbf{r_i}}\Phi\left(\mathbf{r}_{(1,t)}, \ldots, \mathbf{r}_{(N,t)}\right) \qquad i = 1, \ldots, N$$

- Φ is the potential (or force-field), defined *before* starting the simulation.
- Φ should be tailored to the target application.
- Reliability depends on the accuracy of Φ.
- Continuous development of potentials and experience over the years.
- MD potential represents an important source of uncertainty.

Background and Motivation: MD overview

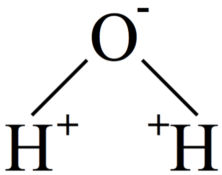
- Classical MD simulation (Frenkel,2001; Allen & Tildesley,1987):

$$\frac{d^2\mathbf{r}_{(i,t)}}{dt^2} = \frac{\mathbf{f}_{(i,t)}}{m_i} \qquad \mathbf{f}_{(i,t)} = -\nabla_{\mathbf{r_i}}\Phi\big(\mathbf{r}_{(1,t)}, \ldots, \mathbf{r}_{(N,t)}\big) \qquad i = 1, \ldots, N$$

- Φ is the potential (or force-field), defined *before* starting the simulation.
- Φ should be tailored to the target application.
- Reliability depends on the accuracy of Φ.
- Continuous development of potentials and experience over the years.
- MD potential represents an important source of uncertainty.

Potential Uncertainty for Water

- Water is the most investigated liquid.
- "Looks" simple...but it is not!
- Behavior of liquid water is quite different from other similar liquids: 41 "anomalies"!



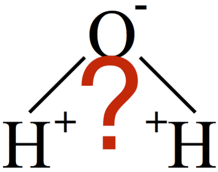
- More that 50 water models have been developed!

✗ NO existing model is able to reproduce with good accuracy *all* its properties.

✗ Many sources of uncertainty to consider: potential form and parameters, thermal noise, molecular geometry, etc...

Potential Uncertainty for Water

- Water is the most investigated liquid.
- "Looks" simple...but it is not!
- Behavior of liquid water is quite different from other similar liquids: 41 "anomalies"!
- More that 50 water models have been developed!

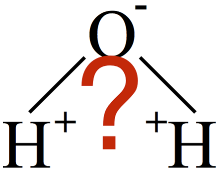


✗ NO existing model is able to reproduce with good accuracy *all* its properties.

✗ Many sources of uncertainty to consider:
potential form and parameters, thermal noise, molecular geometry, etc...

Potential Uncertainty for Water

- Water is the most investigated liquid.
- "Looks" simple...but it is not!
- Behavior of liquid water is quite different from other similar liquids: 41 "anomalies"!
- More that 50 water models have been developed!



✗ NO existing model is able to reproduce with good accuracy *all* its properties.

✗ Many sources of uncertainty to consider:
potential form and parameters, thermal noise, molecular geometry, etc...

Potential Uncertainty for Water

- Review of MD water models Guillot(2002) and Wallqvist(2007).

Acronym	Date	Type	Sites	Reference
SPC	1981	rigid	3	(Berendsen,1981)
TIP3P	1981	rigid	3	(Jorgensen,1983)
SPC/F	1985	flexible	3	(Toukan,1985)
SPC/FP	1991	flexible,polarizable	3	(Zhu,1991)
NSPCE	1998	rigid	3	(Errington,1998)
SPC/Fw	2006	flexible	3	(Wu,2006)
BF	**1933**	rigid	4	(Bernal,1933)
RWK	1982	flexible	4	(Reimers,1982)
TIP4P	1983	rigid	4	(Jorgensen,1983)
PTIP4P	1991	polarizable	4	(Sprik,1991)
TIP4P/FQ	1994	polarizable	4	(Rick,1994)
TIP4P-Ew	2004	rigid	4	(Horn,2004)
TIP4P/2005	2005	rigid	4	(Abascal,2005)
ST2	1973	rigid	5	(Stillinger,1974)
TIP5P	2000	rigid	5	(Mahoney,2000)
TIP5P-Ew	2004	rigid	5	(Rick,2004)
NvdE	2003	rigid	6	(Nada,2003)

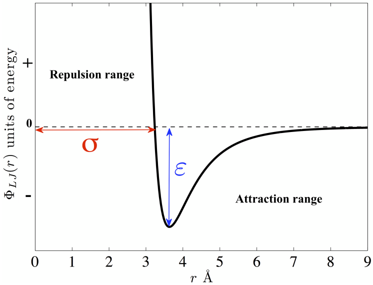
Table: Reduced list of water models developed since 1933.

Potential Uncertainty for Water

a Most water models use Lennard-Jones (LJ) potential to describe Van der Waals forces.

$$\Phi_{LJ}(r) = 4\varepsilon \left\{ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right\}$$

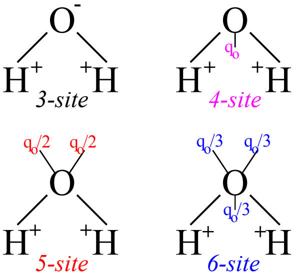
- Different models involve different values of the LJ parameters ε, σ.



b Rigid or flexible molecule.

c H_2O geometry: from 3-site to 6 sites models.

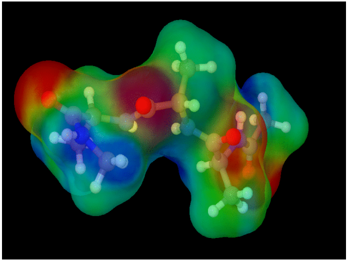
...



- Discussion holds for several other systems: potential and parameters are important sources of uncertainty to consider.

Ab Initio MD versus Classical MD

- 1984: ab initio MD by Car and Parrinello.
  Full quantum mechanical electronic structure problem is solved "on-the-fly" to compute forces.
  - √ No need for the potential.
  - *X* Large computational cost, small-size systems.
  - *X* Practical time scales on the order of picosec.



Ab initio simulation of protein folding.
Isosurface of electrostatic potential -, +
due to the instantaneous configuration.
Source: Pietro Faccioli, Univ. TN, Italy

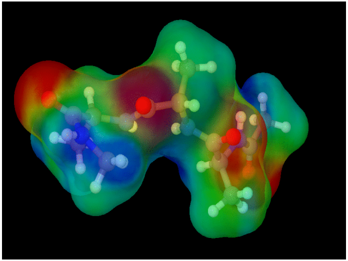
- Classical MD:
  - *X* Need potential.
  - √ Systems of order 10^6 atoms with current supercomputers.
  - √ Practical time-scales on the order of nanosec/microsec.

- Nano(micro) seconds: ideal time-scale to explore atomistic systems.

⇒ Feasible time scales still makes classical MD the preferred setting.

Ab Initio MD versus Classical MD

- 1984: ab initio MD by Car and Parrinello.
  Full quantum mechanical electronic structure problem is solved "on-the-fly" to compute forces.
  - √ No need for the potential.
  - X Large computational cost, small-size systems.
  - X Practical time scales on the order of picosec.



Ab initio simulation of protein folding.
Isosurface of electrostatic potential -, + due to the instantaneous configuration.
Source: Pietro Faccioli, Univ. TN, Italy

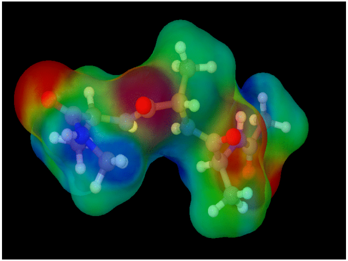
- Classical MD:
  - X Need potential.
  - √ Systems of order 10^6 atoms with current supercomputers.
  - √ Practical time-scales on the order of nanosec/microsec.

- Nano(micro) seconds: ideal time-scale to explore atomistic systems.

⇒ Feasible time scales still makes classical MD the preferred setting.

Ab Initio MD versus Classical MD

- 1984: ab initio MD by Car and Parrinello.
  Full quantum mechanical electronic structure problem is solved "on-the-fly" to compute forces.
  - √ No need for the potential.
  - X Large computational cost, small-size systems.
  - X Practical time scales on the order of picosec.



Ab initio simulation of protein folding.
Isosurface of electrostatic potential -, +
due to the instantaneous configuration.
Source: Pietro Faccioli, Univ. TN, Italy

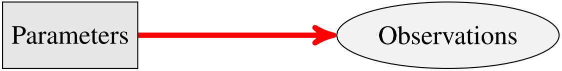
- Classical MD:
  - X Need potential.
  - √ Systems of order 10^6 atoms with current supercomputers.
  - √ Practical time-scales on the order of nanosec/microsec.

- Nano(micro) seconds: ideal time-scale to explore atomistic systems.

⇒ Feasible time scales still makes classical MD the preferred setting.

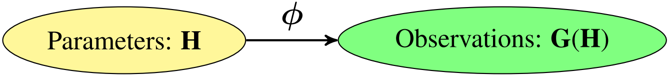
Forward Propagation in MD

- Focus on MD simulations of liquid water at ambient conditions.
- How uncertainties in a set of potential parameters affect MD predictions.



Forward Problem

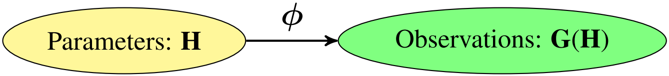
- Consider a generic computational model (ODEs/PDEs):



- Forward problem = uncertainty definition + propagation.
- Aleatoric (intrisic) uncertainty:
  - Physical variability in the system or its environment: e.g. fabrication processes.
  - Not strictly due to lack of knowledge.
  - It cannot be avoided/reduced.
- Epistemic (parametric) uncertainty:
  - Uncertainty solely due to lack of knowledge: e.g. turbulence models.
  - E.g. missing/partial information, simplifications in the model formulation, etc.
  - It can be reduced and is usually important.
- Propagation: characterize the impact on target model observables.

Forward Problem

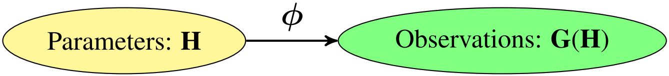
- Consider a generic computational model (ODEs/PDEs):



- Forward problem = uncertainty definition + propagation.
- Aleatoric (intrisic) uncertainty:
  - Physical variability in the system or its environment: e.g. fabrication processes.
  - Not strictly due to lack of knowledge.
  - It cannot be avoided/reduced.
- Epistemic (parametric) uncertainty:
  - Uncertainty solely due to lack of knowledge: e.g. turbulence models.
  - E.g. missing/partial information, simplifications in the model formulation, etc.
  - It can be reduced and is usually important.
- Propagation: characterize the impact on target model observables.

Forward Problem

- Consider a generic computational model (ODEs/PDEs):



- Forward problem = uncertainty definition + propagation.
- Aleatoric (intrisic) uncertainty:
  - Physical variability in the system or its environment: e.g. fabrication processes.
  - Not strictly due to lack of knowledge.
  - It cannot be avoided/reduced.
- Epistemic (parametric) uncertainty:
  - Uncertainty solely due to lack of knowledge: e.g. turbulence models.
  - E.g. missing/partial information, simplifications in the model formulation, etc.
  - It can be reduced and is usually important.
- Propagation: characterize the impact on target model observables.

Forward Problem for Intensive Simulations

- Assume ϕ is **expensive**, i.e. we can afford limited number of runs.
- Propagation methods:
  1. "Pure sampling": Monte Carlo, importance or adaptive sampling.
     - ✗ Require many runs, yields limited information, slow convergence.
  2. "Local methods": Taylor series, perturbation method.
     - ✗ Local variability of an output with respect to inputs.
     - ✗ Only local information, no PDFs of **G** can be obtain.
  3. "Functional methods": **polynomial chaos expansion (PCe)**.
     - ✓ "Global" representation with respect to the input space.
     - ✓ Allows estimation of PDFs and moments of observable (**G**) easily and efficiently.

PCe: one uncertain input, one output

- Forward model: $G(x,t,h) = \phi(x,t;h)$; h = parameter; G = observable.
- Suppose uncertainty on h in the form:

$$h = \mu + \sigma\,\xi, \qquad p(\xi) = \mathcal{N}(0,1)$$

- h is a RV $\Rightarrow$ output $G(x,t,h)$ be considered as a RV.
- Wiener (1938): if G has finite variance, it can be expressed as a spectral expansion of the uncertain variable (or "germ") ξ:

$$G(x,t,\xi) = \sum_{\ell=0}^{\infty} \underbrace{c_\ell(x,t)}_{deterministic} \underbrace{\psi_\ell(\xi)}_{stochastic}$$

- $\psi_\ell(\xi)$ are Hermite polynomials: $\psi_0 = 1$, $\psi_1 = \xi$, $\psi_2 = \xi^2 - 1$, ...
- The basis form a complete set of **orthogonal** functions in prob space:

$$< \psi_k, \psi_\ell >= \int \psi_k\, \psi_\ell\, \mathrm{p}(\xi) d\xi = h_k \delta_{k\ell}$$

PCe: one uncertain input, one output

- Forward model: $G(x,t,h) = \phi(x,t;h)$; h = parameter; G = observable.
- Suppose uncertainty on h in the form:

$$h = \mu + \sigma\,\xi, \qquad p(\xi) = \mathcal{N}(0,1)$$

- h is a RV $\Rightarrow$ output $G(x,t,h)$ be considered as a RV.
- Wiener (1938): if G has finite variance, it can be expressed as a spectral expansion of the uncertain variable (or "germ") ξ:

$$G(x,t,\xi) = \sum_{\ell=0}^{\infty} \underbrace{c_\ell(x,t)}_{deterministic} \underbrace{\psi_\ell(\xi)}_{stochastic}$$

- $\psi_\ell(\xi)$ are Hermite polynomials: $\psi_0 = 1$, $\psi_1 = \xi$, $\psi_2 = \xi^2 - 1$, ...
- The basis form a complete set of **orthogonal** functions in prob space:

$$< \psi_k, \psi_\ell >= \int \psi_k\, \psi_\ell\, \mathrm{p}(\xi) d\xi = h_k \delta_{k\ell}$$

PCe: one uncertain input, one output

- Forward model: $G(x,t,h) = \phi(x,t;h)$; h = parameter; G = observable.
- Suppose uncertainty on h in the form:

$$h = \mu + \sigma\,\xi, \qquad p(\xi) = \mathcal{N}(0,1)$$

- h is a RV $\Rightarrow$ output $G(x,t,h)$ be considered as a RV.
- Wiener (1938): if G has finite variance, it can be expressed as a spectral expansion of the uncertain variable (or "germ") ξ:

$$G(x,t,\xi) = \sum_{\ell=0}^{\infty} \underbrace{c_\ell(x,t)}_{deterministic} \underbrace{\psi_\ell(\xi)}_{stochastic}$$

- $\psi_\ell(\xi)$ are Hermite polynomials: $\psi_0 = 1$, $\psi_1 = \xi$, $\psi_2 = \xi^2 - 1$, ...
- The basis form a complete set of **orthogonal** functions in prob space:

$$< \psi_k, \psi_\ell >= \int \psi_k\, \psi_\ell\, \mathrm{p}(\xi) d\xi = h_k \delta_{k\ell}$$

PCe: one uncertain input, one output

- Forward model: $G(x,t,h) = \phi(x,t;h)$; $h =$ parameter; $G =$ observable.
- Suppose uncertainty on h in the form:

$$h = \mu + \sigma\,\xi, \qquad p(\xi) = \mathcal{N}(0,1)$$

- h is a RV $\Rightarrow$ output $G(x,t,h)$ be considered as a RV.
- Wiener (1938): if G has finite variance, it can be expressed as a spectral expansion of the uncertain variable (or "germ") ξ:

$$G(x,t,\xi) = \sum_{\ell=0}^{\infty} \underbrace{c_\ell(x,t)}_{deterministic}\ \underbrace{\psi_\ell(\xi)}_{stochastic}$$

- $\psi_\ell(\xi)$ are Hermite polynomials: $\psi_0 = 1$, $\psi_1 = \xi$, $\psi_2 = \xi^2 - 1$, ...
- The basis form a complete set of **orthogonal** functions in prob space:

$$< \psi_k, \psi_\ell >= \int \psi_k\, \psi_\ell\, \mathrm{p}(\xi) d\xi = h_k \delta_{k\ell}$$

PCe: moments of the target output variable

- PCe: $G(x,t,\xi) = \sum_{\ell=0}^{\infty} \underbrace{c_\ell(x,t)}_{deterministic} \underbrace{\psi_\ell(\xi)}_{stochastic}$
- PC coefficients fully determine the expansion.
- Orthogonal Polynomials:

$$E[\psi_0] = \int_\Omega \psi_0 \, p(\xi) d\xi = \int_\Omega 1 \, p(\xi) d\xi = 1$$
$$E[\psi_k] = \int_\Omega \psi_k \, p(\xi) d\xi = \int_\Omega \psi_k \, 1 \, p(\xi) d\xi = 0, \quad k >= 1$$

- The moments of G can be **directly** computed:

$$E[G] = \int_\Omega G \, p(\xi) d\xi = \int_\Omega \left[\sum_{l=0}^{\infty} c_l(x,t) \Psi_l(\xi) \right] p(\xi) d\xi = c_0$$

$$Var(G) = E[(G - c_0)^2] = \sum_{\ell=0}^{\infty} c_\ell^2 < \psi_\ell^2 >$$

- We can also easily reconstruct PDF(G) since PCe is very cheap to evaluate.

PCe: multiple uncertain inputs, different germs

- Let $G(x, t, \mathbf{H}) = \phi(x, t; \mathbf{H})$, with $\mathbf{H} = \{h_1, \ldots, h_m\}$ being a vector of parameters.
- Parametrize uncertainty:

$$\mathbf{H} = \mathbf{f}(\boldsymbol{\xi})$$

  where $\boldsymbol{\xi} = \{\xi_1, \ldots, \xi_m\}$ are *i.i.d.* standard RVs.
- Framework holds for various germs, not only Gaussian.

$$G(x, t, \boldsymbol{\xi}) = \sum_{\ell=0}^{\infty} c_\ell(x, t) \Psi_\ell(\boldsymbol{\xi})$$

- $\xi_1, ..., \xi_m \sim \mathcal{U}[-1, 1] \Longrightarrow \Psi_l(\boldsymbol{\xi})$ are Multivariate Legendre polynomials.
- $\xi_1, ..., \xi_m \sim \mathcal{N}[0, 1] \Longrightarrow \Psi_l(\boldsymbol{\xi})$ are Multivariate Hermite polynomials.
- ...can be generalized to other probability distributions.
- How to compute the coefficients? Use orthogonality of basis functions!

Orthogonality

- The orthogonality of the basis functions yields:

$$c_\ell = \frac{1}{\langle \Psi_\ell, \Psi_\ell \rangle} \int_\Omega G(\boldsymbol{\xi}) \Psi_\ell(\boldsymbol{\xi}) p_f(\boldsymbol{\xi}) d\boldsymbol{\xi}, \qquad \ell = 0, \ldots, P,$$

where Ω is the support of $\boldsymbol{\xi}$, and $<>$ denotes the inner product.

1. Intrusive Spectral Projection (ISP):
   - Galerkin procedure to the governing equations: original governing equations are replaced with equations for the PC coefficients.
   - Not applicable in the absence of a deterministic forward model.

2. Non-Intrusive Spectral Projection (NISP):
   - No reformulation of the governing equations.
   - Based on independent sampling of G and $\boldsymbol{\xi}$ to compute the projection integral.
   - **Numerical integration**, collocation methods, least-square fitting.

Orthogonality

- The orthogonality of the basis functions yields:

$$c_\ell = \frac{1}{\langle \Psi_\ell, \Psi_\ell \rangle} \int_\Omega G(\boldsymbol{\xi}) \Psi_\ell(\boldsymbol{\xi}) p_f(\boldsymbol{\xi}) d\boldsymbol{\xi}, \qquad \ell = 0, \dots, P,$$

where Ω is the support of $\boldsymbol{\xi}$, and $<>$ denotes the inner product.

1. Intrusive Spectral Projection (ISP):
   - Galerkin procedure to the governing equations: original governing equations are replaced with equations for the PC coefficients.
   - Not applicable in the absence of a deterministic forward model.

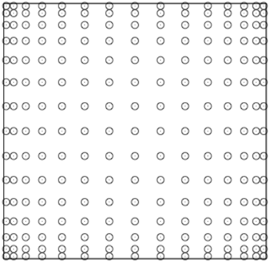
2. Non-Intrusive Spectral Projection (NISP):
   - No reformulation of the governing equations.
   - Based on independent sampling of G and $\boldsymbol{\xi}$ to compute the projection integral.
   - **Numerical integration**, collocation methods, least-square fitting.

NISP approach

- Gauss quadrature with n nodes along each dimension yields:

$$c_\ell = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_m=1}^{n} G(\xi_{i_1},\xi_{i_2},\ldots,\xi_{i_m})\frac{\Psi_\ell(\xi_{i_1},\xi_{i_2},\ldots,\xi_{i_m})}{\langle\Psi_\ell,\Psi_l\rangle}\left[\prod_{q=1}^{m} w_{i_q}\right], \quad \ell = 0,\ldots,P$$

- $\{\xi_j\}_{j=1}^{n}$: nodes
- $\{w_j\}_{j=1}^{n}$: weights
- $G(\xi_{i_1},\xi_{i_2},\ldots,\xi_{i_m})$ is the observable value
- Regularity of G with respect to $\boldsymbol{\xi}$.
- Feasible for low-dimensional problems.
- Sparse tensorization approaches can mitigate the curse of dimensionality, but issues may arise due to negative weights in the corresponding quadrature rules.

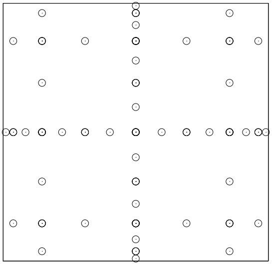


NISP approach

- Gauss quadrature with n nodes along each dimension yields:

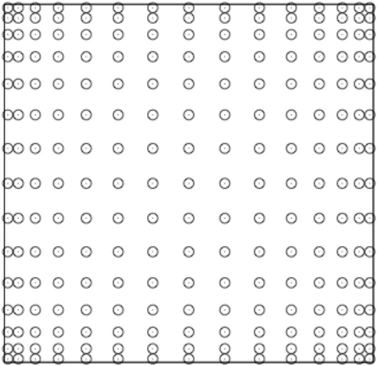
$$c_\ell = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \dots \sum_{i_m=1}^{n} G(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_m}) \frac{\Psi_\ell(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_m})}{\langle \Psi_\ell, \Psi_l \rangle} \left[\prod_{q=1}^{m} w_{i_q} \right], \quad \ell = 0, \dots, P$$

- $\{\xi_j\}_{j=1}^{n}$: nodes
- $\{w_j\}_{j=1}^{n}$: weights
- $G(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_m})$ is the observable value
- Regularity of G with respect to $\boldsymbol{\xi}$.
- Feasible for low-dimensional problems.
- Sparse tensorization approaches can mitigate the curse of dimensionality, but issues may arise due to negative weights in the corresponding quadrature rules.



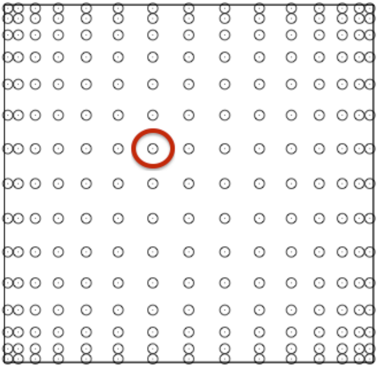
NISP details

- Procedure:
  1. Construct the quadrature grid.
  2. Each node $\boldsymbol{\xi}^{(i)}$ corresponds to a set of driving parameters, $\mathbf{H}^{(i)}$: $\mathbf{H}^{(i)} = \mathbf{f}(\boldsymbol{\xi}^{(i)})$
  3. Forward model yields: $G^i = \phi(\mathbf{H}^{(i)})$.
  4. Collect $\{G^i\}_{i=1}^N$ and evaluate quadrature integral for each PC coefficient.



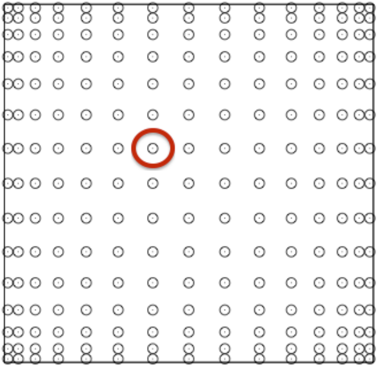
NISP details

- Procedure:
  1. Construct the quadrature grid.
  2. Each node $\boldsymbol{\xi}^{(i)}$ corresponds to a set of driving parameters, $\mathbf{H}^{(i)}$: $\mathbf{H}^{(i)} = \mathbf{f}(\boldsymbol{\xi}^{(i)})$
  3. Forward model yields: $G^i = \phi(\mathbf{H}^{(i)})$.
  4. Collect $\{G^i\}_{i=1}^N$ and evaluate quadrature integral for each PC coefficient.



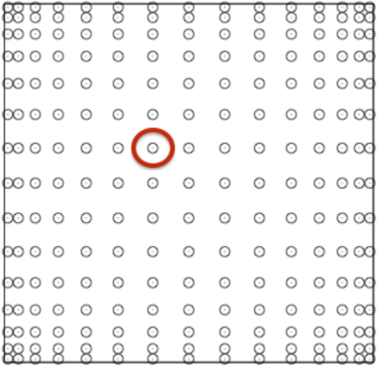
NISP details

- Procedure:
  1. Construct the quadrature grid.
  2. Each node $\boldsymbol{\xi}^{(i)}$ corresponds to a set of driving parameters, $\mathbf{H}^{(i)}$: $\mathbf{H}^{(i)} = \mathbf{f}(\boldsymbol{\xi}^{(i)})$
  3. Forward model yields: $G^i = \phi(\mathbf{H}^{(i)})$.
  4. Collect $\{G^i\}_{i=1}^{N}$ and evaluate quadrature integral for each PC coefficient.



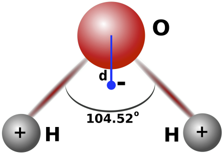
NISP details

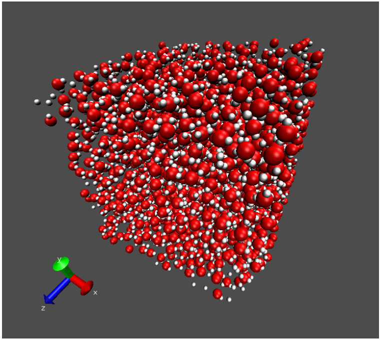
- Procedure:
  1. Construct the quadrature grid.
  2. Each node $\boldsymbol{\xi}^{(i)}$ corresponds to a set of driving parameters, $\mathbf{H}^{(i)}$: $\mathbf{H}^{(i)} = \mathbf{f}(\boldsymbol{\xi}^{(i)})$
  3. Forward model yields: $G^i = \phi(\mathbf{H}^{(i)})$.
  4. Collect $\{G^i\}_{i=1}^{N}$ and evaluate quadrature integral for each PC coefficient.



Foward Problem for MD: Computational System

- *Isothermal*, *isobaric* MD simulations of liquid water: $T = 298$ K, $P = 1$ atm.
- Domain: periodic cubic box of volume $\sim 64\ nm^3$ with 1728 molecules.
- Water molecule: four-site rigid model (TIP4P): widely used for liquid water.





- Potential: Lennard-Jones + Coulombic interaction.
- Simulations: MPI-C++ code adapted from LAMMPS (lammps.sandia.gov).
- Long simulation time to ensure steady state and proper time-averaged observables.

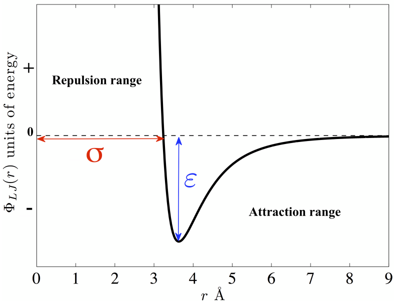
Forward Problem: Parametric Uncertainty

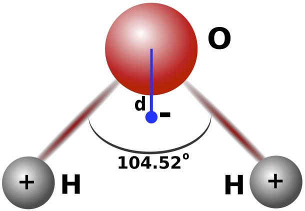
Introduce **parametric uncertainty** on three potential parameters: σ, ε, d

$\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ in the LJ potential:

$$\Phi_{LJ} = 4\varepsilon \left\{ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right\}$$

The distance, **d**, from the oxygen to the massless point where negative charge is placed in the TIP4P model.





Formulation

- Parametric uncertainty (PU) expressed as (values extracted from literature):

$$\begin{aligned} \varepsilon(\xi_1) &= 0.1470 + 0.043\,\xi_1 \ \text{(kcal/mol)} \\ \sigma(\xi_2) &= 3.1506 + 0.021\,\xi_2 \ (\text{Å}) \qquad \text{where } \{\xi_i\}_{i=1}^{3} \sim \mathcal{U}(-1,1) \\ d(\xi_3) &= 0.1400 + 0.035\,\xi_3 \ (\text{Å}) \end{aligned}$$

- MD intrinsic noise (IN): single sample, $\{\varepsilon^j, \sigma^j, d^j\}$, yields multiple predictions of the target observable, $G_1^j, G_2^j, \ldots, G_n^j$.
- PU + IN $\Rightarrow$ non-deterministic, noisy MD predictions of the water observables.
- Account for PU and IN using PCe:

$$\overline{G} \approx M(\xi_1, \xi_2, \xi_3) \equiv \sum_{k=0}^{P} c_k \Psi_k(\xi_1, \xi_2, \xi_3)$$

- ⋆ $\overline{G}$: quantity at steady-state averaged over m MD realizations (and time).
- ⋆ $\mathbf{c} = \{c_0, \ldots, c_P\}$: deterministic PC coefficients; $\Psi_k(\boldsymbol{\xi})$: multi-d Legendre poly.

Formulation

- Parametric uncertainty (PU) expressed as (values extracted from literature):

$$\begin{aligned} \varepsilon(\xi_1) &= 0.1470 + 0.043\,\xi_1 \ \text{(kcal/mol)} \\ \sigma(\xi_2) &= 3.1506 + 0.021\,\xi_2 \ (\text{Å}) \qquad \text{where } \{\xi_i\}_{i=1}^{3} \sim \mathcal{U}(-1,1) \\ d(\xi_3) &= 0.1400 + 0.035\,\xi_3 \ (\text{Å}) \end{aligned}$$

- MD intrinsic noise (IN): single sample, $\{\varepsilon^j, \sigma^j, d^j\}$, yields multiple predictions of the target observable, G_1^j, G_2^j, ..., G_n^j.
- PU + IN ⇒ non-deterministic, noisy MD predictions of the water observables.
- Account for PU and IN using PCe:

$$\overline{G} \approx M(\xi_1, \xi_2, \xi_3) \equiv \sum_{k=0}^{P} c_k \Psi_k(\xi_1, \xi_2, \xi_3)$$

⋆ $\overline{G}$: quantity at steady-state averaged over m MD realizations (and time).

⋆ $\mathbf{c} = \{c_0, \ldots, c_P\}$: deterministic PC coefficients; $\Psi_k(\boldsymbol{\xi})$: multi-d Legendre poly.

Formulation

- Parametric uncertainty (PU) expressed as (values extracted from literature):

$$\begin{aligned} \varepsilon(\xi_1) &= 0.1470 + 0.043\,\xi_1 \;\; \text{(kcal/mol)} \\ \sigma(\xi_2) &= 3.1506 + 0.021\,\xi_2 \;\; (\text{Å}) \qquad \text{where } \{\xi_i\}_{i=1}^{3} \sim \mathcal{U}(-1,1) \\ d(\xi_3) &= 0.1400 + 0.035\,\xi_3 \;\; (\text{Å}) \end{aligned}$$

- MD intrinsic noise (IN): single sample, $\{\varepsilon^j, \sigma^j, d^j\}$, yields multiple predictions of the target observable, G^j_1, G^j_2, ..., G^j_n.
- PU + IN $\Rightarrow$ non-deterministic, noisy MD predictions of the water observables.
- Account for PU and IN using PCe:

$$\overline{G} \approx M(\xi_1, \xi_2, \xi_3) \equiv \sum_{k=0}^{P} c_k \Psi_k(\xi_1, \xi_2, \xi_3)$$

⋆ $\overline{G}$: quantity at steady-state averaged over m MD realizations (and time).

⋆ $\mathbf{c} = \{c_0, \ldots, c_P\}$: deterministic PC coefficients; $\Psi_k(\boldsymbol{\xi})$: multi-d Legendre poly.

NISP implementation

- "Non-intrusive spectral projection" exploits the orthogonality of the basis functions:

$$c_k = \frac{1}{\langle \Psi_k, \Psi_k \rangle} \int_{-1}^{+1} \int_{-1}^{+1} \int_{-1}^{+1} \overline{G}(\xi_1, \xi_2, \xi_3) \Psi_k(\xi_1, \xi_2, \xi_3) \frac{1}{8} \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}\xi_3, \quad k = 0, ..., P.$$

- Gauss-Legendre (GL) quadrature: regularity with respect to $\boldsymbol{\xi}$ is assumed.

1. Quadrature grid: fully tensorized 7-node grid.
2. Each quadrature point yields specific set of parameters:

$$\begin{aligned} \varepsilon(\xi_1^{(i)}) &= 0.1470 + 0.043\, \xi_1^{(i)}, \\ \sigma(\xi_2^{(i)}) &= 3.1506 + 0.021\, \xi_2^{(i)}, \\ d(\xi_3^{(i)}) &= 0.1400 + 0.035\, \xi_3^{(i)}, \end{aligned}$$

3. 4 realizations of the MD system at each node.
4. $\overline{G}$: mean of steady-state values of G: $\overline{G}(\xi^{(i)}) = \mathrm{mean}(G_1^{(i)}, G_2^{(i)}, G_3^{(i)}, G_4^{(i)})$

NISP implementation

- "Non-intrusive spectral projection" exploits the orthogonality of the basis functions:

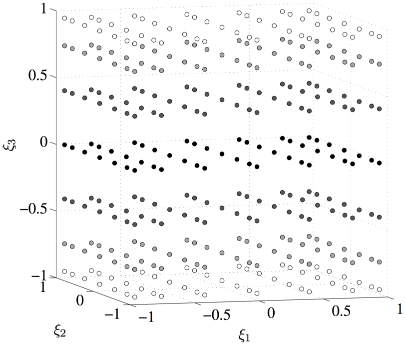
$$c_k = \frac{1}{\langle \Psi_k, \Psi_k \rangle} \int_{-1}^{+1} \int_{-1}^{+1} \int_{-1}^{+1} \overline{G}(\xi_1, \xi_2, \xi_3) \Psi_k(\xi_1, \xi_2, \xi_3) \frac{1}{8} \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}\xi_3, \quad k = 0, ..., P.$$

- Gauss-Legendre (GL) quadrature: regularity with respect to $\boldsymbol{\xi}$ is assumed.

1. Quadrature grid: fully tensorized 7-node grid.
2. Each quadrature point yields specific set of parameters:

$$\begin{aligned} \varepsilon(\xi_1^{(i)}) &= 0.1470 + 0.043\, \xi_1^{(i)}, \\ \sigma(\xi_2^{(i)}) &= 3.1506 + 0.021\, \xi_2^{(i)}, \\ d(\xi_3^{(i)}) &= 0.1400 + 0.035\, \xi_3^{(i)}, \end{aligned}$$

3. 4 realizations of the MD system at each node.
4. $\overline{G}$: mean of steady-state values of G: $\overline{G}(\boldsymbol{\xi}^{(i)}) = \mathrm{mean}(G_1^{(i)}, G_2^{(i)}, G_3^{(i)}, G_4^{(i)})$



NISP implementation

- "Non-intrusive spectral projection" exploits the orthogonality of the basis functions:

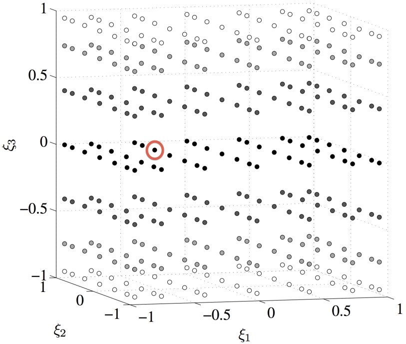
$$c_k = \frac{1}{\langle \Psi_k, \Psi_k \rangle} \int_{-1}^{+1} \int_{-1}^{+1} \int_{-1}^{+1} \overline{G}(\xi_1, \xi_2, \xi_3) \Psi_k(\xi_1, \xi_2, \xi_3) \frac{1}{8} \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}\xi_3, \quad k = 0, \ldots, P.$$

- Gauss-Legendre (GL) quadrature: regularity with respect to $\boldsymbol{\xi}$ is assumed.

1. Quadrature grid: fully tensorized 7-node grid.
2. Each quadrature point yields specific set of parameters:

$$\begin{aligned} \varepsilon(\xi_1^{(i)}) &= 0.1470 + 0.043\, \xi_1^{(i)}, \\ \sigma(\xi_2^{(i)}) &= 3.1506 + 0.021\, \xi_2^{(i)}, \\ d(\xi_3^{(i)}) &= 0.1400 + 0.035\, \xi_3^{(i)}, \end{aligned}$$

3. 4 realizations of the MD system at each node.
4. $\overline{G}$: mean of steady-state values of G: $\overline{G}(\boldsymbol{\xi}^{(i)}) = \mathrm{mean}(G_1^{(i)}, G_2^{(i)}, G_3^{(i)}, G_4^{(i)})$



NISP implementation

- "Non-intrusive spectral projection" exploits the orthogonality of the basis functions:

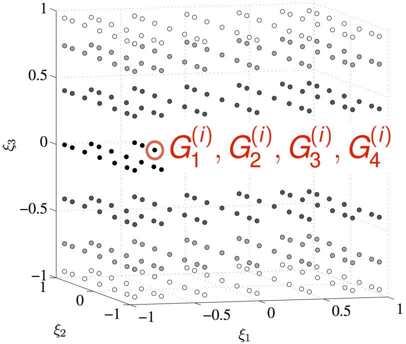
$$c_k = \frac{1}{\langle \Psi_k, \Psi_k \rangle} \int_{-1}^{+1} \int_{-1}^{+1} \int_{-1}^{+1} \overline{G}(\xi_1, \xi_2, \xi_3) \Psi_k(\xi_1, \xi_2, \xi_3) \frac{1}{8} d\xi_1 d\xi_2 d\xi_3, \quad k = 0, ..., P.$$

- Gauss-Legendre (GL) quadrature: regularity with respect to $\boldsymbol{\xi}$ is assumed.

1. Quadrature grid: fully tensorized 7-node grid.
2. Each quadrature point yields specific set of parameters:

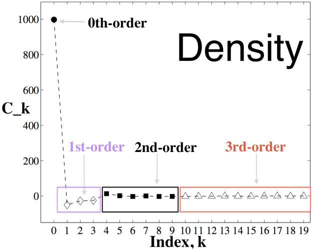
$$\varepsilon(\xi_1^{(i)}) = 0.1470 + 0.043\, \xi_1^{(i)},$$
$$\sigma(\xi_2^{(i)}) = 3.1506 + 0.021\, \xi_2^{(i)},$$
$$d(\xi_3^{(i)}) = 0.1400 + 0.035\, \xi_3^{(i)},$$

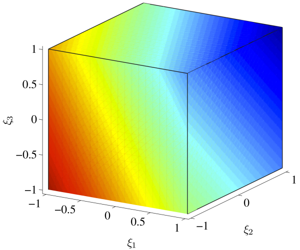
3. 4 realizations of the MD system at each node.
4. $\overline{G}$: mean of steady-state values of G: $\overline{G}(\boldsymbol{\xi}^{(i)}) = \text{mean}(G_1^{(i)}, G_2^{(i)}, G_3^{(i)}, G_4^{(i)})$



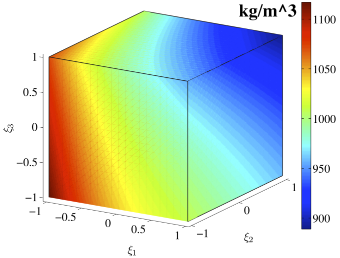
Response surface

- Recall: $\varepsilon(\xi_1), \sigma(\xi_2), d(\xi_3)$
- Result: $\overline{G} \approx \sum_{k=0}^{P} c_k \Psi_k(\xi_1, \xi_2, \xi_3)$
- Rapidly decaying PC spectrum.
- **Linear** and **cubic** response surface.
- Potential uncertainty has large impact.





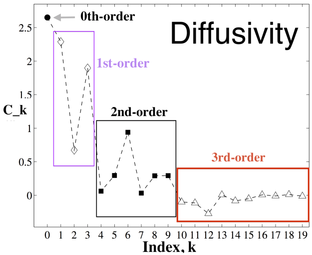
Linear

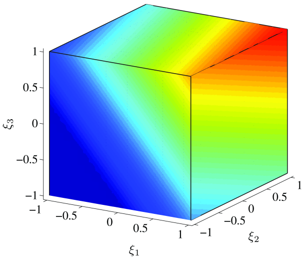


Cubic

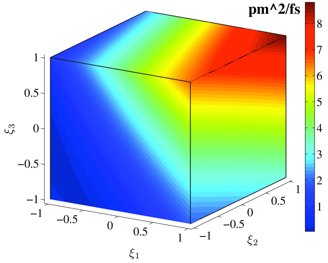
Response surface

- Recall: $\varepsilon(\xi_1)$, $\sigma(\xi_2)$, $d(\xi_3)$
- Result: $\overline{G} \approx \sum_{k=0}^{P} c_k \Psi_k(\xi_1, \xi_2, \xi_3)$
- Rapidly decaying PC spectrum.
- **Linear** and **cubic** response surface.
- Potential uncertainty has large impact.





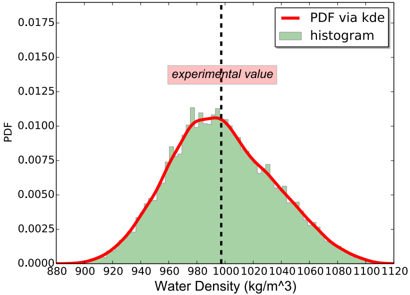
Linear



Cubic

PDF of Water Density

- Recall: $\varepsilon(\xi_1)$, $\sigma(\xi_2)$, $d(\xi_3)$
- $\overline{G} \approx \sum_{k=0}^{P} c_k \Psi_k(\xi_1, \xi_2, \xi_3)$
- Reconstruct the PDF of the observable.



(ε, σ, d)

Conclusions

- UQ is important for simulations where high-fidelity and risk assessment are key.
- It is being used in science and engineering increasingly more.
- Two main parts: forward and inverse problem.
- Bayesian inference provides a suitable setting for inverse problems since it accounts for all the noise present in the data.
- Parameters' correlation can be a key information, but it is often neglected.
- The push towards exascale is turning the paradigm of how we do simulations:
  - *from* single, deterministic runs,
  - *to* stochastic frameworks and ensembles of runs.
- Very active field of research.
- Exciting future to see how UQ will be applied to multi-physics problems.

References

- **F. Rizzi**, R. Jones, B. Debusschere and O. Knio
  "Uncertainty Quantification in MD simulations of concentration driven ionic flow through a silica nanopore. Part I: sensitivity to physical parameters of the pore.",
  J. Chem. Phys., 138:194104, 2013.

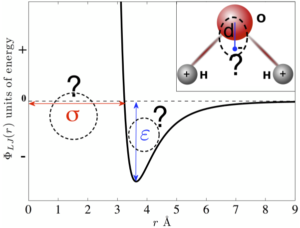
- **F. Rizzi**, R. Jones, B. Debusschere and O. Knio
  "Uncertainty Quantification in MD simulations of concentration driven ionic flow through a silica nanopore. Part II: uncertain potential parameters.",
  J. Chem. Phys., 138:194105, 2013.

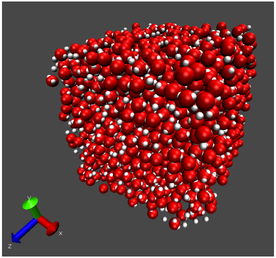
- **F. Rizzi**, H. Najm, B. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson and O. Knio
  "Uncertainty quantification in MD simulations of Water. Part I: Forward propagation", *SIAM MMS*, 10(4), 1428-1459, 2012.

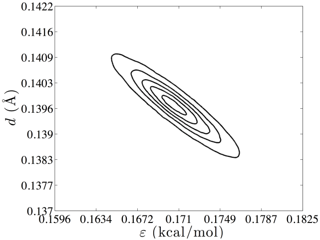
- **F. Rizzi**, H. Najm, B. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson and O. Knio
  "Uncertainty quantification in MD simulations of Water. Part II: Bayesian Inference of force-field parameters", *SIAM MMS*, 10(4), 1460-1492, 2012.

Acknowledgments

Thank you for your attention!







Research supported by:
US, DOE - Office of Advanced Scientific Computing Research

Questions?

- How to approach inference in high-dimensional parameter spaces?
- How about when the forward problem is very expensive?
- What if you don't need UQ?
- ...