Uncertainty Quantification for Xolotl

Sophie Blondel[1], Crystal Jernigan[1], David Higdon[2], Omar Knio[3], Habib Najm[4], Khachik Sargsyan[4], Jay Jay Billings[1]

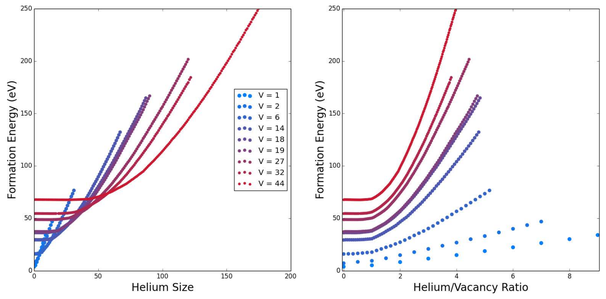
[1]Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[2]Statistical Sciences Group, Los Alamos National Laboratory, Los Alamos, NM, USA
[3]Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC, USA
[4]Sandia National Laboratories, Livermore, CA, USA

Formation Energies

Xolotl takes binding energies as an input to compute the dissociation rates for the clusters present in the reaction network. They are calculated from formation energies with the formula

$$E_A(A_iB_j) = E_f(A_{i-1}B_j) + E_f(A_1) - E_f(A_iB_j)$$

where $E_A(A_iB_j)$ is the "A" (He, V, or I) binding energy of the A_iB_j cluster, and E_f are formation energies.



Formation energy data is obtained from MD simulations and uncertainty quantification methods are used to model this data and propagate its uncertainties through Xolotl.

Polynomial Chaos Expansion

PCE is a method used to represent a random variable (RV), mapping it from standard RV in terms of orthogonal polynomials.

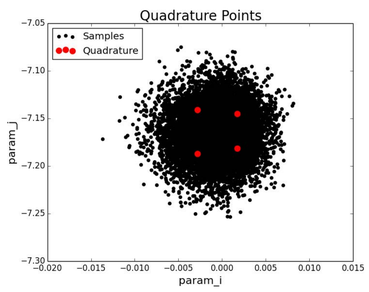
$$u \sim \sum_{k=0}^{P} u_k \varphi_k(\xi)$$

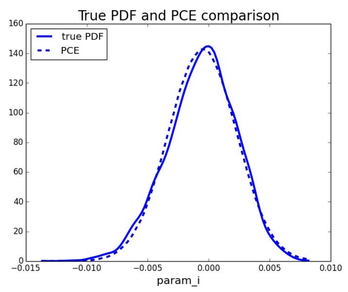
where u represents the RV, P the order of the expansion, f_k the PC coefficients, φ_k the basis functions, and ξ the standard RV.

In our case, the PCEs are built on the quantities of interest that are the model coefficients, $f_{i,j}$.

Using UQ Toolkit, u_k is computed projecting u onto the space spanned by the basis evaluated at quadrature points.

Here is a quadrature example and an example comparing the expansion to the original PDF obtained with bayesian inference.



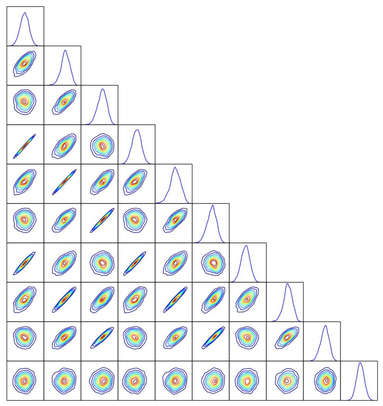


Bayesian Inference

Bayesian Inference Principle:

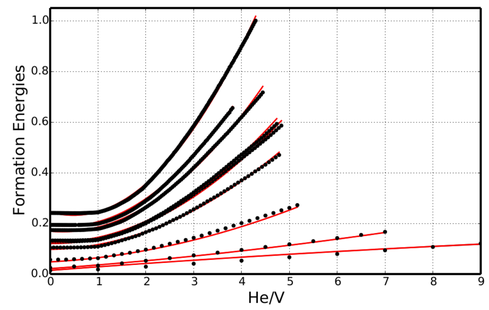
$$P(H|E) \propto P(E|H) \cdot P(H)$$

- the **posterior** $P(H|E)$ (probability of the hypothesis H given the evidence E) that is infered
- the **likelihood** $P(E|H)$ (probability of the evidence E given the hypothesis H)
- and the **prior** $P(H)$ that gathers all the information one had before the evidence E was observed



Joint densities of all the inferred parameters.

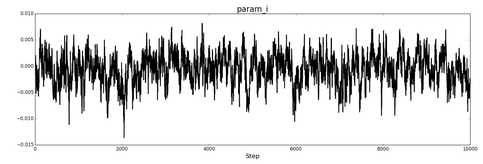
Use Markov Chain Monte Carlo (MCMC) method to estimate the posterior distribution with the UQ Toolkit software (`https://www.sandia.gov/UQToolkit/`).



Model plugged in UQ Toolkit:

$$f(x,y) = \sum_{i=0}^{2}\sum_{j=0}^{2} f_{i,j} P_i(x) \times P_j(y) + \sigma$$

where P_i is the Legendre polynomial of order i.
$f_{i,j}$ and σ are the inferred parameters.



Global Sensitivity Analysis

GSA is used to help reduce model dimensionality.
For Xolotl, it means determining which input parameters really influence the output results, and setting the ones that don't affect the results to constant values.

First order sensitivity indices are defined by

$$S_i = \frac{Var(f_i)}{Var(f)}$$

where f_i is the conditional expectation, and describes the effect on the model output that results from varying the i-th input parameter.
Thus, S_i characterize the fraction of the variance due to the i-th input parameter only.

One way to perform GSA is to represent the output with PCE, and the first order indices come as a direct result

$$S_i = \frac{\sum_{k \in \mathcal{I}_i} u_k^2 ||\varphi_k||^2}{Var(f)}$$

where $\mathcal{I}_i$ are the terms involving ξ_i only.

$S(f_{i,j})$ for $C(He)$:

0.0004	0.4616	0.2123
0.0117	0.2937	0.0053
0.0064	0.0052	0.0033

Forward UQ

The uncertainties on the input parameters can be easily propagated through Xolotl thanks to the PCE representation, because the use of quadrature rules limit the number of output evaluations needed.

For instance, define the ouput to be the concentration of helium clusters in the material with no incoming flux, after a given time. Then using forward propagation one obtains

$$C(He) = 2.2188 \cdot 10^{-10} \pm 0.0045 \cdot 10^{-10}$$

Next Steps

- Utilizing a higher order model for the formation energies to obtain a better fit.
- Studying the affects of other uncertain input parameters: diffusion factor, migration energy, temperature, ...
- Performing a Monte Carlo sampling based Global Sensitivity Analysis on multiple uncertain input parameters in order to reduce the overall dimensionality of the problem.