

Uncertainty Quantification in ion-solid interaction simulations

R. Preuss, U. von Toussaint
Max-Planck-Institute for Plasma Physics, Boltzmannstr. 2,
85748 Garching, Germany

September 6, 2017

Abstract

Within the framework of Bayesian uncertainty quantification we propose a non-intrusive reduced-order spectral approach (polynomial chaos expansion) to the simulation of ion-solid interactions. The method not only reduces the number of function evaluations but provides simultaneously a quantitative measure for which combinations of inputs have the most important impact on the result. It is applied to SDTRIM-simulations [1] with several uncertain and Gaussian distributed input parameters (i.e. angle, projectile energy, surface binding energy, target composition) and the results are compared to full-grid based approaches and sampling based methods with respect to reliability, efficiency and scalability.

Keywords: Polynomial chaos, spectral expansion, discrete projection, uncertainty quantification, ion-solid interaction, SDTRIM

PACS: 02.50.-r, 52.65.-y

1 Introduction

Almost all computer codes for the simulation of ion-solid interaction [2] rely on a large number of input parameters, e.g. surface binding energies, composition, energy distribution etc. However, many of these parameters are uncertain and a proper comparison with experimental data or other models requires the quantification of the uncertainty of the result. Unfortunately, the computational demand of single simulation runs often severely restricts the quantification of output uncertainties by full-grid or simple sampling (e.g.

Monte Carlo sampling) based approaches due to the curse of dimensionality for more than a very limited number of uncertain input parameters. To reduce the computational effort we propose a non-intrusive reduced-order model approach (polynomial chaos expansion), which not only reduces the number of function evaluations but provides simultaneously a quantitative measure of which combinations of inputs have the most important impact on the result, i.e. it yields a sensitivity analysis and the associated Sobol coefficients.

2 Bayesian uncertainty quantification

Based on the Bayesian framework we employ a spectral expansion to quantify the propagation of uncertainty through the model. First introduced by Wiener [3] in the context of Hermite basis functions it was termed ‘polynomial chaos expansion’ at his time. Nowadays the notion of ‘chaos’ has shifted and the use of the term ‘spectral expansion’ is more appropriate. Once successfully achieved, the spectral representation is capable of quantifying the uncertainty for any point in model space or to serve as a surrogate model.

Since we calculate the sought-for spectral coefficients from a discrete set of collocation points in the space of the random variable, our approach is non-intrusive, but approximate. The emerging integrals in the calculation of the coefficients are evaluated by Gaussian quadrature which identifies the collocation points with those of the quadrature. Moreover, we assume mutually independent normally distributed random variables. The adjunctive set of orthonormal basis functions in such a case are Hermite polynomials.

To quantify the uncertainty of a result R we seek the appropriate function $g(\xi)$, such that R will have the required distribution of the model response, $R = g(\xi)$. As for all random variables with finite variance it is possible to find an infinite expansion

$$g(\xi) = \sum_{k=0}^{\infty} a_k \psi_k(\xi) \approx \sum_{k=0}^P a_k \psi_k(\xi) \quad , \quad (1)$$

which we limit to polynomial order P since the contributions of higher orders become numerically insignificant. The coefficients are given by

$$a_k = \frac{\langle g(\xi), \psi_k(\xi) \rangle}{\langle \psi_k(\xi), \psi_k(\xi) \rangle} \quad , \quad \text{with} \quad \langle g(\xi), \psi(\xi) \rangle = \int g(\xi) \psi(\xi) p(\xi) d\xi \quad . \quad (2)$$

We assume Gaussian character for the random variable, so the density $p(\xi)$

is distributed according to the normal (probability) distribution

$$p(\xi) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{\xi^2}{2} \right\} . \quad (3)$$

The adjunctive set of orthonormal basis functions is given by the so-called *probabilist* Hermite functions, which read up to fourth order

$$\begin{aligned} \psi_0(\xi) &= 1 \quad , \\ \psi_1(\xi) &= \xi \quad , \\ \psi_2(\xi) &= \xi^2 - 1 \quad , \\ \psi_3(\xi) &= \xi^3 - 3\xi \quad , \\ \psi_4(\xi) &= \xi^4 - 6\xi^2 + 3 \quad . \end{aligned} \quad (4)$$

It turns out that for the model simulations under consideration this polynomial order is sufficient since contributions from higher orders become numerically insignificant for the result. With these definitions the normalization constants in Eq. (2) are readily

$$\langle \psi_k, \psi_k \rangle = \int \psi_k(\xi) \psi_k(\xi) p(\xi) d\xi = k! \quad . \quad (5)$$

Due to the Gaussian nature of the probability function omnipresent in the integrals above, it is beneficial to use Gauss-Hermite quadrature for the evaluation

$$\langle g(\xi), \psi(\xi) \rangle \stackrel{\text{G.H.}}{=} \sum_{l=0}^L g(\xi_l) \psi(\xi_l) w_l \quad , \quad (6)$$

where the weights w_l and the abscissas ξ_l are for instance provided by Numerical Recipes [4]. Eventually, by exploiting the properties of the orthogonal Hermite polynomials the expectation value of the model outcome and its variance can be assigned to the spectral coefficients in Eq. (2)

$$\langle R \rangle = a_0 \quad , \quad \text{var}(R) = \langle R^2 \rangle - \langle R \rangle^2 = \sum_{k=1}^P a_k^2 k! \quad . \quad (7)$$

In order to provide a measure for the influence of the uncertainty of input variables on the above variance we employ Sobol coefficients [5]. They are defined by

$$S_i = \frac{D_i}{\text{var}(R)} \quad , \quad S_{ij} = \frac{D_{ij}}{\text{var}(R)} \quad , \quad \dots \quad , \quad (8)$$

where the evaluation of the integrals

$$\begin{aligned}
D_i &= \int g_i^2(\xi_i) d\xi_i \quad , \\
D_{ij} &= \int \int g_{ij}^2(\xi_i, \xi_j) d\xi_i d\xi_j \quad , \\
&\dots \quad ,
\end{aligned} \tag{9}$$

results in combinations of the coefficients of Eq. (2) (the index of the function $g_{index}(\dots)$ relates to the specific variable(s) ξ_{index} which are omitted in the integral $g_{index} = \int \dots \int g(\boldsymbol{\xi}) d\xi_{\{index\}}$). The higher the value of a Sobol coefficient with respect to the others is, the more it is advantageous to reduce the uncertainty of its associated variable in order to reduce the uncertainty of the quantity of interest.

For further readings about our implementation of the uncertainty quantification we refer to [6].

3 Results and discussion

The above analysis is applied to ion-solid interaction simulations with the program SDTRIM.SP [1] for the case of incident deuterium ions with an energy of $E_0 = 200$ eV at $\alpha = 45$ degrees to a surface consisting of iron and tungsten with a surface binding energy of $E_{SB} = 4.28$ eV. We assume the parameters to be normally distributed within a standard deviation of roughly 10%, i.e. $\sigma_{E_0} = 20$ eV, $\sigma_{E_{SB}} = 0.4$ and $\sigma_\alpha = 4$ eV. Extending the formulas of section 2 to three random variables $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$ with $\hat{E}_0 = E_0 + \xi_1 \sigma_{E_0}$, $\hat{E}_{SB} = E_{SB} + \xi_2 \sigma_{E_{SB}}$ and $\hat{\alpha} = \alpha + \xi_3 \sigma_\alpha$, the summation of the terms in Eq. (6) runs over three indices l_1, l_2 and l_3 with an upper boundary of $P + 1 = 5$ in the present setup of fourth order polynomials (for numerical accuracy of the Gaussian quadrature it is expedient to be one order higher than the polynomial order of the spectral expansion). This results in a total of 216 terms (three nested summations, each running from $l_i=0$ to 5 with $i=1,2,3$) over the collocation points composed of 6 Gaussian quadrature abscissas assigned to ξ_{l_i} and 6 weights w_{l_i} (the specific values are obtained from numerical routines in [4]):

$$\langle g(\boldsymbol{\xi}), \Psi_k(\boldsymbol{\xi}) \rangle = \sum_{l_1=0}^5 \sum_{l_2=0}^5 \sum_{l_3=0}^5 g(\xi_{l_1}, \xi_{l_2}, \xi_{l_3}) \Psi_k(\xi_{l_1}, \xi_{l_2}, \xi_{l_3}) w_{l_1} w_{l_2} w_{l_3} \quad , \tag{10}$$

while the $\Psi_k(\boldsymbol{\xi})$ consist of all possible 35 combinations of the Hermite polynomials of Eq. (4) up to fourth order (e.g. $\Psi_0 = 1, \Psi_1 = \xi_{l_1}, \Psi_2 = \xi_{l_2}, \Psi_3 =$

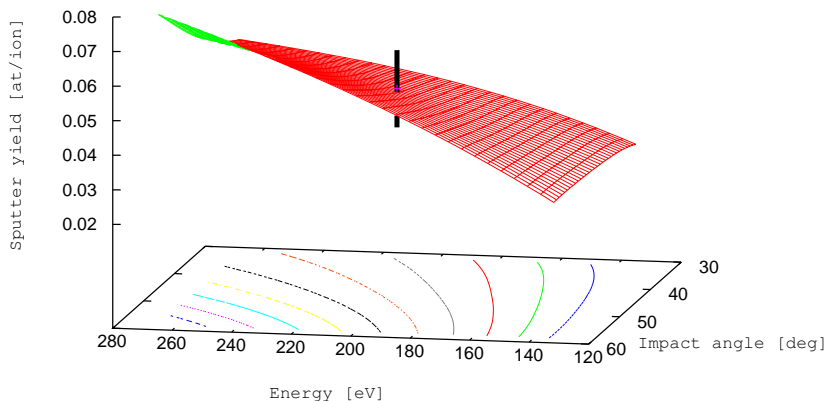


Figure 1: Sputter yield reproduced by the surrogate model from the uncertainty quantification of SDTRIM calculations for $E_0 = 200$ eV, $E_{SB} = 4.28$ eV and an incident angle of 45 degrees, with contour lines in the base. The line in the center shows a yield of 0.0497 for these settings with a standard deviation of 0.0112 (dark line). For reason of portrayal the surrogate model was varied only in two dimensions for E_0 and α , while E_{SB} was kept fixed at 4.28 eV.

$\xi_{l_3}, \Psi_4 = \xi_{l_1} \xi_{l_2}, \Psi_5 = \xi_{l_1} \xi_{l_3}, \Psi_6 = \xi_{l_2} \xi_{l_3}, \Psi_7 = \xi_{l_1}^2 - 1, \dots, \Psi_{34} = \xi_{l_3}^4 - 6\xi_{l_3}^2 + 3$). The value for the function $g(\xi_{l_1}, \xi_{l_2}, \xi_{l_3})$ is obtained from a SDTRIM.SP run, which takes roughly 3 minutes on a modern CPU. However, the complete run for the 216 terms can be speeded up enormously since the calculations are independent and can be done in parallel. Once calculated, the 35 coefficients of Eq. (10) establish a fast surrogate model, which is simply the evaluation of a polynomial. Without the need to do any further simulations, various quantities may be inferred from the coefficients, e.g. the variance as in Eq. (2), or the Sobol coefficients, which allow to investigate the sensitivity of the result on the uncertainty of the input variables. For the above variables E_0 , α and E_{SB} we get a relationship of 20:10:70 in the Sobol coefficients (only first order is numerically significant) for $\{E_0, \alpha, E_{SB}\}$, indicating that the improvement of the knowledge of E_{SB} is most rewarding if one wants to reduce the uncertainty of the sputter yield.

4 Acknowledgement

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

- [1] W. Möller, W. Eckstein, J. Biersack, Tridyn - binary collision simulation of atomic collisions and dynamic composition changes in solids, *Comput. Phys. Commun.* 51 (3) (1988) 355–368.
- [2] W. Eckstein, *Computer Simulation of Ion-Solid Interactions*, Vol. 10 of Springer Series in Materials Science, Springer-Verlag, Berlin, Heidelberg, New York, 1991.
- [3] N. Wiener, *Am. J. Math.* 60 (1938) 897.
- [4] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd Edition, Cambridge University Press, 2007.
- [5] R. C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM, Philadelphia, 2014.
- [6] R. Preuss, U. von Toussaint, Investigations on bayesian uncertainty quantification with two examples, in: A. Giffin, K. Knuth (Eds.), *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Vol. 1757, AIP Publishing, Melville, NY, 2016.