Investigations on Bayesian uncertainty quantification with two examples

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

April 20, 2017

Abstract

Input quantities for the numerical simulation of fusion plasmas involve field quantities which are hampered by noise. In order to compare data from experiment to model results, or to have an estimation of the fluctuation margin of a model prediction, a quantification of the uncertainties is necessary. Within a discrete projection framework we employ a spectral expansion to represent the random process responsible for the noise. Since Gaussian distributed noise is assumed Hermite polynomials are chosen for the orthonormal basis system. The coefficients are calculated from collocation points defined by Gaussian quadrature in a non-intrusive approach. An instructive example of absorption in media serves for the validation of the procedure. Finally the method is applied to the Vlasov-Poisson model describing electrostatic plasmas, which will be influenced by an uncertain external field.

Keywords: Polynomial chaos, spectral expansion, discrete projection, uncertainty quantification, Vlasov-Poisson model
PACS: 02.50.-r, 52.65.-y

1 Introduction

Fusion plasmas contain a macroscopic number of particles (about 10^{21}) with temperatures in the order of several million Kelvin in the plasma core. Together with electromagnetic fields exposed from the outside and induced by

the plasma itself the emerging processes show highly complex non linear behavior. Both the experimental implementation of parameter settings and measured data from diagnostics suffer from noise and lead to input quantities and data bases with uncertainties. If calculations are based on such input there is interest in quantifying the effect of the uncertainties on the model response.

Comprising the ongoing effects on the theoretical side plasma models are sets of coupled differential or partial differential equations. In most cases no analytical solution is possible and one depends on numerical approaches which end up with approximate solutions to problems for which we do not know the exact answer. Depending on the level for the numerical description of the plasma the model results suffer from simplifications, limitations to certain regimes, or the insufficient number of particles when trying to simulate a macroscopic state from (microscopic) kinetic equations of motion. Again this raises the question for the confidence of the computational predictions, especially when switching on/off model components to study plasma phenomena.

In view of the addressed peculiarities above it comes with no surprise that classical error propagation misses this task entirely when time dependence enters the stage. Moreover, when model results shall be compared with experimental data we have to deal with different sources of uncertainties either of epistemic (lack of knowledge, basically no variability) or aleatoric (noise, basically exact) origin. While being strictly obedient to orthodox statistics it is not possible to cope with the latter [Jaynes(2003)]. No such problems arise in Bayesian probability theory which deals with uncertainties of different origin or character in assigning a probability density function to all kinds of uncertain quantities [Najm(2009)].

Based on the Bayesian framework we employ a spectral expansion to quantify the propagation of uncertainty through the model. First introduced by Wiener in the context of Hermite basis functions [Wiener(1938)] it was termed 'polynomial chaos expansion' at his time. Nowadays the notion of 'chaos' has shifted and in our context the use of the term 'spectral expansion' is more appropriate. The spectral approach seems expedient as for most cases high-dimensional parameter spaces are sufficiently smooth to facilitate the calculation of the coefficients needed to determine the moments for the quantity of interest. It is embedded in the *Galerkin* framework, which projects weighted residuals onto a finite-dimensional space spanned by appropriate basis functions [Smith(2014)]. Once successfully achieved, the spectral representation is capable of quantifying the uncertainty for any time in model space. However, this procedure requires to solve the (differential) model equations for all of the spectral coefficients, which usually necessitates

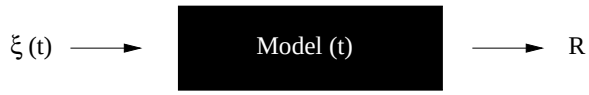

Figure 1: Random variable ξ is input to a time-dependent model with output R.

to change the model solving part of a numerical program completely and therefore is termed intrusive. Non-intrusive, but approximate is *stochastic collocation* which calculates the sought-for coefficients from a discrete set of collocation points in the space of the random variable. For the matters of this paper we assume mutually independent normally distributed random variables for which Hermite polynomials are the adjunctive set of orthonormal basis functions. Additionally, we assume the model not to change the functional properties of the probability distribution of the input quantity. If we approximate the emerging integrals for the coefficients by Gaussian quadrature we end up by *discrete projection* which identifies the collocation points with those of the quadrature.

Fig. 1 depicts the problem we want to tackle in a simplified scheme. A model input parameter is hampered by generally time-dependent noise described by random variable ξ. The model either propagates over time or responses directly with result R which uncertainty distribution we would like to know. Since our approach is non-intrusive we are forced to think of the model as a black box. In the following we present our implementation for two examples. The first is used to validate the procedure with an analytically fully accessible model of absorption in media. The second example sets up a reference case for tackling plasma physics models with the non-trivial Vlasov-Poisson system as an example for a nonlinear partial differential model equation. Altogether this paper presents a rough sketch of the analysis for uncertainty quantification only. An elaborated view may be found in, e.g. [Smith(2014), Maître and Knio(2010)].

2 Sampling approach

In order to have a calibration standard to compare with the results of the discrete projection procedure we employ random sampling of the model response. For each realization of the random variable $\{\xi_1, ..., \xi_N\}$ there exists a model response $R_i = R(\xi_i)$ constituting the sample solution set $\{R_1, ..., R_N\}$

from which moments can be computed. The expected mean is

$$\langle R \rangle = \frac{1}{N} \sum_{i=1}^{N} R(\xi_i) \quad , \tag{1}$$

and its variance reads

$$\mathrm{var}(R) = \langle R^2 \rangle - \langle R \rangle^2 \quad . \tag{2}$$

Even the full uncertainty distribution may be established with help of a histogram if the sample solution set is sufficiently large ($N \gtrsim 1000$). Though this procedure is straightforward and automatically contains the full model answer with all correlations, it has the vital drawback of a comparatively low convergence rate. If the computation time of a single model output is not in the order of seconds or becomes more sophisticated with a higher number of variables (curse of dimension), the mere accumulation of sample point densities to infer the complete distribution is futile. In this respect more promising are spectral approaches which will be discussed next.

3 Uncertain initial value problem

Recall the situation of Fig. 1: a single parameter hampered by Gaussian noise represented by random variable ξ is input to a model. Setting the parameter constant throughout model propagation defines the scalar initial value problem. The model (i.e. the (system of) differential equations) is not *harmful* to the character of the random variable which shall mean that the functional properties of the stochastic process are kept the same. So the result will itself be hampered by noise of Gaussian character. To quantify the uncertainty of the result we seek the appropriate function $g(\xi)$, such that R will have the required distribution of the model response

$$R = g(\xi) \quad . \tag{3}$$

As for all random variables with finite variance it is possible to find an infinite expansion

$$g(\xi) = \sum_{k=0}^{\infty} a_k \psi_k(\xi) \approx \sum_{k=0}^{P} a_k \psi_k(\xi) \quad , \tag{4}$$

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

$$a_k = \frac{\langle g(\xi), \psi_k(\xi) \rangle}{\langle \psi_k(\xi), \psi_k(\xi) \rangle} \quad , \qquad \text{with} \quad \langle g(\xi), \psi(\xi) \rangle = \int g(\xi) \psi(\xi) p(\xi) \mathrm{d}\xi \quad . \tag{5}$$

We assume Gaussian character for the random variable, so the density $p(\xi)$ is distributed according to the normal (probability) distribution

$$p(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{\xi^2}{2}\right\} \quad . \tag{6}$$

The adjunctive set of orthonormal basis functions for the distribution Eq. (6) is given by the so-called *probabilist* Hermite functions which we employ in this paper up to forth order

$$\begin{aligned} \psi_0(\xi) = 1 \quad , \quad \psi_1(\xi) = \xi \quad , \quad \psi_2(\xi) = \xi^2 - 1 \quad , \\ \psi_3(\xi) = \xi^3 - 3\xi \quad , \quad \psi_4(\xi) = \xi^4 - 6\xi^2 + 3 \quad . \end{aligned} \tag{7}$$

The normalization constants are readily

$$\langle \psi_k, \psi_k \rangle \;=\; \int \psi_k(\xi)\psi_k(\xi)p(\xi)\mathrm{d}\xi \;=\; k! \quad . \tag{8}$$

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

$$\langle g(\xi), \psi(\xi) \rangle \overset{\text{G.H.}}{=} \sum_{l=1}^{L} g(\xi_l)\psi(\xi_l)w_l \quad , \tag{9}$$

where the weights w_l and the abscissas ξ_l are for instance provided by Numerical Recipes [Press et al.(2007)].

Eventually, by exploiting the properties of the orthogonal Hermite polynomials the expectation value for the model outcome is

$$\langle R \rangle = a_0 \quad , \tag{10}$$

and its variance

$$\mathrm{var}(R) \;=\; \langle R^2 \rangle - \langle R \rangle^2 \;=\; \sum_{k=1}^{P} a_k^2 \langle \psi_k, \psi_k \rangle \;=\; \sum_{k=1}^{P} a_k^2 k! \quad . \tag{11}$$

To avoid numerical errors the order L of the Hermite polynomials used for the calculation of the coefficients in Eq. (5), i.e. Gauss-Hermite quadrature of Eq. (9), should be at least one higher than the order P of the Hermite polynomials used for the calculation of the model response in Eq. (4), L=P+1.

4 Parameter uncertain throughout model propagation

Above we examined the case of a model starting with an uncertain parameter which does not vary throughout time propagation. In this section we consider the model parameter uncertain at every of N time steps thereby establishing a random field of N variables. The mutual independence of the field variables implies that we can compile the model response from a tensor product of the univariate polynomials of the previous section. However, already for a decent number of time steps the dimension of the random field gets rapidly too large to be treatable and has to be reduced. Since the time step has to be chosen small enough to cover the model evolution correctly a correlation of the model responses within a few time slices is to be expected. The impact of a random variable at a certain time slice will persist in the model response for some more time slices, but then gets blurred by the influences of more and more random variables on the consecutive time slices. We mimic this circumstance by introducing a correlation length λ via a covariance matrix

$$C_{ij} = \exp\left\{-\frac{1}{2}\frac{(t_i - t_j)^2}{\lambda^2}\right\} \quad , \tag{12}$$

and let a Karhunen-Loève expansion reduce the N random representations of R to $M \ll N$ informative ones.

After performing a singular value decomposition $\boldsymbol{C} = \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^T$ the orthogonal matrix $\boldsymbol{U}$ consists of N eigenvectors $\boldsymbol{u}_m$ with corresponding eigenvalues Δ_m (see Fig. 2). Only those M eigenvectors are chosen which correspond to eigenvalues above a certain level (e.g. 5% of maximum eigenvalue). They constitute the new random variable ξ^{KL} in a Karhunen-Loève expansion:

$$\xi^{KL} = \sum_{m=1}^{M} \xi_m \sqrt{\Delta_m}\boldsymbol{u}_m \quad . \tag{13}$$

While we expect the model responses to be correlated the random variables of adjacent time slices are not. Therefore the assumption of mutual independence is still valid and we can construct the multidimensional basis functions as a tensor product of univariate Hermite polynomials. Analogously to Eq. 5 the coefficients for the reduced M-dimensional random field read

$$a_k = \frac{1}{k!}\int ... \int \mathrm{d}\xi_1 ... \mathrm{d}\xi_M g(\xi_1, ..., \xi_M)\psi_k(\xi_1, ..., \xi_M)p(\xi_1, ..., \xi_M) \quad . \tag{14}$$

Gauss-Hermite quadrature is employed again

$$a_k = \frac{1}{k!}\sum_{l_1=1}^{L} ... \sum_{l_M=1}^{L} g(\xi_{l_1}, ..., \xi_{l_M})\psi_k(\xi_{l_1}, ..., \xi_{l_M})w_{l_1} \cdot ... \cdot w_{l_M} \quad , \tag{15}$$

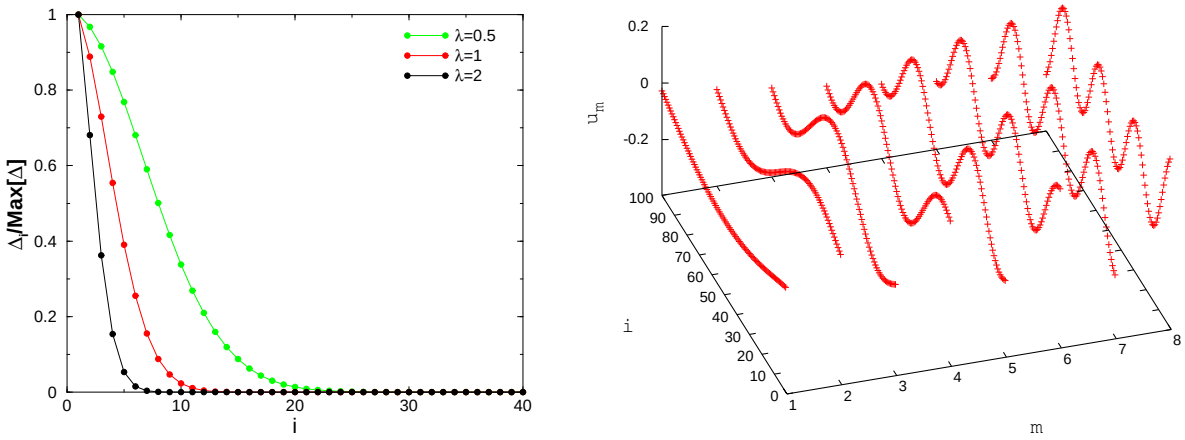

Figure 2: Left: eigenvalues of the covariance matrix for different settings of the correlation length λ (lines are guide to the eye). Right: corresponding eigenvectors with eigenvalues above 5% of maximum eigenvalue for the case of λ=1.

and inserted in Eq. (10) and Eq. (11) to describe the model response.

The number of terms N_{int} entering in Eq. (15) increases extensively with the number of eigenvalues and polynomial order:

$$N_{\text{int}} = \frac{(M+P)!}{M!P!} \quad . \tag{16}$$

Note, that for an efficient computing of Eq. (15) one has to deal with two independent index fields: one for the model response $g(\xi_{j_1}, ..., \xi_{j_M})$ and one for the Gauss-Hermite quadrature. However, these index fields increase to an enormous extent with the number of variables and polynomial order, so that in the end one is left behind solely with parallelization of the computer runs to narrow running time.

5 Absorption in media

For introductory reasons and to validate our implementation we start with a theoretically fully accessible problem. The ordinary differential equation for the absorption of light with intensity I_0 in media reads

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \alpha I \tag{17}$$

or in discrete form

$$I_{i+1} - I_i = \alpha I_i \Delta t \quad . \tag{18}$$

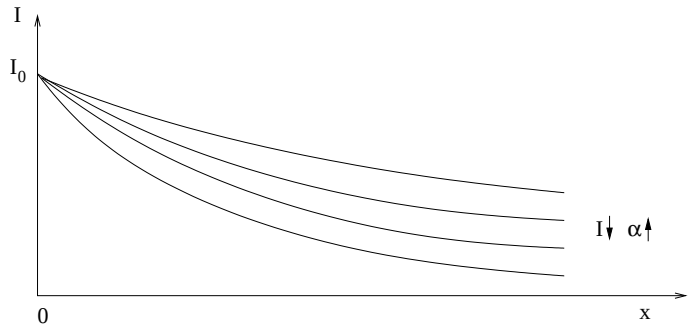

Figure 3: Absorption: light of intensity I_0 gets absorbed with coefficient α by passing through media. The higher the absorption coefficient, the lower the intensity of the light at a certain depth x. If the coefficient shows an uncertainty distribution the intensity will be distributed in the same way.

Let us think of an uncertain absorption coefficient $\alpha = \alpha_0 + \xi\sigma_\alpha$ with $\xi \sim \mathcal{N}(0,1)$ (see Fig. 3). The analytic random solution is

$$I(x,\xi) = I_0 e^{-\alpha x} = I_0 \mathrm{e}^{-(\alpha_0+\sigma_\alpha\xi)x} \quad , \tag{19}$$

with which we can compute exact mean

$$\langle I(x)\rangle = \int I(x,\xi)\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\xi^2}\mathrm{d}\xi = I_0 e^{-\alpha_0 x} e^{\frac{1}{2}\sigma_\alpha^2 x^2} \tag{20}$$

and variance [1]

$$\mathrm{var}(I(x)) = \langle I(x)^2\rangle - \langle I(x)\rangle^2 = e^{-2\alpha_0 x} I_0^2 \left(e^{2\sigma_\alpha^2 x^2} - e^{\sigma_\alpha^2 x^2}\right) \quad . \tag{21}$$

Here we deal with a somewhat artificial problem: both quantities in Eq. (20) and (21) may increase exponentially with large x because the random variable is not limited to ensure the absorption coefficient to be positive. Even very rare realizations of the negative α will at some x overwhelm the huge number of "sober" solutions and drive the integral sum to infinity. On the physics side it is clear that negative absorption coefficients do not exist and in the analysis it is clear how to deal with this: one would simply restrict the α to be positive or choose a different PDF to draw from. But for our introductory approach it is a perfect example to study the properties of the uncertainty quantification.

The results are depicted in Fig. 4. As mentioned above the rare events with negative α will be responsible for the exponential growth at large depths

[1] There is a misprint in the book of R. Smith [Smith(2014)] which we correct for the plus sign in the exponent of the second exponential term.

x. For the sampling approach it takes more than 10^7 runs to unveil this behaviour with our choice of model parameters, i.e. I_0=10, α_0=1 and within a depth of x=12 (left panel of Fig. 4). Similarly we have to get to 16^{th} order in Hermite polynomials to see the same effect in the uncertainty quantification. While for simple sampling it depends somewhat on the (arbitrary) choice of the random seed until negative values for α come up, the situation here is more deterministic. For reasons of numerical accuracy the order for the Gaussian-Hermite quadrature should be at least one higher than that of the polynomials. Since the values of the quadrature points spread with the size of the order to both sides of the real axis it takes this high order to get a such negative quadrature point which then produces the exponential growth in the result (right panel of Fig. 4).

6 Vlasov-Poisson system

Eventually, we examine the Vlasov-Poisson system, a kinetic plasma model which characterizes the particles of mass m and charge q in the plasma by a distribution function $f(\boldsymbol{x}, \boldsymbol{v}, t)$ in phase space. It is based on the assumption that in a small volume element $\mathrm{d}x\mathrm{d}y\mathrm{d}z$ all plasma particles with the same energy react in the same way to the electromagnetic forces present at this point in space. For the collisionless case the Vlasov equation describes the evolution of charged particles in an electromagnetic field $\boldsymbol{E}(\boldsymbol{x}, t)$ with a neutralizing background

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f}{\partial x} + \frac{q}{m} \left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right) \cdot \frac{\partial f}{\partial v} = 0 \quad . \tag{22}$$

We are interested in the self-consistent electromagnetic field resulting from external fields applied to the system and the internal distribution of the particles. It can be inferred by coupling Maxwell's equations with sources that are the charge densities and current calculated from the particles, (electric and magnetic fields shall be constant with respect to time) with the charge density ρ and current J expressed by the distribution functions

$$\rho(\boldsymbol{x}, t) = q \int f(\boldsymbol{x}, \boldsymbol{v}, t) d\boldsymbol{v} \quad , \qquad \boldsymbol{J}(\boldsymbol{x}, t) = q \int f(\boldsymbol{x}, \boldsymbol{v}, t) \boldsymbol{v} d\boldsymbol{v} \quad . \tag{23}$$

Furthermore, we assume either the magnetic field $\boldsymbol{B}$ or its contribution in the Lorentz force $\boldsymbol{v} \times \boldsymbol{B}$ to be negligible compared to the electric field $\boldsymbol{E}$. Finally we state the Poisson equation (relation between electric field and potential ϕ)

$$-\Delta\phi = 1 - \rho(t, \boldsymbol{x}) = 1 - \int f(t, \boldsymbol{x}, \boldsymbol{v}) \mathrm{d}\boldsymbol{v} \quad , \qquad \boldsymbol{E} = -\nabla\phi \quad , \tag{24}$$

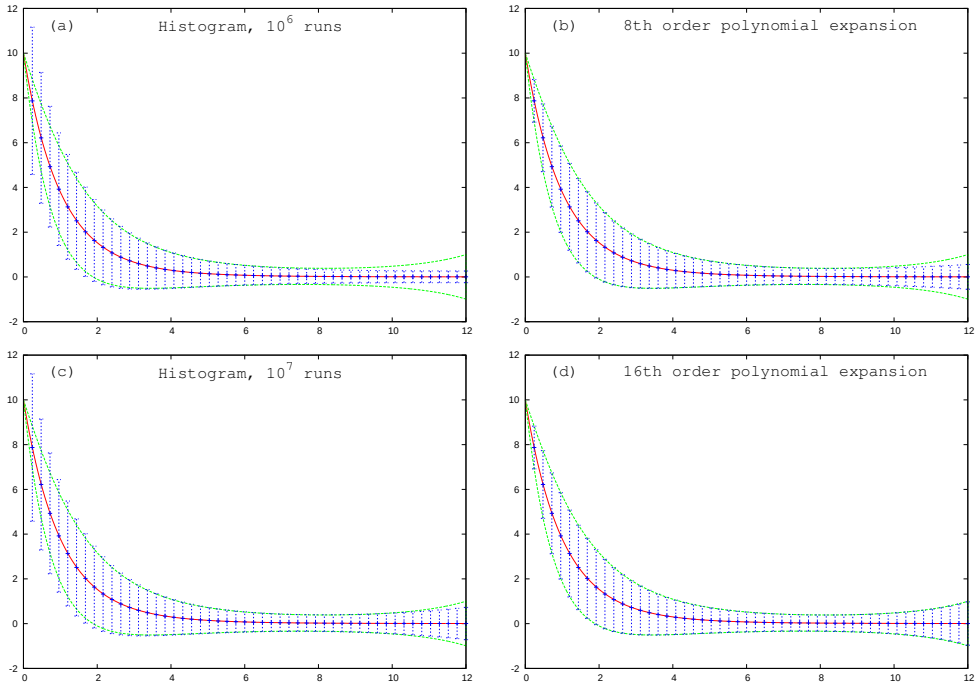

Figure 4: Absorption: light of intensity I_0=10 enters media with uncertain absorption coefficient initially guessed to be α_0=1. The line through the points is the exact solution, accompanying lines show the 2σ region of the analytic random solution Eq. (21). System runs with drawings of the absorption constant α around 1 with $\sigma_\alpha = 0.25$ give the errorbars. Left panel: sampling approach. Top: 10^6 runs are not sufficient to show the increase of the σ region at the end of the interval. Obviously more runs are needed to explore the parameter space to a larger extent. Bottom: 10^7 runs reproduce the spread nicely. Indeed these are a lot of runs! Right panel: uncertainty quantification with α varied at the start and then kept constant throughout model propagation. Top: expansion with 8 Hermite polynomials. Bottom: expansion with 16 Hermite polynomials. Only a higher order in Hermite polynomials/Gaussian-Hermite quadrature is capable of reproducing the spread.

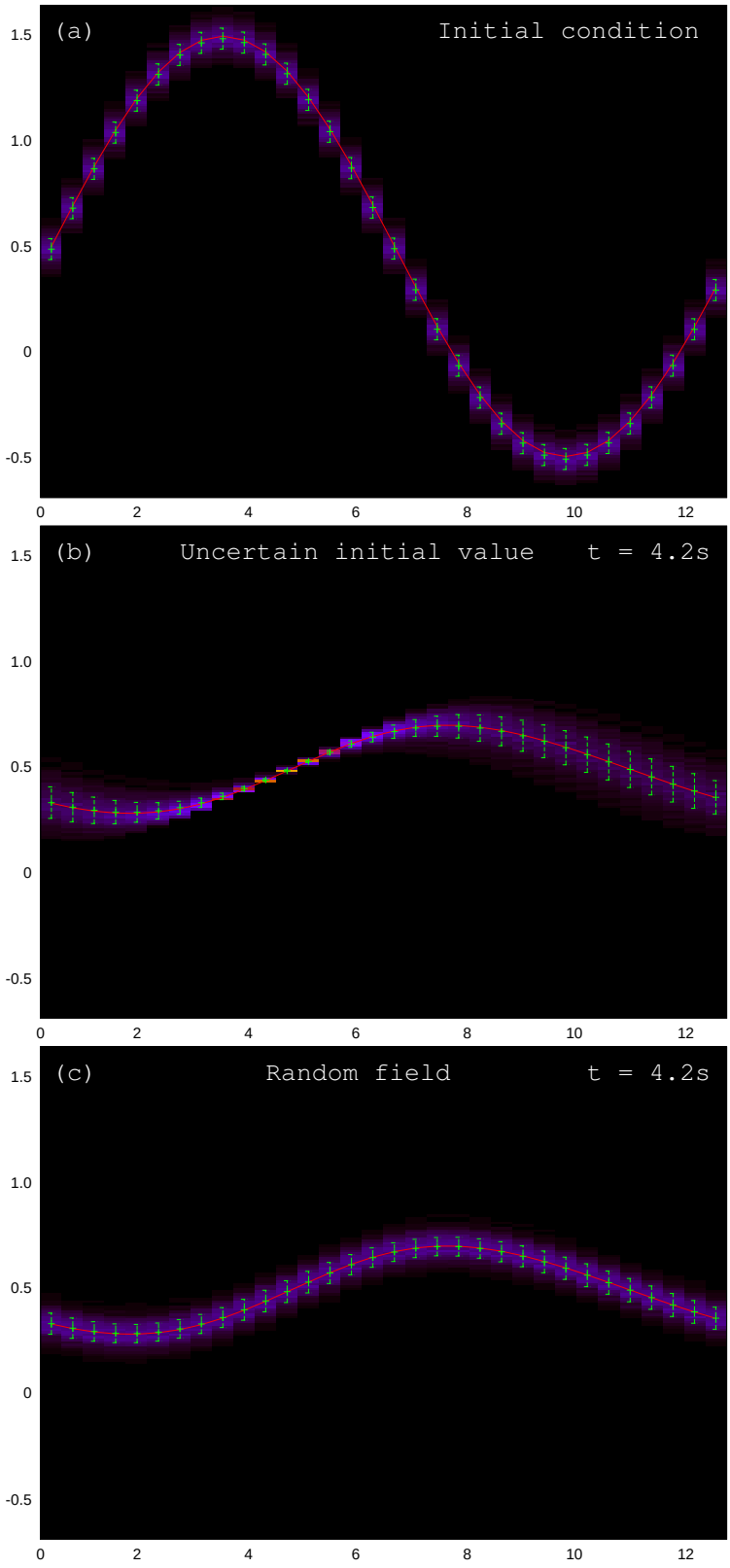

Figure 5: Electric field at the 32 sites obtained from the Vlasov-Poisson model: (a) initial condition; (b) uncertain initial value problem; (c) multiple random variables.

to complete the model equations. Since $\boldsymbol{E}$ depends on f the Vlasov-Poisson system is non linear. The numerical solution is obtained by the semi-Lagrangian method [Sonnendrücker et al.(1999)] which employs a phase space mesh in $\boldsymbol{x}$ and $\boldsymbol{v}$. The specificity of this method is that it uses the characteristics of the scalar hyperbolic equation, along with an interpolation method (mostly cubic spline) to update the unknown from one time step to the next.

Consider a one-dimensional Vlasov-Poisson system of 32 sites evenly spread over 4π in space with periodic boundary conditions. Its distribution function is initialized according to a sin-function with a period of 4π in phase space. Beginning at t=0 the system starts to propagate in time with step Δt=0.01, while being exposed to an uncertain external electric field $E^0_{\text{ext}} = 0.5$ with standard deviation of 10% ($\sigma_{E_{\text{ext}}} = 0.05$) and ξ drawn from $\mathcal{N}(0,1)$:

$$E_{\text{ext}} = E^0_{\text{ext}} + \xi\sigma_{E_{\text{ext}}} \quad . \tag{25}$$

In Fig. 5(a) one can see the electric field resulting from the initial conditions. It reproduces the sin-function of the distribution function, but shows already a spread due to the uncertainty of its value. Notice the overall elevation of 0.5 due to the external field.

First we study the univariate case of the uncertain initial value problem and let the external electric field, same at all sites, vary at the system start, but kept constant throughout system propagation. Fig. 5(b) depicts the result at t=4.2s. While the exact solution is shown by a line, the shading represents the uncertainty distribution derived from the sampling of 1000 system runs. This compares excellent with the result of the uncertainty quantification. given by plus signs with errorbars. We want to point out the fluctuating behaviour of the errorbars, especially with respect to the initial condition. In contrast to the increase around x=11 the errorbars around x=5 even decreased. This behaviour, which is not possible to reproduce following the error propagation law, can be pushed to the non-linear character of the Vlasov-Poisson system.

Finally, we consider the external electric field (still kept the same at all sites) to be uncertain at every time slice which results in a field of random variables. However, with a time step of 0.01s required for stable propagation and a choice of $\lambda = 1$ validated by screening over various settings, at t=4.2s its dimension would reach N=420 – far too large to deal with (according to Eq. (16) N_{int}=424!/420!/4!=1327668126 model calculations for 4th order Hermite polynomials). Therefore we follow the treatment above and employ the Karhunen-Loève expansion from Eq. (13) to reduce the number of random variables. A cutoff below 5% of the maximum eigenvector results in four remaining eigenvectors, which still means 1296 system runs. Fig. 5(c) shows

again an excellent agreement with the simple sampling approach and thereby validates our implementation of uncertainty quantification.

7 Summary

In a discrete projection framework we let a spectral expansion quantify the propagation of uncertainty through a model. As demonstrated for the uncertain initial value problem already the univariate case shows fluctuating behaviour of the uncertainty which is beyond the scope of the error propagation law. For the multivariate case we employed the Karhunen-Loève expansion to reduce the number of random variables to a numerically treatable size. The resemblance with analytic and sampling results validates our procedure.

8 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

[Jaynes(2003)] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, Boston, 2003.

[Najm(2009)] H. N. Najm, *Annu. Rev. Fluid Mech.* **41**, 35 (2009).

[Wiener(1938)] N. Wiener, *Am. J. Math.* **60**, 897 (1938).

[Smith(2014)] R. C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM, Philadelphia, 2014.

[Maître and Knio(2010)] O. P. L. Maître, and O. M. Knio, *Spectral Methods for Uncertainty Quantification*, Springer, Heidelberg, 2010.

[Press et al.(2007)] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, 2007, 3rd edn.

[Sonnendrücker et al.(1999)] E. Sonnendrücker, J. Roche, P. Bertrand, and A. Ghizzo, *J. Comput. Phys* **149**, 201 (1999).