

---

# UNCERTAINTY QUANTIFICATION OF HAMILTONIAN MAPS USING INTRUSIVE POLYNOMIAL CHAOS EXPANSION

---

## MASTER THESIS

in Computational Science and Engineering

Department of Mathematics

ETH Zurich

Written by

FRANCESCO FORCHER  
BSc PHYSICS

Supervised by

Dr. Andreas Adelmann (ETH)

Scientific advisers

Dr. Matthias Frey (ETH)  
Li Sichen (ETH)

January 25, 2021

## Abstract

In the field of uncertainty quantification, polynomial chaos expansion is a technique where the random variables describing the output of stochastic simulations are approximated as a sum of polynomials of random variables of known distribution. In the intrusive method, the original system equations are modified to solve for the coefficients of this expansion. In this thesis, the focus is on the application of polynomial chaos to Hamiltonian equations, with a focus on systems describing accelerator dynamics.

A set of scripts have been developed in Python, using the libraries `chaospy`, `sympy` and `numpy`, which create in a fully procedural way the intrusive polynomial chaos system from a given Hamiltonian. The resulting methods are then applied to a set of concrete problems: linear and non linear oscillators, and a multipole expansion of a relativistic Hamiltonian. The consequences and limitations of the application of polynomial chaos in Hamiltonian systems are explored.

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction and theory</b>   | <b>1</b>  |
| 1.1      | Polynomial Chaos Expansion . . . . .   | 1         |
| 1.1.1    | Intrusive PCE for ODEs . . . . .   | 2         |
| 1.2      | Basic notions of Hamiltonian mechanics . . . . .                               | 6         |
| <b>2</b> | <b>Hamiltonian polynomial chaos expansion theory</b>                           | <b>11</b> |
| 2.1      | The PCE of Hamiltonian equations is Hamiltonian . . . . .                      | 11        |
| 2.1.1    | Outline of derivation . . . . .  | 11        |
| 2.2      | The intrusive solution of some Hamiltonian systems is not optimal . . . . .    | 15        |
| 2.2.1    | Outline of derivation . . . . .  | 15        |
| <b>3</b> | <b>Code overview</b>   | <b>18</b> |
| 3.1      | Polynomial orthogonalization in <code>sympy</code> . . . . .                   | 21        |
| 3.2      | Monte Carlo samples . . . . .  | 21        |
| 3.3      | Non-intrusive fit . . . . .  | 23        |
| 3.4      | Exponential map . . . . .  | 23        |
| 3.5      | Symbolical and numerical quadrature . . . . .                                  | 23        |
| <b>4</b> | <b>Results</b>   | <b>27</b> |
| 4.1      | Types of plots . . . . .   | 27        |
| 4.1.1    | Comparison of uncertainty with Monte Carlo and non-intrusive methods . . . . . | 27        |
| 4.1.2    | Energy comparison . . . . .  | 27        |
| 4.1.3    | Exponential maps . . . . .   | 28        |
| 4.2      | General observations . . . . .   | 28        |
| 4.2.1    | The initial conditions and the order of expansion . . . . .                    | 28        |
| 4.2.2    | Long term analysis and symplectification . . . . .                             | 32        |
| 4.3      | Harmonic oscillator . . . . .  | 33        |
| 4.3.1    | First-order . . . . .  | 34        |
| 4.3.2    | Second order . . . . .   | 37        |
| 4.3.3    | Fourth order . . . . .   | 39        |

|  |   |           |
|--|---|-----------|
| 4.4                                      | Duffing oscillator . . . . .                  | 41        |
| 4.4.1                                    | First-order . . . . .                         | 42        |
| 4.4.2                                    | Second order . . . . .                        | 44        |
| 4.4.3                                    | Third order . . . . .                         | 47        |
| 4.5                                      | Non-linear quadrupole . . . . .               | 48        |
| 4.5.1                                    | Comparison plots, second order . . . . .      | 50        |
| 4.5.2                                    | Exponential map example . . . . .             | 56        |
| 4.5.3                                    | Numerical quadrature example . . . . .        | 56        |
| <b>5</b>                                 | <b>Conclusions</b>                            | <b>58</b> |
| <b>Appendix A Hamiltonian PCE theory</b> |   | <b>62</b> |
| A.1                                      | PCE of harmonic oscillator solution . . . . . | 62        |
| <b>Appendix B PC Hamiltonians</b>        |   | <b>63</b> |
| B.1                                      | Harmonic oscillator . . . . .                 | 63        |
| B.2                                      | Duffing oscillator . . . . .                  | 64        |

# List of Figures

|      |   |    |
|------|---|----|
| 3.1  | Time average of the error of the mean by number of MC samples . . . . .                 | 22 |
| 3.2  | The <code>dask</code> computation graph of the parallel integration . . . . .           | 25 |
| 4.1  | Harmonic Hamiltonian, error of standard deviation. . . . .                              | 29 |
| 4.2  | Harmonic Hamiltonian, $O = 1$ , non-intrusive comparison . . . . .                      | 34 |
| 4.3  | Harmonic Hamiltonian, $O = 1$ , energy comparison . . . . .                             | 35 |
| 4.4  | Harmonic Hamiltonian, $O = 1$ , exponential map . . . . .                               | 36 |
| 4.5  | Harmonic Hamiltonian, $O = 2$ , non-intrusive comparison . . . . .                      | 37 |
| 4.6  | Harmonic Hamiltonian, $O = 2$ , exponential map . . . . .                               | 38 |
| 4.7  | Harmonic Hamiltonian, non-intrusive comparison, stochastic initial conditions . . . . . | 39 |
| 4.8  | Harmonic Hamiltonian, exponential map, stochastic initial conditions                    | 40 |
| 4.9  | Duffing Hamiltonian, $O = 1$ , non-intrusive comparison . . . . .                       | 42 |
| 4.10 | Duffing Hamiltonian, $O = 1$ , energy . . . . .   | 43 |
| 4.11 | Duffing Hamiltonian, $O = 2$ , non-intrusive comparison . . . . .                       | 44 |
| 4.12 | Duffing Hamiltonian, $O = 2$ , exponential map . . . . .                                | 45 |
| 4.13 | Duffing Hamiltonian, $O = 2$ , energy . . . . .   | 46 |
| 4.14 | Duffing Hamiltonian, $O = 3$ , non-intrusive comparison . . . . .                       | 47 |
| 4.15 | Quadrupole Hamiltonian, $x(s)$ . . . . .  | 50 |
| 4.16 | Quadrupole Hamiltonian, $y(s)$ . . . . .  | 51 |
| 4.17 | Quadrupole Hamiltonian, $z(s)$ . . . . .  | 52 |
| 4.18 | Quadrupole Hamiltonian, $p_x(s)$ . . . . .  | 53 |
| 4.19 | Quadrupole Hamiltonian, $p_y(s)$ . . . . .  | 54 |
| 4.20 | Quadrupole Hamiltonian, $\delta(s)$ . . . . .   | 55 |
| 4.21 | Quadrupole Hamiltonian, $x(s)$ , exponential map . . . . .                              | 56 |
| 4.22 | Quadrupole Hamiltonian, $x(s)$ . . . . .  | 57 |

# Chapter 1

## Introduction and theory

### 1.1 Polynomial Chaos Expansion

Polynomial chaos expansion (PCE) is a technique used to estimate uncertainty in stochastic systems and functions. In this framework, the system's random variables  $X$  are approximated as a sum of polynomials  $\Psi_i$  of other random variables  $\Xi$ , called *germs*, with known distribution  $P_{\Xi}(\xi)$ ,  $\xi \in \Omega$  [OHa13], where usually  $\Omega \subseteq \mathbb{R}^n$ :

$$\begin{aligned} X &\sim \sum_{i=0}^{\infty} x_i \Psi_i(\Xi) \\ &\approx \sum_{i=0}^{N_O} x_i \Psi_i(\Xi) . \end{aligned} \tag{1.1}$$

Here  $N_O$  is the number of polynomials in  $d$  variables of degree  $O$ :  $N_O = \binom{O+d}{d}$ , and  $x_i$  represent the numerical weighting coefficients of the various polynomials.  $X \sim Y$  means “ $X$  distributed as  $Y$ ”, equality in distribution.

It can be shown[OHa13] that it is always possible to find a function  $f$  expressing  $X \sim f(\Xi)$ , for any distribution of  $X$  and  $\Xi$ . A canonical example uses the inverse cumulative distribution function (CDF) transform ( $U_{0,1}$  being the uniform distribution on the interval  $[0, 1]$ ):

$$X \sim \text{CDF}_X^{-1}(U_{0,1}) = \text{CDF}_X^{-1}(\text{CDF}_{\Xi}(\Xi)) = f(\Xi) .$$

Then a PC expansion could in principle be obtained by applying Taylor expansion to  $f = \text{CDF}_X^{-1} \circ \text{CDF}_{\Xi}$ . However the order of expansion required to attain sufficient accuracy may vary greatly, for different  $f$  and germs [OHa13]. As an example, the  $\chi_k^2$  distribution is defined as the sum of  $k$  Gaussian variables  $\mathcal{N}$  squared, so a second degree polynomial in  $k$  Gaussian variables represents  $\chi_k^2$  exactly at all orders:

$$\chi_k^2 = \mathcal{N}_1^2 + \dots + \mathcal{N}_k^2 .$$

But a higher degree of expansion could be required, if  $\chi_k^2$  is to be accurately represented as a PC expansion of other germ distributions, or even of a different number of Gaussian variables [OHa13]. This exemplifies how choosing the most appropriate germ variables may not be straightforward, and may greatly influence the accuracy or computational complexity of the result.

To simplify calculations, the polynomials  $\Psi_i(\Xi)$  are usually chosen as to be orthogonal or even orthonormal with respect to the inner product  $\langle \Psi_i \Psi_j \rangle_{\Xi}$ , defined as taking the expected value over  $\Xi$  of their multiplication:

$$\begin{aligned}\langle \Psi_i \Psi_j \rangle_{\Xi} &= \int_{\Omega} \Psi_i(\xi) \Psi_j(\xi) P_{\Xi}(\xi) d\xi \\ &= \delta_{ij}.\end{aligned}\tag{1.2}$$

The task becomes now to find the coefficients  $x_i$  of Eq. 1.1, such that the resulting distribution best approximates the original variable  $X$  [Feia; OHa13]. There are two main approaches, called intrusive and non-intrusive respectively. In the latter, the coefficients are fitted over some empirical dataset representing samples of the original variable, as described in 3.3. Often in simulations one can generate such dataset using Monte Carlo methods.

### 1.1.1 Intrusive PCE for ODEs

In the intrusive approach on the other hand, the system itself is modified to solve for the coefficients instead of the original variables [Feia] [OHa13]. A concrete example [Feia]: the starting point is an ordinary differential equation in normal form, which depends on some random parameters  $\alpha$ :

$$\frac{dY}{dt}(t) = f(t, Y; \alpha),\tag{1.3}$$

with initial conditions

$$Y(0) = Y_0\tag{1.4}$$

where  $Y_0$  is some arbitrary initial distribution of the possible initial values  $y_0$ .

The parameter values in Eq. 1.3, Eq. 1.4 are assumed to be uncertain but fixed, and the time evolution is thus deterministic. Consequently, 1.3 represents an ensemble of deterministic equations of the form 1.5, for each value  $a$ ,  $y_0$  weighted by its relative probability  $p_{\alpha, Y_0}(a, y_0)$ :

$$y(t) = f(t, y; a), \quad y(0) = y_0.\tag{1.5}$$

$Y(t)$ , the solution distribution of Eq. 1.3, is a random variable representing the probability distribution of the possible solution values  $y$  at time  $t$ .

## The main equation

Approximating the  $Y(t)$  distribution using polynomial chaos we obtain:

$$Y(t) \sim \sum_{i=0}^{\infty} y_i(t) \Psi_i(\Xi) \approx \sum_{i=0}^{N_O} y_i(t) \Psi_i(\Xi) \quad (1.6)$$

Usually, the expansion is chosen such that only the coefficients vary in time, while the germs distribution and polynomials do not. The intrusive PCE approach results in a new system of equations for the coefficients  $y_i(t)$ , obtained using Galerkin projection [Feia; OHa13].

First, one substitutes the PCE approximation 1.6 in the equation 1.3 [Feia]:

$$\frac{d}{dt} \left( \sum_{i=0}^{N_O} y_i(t) \Psi_i(\Xi) \right) = f \left( t, \sum_{i=0}^{N_O} y_i(t) \Psi_i(\Xi); \alpha \right).$$

Given that only the coefficients vary in time, this becomes:

$$\sum_{i=0}^{N_O} \frac{dy_i}{dt}(t) \Psi_i(\Xi) = f \left( t, \sum_{i=0}^{N_O} y_i(t) \Psi_i(\Xi); \alpha \right).$$

Finally, a system of  $N_O$  equations is created using the Galerkin method. First both sides are multiplied with the polynomials  $\Psi_k$  (test functions, in Galerkin terminology):

$$\sum_{i=0}^{N_O} \frac{dy_i}{dt}(t) \Psi_i(\Xi) \Psi_k(\Xi) = f \left( t, \sum_{i=0}^{N_O} y_i(t) \Psi_i(\Xi); \alpha \right) \Psi_k(\Xi).$$

Then, one takes the expected value of both sides, which is equivalent to Galerkin projection using the  $L_2$  scalar product  $\langle f, \psi_k \rangle = \int_{\Omega} f \psi_k d\omega$ :

$$\left\langle \sum_{i=0}^{N_O} \frac{dy_i}{dt}(t) \Psi_i(\Xi) \Psi_k(\Xi) \right\rangle_{\Xi} = \left\langle f \left( t, \sum_{i=0}^{N_O} y_i(t) \Psi_i(\Xi); \alpha \right) \Psi_k(\Xi) \right\rangle_{\Xi}.$$

The left hand side simplifies due to the orthonormality condition:

$$\begin{aligned} \sum_{i=0}^{N_O} \frac{dy_i}{dt}(t) \langle \Psi_i(\Xi) \Psi_k(\Xi) \rangle_{\Xi} &= \sum_{i=0}^{N_O} \frac{dy_i}{dt}(t) \delta_{ik} \\ &= \frac{dy_k}{dt}(t). \end{aligned}$$

We have thus obtained the general form of the intrusive PCE system:

$$\begin{aligned}\frac{dy_k}{dt}(t) &= \left\langle f \left( t, \sum_{i=0}^{N_O} y_i(t) \Psi_i(\Xi); \alpha \right) \Psi_k(\Xi) \right\rangle_{\Xi} \\ &= \int_{\Omega} f \left( t, \sum_{i=0}^{N_O} y_i(t) \Psi_i(\xi); \alpha(\xi) \right) \Psi_k(\xi) P_{\Xi}(\xi) d\xi .\end{aligned}\quad (1.7)$$

### The initial conditions

Similarly, we insert the PC expansion into the initial conditions, then apply Galerkin projection [Feia].

$$\begin{aligned}Y(0) &= Y_0 \\ \left\langle \sum_{i=0}^{N_O} y_i(0) \Psi_i(\Xi) \Psi_k(\Xi) \right\rangle &= \langle Y_0 \Psi_k(\Xi) \rangle \\ \sum_{i=0}^{N_O} y_i(0) \langle \Psi_i(\Xi) \Psi_k(\Xi) \rangle &= \langle Y_0 \Psi_k(\Xi) \rangle \\ y_k(0) &= \langle Y_0(\Xi) \Psi_k(\Xi) \rangle \\ &= \int_{\Omega} Y_0(\xi) \Psi_k(\xi) P_{\Xi}(\xi) d\xi .\end{aligned}\quad (1.8)$$

This gives us the corresponding initial conditions for the coefficients  $y_k(0)$ .

### Choice of germ variables

To perform the expected value integrals, the distributions of the stochastic parameters such as  $Y_0$  and  $\alpha$  have to be expressed as functions of the germs  $\Xi$ :

$$\alpha \sim \sum_{i=0}^{N_O} a_i \Psi_i(\Xi) \quad (1.9)$$

$$Y_0 \sim \sum_{j=0}^{N_O} y_{0j} \Psi_j(\Xi) . \quad (1.10)$$

In fact, Eq. 1.8 represents the same Galerkin projection as Eq. 1.10, for the initial conditions  $Y_0$ .

However, this introduces yet another layer of approximation. Moreover, often spurious correlations appear between them, if the same germs are used to generate both variables. For example, both  $\alpha, Y_0$  could be uniformly distributed:

$$\begin{aligned}\alpha &\sim U(a_{\text{Min}}, a_{\text{Max}}) \\ Y_0 &\sim U(y_{0\text{Min}}, y_{0\text{Max}}).\end{aligned}$$

Then, a simple PC solution is a simple linear rescaling of a single uniform germ  $U_0(0, 1)$ :

$$\begin{aligned}\alpha &\sim a_{\text{Min}} + (a_{\text{Max}} - a_{\text{Min}})U_0(0, 1) \\ Y_0 &\sim y_{0\text{Min}} + (y_{0\text{Max}} - y_{0\text{Min}})U_0(0, 1).\end{aligned}$$

This is a first order polynomial expansion, but it is exact (all other coefficients would be zero). However, it implies that  $\alpha$  and  $Y_0$  are perfectly correlated with each other: a specific value of one implies one specific value of the other, for example:

$$y_{0\text{Val}}(a_{\text{Val}}) = y_{0\text{Min}} + (y_{0\text{Max}} - y_{0\text{Min}}) \frac{a_{\text{Val}} - a_{\text{Min}}}{a_{\text{Max}} - a_{\text{Min}}}.$$

While this is certainly a valid approximation, it often fails to properly capture all the variability intrinsic in systems with independent parameters. A solution commonly adopted is to include the random parameters of the ODE themselves as independent germs [Feia]:  $\{\alpha_i\} \subseteq \{\Xi_i\}$ .

While the parameter distributions themselves may be arbitrary and/or data-derived, it is still possible to construct the appropriate orthogonal polynomials. A direct way would be to use a histogram or kernel density estimation with a known kernel, to approximate the probability distribution function  $P_{\Xi}(\xi)$ , then use an orthogonalization procedure such as Gram–Schmidt on the polynomials. More sophisticated methods [ON12] construct explicitly the polynomials requiring only the knowledge of the first  $n$  statistical moments of the parameter, for an order- $n$  PC expansion. The full knowledge or even existence of the pdf  $P_{\Xi}(\xi)$  is not required.

These approaches, however, may exacerbate one of the main drawbacks of polynomial chaos: the number of polynomials of degree  $O$  in  $d$  variables is  $N_O = \binom{O+d}{d}$ , which increases exponentially with the number of germs  $d$  and degree  $O$ . This means that in systems with many parameters, Monte Carlo methods typically converges faster than intrusive polynomial chaos.

The equations 1.8 and 1.7 represent a standard deterministic ODE system, without any explicit stochastic variables. Its solution are the coefficients  $y_i(t)$  of the polynomial chaos approximation of  $Y(t)$ , denoted  $\hat{Y}(t) = \sum_i^{N_O} y_i(t)\Psi_i(\Xi)$ .

Some useful properties of  $\hat{Y}(t)$  can be described in closed form from the coefficients. In particular, we have a closed form for mean and variance (first and second moments) thanks to  $\Psi_i$  orthonormality:

$$\mathbb{E}(\hat{Y})(t) = y_0(t) \quad (1.11)$$

$$\mathbb{E}(\hat{Y}^2)(t) = \sum_{j=0}^{N_O} y_j^2(t) \langle \Psi_j, \Psi_j \rangle = \sum_{j=0}^{N_O} y_j^2(t)$$

$$\text{Var}(\hat{Y})(t) = \mathbb{E}\left(\left((\hat{Y} - \mathbb{E}(\hat{Y}))\right)^2\right)(t) = \sum_{j=0}^{N_O} y_j^2(t) - y_0^2(t) \quad (1.12)$$

Higher-order moments are computable as well but involve three-terms integrals. While often formulas exists for common families of orthogonal polynomials, often it is quicker to sample the PC expansion and do a Monte Carlo estimation [OHa13].

The integrals of the equations right hand sides may be very complex to evaluate in closed form. A numerical quadrature scheme can be attempted, but often it results in significantly degraded performance, as it needs to be evaluated at each time step: a non-intrusive approach would probably be preferred in this case.

A few simplifications are possible though:  $f$  can typically be expanded as a power series, so that the resulting expression is easier to integrate.

The approach outlined above of including the equation parameters directly among the germs typically simplifies the  $f$  expression evaluation. Choosing germs with known  $P_\Xi(\xi)$  distributions such as Gaussians, uniforms, exponentials, often leads to integrable analytical expressions. For arbitrary parameter distributions, the approximate pdf method and the moments methods outlined in [ON12] can be applied to simple expressions, such as polynomial ones as obtained from the power series.

## 1.2 Basic notions of Hamiltonian mechanics

The three main mathematical frameworks used to study the motion of objects and particles are Newtonian, Lagrangian, and Hamiltonian mechanics. The three approaches have their own strengths and drawbacks, but in classical mechanics (including relativistic), the Hamiltonian approach is typically favored.

Hamiltonian differential equations are a class of ordinary differential equations, which represent the equations of motion describing the trajectory of some physical

system. Their general form is:

$$\begin{cases} \frac{dq_i}{dt} = +\frac{\partial H}{\partial p_i}(q_i, p_i) \\ \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}(q_i, p_i) \end{cases} \quad (1.13)$$

for a given function  $H(q_i, p_i)$  called the Hamiltonian function of the system. The Hamiltonian is a scalar function of two sets of  $n$  coordinates,  $q_i$  and  $p_i$ . The variables  $q_i \leftrightarrow p_i$  have a special pairwise relationship: they are called *conjugated* to each other. The  $q_i$  are typically identified as the system positions, and their conjugated  $p_i$  are the conjugate momenta.

However, this identification is not as rigid as the distinction between positions  $q_i$  and their velocities  $\dot{q}_i$  (the dot denotes the time derivative  $\dot{q}_i = \frac{\partial q_i}{\partial t}$ ) in Lagrangian mechanics. In Lagrangian dynamics, the admissible coordinate changes can only depend on the positions  $q_i$ :

$$(q_i, \dot{q}_i) \rightarrow \left( f(q_i), \frac{df}{dq_i}(q_i)\dot{q}_i \right).$$

The coordinate change for the velocities  $\dot{q}_i$  is completely determined by the position map  $q_i \rightarrow f(q_i)$ .

In contrast, one of the major advantages of Hamiltonian mechanics is the ability to perform much more general changes of coordinates in the whole phase space of  $(q, p) \rightarrow C(q, p)$ , while still preserving the Hamiltonian character of the ODE system: these are the so-called symplectic diffeomorphisms. This is a consequence of the deep geometrical implications of Hamiltonian mechanics, which are realized in the field of symplectic geometry [Arn89].

### Symplectic unit $\mathbb{J}$

The  $n$  coordinates  $q$  and  $p$  can therefore be unified in a single  $2n$  coordinate system for the whole phase space: for example

$$z = (q_0, \dots, q_i, p_0, \dots, p_i). \quad (1.14)$$

The Hamilton equations 1.13 will then be written as

$$\frac{dz_i}{dt}(t) = \mathbb{J}_{ij} \frac{\partial H}{\partial z_j}(z(t)), \quad (1.15)$$

where  $\mathbb{J}_{ij} \in \mathbb{R}^{2n} \times \mathbb{R}^{2n}$  is called the symplectic matrix. Given the identification 1.14 it follows that in order to obtain the equations 1.13 from 1.15,  $\mathbb{J}$  must have

the block form:

$$\mathbb{J} = \begin{bmatrix} \mathbb{O} & \mathbb{I} \\ -\mathbb{I} & \mathbb{O} \end{bmatrix} .$$

The symplectic matrix  $\mathbb{J}$  can have other concrete realizations, for example for an identification  $z = (q_0, p_0, \dots, q_i, p_i)$ ,  $\mathbb{J}$  will have the block form

$$\mathbb{J} = \begin{bmatrix} 0 & 1 & & & 0 \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ 0 & & & -1 & 0 \end{bmatrix} .$$

But they all share the following properties:

$$\begin{aligned} \mathbb{J}^T &= \mathbb{J}^{-1} = -\mathbb{J} \\ \mathbb{J}^2 &= -\mathbb{I} \\ \det(\mathbb{J}) &= 1 . \end{aligned}$$

## Symplectic matrices

A matrix  $S$  is said to be *symplectic* if it satisfies

$$S\mathbb{J}S^T = \mathbb{J} .$$

A diffeomorphic map (such as coordinate changes)  $C(z) : W \rightarrow \hat{W}$  between two open subsets  $\subseteq \mathbb{R}^{2n}$  is said to be symplectic if its Jacobian matrix  $C'(z)$  is symplectic, in every point.

$$C'(z)\mathbb{J}C'(z)^T = \mathbb{J}, \quad \forall z \in W .$$

The composition of two symplectic maps  $C = C_2 \circ C_1$  is still symplectic, and the inverse  $C^{-1}$  is symplectic too. Symplectic maps preserve the signed volume of phase space.

## Poisson brackets

Given a generic ODE  $y' = f(y)$ , one can identify with  $X_f$  its corresponding vector field, whose integral curves are the orbits of the system. The Lie derivative of a scalar function  $g$  with respect to a vector field  $X_f$  is defined as the directional derivative of  $g$ , in the direction of  $X_f$ :

$$\mathcal{L}_{X_f} g = X_f \cdot \nabla g .$$

A function  $g$  which has zero Lie derivative with respect to some field  $X_f$ , is called a first integral (also integral of motion, constant of motion) of the ODE, meaning that its value is constant along the system trajectories.

From equation 1.15, written in compact form as  $\dot{z} = \mathbb{J}\nabla H$ , one can identify the corresponding vector field

$$X_H = \mathbb{J}\nabla H = \sum_j \mathbb{J}_{ij} \frac{\partial H}{\partial z_j} .$$

This vector field  $X_H$  is called the *Hamiltonian vector field* associated with the (scalar) function  $H$ . Now taking two generic scalar functions in the phase space  $f, g$ , we can create a third function, their *Poisson brackets*  $[f, g]$ . This is done by taking the (minus) Lie derivative of  $g$  along  $X_f$ , here denoted as  $L_{X_f}g$ :

$$\begin{aligned} \{f, g\} &= -L_{X_f}g \\ &= -\mathbb{J}\nabla f \cdot \nabla g = \nabla f \cdot \mathbb{J}\nabla g \\ &= \sum_i^n \left( \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right) . \end{aligned} \quad (1.16)$$

The Poisson bracket characterizes the first integrals of a Hamiltonian ODE system: a function  $f$  is a first integral of a Hamiltonian system of Hamiltonian  $H$  if and only if  $[H, f] = 0$ . It is immediate to see that  $[H, H] = 0$ , so the Hamiltonian is always a first integral of the system. In fact,  $H$  in physical systems corresponds to the conserved total energy.

### Liouville's theorem

It can be proven that the map given at time  $t$  by the flux  $\Phi_t^{X_H}$  of an Hamiltonian vector field is symplectic. This implies that the flux preserves the volume of phase space. This has important consequences, for example dissipative systems, tend to decrease the phase space volume over time, so they cannot be represented directly, within this framework.

### Lie operators

Given a phase space function  $f$ , a differential operator called *Lie operator*, denoted as  $:f:$ , can be defined using the Poisson brackets 1.16 [Dra13]:

$$:f: = \sum_i^n \left( \frac{\partial}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial}{\partial p_i} \frac{\partial g}{\partial q_i} \right) ,$$

such that its application to a function  $g$  results in the Poisson brackets between  $f$  and  $g$

$$:f:g = [f, g] .$$

Consequently, the powers of a Lie operator are:

$$\begin{aligned} :f:^0 g &= g \\ :f:^1 g &= [f, g] \\ :f:^2 g &= [f, [f, g]] \\ &\dots \\ :f:^n g &= [f, [...[f, g]]] , \end{aligned}$$

and so on recursively.

### Exponential maps

The Lie operator is used in the definition of the exponential map form of a Hamiltonian system solution. It is easy to verify that Hamilton equations 1.15 can be written in the form:

$$\dot{z} = - :H: z .$$

The flux (or transport) map  $M_t = \Phi_{X_H}(t)$  which sends the starting values  $z(0)$  to  $z(t)$  can be defined as:

$$z(t) = M_t(z(0)) = \exp(-t :H:) z(0) \quad (1.17)$$

Using the identities of the Lie operator and truncating the Taylor expansion of the exponential function, this map can then be approximated at any order  $T$ :

$$\begin{aligned} z(t) &= M_t(z(0)) = \exp(-t :H:) z(0) \\ &\approx \left( :H:^0 + \frac{(-t)^1}{1!} :H:^1 + \dots + \frac{(-t)^T}{T!} :H:^T \right) z(0) \\ &= z(0) + \frac{(-t)^1}{1!} [H, z(0)] + \dots + \frac{(-t)^T}{T!} [H, [H, \dots, z(0)]] . \quad (1.18) \end{aligned}$$

A phase space function  $f(z, t = 0)$  such as a probability density function can be transported by the map along the Hamiltonian flux, thus giving the distribution at time  $f(z, t)$ .

$$\frac{\partial f}{\partial t} = - :H: f \quad (1.19)$$

These maps are often used in accelerator physics, where even non-linear particle accelerator dynamics of distributions of particles can be efficiently described as a composition of the transport maps of the various accelerator elements.

# Chapter 2

## Hamiltonian polynomial chaos expansion theory

In this chapter the principal results on intrusive Hamiltonian PC expansion theory, described in [PS13], are reviewed.

### 2.1 The PCE of Hamiltonian equations is Hamiltonian

The special structure of Hamiltonian differential equations leads to a special structure for their PCE approximation as well. The resulting equations for the coefficients are themselves Hamiltonian, with a new Hamiltonian resulting from the averaging of the original system ones over the germ distributions: Denoting the original Hamiltonian with  $H(q_i, p_i; \alpha_i)$ , which has a stochastic distribution dependent on the random parameters  $\alpha_i$ , a new Hamiltonian  $\hat{H}$  is obtained:

$$\begin{aligned}\hat{H}(P_{ij}, Q_{kl}) &= \left\langle H \left( \sum_j^{N_O} P_{ij} \Psi_j, \sum_l^{N_O} Q_{kl} \Psi_l, \sum_n^{N_O} A_{mn} \Psi_n \right) \right\rangle_{\Xi} \\ &= \int_{\Omega} H \left( \sum_j^{N_O} P_{ij} \Psi_j, \sum_l^{N_O} Q_{kl} \Psi_l, \sum_n^{N_O} A_{mn} \Psi_n \right) P_{\Xi}(\xi) d\xi \quad (2.1)\end{aligned}$$

#### 2.1.1 Outline of derivation

##### Notation

Unless specified otherwise Einstein notation will be used from now on, where a sum over repeated indices in expressions is implied

$$w_i = A_{ij} v_j \equiv \sum_j A_{ij} v_j .$$

The symbol  $\delta_{ij\dots k}$  will denote a multi index Kronecker delta:  $\delta_{ij\dots k} = 1$  if  $i = j = \dots = k$  and 0 otherwise. It is useful to remember the main identity of the  $\delta_{ij}$  on a vector  $v_i$ , and the expression for a diagonal matrix  $D$  with  $D_{kk} = d_k$  as diagonal elements, in index notation:

$$\begin{aligned} v_i &= \delta_{ij}v_j \\ D_{ij} &= \delta_{ijk}d_k \end{aligned} \quad (2.2)$$

Sometimes a bold font is used to denote a vector (or matrix, or tensor)  $\mathbf{v}$ , when referring to it as the whole object rather than to its parametrized component  $v_i$ . For example, for a scalar function  $f$ , compare the expression for the components of vector  $\mathbf{w}$  of elementwise application,  $w_i = f(v_i)$ , to some scalar-valued expression  $C = f(\mathbf{v}) = f(v_0, \dots, v_N)$ .

We generally denote the germ variables as  $\Xi = (\Xi_0, \dots, \Xi_{N_\xi})$ , assuming values  $\xi$  with probability distribution  $\mathcal{P}_\Xi(\xi)$  over domain  $\Omega$ . The angle brackets  $\langle \dots \rangle$  are used to denote the expected value of an expression  $f(\Xi)$  over the  $\Xi$  distribution.

$$\langle f(\Xi) \rangle = \int_{\Omega} f(\xi) P_\Xi(\xi) d\xi$$

Conveniently, this is consistent with the use of the brackets to denote also the inner product  $\langle \cdot, \cdot \rangle$  between two  $\Xi$  polynomials given the  $\Xi$  distribution (see Eq. 1.2).

For example, the polynomials orthogonality property, but not necessarily normality, as shown by the constants  $\|\Psi_i\|^2 = \langle \Psi_i \Psi_i \rangle$ , is written (using the diagonal matrix formula Eq. 2.2) as:

$$\langle \Psi_i, \Psi_j \rangle = \langle \Psi_i \Psi_j \rangle = \int_{\Omega} \Psi_i(\xi) \Psi_j(\xi) \mathcal{P}_\Xi(\xi) d\xi = \delta_{ij} \|\Psi_i\|^2.$$

## Calculations

Following [PS13], we start from some Hamiltonian system, with Hamiltonian  $H(\mathbf{q}, \mathbf{p}; \boldsymbol{\alpha})$ , which depends on some random parameters  $\alpha_i$ . We collect all variables, positions  $q_i$  and momenta  $p_i$ , in the vector  $z_i$ , for example  $\mathbf{z}(t) = (\mathbf{q}(t), \mathbf{p}(t))$ . The standard Hamilton equations of the system will be:

$$\frac{dz_i}{dt}(t) = \mathbb{J}_{ij} \frac{\partial H}{\partial z_j}(\mathbf{z}(t); \boldsymbol{\alpha}) \quad (2.3)$$

Then we approximate  $\mathbf{z}$  and  $\boldsymbol{\alpha}$  using polynomial chaos expansion, in the polynomial basis  $\Psi_i(\Xi_1, \dots, \Xi_O)$ , for some set of  $N_\xi$  variables  $\Xi$ :

$$\begin{cases} z_m(t) \approx Z_{mn}(t) \Psi_n \\ \alpha_k \approx A_{kl} \Psi_l \end{cases} \quad (2.4)$$

It should be noted that the “parameters as germs” technique is a subcase of this general expansion, where the coefficients matrix  $A_{kl}$  has a simple form of zeros and ones. From this definition follows trivially that the derivative of  $z_i$  with respect to  $Z_{jk}$  is:

$$\frac{\partial z_i}{\partial Z_{jk}} = \delta_{ij} \Psi_k \quad (2.5)$$

The average Hamiltonian  $\hat{H}(Z_{mn}(t); \boldsymbol{\alpha})$  from 2.1 can be written more compactly as:

$$\hat{H}(\mathbf{Z}; \mathbf{A}) = \int_{\Omega} H(Z_{mn}(t) \Psi_n; A_{kl} \Psi_l) \mathcal{P}_{\Xi}(\xi) d\xi. \quad (2.6)$$

### The left-hand side

The new system of equations for the coefficients  $Z_{mn}(t)$  is obtained by substituting and applying Galerkin projection on the left side of Eq. 2.3. Since the  $\Psi_n$  do not depend on time, we can take them out of the derivative. Then, we can simplify the polynomial product using orthogonality:

$$\begin{aligned} \left\langle \frac{dZ_{mn} \Psi_n}{dt}(t), \Psi_s \right\rangle &= \frac{dZ_{mn}}{dt}(t) \langle \Psi_n \Psi_s \rangle \\ &= \frac{dZ_{mn}}{dt}(t) \delta_{nsk} \|\Psi_k\|^2 = \frac{dZ_{ms}}{dt}(t) \|\Psi_s\|^2 \end{aligned}$$

This shows how we have reduced the system of equations of coefficients  $Z_{mn}(t)$  system in standard “vector field” form  $\mathbf{y}'(t) = f(\mathbf{y}(t), t)$ : after dividing by  $\|\Psi_s\|^2$ , we will have now on the left side only the derivative of  $Z_{ms}(t)$  with respect to time (and not a sum), and the corresponding right side will then show how  $Z_{ms}(t)$  evolves.

### The right-hand side

To simplify a bit the notation, lets impose without loss of generality, orthonormality of the polynomials so that  $\|\Psi_s\|^2 = 1$ . Now after substituting and applying Galerkin projection, and simplifying the left side, the original equations Eq. 2.3 look like this:

$$\begin{aligned} \frac{dZ_{ms}}{dt}(t) &= \left\langle \mathbb{J}_{mj} \frac{\partial H}{\partial z_j}(\mathbf{z}(t); \boldsymbol{\alpha}), \Psi_s \right\rangle \\ &= \int_{\Omega} \mathbb{J}_{mj} \frac{\partial H}{\partial z_j}(\mathbf{z}(t); \boldsymbol{\alpha}) \Psi_s(\xi) \mathcal{P}_{\Xi}(\xi) d\xi. \end{aligned} \quad (2.7)$$

These equations for the coefficients  $Z_{ms}(t)$  can be written as standard Hamiltonian equations, using the Hamiltonian  $\hat{H}(\mathbf{Z})$  defined in Eq. 2.6. This means it will have the form

$$\frac{dZ_{ms}}{dt}(t) = \mathbb{J}_{mj} \frac{\partial \hat{H}}{\partial Z_{js}}(\mathbf{Z}(t); \mathbf{A}) \quad (2.8)$$

To do this, we need to show how the r.h.s. of Eq. 2.7 is the same as the r.h.s. of Eq. 2.8. The starting point is the right hand side of Eq. 2.8:

$$\mathbb{J}_{mj} \frac{\partial \hat{H}}{\partial Z_{js}}(\mathbf{Z}; \mathbf{A}) = \mathbb{J}_{mj} \frac{\partial}{\partial Z_{js}} \int_{\Omega} H(\mathbf{Z}; \mathbf{A}) \mathcal{P}_{\Xi}(\xi) d\xi.$$

First, the integral and the derivative act with respect to different variables, so we can exchange them (all necessary regularity conditions such as being continuously differentiable are assumed to be satisfied):

$$= \int_{\Omega} \mathbb{J}_{mj} \frac{\partial H}{\partial Z_{js}}(\mathbf{Z}; \mathbf{A}) \mathcal{P}_{\Xi}(\xi) d\xi$$

Subsequently, the chain rule is applied to the derivative:  $\frac{\partial \hat{H}}{\partial Z_{js}} = \frac{\partial H}{\partial z_i} \frac{\partial z_i}{\partial Z_{js}}$

$$= \int_{\Omega} \mathbb{J}_{mj} \frac{\partial H}{\partial z_i}(\mathbf{z}; \boldsymbol{\alpha}) \frac{\partial z_i}{\partial Z_{js}}(\mathbf{Z}) \mathcal{P}_{\Xi}(\xi) d\xi$$

Now the result of Eq. 2.5,  $\frac{\partial z_i}{\partial Z_{jk}} = \delta_{ij} \Psi_k$ , is substituted (using the right indices):

$$= \int_{\Omega} \mathbb{J}_{mj} \frac{\partial H}{\partial z_i}(\mathbf{z}; \boldsymbol{\alpha}) \delta_{ij} \Psi_s \mathcal{P}_{\Xi}(\xi) d\xi$$

The Kronecker delta  $\delta_{ij}$  can be simplified leaving only the terms where  $i = j$ :

$$= \int_{\Omega} \mathbb{J}_{mj} \frac{\partial H}{\partial z_j}(\mathbf{z}; \boldsymbol{\alpha}) \Psi_s \mathcal{P}_{\Xi}(\xi) d\xi.$$

Here in the last step again the r.h.s. of Eq. 2.7 is obtained, therefore proving in general terms the ability of  $\hat{H}$  to produce the PCE approximation of the original equations.  $\square$

It is noteworthy how the finiteness of the PC expansion never interferes: at any order, the new system is exactly Hamiltonian [PS13].

## 2.2 The intrusive solution of some Hamiltonian systems is not optimal

However, this Hamiltonian character of the new equations is not necessarily an advantage: in fact, it is shown in [PS13] how for a simple harmonic oscillator system, the coefficients calculated by the intrusive method will necessarily diverge from the best solution, defined as the one with the lowest Galerkin  $L_2$  error norm. This is a consequence of Liouville theorem, which forces an Hamiltonian flux to preserve volume.

### 2.2.1 Outline of derivation

The Hamiltonian of an harmonic oscillator is:

$$H(q, p) = \frac{p^2}{2} + \frac{\omega^2 q^2}{2}. \quad (2.9)$$

This simple system has an exact solution, which for initial conditions  $q(0) = 1, p(0) = 0$  is:

$$\begin{aligned} q_{\text{Sol}}(t; \omega) &= \cos \omega t \\ p_{\text{Sol}}(t; \omega) &= \omega \sin \omega t. \end{aligned}$$

The frequency  $\omega$  is now assumed to be a random variable, uniformly distributed in  $(\omega_0, \omega_1)$ . To generate such a distribution,  $\Lambda$ , a uniformly distributed germ over  $(-1, 1)$ , is used (an example of PC expansion of stochastic coefficents):

$$\begin{aligned} \omega &= \frac{\omega_0 + \omega_1}{2} + \frac{\omega_0 - \omega_1}{2}\Lambda \\ &= \bar{\omega} + L\Lambda. \end{aligned}$$

A PC expansion of  $q(t)$  will look as follows:

$$q(t; \Lambda) = \sum_{k=0}^{N_O} Q_k(t) \Psi_k(\Lambda),$$

where  $\Psi(\Lambda)$  are the orthonormal Legendre polynomials (including the sum symbol explicitly, to avoid confusion with powers, in Einstein notation):

$$\begin{aligned} \Psi_k(\Lambda) &= \sum_{l=0}^k B_{kl} \Lambda^l \\ B_{kl} &= \sqrt{2k+1} 2^k \binom{k}{l} \binom{(k+l-1)/2}{k}. \end{aligned} \quad (2.10)$$

The optimal PC coefficients of the Hamiltonian system,  $Q_k^{\text{Best}}$  can be obtained directly by Galerkin projection of the explicit solution  $q_{\text{Sol}}$  on  $\Psi_k$ :

$$\begin{aligned} Q_k^{\text{Best}} &= \langle q_{\text{Sol}} \Psi_k \rangle / \|\Psi_k\|^2 \\ &= \int_{\Omega} q_{\text{Sol}}(\lambda) \Psi_k(\lambda) P_{\Lambda}(\lambda) d\lambda / 1 \\ &= \int_{-1}^1 \cos((\bar{\omega} + L\lambda)t) \Psi_k(\lambda) \cdot \frac{1}{2} d\lambda . \end{aligned}$$

By linearity, the powers  $\lambda^l$  in the polynomial  $\Psi(\lambda)$  can be integrated separately. Recalling 2.10:

$$\int_{-1}^1 \cos(\bar{\omega}t + L\lambda t) \Psi_k(\lambda) d\lambda = \frac{1}{2} \sum_{l=0}^k B_{kl} \int_{-1}^1 \cos(\bar{\omega}t + L\lambda t) \lambda^l d\lambda$$

Integration by parts of the latter integral returns a recursive formula (Appendix A.1). Denoting with  $I_l$  the projection integral on  $\lambda^l$ :

$$\begin{aligned} I_l &= \int_{-1}^1 \cos(\bar{\omega}t + L\lambda t) \lambda^l d\lambda \\ &= \frac{1}{Lt} [\sin \omega_2 t - (-1)^l \sin \omega_1 t] + \\ &\quad + \frac{l}{(Lt)^2} [\cos \omega_2 t - (-1)^{l-1} \cos \omega_1 t] - \frac{l(l-1)}{(Lt)^2} I_{l-2} . \end{aligned} \tag{2.11}$$

It is easy to verify that  $I_0$  and  $I_1$  go to 0 for  $t \rightarrow \infty$  (Appendix A.1). Therefore, by recursion all the PC coefficients  $Q_k^{\text{Best}}(t) = B_{kl} I_l$  will tend to zero when  $t \rightarrow \infty$ . An analog calculation reveal that  $P_k^{\text{Best}}(t)$  tend to zero too, for  $t \rightarrow \infty$ .

It is important to remark that  $Q_k^{\text{Best}}, P_k^{\text{Best}}$  are not the same coefficients as obtained with the intrusive ODE method outlined before. Rather, the existence of a closed-form solution allows the direct application of Galerkin projection to find the lowest error representation in the polynomial space with respect to the given norm, in this case  $L_2$ .

Instead, the intrusive PC expansion coefficients, denoted  $Q_k^{\text{Intr}}, P_k^{\text{Intr}}$ , are the solution of the Hamiltonian system of Hamiltonian  $\hat{H}$  (2.1). Their trajectory in phase space is a Hamiltonian trajectory: it will be an integral curve of the Hamiltonian field  $X_{\hat{H}} = \mathbb{J} \nabla \hat{H}$ . Liouville's theorem implies that the flux of this field preserves the volume and is divergence-free. Thus, it does not admit asymptotically stable solutions, “sinks” that require  $\nabla \cdot X_{\hat{H}} < 0$  in a neighbourhood of the equilibrium.

Therefore, the behaviour of  $Q_k^{\text{Best}}, P_k^{\text{Best}}$  tending to zero is incompatible with the properties of the solution of any Hamiltonian system. Consequently, in certain Hamiltonian systems, the intrusive solution  $Q_k^{\text{Intr}}(t), P_k^{\text{Intr}}(t)$  cannot achieve the lowest error PC representation  $Q_k^{\text{Best}}(t), P_k^{\text{Best}}(t)$ .

# Chapter 3

## Code overview

In this work, a set of Python scripts has been developed, enabling the user to automatically perform polynomial chaos expansion of an Hamiltonian set of equations. The code is written in Python using the `chaospy` [FL15], `sympy` [Meu+17], `scipy` [Vir+20], and `numpy` [Har+20] Python packages. A high level outline of the script is delineated in Alg. 1.

The user defines a set of germs  $\Xi_i$ , their multivariate probability distribution  $P_\Xi(\xi)$  and their domain  $\Omega$ . The script then uses the Gram–Schmidt algorithm to generate appropriate orthogonal polynomials  $\Psi_k(\Xi)$  up to a given degree  $O$ , as described in section 3.1. The user subsequently defines some Hamiltonian  $H(q, p; \alpha)$  depending on some random parameters  $\alpha$ , usually defined to be a subset of the germs  $\Xi$ . If the Hamiltonian  $H$  has a complex expression, a function is provided which can approximate it as a Taylor series expansion, which easier to integrate.

Then, following the results in chapter 2, the script generates the new  $\hat{H}(Z_{ij})$  and the respective Hamilton equations for the coefficients  $Z_{ij}$ . All the previous steps were performed using `sympy` symbolical algebra capabilities. Subsequently `sympy.lambdify` is used to transform the `sympy` expressions into standard Python/`numpy` functions. The initial conditions  $Z_{ij}(0)$  are calculated by integration of the initial distribution  $z_i(0)$  as shown in equation 1.8:  $Z_{ij}(0) = \int_\Omega z_i(0)\Psi_j P_\Xi d\xi$ . The resulting equations can now be solved directly by a numerical solver. In this script, `scipy`'s `integrate.solve_ivp`, using a 8th order Runge-Kutta algorithm `DOP853`. While the latter it is not a symplectic algorithm, higher order RK methods have better performance, for short to medium times [Rac]. A posteriori, this is justified by the fact that the empirical relative standard deviation of the Hamiltonian over time is  $\frac{\text{std}(H(t))}{\langle H(t) \rangle} \approx 10^{-15}$ . The coefficients  $Z_{ij}(t)$  are used to calculate the average and standard deviation from the formulas 1.11 and 1.12.

Subsequently, Monte Carlo (MC) sampling is used to generate an empirical solution distribution, approximating the true distribution (sec. 3.2). A non-intrusive

polynomial chaos approximation is then fitted over the Monte Carlo samples, using the least-squares solution for the PC coefficients (sec. 3.3). These are then plotted against the intrusive solution for comparison. The energy over time of the two solutions, the intrusive and the non-intrusive, is also plotted.

Lastly, the exponential map 1.18 of the system is generated, truncated at order  $T$ . The exponential map represents the  $T$ -degree polynomial in  $t$  which can best approximate the solution for given initial conditions.

An investigation (sec. 3.5) has been performed for alternative methods to symbolical `sympy` integration, and an alternative version of the script has been implemented. However, the runtime performance suffers compared with symbolic integration.

---

**Algorithm 1** Script high level outline

---

**Input:** Hamiltonian  $H(z, \alpha)$ ; initial conditions  $z(0)$ ; choice of germs  $\Xi_i$ : domain  $\Omega$  and probability  $P_\Xi$ ; PC order  $D$

- 1: (Optional) approximate Hamiltonian to simpler Taylor polynomial
  - 2: Create orthogonal polynomial from germs distributions
  - 3: Generate PC expansion and `sympy` variables
  - 4:  $H_{\text{PCE}} \leftarrow$  Substitute PC expansion into Hamiltonian:  $H(Z_{ij}\Psi_j)$
  - 5:  $\hat{H} \leftarrow$  Expected value of  $H_{\text{PCE}}$  over germs domain:  $\int_\Omega H_{\text{PCE}} P_\Xi(\xi) d\xi$
  - 6:  $X_{\hat{H}} \leftarrow$  Calculate  $\hat{H}$  derivatives field:  $\mathbb{J}\nabla_{Z_{ij}}\hat{H}$
  - 7: Generate PC initial conditions  $Z_{ij}(0)$  from  $z(0)$
  - 8: Transform  $X_{\hat{H}}$  from `sympy` expression to `numpy` function with `sympy.lambdify`
  - 9: Solve resulting numerical ODE
  - 10: Use coefficients  $Z_{ij}(t)$  to calculate intrusive average  $\mu^{\text{PCE}}(t)$ , standard deviation  $\sigma^{\text{PCE}}(t)$
  
  - 11: Monte Carlo:
  - 12: Generate  $N_{\text{MC}}$  Monte Carlo germ samples  $a_i$  and  $H_i^{\text{MC}}(z; a_i)$
  - 13: Solve ODE  $N_{\text{MC}}$  times for solutions  $z_i^{\text{MC}}(t)$
  - 14: Calculate average  $\mu^{\text{MC}}(t)$ , standard deviation  $\sigma^{\text{MC}}(t)$
  
  - 15: Non-intrusive fit:
  - 16: Fit  $Z_{ij}^{\text{NI}}(t)$  to MC samples  $z_i^{\text{MC}}(t)$
  - 17: Use coefficients  $Z_{ij}^{\text{NI}}(t)$  to calculate average  $\mu^{\text{NI}}(t)$ , standard deviation  $\sigma^{\text{NI}}(t)$
  
  - 18: Exponential map:
  - 19: Calculate Poisson brackets recursively:  $g, [H, g], [H, [H, g]], \dots$
  - 20: Sum them to calculate  $M$
  
  - 21: Make plots
-

### 3.1 Polynomial orthogonalization in sympy

---

**Algorithm 2** Gram–Schmidt orthonormalization for polynomials

---

**Input:**  $D$ : Max degree;  $N$ : number of variables; inner product definition: domain  $\Omega$  and pdf  $P_\Xi$

```

basis  $\leftarrow$  Monomial, non-orthogonal, basis, ex:  $[1, x, y, x^2, xy, y^2, \dots, xy^{D-1}, y^D]$ 
polynomials  $\leftarrow$  Orthogonalized basis: [ basis[0] ]
snorms  $\leftarrow$  squared norms:  $[1, 1, \dots]$ 
inner( $\Psi_1, \Psi_2$ )  $\leftarrow$  Expected value of product:  $\int_{\Omega} \Psi_1(\xi) \Psi_2(\xi) P_\Xi(\xi) d\xi$ 
for  $i \leftarrow 1$  to  $\text{length}(\text{basis})$  do ▷ For each monomial...
    for  $j \leftarrow 0$  to  $i$  do
        projcoeff  $\leftarrow$  inner(basis[ $i$ ], polynomial[ $j$ ]) ▷ Project on previous elements
        basis[ $i$ ]  $\leftarrow$  basis[ $i$ ]  $-$  projcoeff  $* \frac{\text{polynomial}[j]}{\text{snorms}[j]}$  ▷ Subtract parallel parts
    end for
    snorms[ $i+1$ ]  $\leftarrow$  inner(basis[ $i$ ], basis[ $i$ ])
    basis[ $i$ ]  $\leftarrow \frac{\text{basis}[i]}{\sqrt{\text{snorms}[i+1]}}$  ▷ Renormalize
    polynomial[ $i+1$ ]  $\leftarrow$  basis[ $i$ ]
end for

```

---

**Output:** tuple (polynomials, snorms)

---

### 3.2 Monte Carlo samples

The true empirical distribution of the values at time  $t$   $z_i(t)$  can be estimated by Monte Carlo sampling. First,  $N_{MC}$  samples  $\Xi[n]$  of the germs are generated (array index notation [ $n$ ] is used to distinguish independent Monte Carlo sample sets from the proper mathematical indices of the quantities in play). They are substituted into the Hamiltonian to generate several possible sets of equations  $z'[n] = \mathbb{J}\nabla(H(z; \Xi[n]))$  for each sample  $n = 1 \dots N_{MC}$ . The respective set of solutions,  $z_i[n](t)$ , represents an empirical sample of the distribution  $z(t)$ . From these, the average  $\mu^{MC}(t)$  and standard deviation  $\sigma^{MC}(t)$  can be calculated.

The number of MC samples  $N_{MC} = 600$  has been chosen empirically as to have a low enough standard deviation error of the mean  $\mu_{MC}$  averaged in time. In figure Fig. 3.1, this averaged  $\bar{\sigma}_\mu = \bar{\sigma}_{MC}/\sqrt{N_{MC}}$  has been plotted with respect to the number of MC samples. At 600 samples MC error is less than 1% of the characteristic scale of the system, and this is deemed an acceptable compromise.

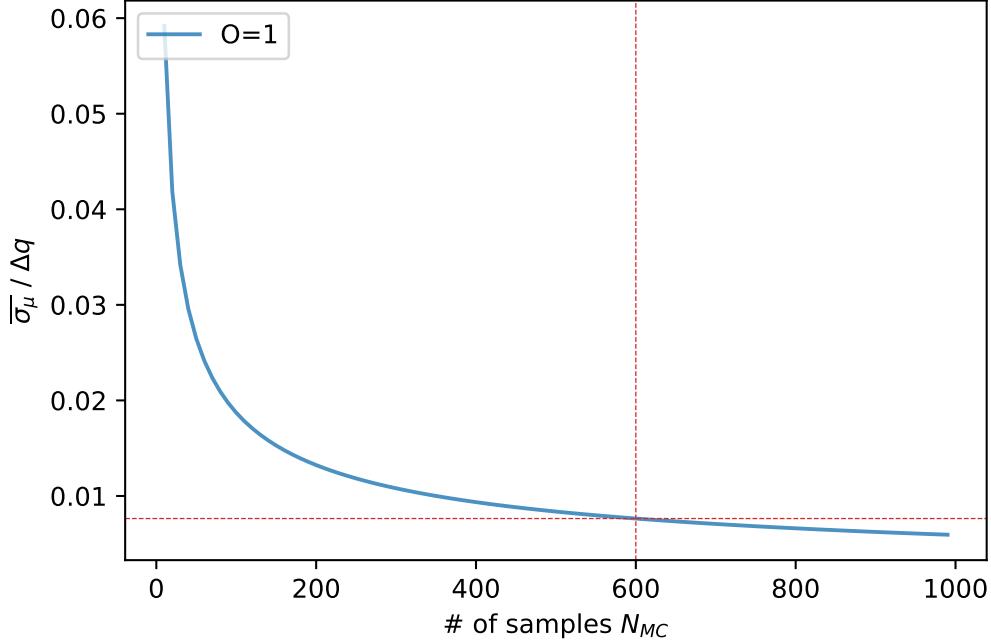


Figure 3.1: Harmonic Hamiltonian  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ , runtime 15 seconds. 1000 MC samples are used to calculate the time averaged empirical standard deviation  $\bar{\sigma}_{\text{MC}}$  of the distribution. This is then used to estimate the time averaged standard deviation of the mean  $\bar{\sigma}_\mu = \bar{\sigma}_{\text{MC}}/\sqrt{N_{\text{MC}}}$ . The result  $\bar{\sigma}_\mu$  is then divided by the characteristic scale of the system  $\Delta q$ , in this case the average oscillation amplitude  $\Delta q \approx 2$ . At 600 samples, the error of the mean is thus  $\approx 0.8\%$  of the characteristic scale.

---

**Algorithm 3** Monte Carlo empirical sampling

**Input:**  $N_{\text{MC}}$ : number of samples; *distr*: germs distribution, *timesteps*: solution evaluation times

```

samples[NMC] ← Sample  $N_{\text{MC}}$  germs from distribution
HMC[i] ←  $H(z; \alpha = \text{samples}[i])$ 
y0MC[i] ←  $y_0(\alpha = \text{samples}[i])$ 
for  $i \leftarrow 0$  to  $N$  do
    sol[:, :, i] ← solve_ODE( $\mathbb{J}\nabla(H^{\text{MC}}[i]), y_0^{\text{MC}}[i], \text{timesteps}$ )
end for

```

---

**Output:** array *sol* (dimensions=[time,  $Z_{ij}$ , samples])

---

### 3.3 Non-intrusive fit

The non-intrusive coefficients fitting is done using the point collocation method [Feib]. The germ samples  $\Xi[n]$  generated by Monte Carlo are substituted in the polynomials  $\Phi_j(\Xi[n])$ . Then, the coefficients  $Z_{ij}^{\text{NI}}(t)$  are the least-squares solution of the (overdetermined) linear system of equations [Feib]:

$$z_i[n](t) = \sum_j Z_{ij}^{\text{NI}}(t) \Phi_j(\Xi[n]) \quad \text{for } n = 1, \dots, N . \quad (3.1)$$

Defining the following matrices 3.2

$$\begin{aligned} \hat{Z}^{\text{NI}} &= Z_{ij}^{\text{NI}} \\ \hat{P} &= [\Phi_j(Q[n])]_{jn} \\ \hat{z} &= [z_i[n]]_{in} , \end{aligned} \quad (3.2)$$

the least squares solution, at each time  $t$ , can be expressed in matrix notation as [Feib]:

$$\hat{Z}^{\text{NI}}(t) = (\hat{P}^T \hat{P})^{-1} \hat{P}^T \hat{z}(t) \quad (3.3)$$

### 3.4 Exponential map

The approximated exponential map is calculated by recursively calculating the poisson brackets then summing them, as described in equation 1.18.

### 3.5 Symbolical and numerical quadrature

It is very advantageous if the integral in 2.6 can be done symbolically, as this presents the biggest advantage in terms of performance. However, it can be really difficult to find a closed form expression. Even if this is theoretically possible, the number of variables, terms, and time necessary all increase exponentially as the polynomial degree and number of variables increases: the number of  $n$ -degree polynomial terms in  $d$  variables is of the order of  $\binom{n+d}{d}$ .

As a consequence, performing the integral symbolically using computer algebra may take a long time, even though it would be necessary only once, for a given  $H$  and germs. The integration of the simple harmonic oscillator Hamiltonian described in section 4.3 on a laptop ranges from a few seconds for PC expansion order  $O = 1$ , to 2 hours for  $O = 5$ . More complex Hamiltonians scale even worse: the

Duffing non linear oscillator Hamiltonian (Sec. 4.4) requires 5 hours at only  $O = 3$ .

An attempt has been made to parallelize the symbolic computation using the Python library `dask` [Das16], by exploiting the linearity of the integral to integrate different parts of a sum separately. The procedure is outlined in Alg. 4 (using Python array slicing notation  $[a : b]$ ) and a flowchart is shown in Fig. 3.2.

While in some cases this delivered up to a 2x speedup (using 4 threads), often there was no clear speed benefit, and sometimes the isolated integration of some parts failed to derive a closed form (while the whole expression would have one). This approach seems thus to run counter certain `sympy` assumptions.

---

**Algorithm 4** dask parallel integration of  $F(Z) = \int_{\Omega} f(Z, \xi) d\Xi$ 


---

**Input:**  $N_{\text{threads}}$ : number of threads to use;  $f$ : expression to integrate

```

1: single_pieces_num  $\leftarrow \text{length}(f.\text{args})$ 
2: pieces_per_thread  $\leftarrow \text{single\_pieces\_num} \div N_{\text{threads}}$ 
3: remaining_pieces  $\leftarrow \text{single\_pieces\_num \% } N_{\text{threads}}$ 

4: for  $i \leftarrow 0$  to  $N_{\text{threads}}$  do                                 $\triangleright$  Divide for  $N_{\text{threads}}$  threads
5:    $f\_pieces[i] \leftarrow \text{sum}(f.\text{args}[ i*pieces\_per\_thread : (i+1)*pieces\_per\_thread ])$ 
6: end for

7: for  $r \leftarrow 0$  to  $remaining\_pieces$  do           $\triangleright$  Sum remaining addends
8:    $f\_pieces[r] \leftarrow f\_pieces[r] + f.\text{args}[ N_{\text{threads}}*pieces\_per\_thread+r ]$ 
9: end for

10:  $F\_pieces[] \leftarrow \text{Empty list } []$ 
11: for  $i \leftarrow 0$  to  $\text{length}(f\_pieces)$  do            $\triangleright$  Construct compute graph
12:    $F\_piece \leftarrow \text{dask.delayed}(\text{sympy.integrate})(f\_pieces[i], \Omega)$ 
13:    $F\_pieces.append(F\_piece)$ 
14: end for
15:  $F \leftarrow \text{dask.delayed}(\text{sum})(F\_pieces)$             $\triangleright$  Create summation node
16:  $\text{output} \leftarrow F.\text{compute}()$                        $\triangleright$  Start computation

```

**Output:**  $F(Z) = \int_{\Omega} f(Z, \xi) d\Xi$

---

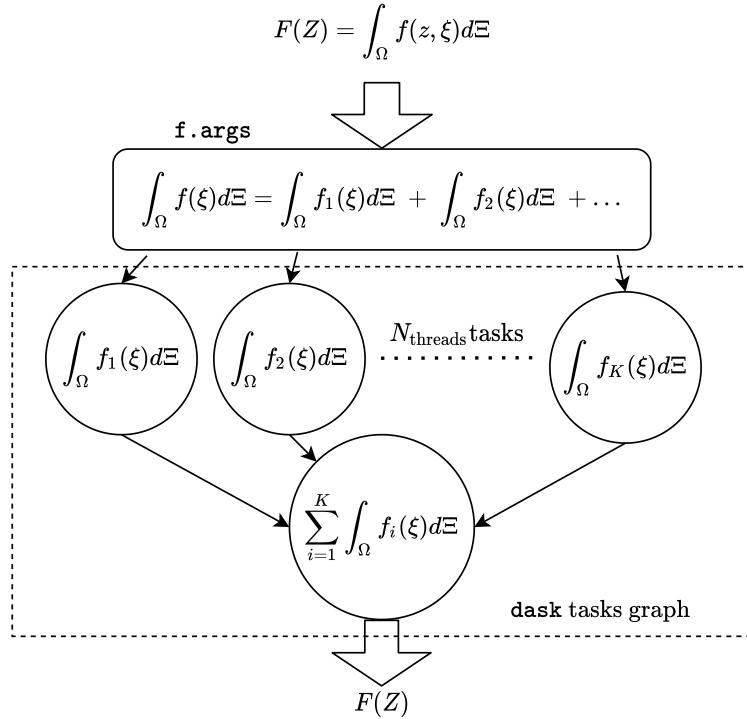


Figure 3.2: Flowchart representing the `dask` computation graph of the attempted parallel integration algorithm. When the integrand  $f$  is a `sympy.add` expression, its summands can be recovered using its `f.args` method. These can then be subdivided over  $N_{\text{threads}}$  independent `dask` nodes to be integrated. The results are used by the final node, which sums them to recover the integral function  $F(Z) = \int_{\Omega} f(Z, \xi) d\xi$ .

### Numerical quadrature

An alternative numerical quadrature implementation has been explored. The expected value integral is approximated by an appropriately weighted sum of integrand evaluations at a grid of points:

$$\langle f(\Xi) \rangle_{\Xi} = \int_{\Omega} f(\xi) p_{\Xi}(\xi) d\xi \approx \sum_i w_i f(\xi_i) .$$

The Hamilton equations  $\dot{Z} = \mathbb{J} \frac{\partial \hat{H}}{\partial Z}$  will have the form

$$\dot{Z} = \sum_i w_i \mathbb{J} \frac{\partial H_{\text{PCE}}}{\partial Z}(\xi_i) \quad (3.4)$$

where  $H_{\text{PCE}}$  represents the original Hamiltonian  $H$  after the PC expansions have been substituted, but before the germ random variables  $\Xi$  are averaged out:

$$\begin{aligned} H_{\text{PCE}}(Z, \Xi) &= H(Z_{ij}\Psi_j(\Xi); A_{mn}\Psi_n(\Xi)) \\ \hat{H}(Z) &= \langle H_{\text{PCE}} \rangle_{\Xi} . \end{aligned}$$

It is important to notice how, integrating numerically the equations 3.4, it is necessary to perform a sum over many terms at each time step. This degrades the performance of the intrusive PCE method, and often loses the speed advantage over Monte Carlo methods.

However, it may be useful to study the properties of intrusive PCE approximations when the expected value integral is too complex to evaluate in closed form, or if the evaluation would require too much time to be practical. In Subsec. 4.5.3 the numerical quadrature method is applied to a three-germ Hamiltonian, showing an example where the method bypasses the cumbersome calculation of a 6-dimensional 3-germs Hamiltonian.

Further explorations were performed in utilizing code transformation libraries such as `zygote.jl` and Google’s `jax`, that enable the user to perform automatic differentiation of general functions. The goal was to explore an analogue to automatic differentiation for integration, however, this exploration was not fruitful. A successful development of a procedural (as opposed to numeric or symbolic) integration may greatly facilitate the integration with numerical/automatic algebra frameworks such as `DA`.

# Chapter 4

## Results

The script has been run on several typical Hamiltonians with some random parameters. Specifically, the first-order harmonic oscillator Hamiltonian, the third-order Duffing oscillator, and the relativistic Hamiltonian of a non-linear quadrupole magnet, from particle accelerators. There are three general types of plots that are produced by the scripts for each Hamiltonian at various orders.

### 4.1 Types of plots

#### 4.1.1 Comparison of uncertainty with Monte Carlo and non-intrusive methods

In this plot, a comparison is made between the uncertainty estimated from the three methods: the intrusive one, the non-intrusive, and Monte Carlo simulation. The latter is assumed to represents the best approximation to the empirical truth, for a high enough number of samples. The plot represent single Monte Carlo trajectories as faint blue lines. From this set a band of uncertainty, composed by the Monte Carlo average over time  $\mu^{\text{MC}}(t)$  plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ , is calculated and plotted in blue.

In the same vein, the intrusive average over time  $\mu^{\text{PCE}}(t)$  plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ , calculated from the identities 1.11 and 1.12, is plotted as a red band. Likewise, from the non-intrusive coefficients estimation a green band  $\mu^{\text{NI}}(t) \pm \sigma^{\text{NI}}(t)$  is plotted.

#### 4.1.2 Energy comparison

The plot shows the values over time of the PC Hamiltonian  $\hat{H}(Z_{ij})$  for the intrusive and non-intrusive methods. The red line uses the intrusive coefficients  $Z_{ij} = Z^{\text{PCE}}$ ,

while the green line uses the non-intrusive ones  $Z_{ij} = Z^{\text{NI}}$  in the calculations.

### 4.1.3 Exponential maps

This plot shows an application of the exponential map generated by the script, and compares it to the numerical solution of the differential equation. The difference between the numerical solution and the exponential map is that the numerical solution is obtained by an ODE integrating algorithm (such as Runge-Kutta) evolving the initial conditions in time steps. On the other hand, an exponential map  $M(t) = M(t; Z_{ij}(0))$  is a polynomial function of time of order  $T$ , parametrized by the phase space coordinates. When applied to some initial conditions, it will approximate the solution of the differential equation over time, with an error of order  $O(T + 1)$ .

The uncertainty bands are plotted from the coefficients as described above: blue is Monte Carlo, red is the intrusive numerical solution, and yellow and grey two possible orders  $T$  of the approximate exponential map applied to the same initial condition as the ODE.

## 4.2 General observations

Even in the simple case of the harmonic oscillator, after a long enough time the solution will deviate from an accurate description of the system, as seen in section 2.2. However, a higher PC expansion order leads to an accurate description of the system for a longer time.

This is confirmed empirically in plots such as 4.2, 4.5, 4.7, describing the PC expansion of an harmonic oscillator at order  $O = 1, 2, 4$  respectively. These plots show how the PC expansion is able to track the average and uncertainty of the system for longer periods of time, for higher order expansions.

In plot 4.1 the absolute difference over time between the standard deviation from intrusive PCE,  $\sigma_{\text{PCE}}^2(t)$ , and the Monte Carlo standard deviation,  $\sigma_{\text{MC}}^2(t)$ , is plotted: higher orders diverge later.

### 4.2.1 The initial conditions and the order of expansion

Increasing the order of the PC expansion does not necessarily improve the approximation accuracy of the resulting coefficients. In fact, if the initial conditions are expressed as low-order PC expansions themselves, many coefficients will be zero, and their derivative will be too, so that  $Z_{ij}(t) = 0$  for some  $i, j, \forall t$ . This means that the trajectory of the system is constrained in a lower dimensional submanifold, precisely the one which contains the trajectory at lower orders.

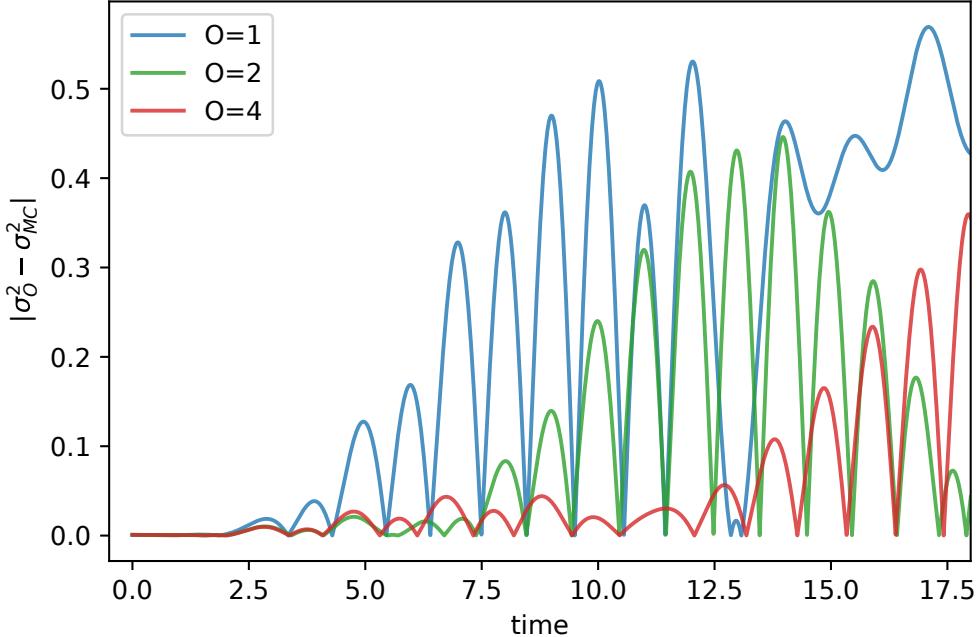


Figure 4.1: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Plotted:  $|\sigma_{\text{PCE}}^2(t) - \sigma_{\text{MC}}^2(t)|$ , for order  $O = 1, 2, 3$ . Higher orders diverge later.

A practical example: the harmonic oscillator Hamiltonian with uncertain frequency,

$$H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2},$$

where the parameter  $k_0$  is distributed uniformly as  $k_0 \sim \mathcal{U}(\text{low} = 3.5, \text{high} = 4.5)$ . This parameter is used as one of the two germs

$$\Xi_i : \left[ \mathcal{U}(l = 3.5, h = 4.5), \mathcal{N}(\mu = 0, \sigma = 1) \right]$$

where the other is a normal Gaussian variable.

Applying PC expansion at order  $O$ , we obtain the Hamiltonians  $\hat{H}_O$  shown in appendix B.1. The initial conditions, as calculated in 1.8, will depend on the PC expression of the initial distribution  $z_i(0)$ . A deterministic single-point initial condition will be for example

$$q(0) = 1, p(0) = 0 \tag{4.1}$$

or using Dirac deltas,  $P_{z(0)}(q, p) = \delta(q - 1)\delta(p)$ . A Gaussian initial distribution around 1 in  $q(t)$  could be constructed as such, using the second, Gaussian, germ:

$$q(0) = 1 + \frac{1}{5}\Xi_1, \quad p(0) = 0 \quad (4.2)$$

These initial conditions are thus represented by polynomial chaos expansion as respectively a zeroth and first-order polynomial. More complex initial distributions would involve higher powers  $\Xi_i^n$  of the germs. These initial distributions results in initial values for the coefficients that are zero after a certain index  $i$ . Using third-degree polynomials, the initial conditions for the  $N_O = 20$  coefficients, from the zeroth order initial conditions 4.3 we get:

$$[Q_0, \dots, Q_9, P_0, \dots, P_9](t = 0) = [1, 0, \dots, 0] . \quad (4.3)$$

Similarly, from the first order initial conditions 4.4 we get:

$$[Q_0, \dots, Q_9, P_0, \dots, P_9](t = 0) = [1, \frac{1}{5}, 0, \dots, 0] . \quad (4.4)$$

By deriving the Hamiltonian of the fourth-order expansion  $\hat{H}_4^{\text{Har}}$  (appendix B.1), the time derivatives of the coefficients are obtained. The resulting Hamilton equations are:

$$\begin{aligned} \frac{dZ_i}{dt}(t) = \mathbb{J} \frac{\partial H_3^{\text{Har}}}{\partial Z_i} = & \left[ P_0, P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9, P_{10}, P_{11}, P_{12}, P_{13}, \right. \\ & P_{14}, -\frac{5Q_0}{2} - \frac{\sqrt{3}Q_2}{3}, -\frac{5Q_1}{2} - \frac{\sqrt{3}Q_4}{3}, -\frac{\sqrt{3}Q_0}{3} - \frac{5Q_2}{2} - \frac{2\sqrt{15}Q_5}{15}, -\frac{5Q_3}{2} - \\ & -\frac{\sqrt{3}Q_7}{3}, -\frac{\sqrt{3}Q_1}{3} - \frac{5Q_4}{2} - \frac{2\sqrt{15}Q_8}{15}, -\frac{2\sqrt{15}Q_2}{5Q_5} - \frac{15}{3\sqrt{35}Q_9}, -\frac{\sqrt{3}Q_{11}}{3} - \\ & \frac{5Q_6}{2} - \frac{2\sqrt{15}Q_{12}}{15} - \frac{\sqrt{3}Q_3}{3} - \frac{5Q_7}{2}, -\frac{3\sqrt{35}Q_{13}}{35} - \frac{2\sqrt{15}Q_4}{5Q_8} - \frac{5Q_8}{2}, -\frac{4\sqrt{7}Q_{14}}{21} - \\ & \frac{3\sqrt{35}Q_5}{2} - \frac{5Q_9}{2}, -\frac{5Q_{10}}{2}, -\frac{5Q_{11}}{2} - \frac{\sqrt{3}Q_6}{3}, -\frac{5Q_{12}}{2} - \frac{2\sqrt{15}Q_7}{15}, -\frac{5Q_{13}}{2} - \\ & \left. -\frac{3\sqrt{35}Q_8}{35}, -\frac{5Q_{14}}{2} - \frac{4\sqrt{7}Q_9}{21} \right] \end{aligned}$$

A generic point in this fourth-order coefficient space is generally  $15 \cdot 2 = 30$ -dimensional. However, for some initial conditions such as 4.3, the solution is constrained in a submanifold of the Hamiltonian surface. More specifically, by setting the subset of the coefficients

$$\begin{aligned} P_1, P_3, P_4, P_6, P_7, P_8, P_{10}, P_{11}, P_{12}, P_{13} &= 0 \\ Q_1, Q_3, Q_4, Q_6, Q_7, Q_8, Q_{10}, Q_{11}, Q_{12}, Q_{13} &= 0 , \end{aligned} \quad (4.5)$$

the derivatives become:

$$\begin{aligned} \frac{dZ_i}{dt}(t) &= \mathbb{J} \frac{\partial H_3^{\text{Har}}}{\partial Z_i} \\ &= \left[ P_0, 0, P_2, 0, 0, P_5, 0, 0, 0, P_9, 0, 0, 0, 0, P_{14}, \right. \\ &\quad -\frac{5Q_0}{2} - \frac{\sqrt{3}Q_2}{3}, 0, -\frac{\sqrt{3}Q_0}{3} - \frac{5Q_2}{2} - \frac{2\sqrt{15}Q_5}{15}, 0, 0, \\ &\quad -\frac{2\sqrt{15}Q_2}{15} - \frac{5Q_5}{2} - \frac{3\sqrt{35}Q_9}{35}, 0, 0, 0, \\ &\quad \left. -\frac{4\sqrt{7}Q_{14}}{21} - \frac{3\sqrt{35}Q_5}{35} - \frac{5Q_9}{2}, 0, 0, 0, 0, -\frac{5Q_{14}}{2} - \frac{4\sqrt{7}Q_9}{21} \right]. \quad (4.6) \end{aligned}$$

It is easy to verify that all the derivatives of the 4.5 coefficients are now zero, for all times. This means that a trajectory whose initial condition satisfy 4.5, such as 4.3, will be constrained in the submanifold that is the intersection of the Hamiltonian manifold defined by the constraint  $\hat{H}_4^{\text{Har}}(Z(t)) = \hat{H}_4^{\text{Har}}(Z(0))$  and the hyperplane 4.5. The effective Hamiltonian of this fourth-order system looks remarkably similar to the second-order one  $\hat{H}_2^{\text{Har}}$  (appendix B.1), albeit with coefficients for different polynomials:

$$\begin{aligned} \hat{H}_{\text{eff}}^{\text{Har}} &= \frac{P_0^2}{2} + \frac{P_{14}^2}{2} + \frac{P_2^2}{2} + \frac{P_5^2}{2} + \frac{P_9^2}{2} + \frac{5Q_0^2}{4} + \frac{\sqrt{3}Q_0Q_2}{3} + \frac{5Q_{14}^2}{4} + \frac{4\sqrt{7}Q_{14}Q_9}{21} + \\ &\quad + \frac{5Q_2^2}{4} + \frac{2\sqrt{15}Q_2Q_5}{15} + \frac{5Q_5^2}{4} + \frac{3\sqrt{35}Q_5Q_9}{35} + \frac{5Q_9^2}{4} \end{aligned} \quad (4.7)$$

Likewise, for first-order Gaussian initial conditions 4.4,  $Q_1(0) \neq 0$ , but still many coefficients will be zero for all times:

$$\begin{aligned} P_3, P_6, P_7, P_{10}, P_{11}, P_{13} &= 0 \\ Q_3, Q_6, Q_7, Q_{10}, Q_{11}, Q_{13} &= 0. \end{aligned} \quad (4.8)$$

The effective Hamiltonian looks similar to the third-order one  $\hat{H}_3^{\text{Har}}$  (appendix B.1):

$$\begin{aligned} \hat{H}_{\text{eff}}^{\text{Har}} &= \frac{P_0^2}{2} + \frac{P_{14}^2}{2} + \frac{P_2^2}{2} + \frac{P_5^2}{2} + \frac{P_9^2}{2} + \frac{5Q_0^2}{4} + \frac{\sqrt{3}Q_0Q_2}{3} + \frac{5Q_{14}^2}{4} + \frac{4\sqrt{7}Q_{14}Q_9}{21} + \\ &\quad + \frac{5Q_2^2}{4} + \frac{2\sqrt{15}Q_2Q_5}{15} + \frac{5Q_5^2}{4} + \frac{3\sqrt{35}Q_5Q_9}{35} + \frac{5Q_9^2}{4} \end{aligned} \quad (4.9)$$

Considering this examples, an observation can be made: equations 4.3 and 4.4 show that the lower the PC order of the physical initial conditions 4.1 and 4.2, the more initial coefficents values  $Z_i(0)$  start at zero.

This can then lead to very sparse systems, where many coefficients are zero. This fact may help mitigate the curse of dimensionality by lessening the number of coefficients needed for some PC order. An effective gain of one or two expansion orders may be possible, if one were to ascertain a priori these submanifolds and use reduced Hamiltonians, as 4.7 and 4.9 show.

### 4.2.2 Long term analysis and symplectification

A proper exponential map of infinite order is symplectic, but this is usually not true of its truncations at finite order. There are procedures to symplectify polynomial Hamiltonian truncated maps (so called “Cremona maps”) [Bla02]. This process is called symplectification.

However, as proven in section 2.2, the best possible solution, obtained through non-intrusive means, may not be Hamiltonian. Moreover, plots such as the ones in Sec. 4.3 show the behaviour of intrusive PC methods on longer time scales: as already evidenced in the literature [PS13], intrusive PC expansion is often not able to track the average of the system after some time. The validity of intrusive PC methods seems therefore limited to short to medium times. For these two reasons, symplectifying the exponential map (which involves a short term accuracy/long term stability tradeoff) may not improve the accuracy, and in practice could even worsen it.

### 4.3 Harmonic oscillator

The harmonic oscillator Hamiltonian with uncertain frequency is  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ , where in our example the parameter  $k_0$  is distributed uniformly as  $k_0 \sim \mathcal{U}(\text{low} = 3.5, \text{high} = 4.5)$ . It is PC expanded in the 2 germs  $\Xi_i : [\mathcal{U}(l = 3.5, h = 4.5), \mathcal{N}(\mu = 0, \sigma = 1)]$ . The first germ is the parameter  $k_0$ , the second is used in the stochastic initial condition. The resulting PC Hamiltonians  $\hat{H}_0^{\text{Har}}$  are shown in appendix B.1.

Comparing the results from different expansion orders is a useful method to discover the maximum valid runtime: observing the divergence from a lower PC order in principle enables the user to bound the approximation error of the higher one.

Note how using a zeroth order pointwise initial condition would be perfectly legitimate, in general there is no need for all the germs to appear in the model explicitly. It may thus be useful to include e.g. a Gaussian germ even if no parameter has such distribution, since given the central limit theorem many empirical distributions are Gaussian-like.

The initial condition in the uncertainty plots such as Fig. 4.2 is the stochastic first-order Eq. 4.2: this shows how PCE can transport a particle distribution. The higher order expansion Fig. 4.7 can track for longer the empirical standard deviation, with respect to the low order one Fig. 4.2.

As evidenced in 4.2.2, the intrusive PCE can not track the average at all times, compared with the non-intrusive method. The non-intrusive energy (Fig. 4.4) seems to oscillate in time, compared with the constant energy of the symplectic intrusive solution. Perhaps this is what enables the non-intrusive fit to better track the empirical average, albeit both PCE methods fail to track the uncertainty after some time.

Fig. 4.6 shows how a high-order exponential map is able to transport a distribution over an entire oscillation. To estimate the order required, comparing the error to a lower-order expansion may be useful. For example, referring to Fig. 4.6, observing the divergence over time between  $T = 18$  and  $T = 20$  may lead to conclude that  $T = 20$  is good up until  $t = 4$  or so. The same technique may be employed at the PC expansion order level, to understand for which times the PC expansion of a certain degree is valid.

### 4.3.1 First-order

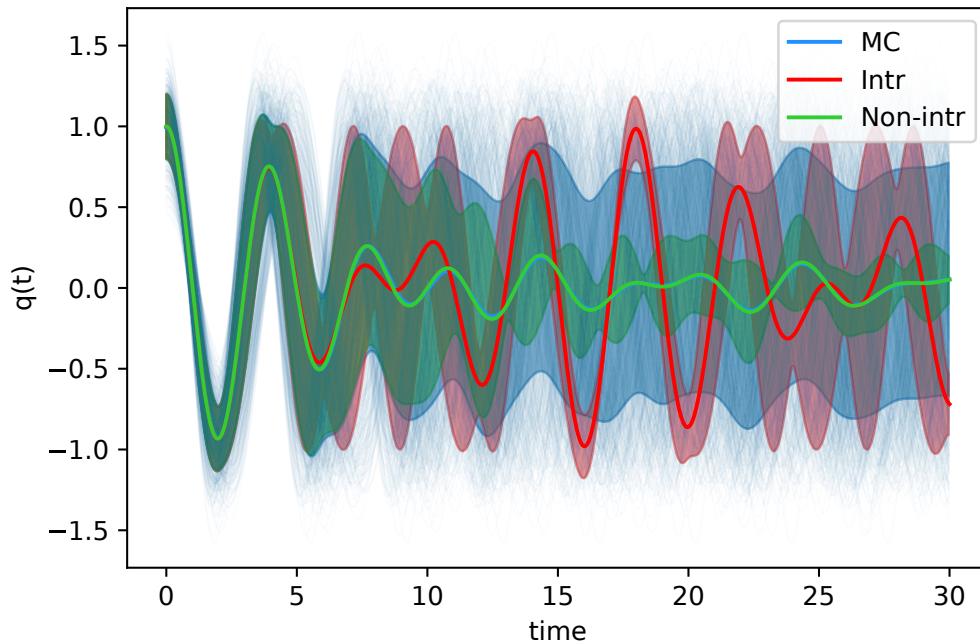


Figure 4.2: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

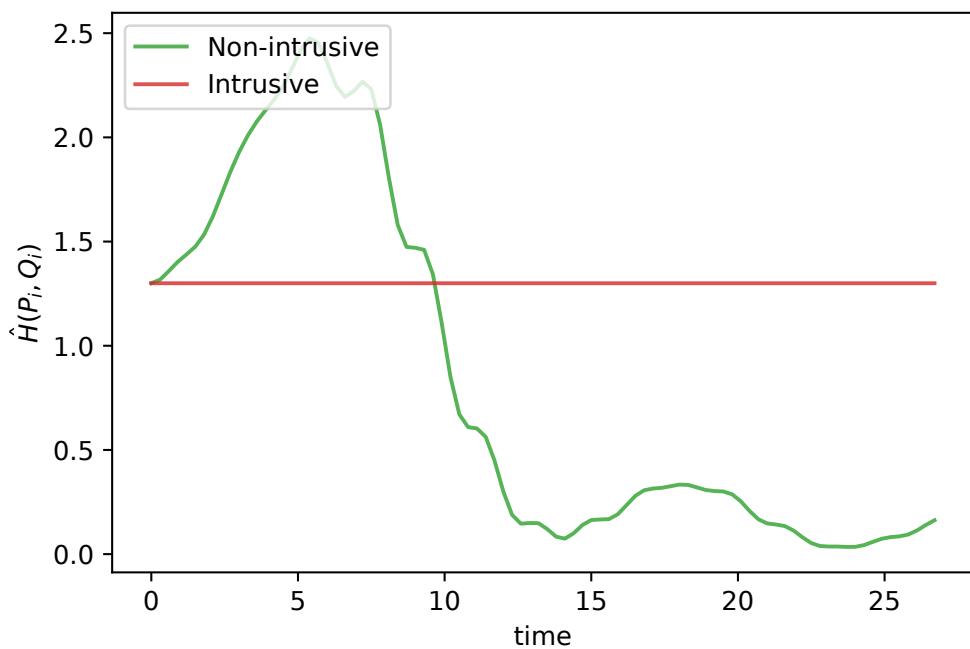


Figure 4.3: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Comparison of the value over time of the PC Hamiltonian  $\hat{H}(t)$ . In red the intrusive solution, where the Hamiltonian is constant. In green, the non-intrusive solution.

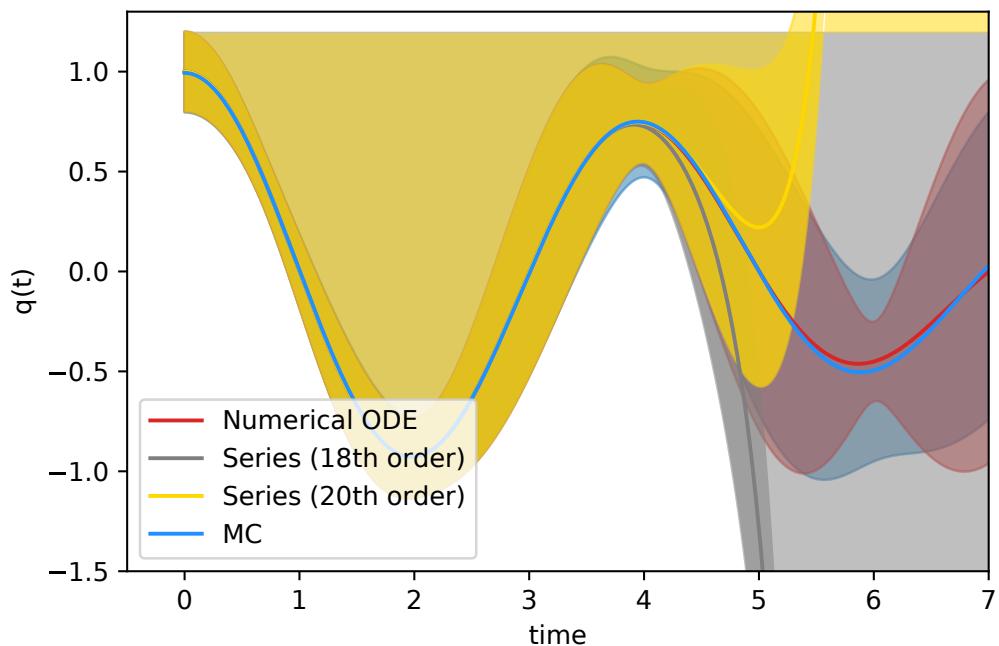


Figure 4.4: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Examples of exponential maps. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Yellow and grey lines and bands: non-intrusive average plus/minus standard deviation from exponential maps applied to the same initial conditions as the ODE.

### 4.3.2 Second order

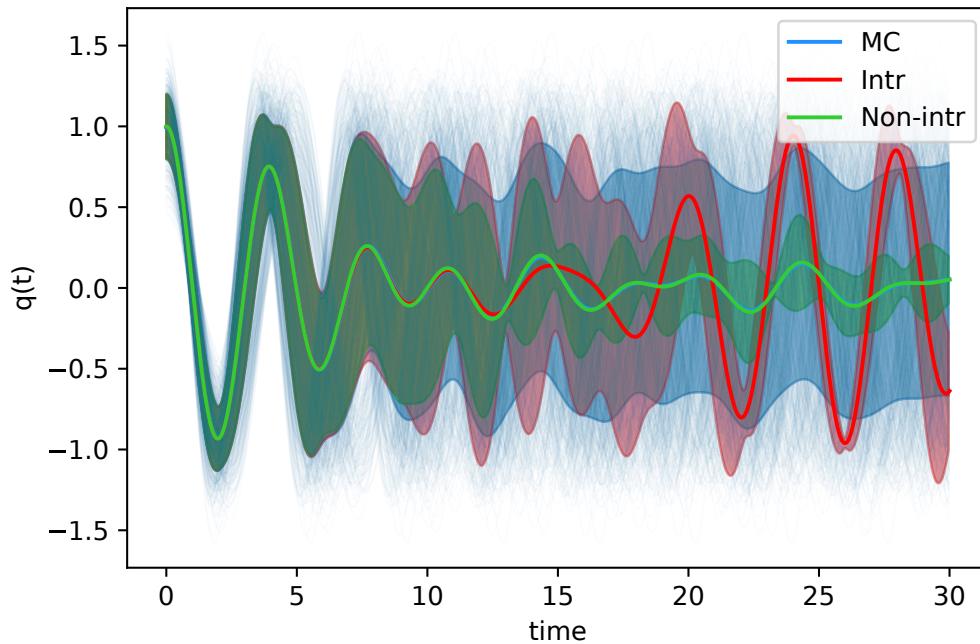


Figure 4.5: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

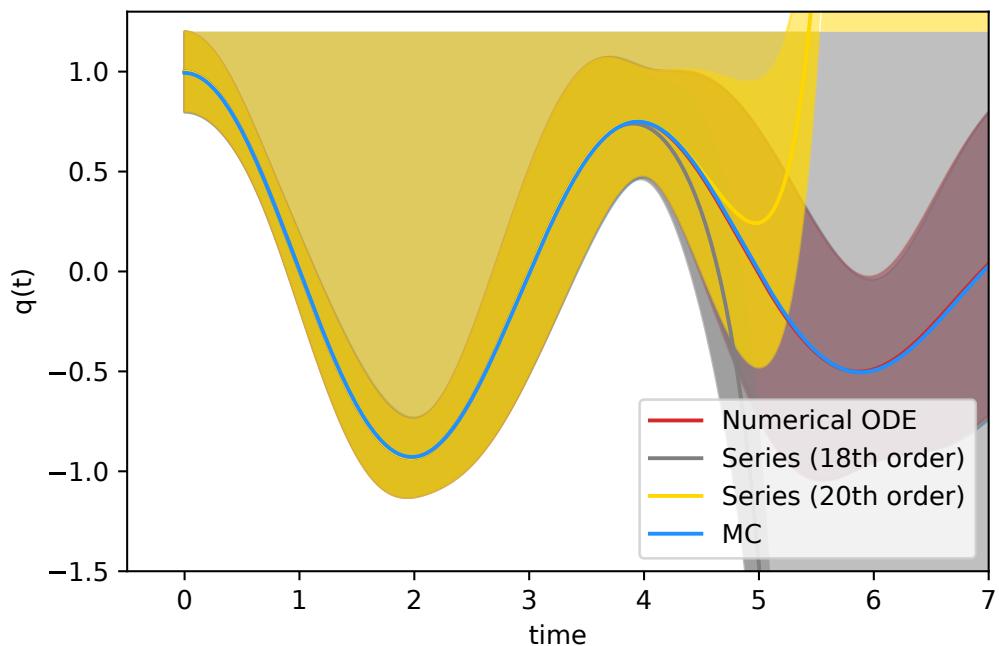


Figure 4.6: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Examples of exponential maps. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Yellow and grey lines and bands: non-intrusive average plus/minus standard deviation from exponential maps applied to the same initial conditions as the ODE.

### 4.3.3 Fourth order

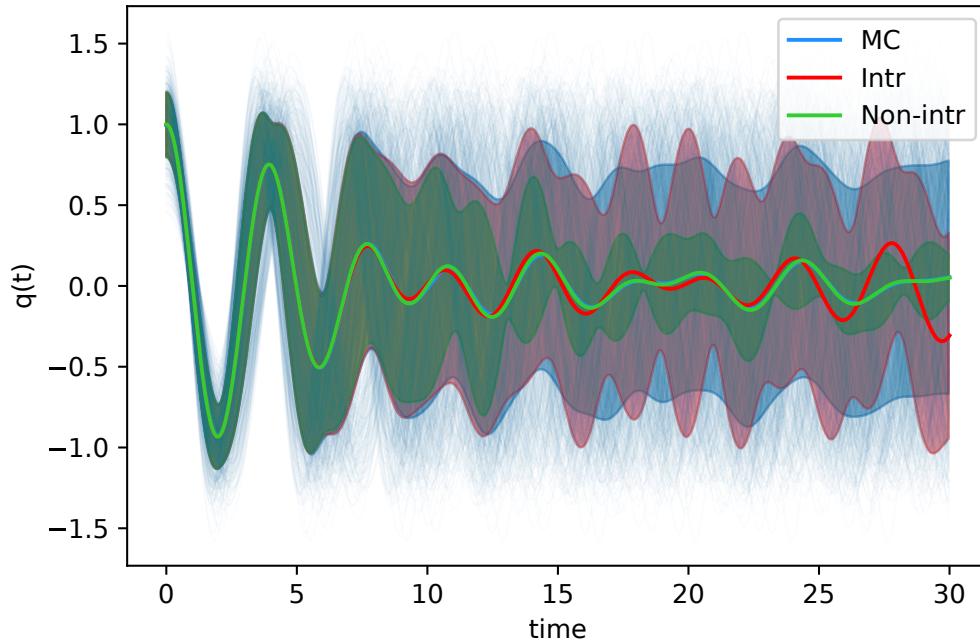


Figure 4.7: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Stochastic initial conditions. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

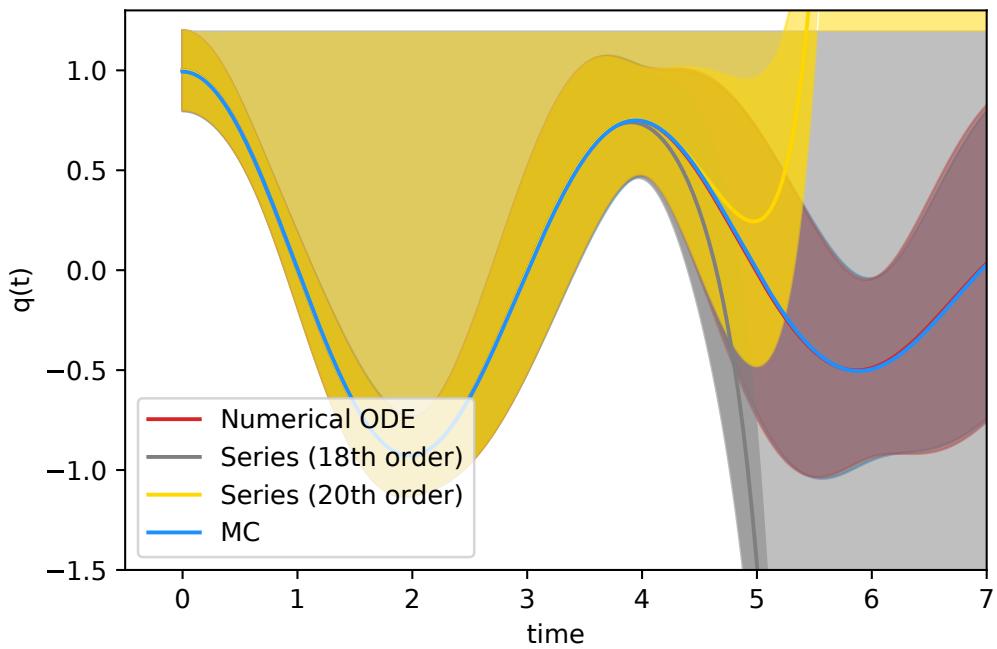


Figure 4.8: Harmonic Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ . Stochastic initial conditions. Examples of exponential maps. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Yellow and grey lines and bands: non-intrusive average plus/minus standard deviation from exponential maps applied to the same initial conditions as the ODE.

## 4.4 Duffing oscillator

The Duffing anharmonic oscillator has a fourth-degree Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ , where the parameter  $k_0$  is distributed uniformly as  $k_0 \sim \mathcal{U}(\text{low} = 3.5, \text{high} = 4.5)$ . It is PC expanded in the 2 germs  $\Xi_i : [\mathcal{U}(l = 3.5, h = 4.5), \mathcal{N}(\mu = 0, \sigma = 1)]$ . The resulting PC Hamiltonians  $\hat{H}_O$  are shown in appendix B.1. The Gaussian initial conditions 4.2 are used.

This is a simple example of a Hamiltonian generating a non-linear ODE system. There is no notable difference in PCE behaviour with respect to the harmonic oscillator. In particular, there is no necessity of a given order for an appropriate description of the system: at short times the first order expansion of Fig. 4.9 is as good as the third order Fig. 4.14 expansion.

#### 4.4.1 First-order

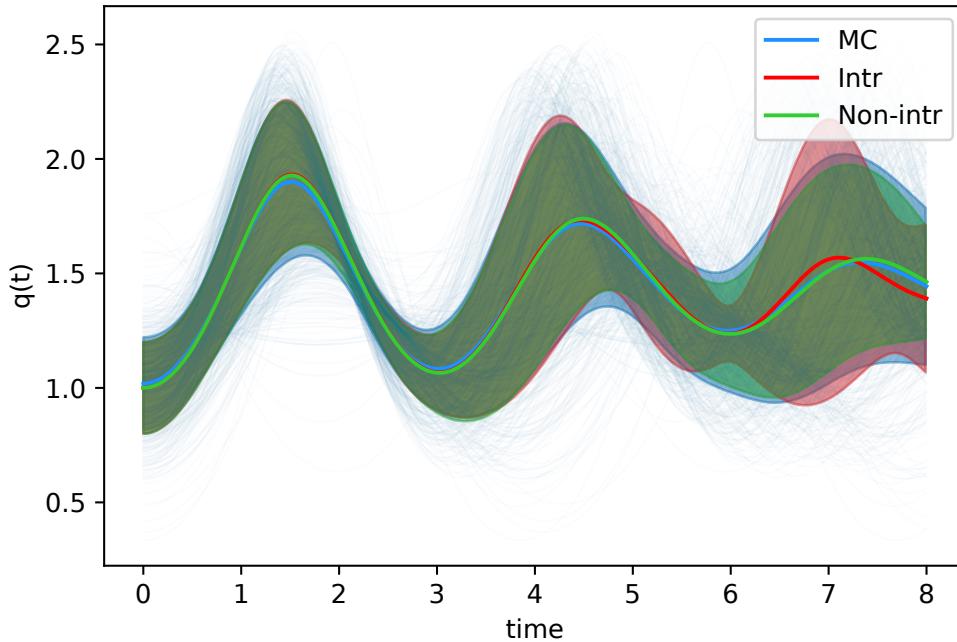


Figure 4.9: Duffing Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

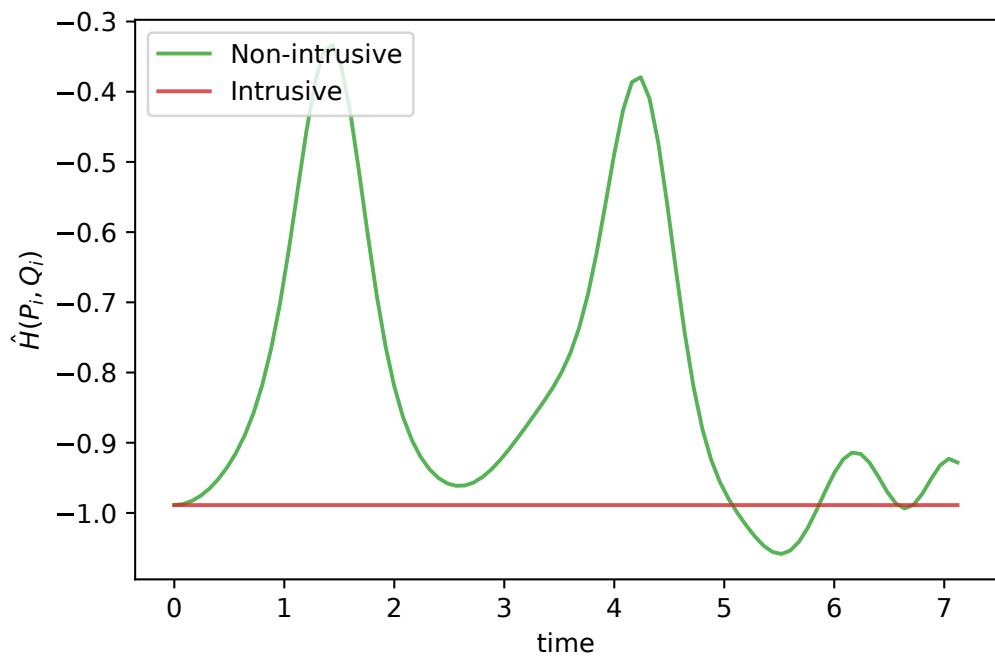


Figure 4.10: Duffing Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ . Comparison of the value over time of the PC Hamiltonian  $\hat{H}(t)$ . In red the intrusive solution, where the Hamiltonian is constant. In green, the non-intrusive solution.

#### 4.4.2 Second order

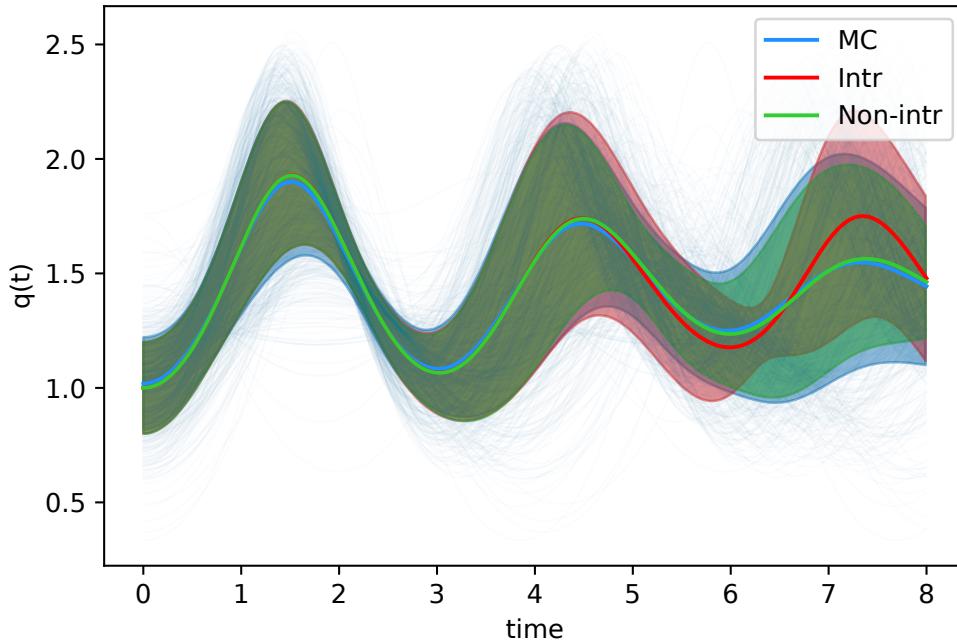


Figure 4.11: Duffing Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

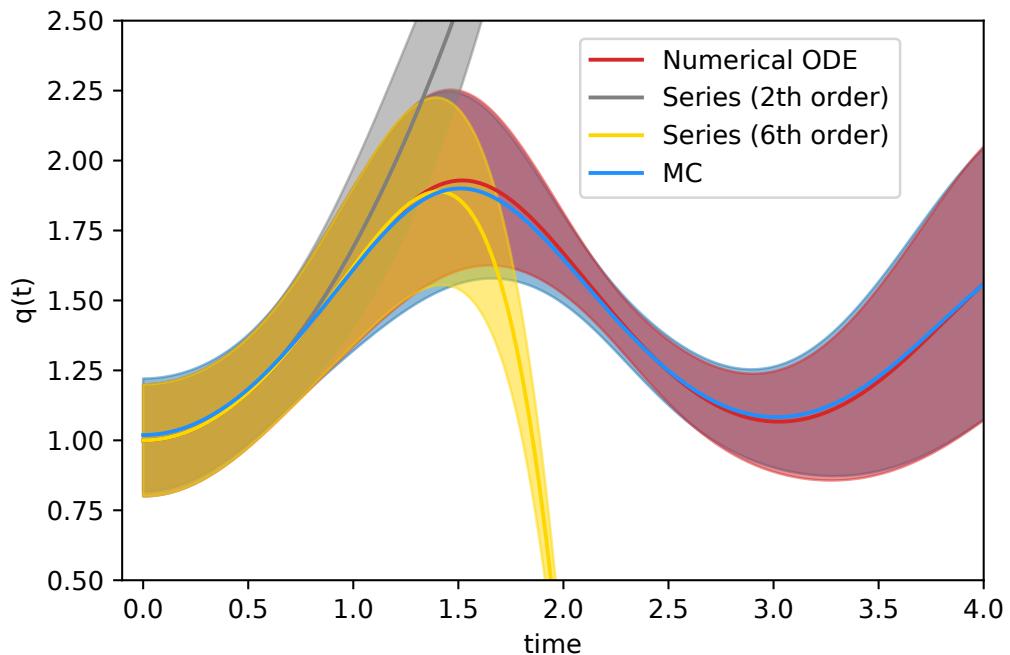


Figure 4.12: Duffing Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ . Examples of exponential maps. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Yellow and grey lines and bands: non-intrusive average plus/minus standard deviation from exponential maps applied to the same initial conditions as the ODE.

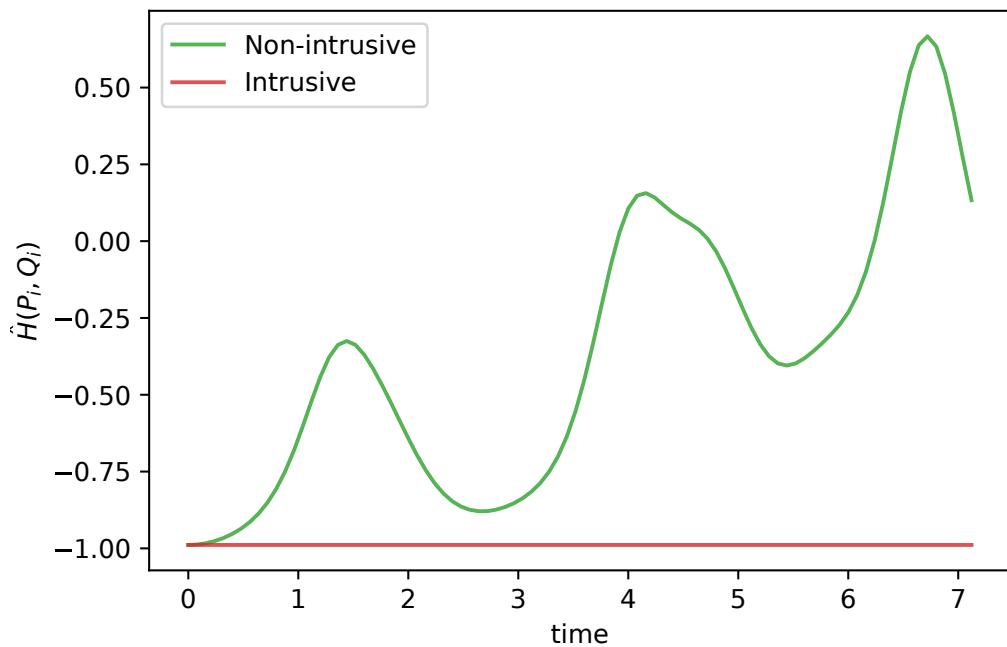


Figure 4.13: Duffing Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ . Comparison of the value over time of the PC Hamiltonian  $\hat{H}(t)$ . In red the intrusive solution, where the Hamiltonian is constant. In green, the non-intrusive solution.

### 4.4.3 Third order

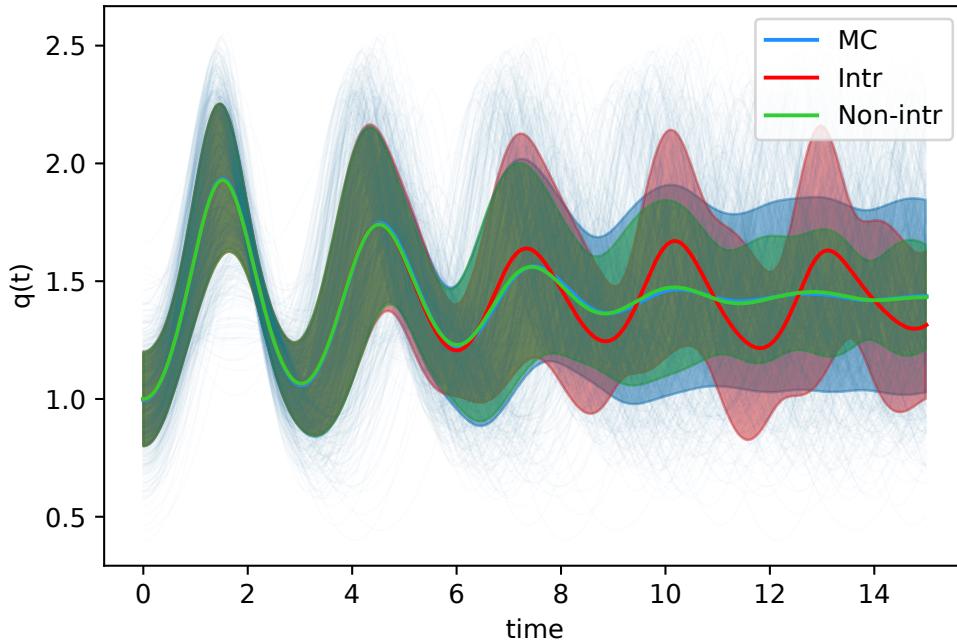


Figure 4.14: Duffing Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

## 4.5 Non-linear quadrupole

In particle accelerator physics, accelerator components are described as multipolar expansions of the Hamiltonian of a particle in an electromagnetic field:

$$H = -\sqrt{(\vec{p} - q\vec{A})^2 c^2 + m^2 c^4} + qV$$

where  $A$  and  $V$  are the vector and scalar potential respectively.

Usually in accelerator physics, a coordinate frame moving along the desired ideal trajectory of a particle with the design speed  $\beta_0$  and  $\gamma_0$  is adopted. Moreover, the independent variable is taken to be the spatial position along this trajectory  $s$ , rather than time  $t$ . However in our analysis this has no impact, and  $s$  and  $t$  can be used interchangeably. Elements such as magnets can be expanded as a multipole series, with the various weighting  $k_i$  indicating the relative force strength.

As case study from this category, the following fully six-dimensional, non-linear quadrupole, Hamiltonian  $H_{\text{Quad}}(x, y, z, p_x, p_y, \delta; k_2, k_3)$  has been analysed:

$$H^{\text{Quad}} = -\sqrt{\left(\frac{1}{\beta_0} + \delta\right)^2 - p_x^2 - p_y^2 - \frac{1}{\beta_0^2 \gamma_0^2}} + \frac{1}{6} (k_3 x^3 - k_2 x y^2) + \frac{\delta}{\beta_0}. \quad (4.10)$$

This Hamiltonian is approximated as a Taylor series, as polynomial Hamiltonians are more amenable to PCE analysis. At the third order, it is:

$$\begin{aligned} H^{\text{Quad}} \approx & -\frac{k_2 x y^2}{6} + \frac{k_3 x^3}{6} + \frac{p_x^2}{2\sqrt{\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}}} - \frac{p_x^2 \delta}{2\beta_0 \left(\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}\right)^{\frac{3}{2}}} + \frac{p_y^2}{2\sqrt{\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}}} - \\ & - \frac{p_y^2 \delta}{2\beta_0 \left(\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}\right)^{\frac{3}{2}}} + \frac{\delta^2 \left(-1 + \frac{1}{\beta_0^2 \left(\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}\right)}\right)}{2\sqrt{\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}}} + \delta \left(\frac{1}{\beta_0} - \frac{1}{\beta_0 \sqrt{\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}}}\right) - \\ & - \sqrt{\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}} + \frac{\delta^3 \left(1 - \frac{1}{\beta_0^2 \left(\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}\right)}\right)}{2\beta_0 \left(\frac{1}{\beta_0^2} - \frac{1}{\beta_0^2 \gamma_0^2}\right)^{\frac{3}{2}}} \end{aligned} \quad (4.11)$$

Here,  $k_2$  and  $k_3$  are assumed uniformly distributed stochastic parameters, and are used as germs.

$$k_2, k_3 = \Xi_i = [\mathcal{U}(l=1.5, h=3.5), \mathcal{U}(l=3, h=4)] \quad (4.12)$$

On the other hand,  $\beta_0$  and  $\gamma_0$  are design parameters, and are fixed. In the following,  $\beta_0 = \frac{99}{100}$  and correspondingly  $\gamma_0 = \frac{1}{\sqrt{1-\beta_0^2}} \approx 7.09$ . Substituting in 4.11, the resulting Hamiltonian will be:

$$H = -\frac{k_2 xy^2}{6} + \frac{k_3 x^3}{6} - \frac{50p_x^2\delta}{99} + \frac{p_x^2}{2} - \frac{50p_y^2\delta}{99} + \frac{p_y^2}{2} - \frac{9950\delta^3}{970299} + \frac{199\delta^2}{19602} - 1 \quad (4.13)$$

For the Taylor series approximation 4.11 to be valid, the values of the variables have to be close to zero. The zeroth order pointwise initial conditions

$$\begin{aligned}[x, y, z](t=0) &= [0.05, 0.05, 0.01] \\ [p_x, p_y, \delta](t=0) &= [0, 0.005, 0.001]\end{aligned}$$

have been used. The relative error between 4.13 and the full Hamiltonian 4.10 is about  $10^{-9}$  at the initial condition.

A fully six-dimensional stochastic initial distribution would require 6 additional independent germs. This exemplifies the curse of dimensionality in PCE when there are many independent stochastic parameters, one of the main drawbacks of the method. In Sec. 4.5.3 the alternative numerical quadrature method is used to explore the FODO hamiltonian system with an initial distribution on  $x(s)$ . This method appears to be slower than Monte Carlo however.

In the following plots the PCE expansion seems to fit better the empirical distribution than in the previous sections, but this is due to the fact that the runtime has been kept pretty short. Some trajectories are bound to exit the validity region of the approximation soon afterwards, and will start to increase exponentially. These would make the overall distribution pretty meaningless and the plots hard to interpret.

Since  $H^{\text{Quad}}$  does not depend on  $z$ , its conjugate momentum  $\delta$  is a constant of motion. In Fig. 4.20, we notice how the PC expansion is able to exactly preserve this quantity.

#### 4.5.1 Comparison plots, second order

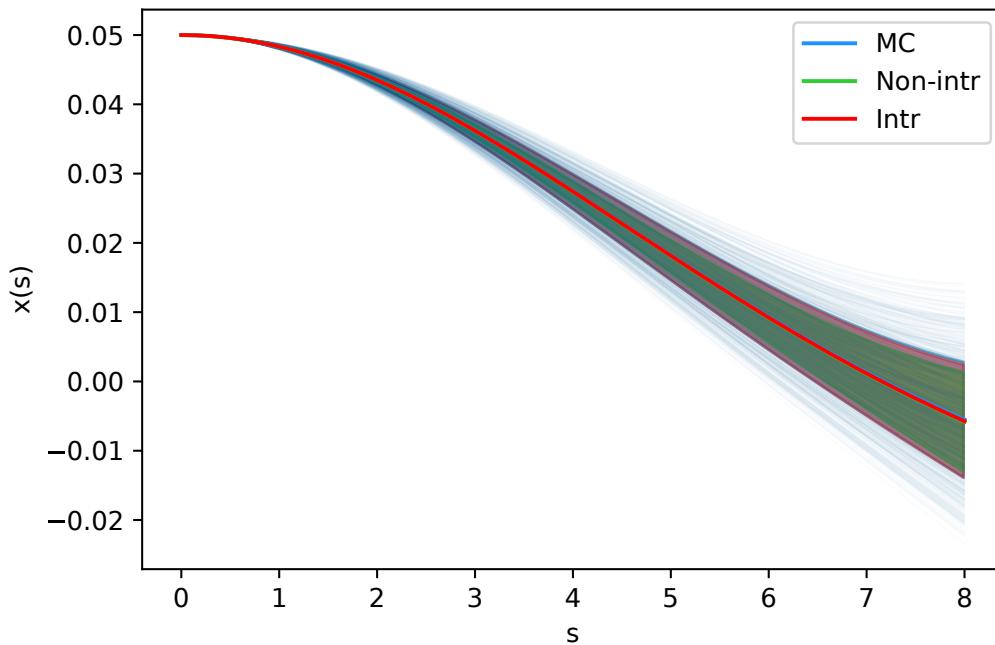


Figure 4.15: Quadrupole Hamiltonian,  $x(s)$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

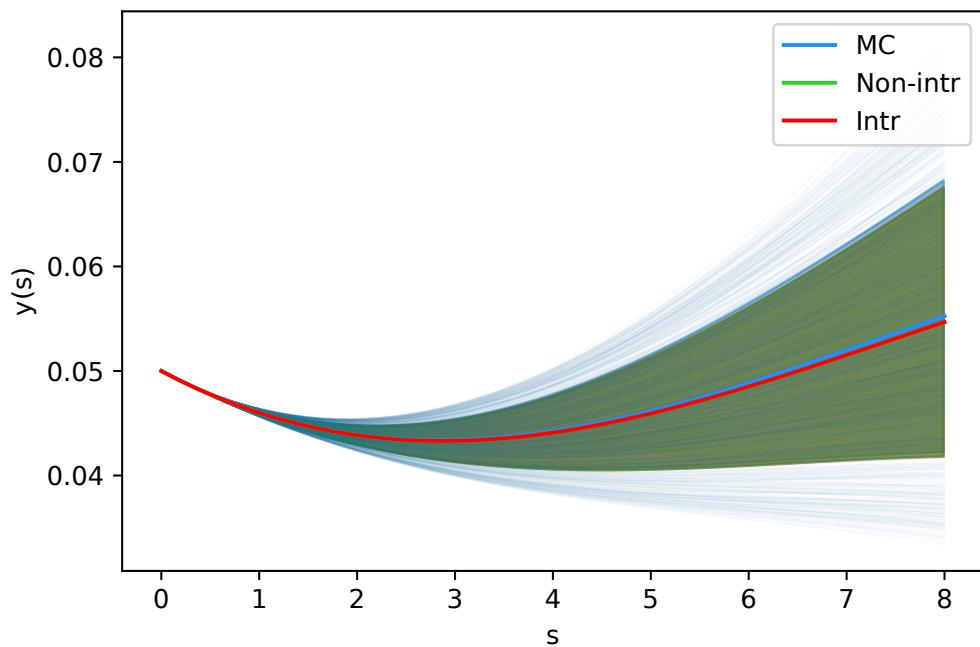


Figure 4.16: Quadrupole Hamiltonian,  $y(s)$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

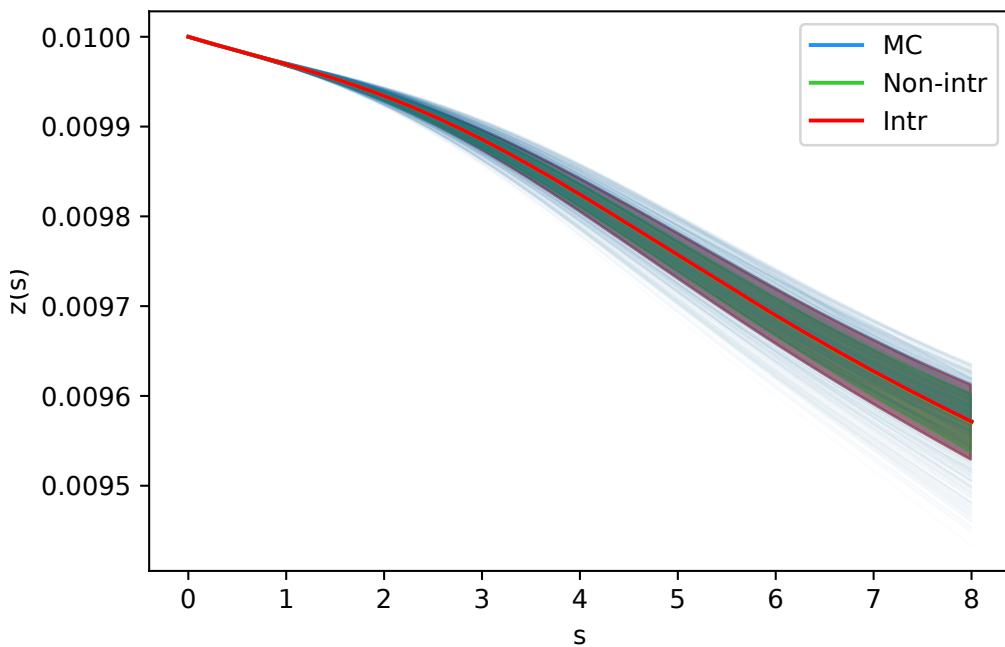


Figure 4.17: Quadrupole Hamiltonian,  $z(s)$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

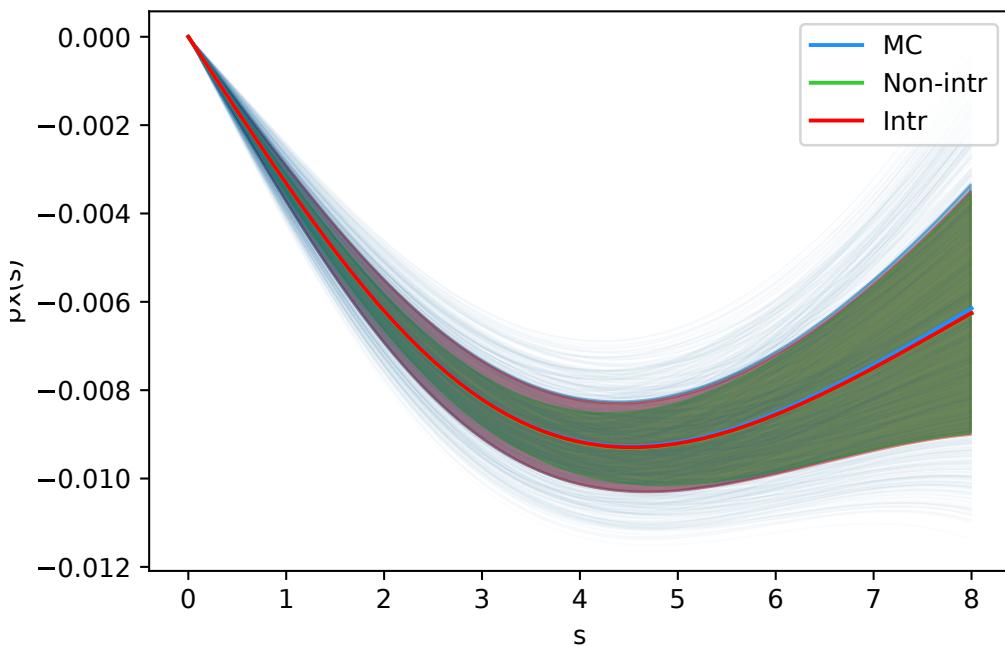


Figure 4.18: Quadrupole Hamiltonian,  $p_x(s)$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

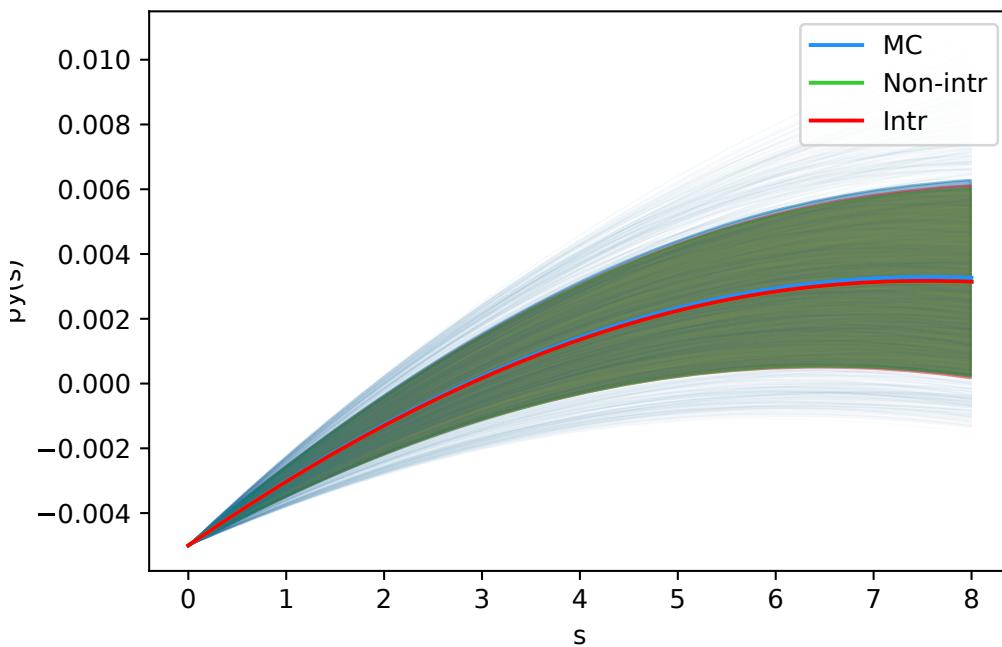


Figure 4.19: Quadrupole Hamiltonian,  $p_y(s)$ . Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

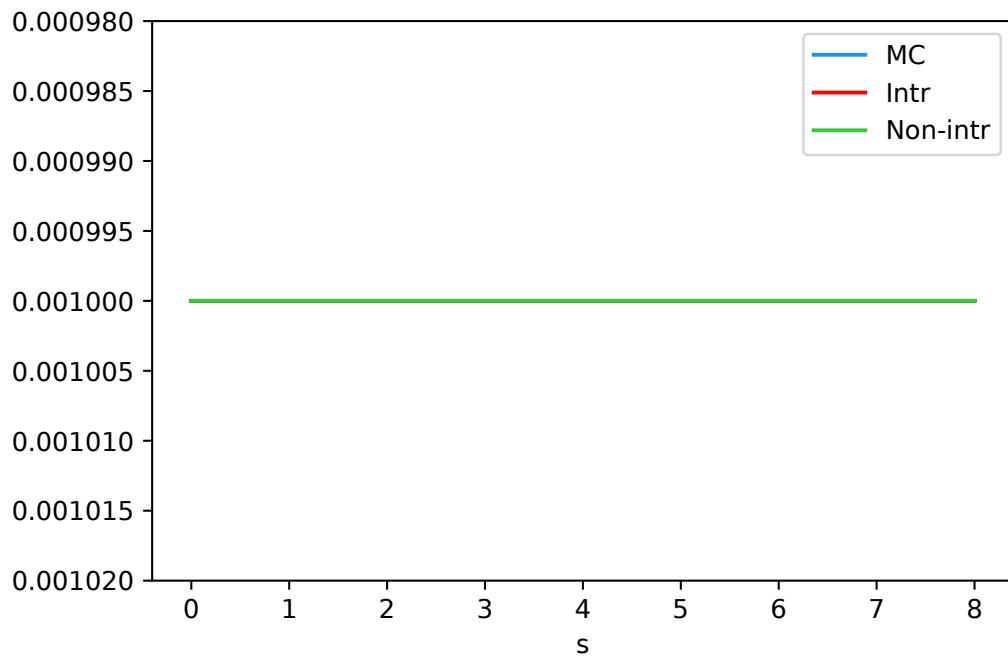


Figure 4.20: Quadrupole Hamiltonian,  $\delta(s)$ . As the Hamiltonian does not depend on  $z$ ,  $\delta$  is constant. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

### 4.5.2 Exponential map example

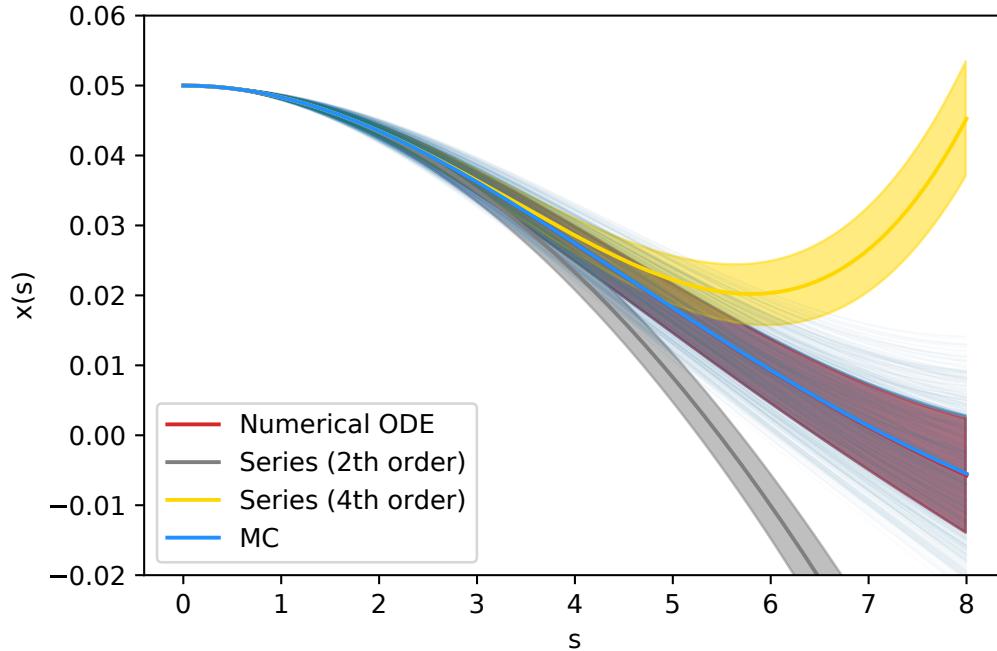


Figure 4.21: Quadrupole Hamiltonian,  $x(s)$ , exponential map. Examples of exponential maps. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Yellow and grey lines and bands: non-intrusive average plus/minus standard deviation from exponential maps applied to the same initial conditions as the ODE.

### 4.5.3 Numerical quadrature example

This example shows the use of numerical quadrature PCE to bypass the cumbersome symbolic integration of an Hamiltonian with three germs. Here, in addition to the two uncertain multipole parameters of section 4.5, an initial distribution germ is present, for the variable  $x(s)$ .

The computation at PC order  $O = 1$  on a laptop required 1 minute and 40 seconds, where (non-parallelized) Monte Carlo required only 1 minute and 20 seconds. Thus, numerical quadrature in intrusive polynomial chaos may have limited practical utility, albeit useful for the study of the method itself.

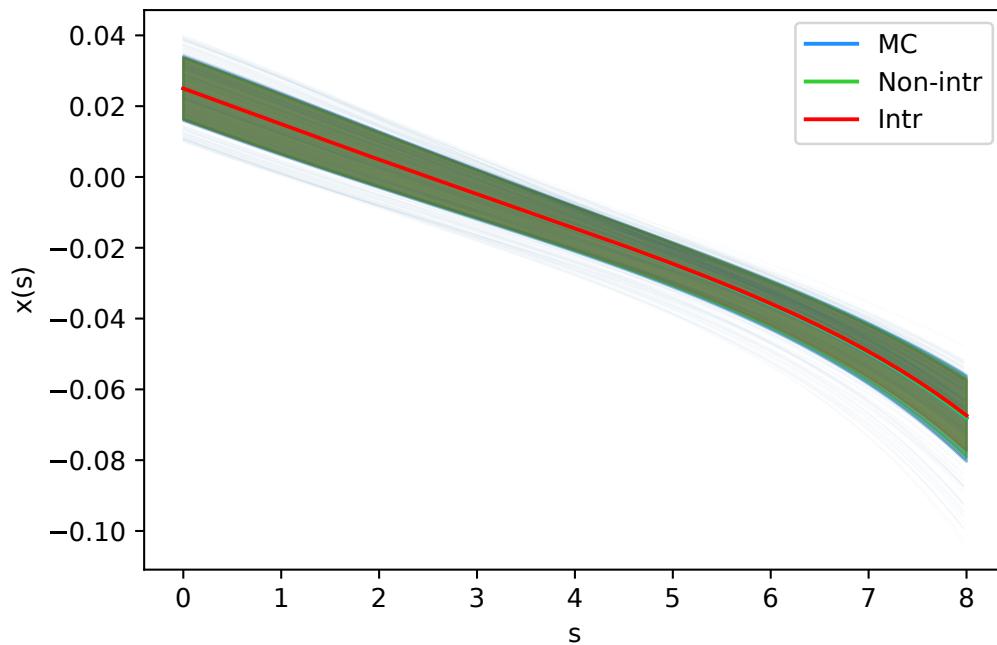


Figure 4.22: Quadrupole Hamiltonian,  $x(s)$ ,  $O = 1$ . Example to show the use of numerical quadrature. Faint blue lines: Monte Carlo sample trajectories. Blue line and band: Monte Carlo average  $\mu^{\text{MC}}(t)$ , plus/minus standard deviation  $\sigma^{\text{MC}}(t)$ . Red line and band: intrusive average  $\mu^{\text{PCE}}(t)$ , plus/minus standard deviation  $\sigma^{\text{PCE}}(t)$ . Green line and band: non-intrusive average  $\mu^{\text{NI}}(t)$ , plus/minus standard deviation  $\sigma^{\text{NI}}(t)$ .

# Chapter 5

## Conclusions

In this work, a script enabling automatic intrusive polynomial chaos expansion of a given Hamiltonian system, for any user defined germ distributions, has been developed, using `chaospy`, `sympy`, `scipy`, and `numpy`. This solves one of the main drawbacks of intrusive polynomial chaos methods, which would otherwise require a complex rewrite of the original system's equations by the user. The user can compare the output of the intrusive system with Monte Carlo simulations and a non-intrusive PC fit, and create the exponential map corresponding to the evolution of the system.

The script has been used to explore a set of Hamiltonians at various expansion orders, enabling empirical comparison with previous theoretical results. For short to medium times, the uncertainty as measured by the standard deviation is close to the empirical one obtained through Monte Carlo methods. To discover this maximum valid runtime, a user can observe the divergence from a lower PC expansion order, which in principle enables the user to bound the approximation error of the higher one. Most importantly, if the new intrusive equations can be derived analytically, as it is the case with most Hamiltonians when expanded in series, the intrusive solution can be obtained in a much shorter time with respect to Monte Carlo methods.

The script is also able to generate a non-linear exponential map for the PC system, which provides a method to generate general transport maps for initial distributions.

The numerical experiments have confirmed [PS13] how sparse the PC systems can be: if the initial distribution is one of the germs, the initial conditions will necessarily be first order with respect to this distribution. This is shown to lead to a PC system equivalent in computational complexity to one of order  $O - 1$  or even  $O - 2$ .

There are some important limitations though. A few were examined in the

literature [PS13] and here confirmed in the numerical experiments: for long times the polynomial chaos expansion will fail to adequately describe the system. In particular, it cannot follow the average of the system as well as the non-intrusive PCE method. Another is that the intrusive system is symplectic, but the ideal solution as returned by the non-intrusive fit is not. This raises questions about the potential disadvantages of symplectifying the approximated exponential map.

The other main obstacle to the use of intrusive PCE is the long time required to generate symbolically the equations: albeit this process needs to be done only once, for high orders or number of germs it needs a much longer time than Monte Carlo. An attempt to parallelize this process has been made using `dask`, however it often runs into `sympy` limitations. A further attempt has been made using numerical integration to bypass the generation of the Hamiltonian  $\hat{H}$ , however this method seems to lose the speed advantage of intrusive PCE.

### Potential improvements

The most straightforward improvement would be understanding and exploiting sparseness. While there are sparse numerical ODE integration methods available, it would be useful to know a priori when some polynomial coefficient will always be zero. This would lead to a better targeted PC expansion of the variables, leading to improved performance in all aspects (symbolical integration, exponential map construction).

Another important possibility is the use of a better symbolical computing framework, better able to perform on multiple cores (perhaps Wolfram `mathematica` [Res]).

A careful study of symplectification and the properties of the PC solution may be fruitful. Perhaps partial symplectification, where only some lower order moments evolve symplectically.

Other possibilities include a hybrid approach, perhaps mixing sampling with PCE. This may also be a path to improve the performance of the fully numerical version, and enabling the connection with other numerical frameworks such as DA.

# Bibliography

- [Arn89] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer New York, 1989. doi: 10.1007/978-1-4757-2063-1. URL: <https://doi.org/10.1007%2F978-1-4757-2063-1>.
- [Bla02] Sergio Blanes. “Symplectic maps for approximating polynomial Hamiltonian systems”. In: *Physical Review E* 65.5 (2002), p. 056703.
- [ON12] S. Oladyshkin and W. Nowak. “Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion”. In: *Reliability Engineering and System Safety* 106 (2012), pp. 179–190. ISSN: 0951-8320. doi: <https://doi.org/10.1016/j.ress.2012.05.002>. URL: <http://www.sciencedirect.com/science/article/pii/S0951832012000853>.
- [Dra13] Alex J. Dragt. *An overview of Lie methods for accelerator physics*. 2013. URL: <https://accelconf.web.cern.ch/pac2013/papers/thap1.pdf> (visited on 11/18/2020).
- [OHa13] Anthony O’Hagan. “Polynomial chaos: A tutorial and critique from a statistician’s perspective”. In: *SIAM/ASA J. Uncertainty Quantification* 20 (2013), pp. 1–20.
- [PS13] J. M. Pasini and T. Sahai. “Polynomial chaos based uncertainty quantification in Hamiltonian and chaotic systems”. In: *52nd IEEE Conference on Decision and Control*. 2013, pp. 1113–1118.
- [FL15] Jonathan Feinberg and Hans Petter Langtangen. “Chaospy: An open source tool for designing methods of uncertainty quantification”. In: *Journal of Computational Science* 11 (2015), pp. 46–57. ISSN: 1877-7503. doi: <https://doi.org/10.1016/j.jocs.2015.08.008>. URL: <http://www.sciencedirect.com/science/article/pii/S1877750315300119>.
- [Das16] Dask Development Team. *Dask: Library for dynamic task scheduling*. 2016. URL: <https://dask.org>.

- [Meu+17] Aaron Meurer et al. “SymPy: symbolic computing in Python”. In: *PeerJ Computer Science* 3 (Jan. 2017), e103. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.103. URL: <https://doi.org/10.7717/peerj-cs.103>.
- [Har+20] Charles R. Harris et al. “Array programming with NumPy”. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: <https://doi.org/10.1038/s41586-020-2649-2>.
- [Vir+20] Pauli Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.
- [Feia] Jonathan Feinberg. *Intrusive Galerkin chaospy tutorial*. URL: [https://chaospy.readthedocs.io/en/master/notebooks/06\\_intrusive\\_galerkin.html](https://chaospy.readthedocs.io/en/master/notebooks/06_intrusive_galerkin.html) (visited on 11/16/2020).
- [Feib] Jonathan Feinberg. *Point Collocation*. URL: [https://chaospy.readthedocs.io/en/master/tutorials/point\\_collocation.html](https://chaospy.readthedocs.io/en/master/tutorials/point_collocation.html) (visited on 11/24/2020).
- [Rac] Chris Rackauckas. *What does symplectic mean in reference to numerical integrators, and does SciPy’s odeint use them?* Computational Science Stack Exchange. URL: <https://scicomp.stackexchange.com/q/29154> (version: 2020-07-19). eprint: <https://scicomp.stackexchange.com/q/29154>. URL: <https://scicomp.stackexchange.com/q/29154>.
- [Res] Wolfram Research. *Mathematica, Version 12.2*. Champaign, IL, 2020. URL: <https://www.wolfram.com/mathematica>.

# Appendix A

## Hamiltonian PCE theory

### A.1 PCE of harmonic oscillator solution

$\hat{I}_l$  denotes the indefinite integral version of  $I_l$  in 2.11. This integral can be performed by integrating by parts 2 times. Knowing that

$$\begin{aligned}\int \cos(\bar{\omega}t + L\lambda t) d\lambda &= \frac{\sin(\hat{\omega}t + L\lambda t)}{Lt} \\ \int \sin(\bar{\omega}t + L\lambda t) d\lambda &= -\frac{\cos(\hat{\omega}t + L\lambda t)}{Lt},\end{aligned}$$

we get

$$\begin{aligned}\hat{I}_l &= \int \lambda^l \cos(\bar{\omega}t + L\lambda t) d\lambda = \frac{\lambda^l}{Lt} \sin(\hat{\omega}t + L\lambda t) - \frac{l}{Lt} \int \lambda^{l-1} \sin(\bar{\omega}t + L\lambda t) d\lambda \\ &= \frac{\lambda^l}{Lt} \sin(\hat{\omega}t + L\lambda t) - \frac{l}{Lt} \left( -\frac{\lambda^{l-1}}{Lt} \cos(\hat{\omega}t + L\lambda t) + \frac{l-1}{Lt} \int \lambda^{l-2} \cos(\hat{\omega}t + L\lambda t) d\lambda \right) \\ &= \frac{\lambda^l}{Lt} \sin(\hat{\omega}t + L\lambda t) + \frac{l\lambda^{l-1}}{L^2 t^2} \cos(\hat{\omega}t + L\lambda t) - \frac{l(l-1)}{L^2 t^2} \hat{I}_{l-2}.\end{aligned}$$

Then, substituting  $\lambda$  to get the definite integral  $I_l = [\hat{I}_l]_{\lambda=-1}^{\lambda=1}$  the required expression 2.11 is recovered. The base cases  $I_0$  and  $I_1$  go to 0 for  $t \rightarrow \infty$  as verified by direct computation:

$$\begin{aligned}I_0 &= \int_{-1}^1 \cos(\bar{\omega}t + L\lambda t) \cdot 1 d\lambda = \frac{\sin(Lt - \bar{\omega}t)}{Lt} + \frac{\sin(Lt + \bar{\omega}t)}{Lt} \xrightarrow[t \rightarrow \infty]{} 0, \\ I_1 &= \int_{-1}^1 \cos(\bar{\omega}t + L\lambda t) \lambda d\lambda = -\frac{\sin(Lt - \bar{\omega}t)}{Lt} + \frac{\sin(Lt + \bar{\omega}t)}{Lt} - \frac{\cos(Lt - \bar{\omega}t)}{L^2 t^2} + \\ &\quad + \frac{\cos(Lt + \bar{\omega}t)}{L^2 t^2} \xrightarrow[t \rightarrow \infty]{} 0.\end{aligned}$$

# Appendix B

## PC Hamiltonians

### B.1 Harmonic oscillator

The base Hamiltonian is  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} + k_0 \frac{q^2}{2}$ , where the parameter  $k_0$  is distributed uniformly as  $k_0 \sim \mathcal{U}(\text{low} = 3.5, \text{high} = 4.5)$ . Expanding in the 2 germs  $\Xi_i : [\mathcal{U}(l = 3.5, h = 4.5), \mathcal{N}(\mu = 0, \sigma = 1)]$  up to order  $O = 0, \dots, 4$ , the resulting PC Hamiltonians  $\hat{H}_O$  are:

- $\hat{H}_0^{\text{Har}} : \frac{P_0^2}{2} + \frac{5Q_0^2}{4}$
- $\hat{H}_1^{\text{Har}} : \frac{P_0^2}{2} + \frac{P_1^2}{2} + \frac{P_2^2}{2} + \frac{5Q_0^2}{4} + \frac{\sqrt{3}Q_0Q_2}{3} + \frac{5Q_1^2}{4} + \frac{5Q_2^2}{4}$
- $\hat{H}_2^{\text{Har}} : \frac{P_0^2}{2} + \frac{P_1^2}{2} + \frac{P_2^2}{2} + \frac{P_3^2}{2} + \frac{P_4^2}{2} + \frac{P_5^2}{2} + \frac{5Q_0^2}{4} + \frac{\sqrt{3}Q_0Q_2}{3} + \frac{5Q_1^2}{4} + \frac{\sqrt{3}Q_1Q_4}{3} + \frac{5Q_2^2}{4} + \frac{2\sqrt{15}Q_2Q_5}{15} + \frac{5Q_3^2}{4} + \frac{5Q_4^2}{4} + \frac{5Q_5^2}{4}$
- $\hat{H}_3^{\text{Har}} : \frac{P_0^2}{2} + \frac{P_1^2}{2} + \frac{P_2^2}{2} + \frac{P_3^2}{2} + \frac{P_4^2}{2} + \frac{P_5^2}{2} + \frac{P_6^2}{2} + \frac{P_7^2}{2} + \frac{P_8^2}{2} + \frac{P_9^2}{2} + \frac{5Q_0^2}{4} + \frac{\sqrt{3}Q_0Q_2}{4} + \frac{5Q_1^2}{4} + \frac{\sqrt{3}Q_1Q_4}{4} + \frac{5Q_2^2}{4} + \frac{2\sqrt{15}Q_2Q_5}{15} + \frac{5Q_3^2}{4} + \frac{\sqrt{3}Q_3Q_7}{3} + \frac{5Q_4^2}{4} + \frac{2\sqrt{15}Q_4Q_8}{15} + \frac{5Q_5^2}{4} + \frac{3\sqrt{35}Q_5Q_9}{35} + \frac{5Q_6^2}{4} + \frac{5Q_7^2}{4} + \frac{5Q_8^2}{4} + \frac{5Q_9^2}{4}$
- $\hat{H}_4^{\text{Har}} : \frac{P_0^2}{2} + \frac{P_{10}^2}{2} + \frac{P_{11}^2}{2} + \frac{P_{12}^2}{2} + \frac{P_{13}^2}{2} + \frac{P_{14}^2}{2} + \frac{P_1^2}{2} + \frac{P_2^2}{2} + \frac{P_3^2}{2} + \frac{P_4^2}{2} + \frac{P_5^2}{2} + \frac{P_6^2}{2} + \frac{P_7^2}{2} + \frac{P_8^2}{2} + \frac{P_9^2}{2} + \frac{5Q_0^2}{4} + \frac{\sqrt{3}Q_0Q_2}{3} + \frac{5Q_{10}^2}{4} + \frac{5Q_{11}^2}{4} + \frac{\sqrt{3}Q_{11}Q_6}{3} + \frac{5Q_{12}^2}{4}$

$$\begin{aligned}
 & \frac{2\sqrt{15}Q_{12}Q_7}{15} + \frac{5Q_{13}^2}{4} + \frac{3\sqrt{35}Q_{13}Q_8}{35} + \frac{5Q_{14}^2}{4} + \frac{4\sqrt{7}Q_{14}Q_9}{21} + \frac{5Q_1^2}{4} + \frac{\sqrt{3}Q_1Q_4}{3} + \\
 & \frac{5Q_2^2}{4} + \frac{2\sqrt{15}Q_2Q_5}{15} + \frac{5Q_3^2}{4} + \frac{\sqrt{3}Q_3Q_7}{3} + \frac{5Q_4^2}{4} + \frac{2\sqrt{15}Q_4Q_8}{15} + \frac{5Q_5^2}{4} + \frac{3\sqrt{35}Q_5Q_9}{35} + \\
 & \frac{5Q_6^2}{4} + \frac{5Q_7^2}{4} + \frac{5Q_8^2}{4} + \frac{5Q_9^2}{4}
 \end{aligned}$$

## B.2 Duffing oscillator

The Duffing Hamiltonian:  $H_{\text{Har}}(p, q; k_0) = \frac{p^2}{2} - k_0 \frac{q^2}{2} + \frac{q^4}{4}$ , where the parameter  $k_0$  is distributed uniformly as  $k_0 \sim \mathcal{U}(\text{low} = 3.5, \text{high} = 4.5)$ . Expanding in the 2 germs  $\Xi_i : [\mathcal{U}(l = 3.5, h = 4.5), \mathcal{N}(\mu = 0, \sigma = 1)]$  up to order  $O = 0, \dots, 2$ , the resulting PC Hamiltonians  $\hat{H}_O$  are:

- $\hat{H}_0^{\text{Duf}} : \frac{P_0^2}{2} + \frac{Q_0^4}{4} - \frac{5Q_0^2}{4}$
- $\hat{H}_1^{\text{Duf}} : \frac{P_0^2}{2} + \frac{P_1^2}{2} + \frac{P_2^2}{2} + \frac{Q_0^4}{4} + \frac{3Q_0^2Q_1^2}{2} + \frac{3Q_0^2Q_2^2}{2} - \frac{5Q_0^2}{4} - \frac{\sqrt{3}Q_0Q_2}{3} + \frac{3Q_1^4}{4} +$   
 $\frac{3Q_1^2Q_2^2}{2} - \frac{5Q_1^2}{4} + \frac{9Q_2^4}{20} - \frac{5Q_2^2}{4}$
- $\hat{H}_2^{\text{Duf}} : \frac{P_0^2}{2} + \frac{P_1^2}{2} + \frac{P_2^2}{2} + \frac{P_3^2}{2} + \frac{P_4^2}{2} + \frac{P_5^2}{2} + \frac{Q_0^4}{4} + \frac{3Q_0^2Q_1^2}{2} + \frac{3Q_0^2Q_2^2}{2} + \frac{3Q_0^2Q_3^2}{2} +$   
 $\frac{3Q_0^2Q_4^2}{2} + \frac{3Q_0^2Q_5^2}{2} - \frac{5Q_0^2}{4} + 3\sqrt{2}Q_0Q_1Q_3 + 6Q_0Q_1Q_2Q_4 + \frac{6\sqrt{5}Q_0Q_2Q_5}{5} -$   
 $\frac{\sqrt{3}Q_0Q_2}{3} + 2\sqrt{2}Q_0Q_3^3 + 3\sqrt{2}Q_0Q_3Q_4^2 + \frac{6\sqrt{5}Q_0Q_4Q_5}{5} + \frac{2\sqrt{5}Q_0Q_5^3}{7} + \frac{3Q_1^4}{4} +$   
 $\frac{3Q_1^2Q_2^2}{2} + \frac{15Q_1^2Q_3^2}{2} + \frac{9Q_1^2Q_4^2}{2} + \frac{3Q_1^2Q_5^2}{2} - \frac{5Q_1^2}{4} + 6\sqrt{2}Q_1Q_2Q_3Q_4 + \frac{12\sqrt{5}Q_1Q_2Q_4Q_5}{5} -$   
 $\frac{\sqrt{3}Q_1Q_4}{3} + \frac{9Q_2^4}{20} + \frac{3Q_2^2Q_3^2}{2} + \frac{27Q_2^2Q_4^2}{10} + \frac{33Q_2^2Q_5^2}{14} - \frac{5Q_2^2}{4} - \frac{2\sqrt{15}Q_2Q_5}{15} + \frac{15Q_3^4}{4} +$   
 $\frac{15Q_3^2Q_4^2}{2} + \frac{3Q_3^2Q_5^2}{2} - \frac{5Q_3^2}{4} + \frac{6\sqrt{10}Q_3Q_4Q_5}{5} + \frac{27Q_4^4}{20} + \frac{33Q_4^2Q_5^2}{14} - \frac{5Q_4^2}{4} + \frac{15Q_5^4}{28} -$   
 $\frac{5Q_5^2}{4}$

## Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

---

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

**Title of work** (in block letters):

UNCERTAINTY QUANTIFICATION OF HAMILTONIAN MAPS USING INTRUSIVE POLYNOMIAL CHAOS EXPANSION

**Authored by** (in block letters):

*For papers written by groups the names of all authors are required.*

**Name(s):**

FORCHER

**First name(s):**

FRANCESCO

With my signature I confirm that

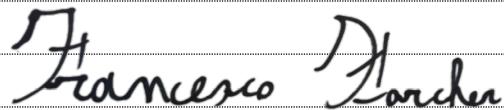
- I have committed none of the forms of plagiarism described in the '[Citation etiquette](#)' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

**Place, date**

Conegliano, 10-12-2020

**Signature(s)**



*For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.*