An efficient algorithm for stochastic differential equations

Philipp Schapotschnikow
Technische Universität München
Siemens AG

March 30, 2005

Abstract

We present an algorithm for numerical solution of differential equations with random input parameters using the approximation by orthogonal polynomials.

1 Introduction

When doing mathematical modeling in science and technique, one usually assumes that all input parameters like physical and material constants are known exactly. But this is not quite true. Since no measurement can be done with an infinite precision, we can be sure up to a certain degree about the range in which the values lie.

This means that when doing an experiment we observe some behaviour of the system corresponding to a set of parameters. They are all fixed numbers, but those numbers are not known exactly. So we write down a model, we predict a behaviour of the system, and in the experiment we observe a different behaviour. Then we say, if the difference is not large, that the experiment corroborates the model, and blame the measurement errors for the deviations. But what we observe in reality is just another possible behaviour, corresponding to the true values, and not to the ones we plugged in the model.

To overcome this obstacle we need the stochastic calculus. Here we assume the input data to be random variables. Those are variables without a fixed value, but with a value lying in the given range with some certain probability.

1.1 Example: the LC-circuit

We consider the LC-circuit from the Figure 1 as an example. Here the current I through the coil and the voltage U at the capacitor obey the following differential

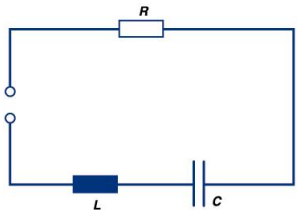

Figure 1: A damped LC-circuit with a capacitor C, induction coil L and resistance R

equations:

$$\dot{U} = -\frac{I}{C} \qquad\qquad \dot{I} = \frac{U - RI}{L} \tag{1}$$

The solution of this system is for a small resistance

$$I(t) = \frac{U_0}{R_0} \cdot \frac{\omega_0}{\omega_e} \cdot e^{-at} \sin(\omega_e t) \tag{2}$$

with

- Decay constant $a = \frac{R}{2L}$
- Characteristic angular frequency $\omega_0 = \frac{1}{\sqrt{LC}}$
- Angular frequency $\omega_e = \sqrt{|\omega_0^2 - a^2|}$
- Characteristic resistance $R_0 = \sqrt{\frac{L}{C}}$.

The system will always behave corresponding to the laws (1). The only problem is - we do not know the exact values. As we can see on the Figure 2,

Figure 2: A capacitor

electronic components usually have a value lying in a certain range. And even this is not sure - there are always some manufacturing errors! Hence we should find a way to plug in this information into the Model 1 and solve the resulting system.

2 Stochastic calculus - a crash course

As first we introduce *stochastic (or random) variables*. Unlike "usual" *deterministic* variables, their values are not numbers, but probability distributions. The density of a distribution shows the probability of finding the value in some certain range. As an example we look at the densities of two typical distributions: the uniform distribution and the Gaussian distribution.

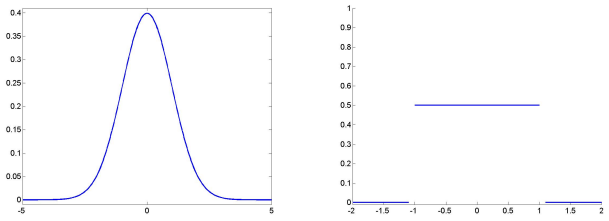

Figure 3: The uniform distribution (right) and the Gaussian distribution (left)

The uniform distribution on $[-1, 1]$ implies, that the values between -1 and 1 are taken with equal probability. In the Gaussian distribution any value is possible, but the values far from the expectation become very improbable. Both distributions are symmetric.

Moments Stochastic variables are characterized by real numbers called power moments and their centered moments. The most important ones are

- Expectation (first power moment) $E[x] = \int x\, \rho(x) dx$.
- Variance (second centered moment) $Var[x] = \int (x - E[x])^2\, \rho(x) dx$.

The expectation describes the mean value after a high number of experiments. The variance describes the deviation from the mean value. Another useful moment is the third centered moment - the *asymmetry* or *skewness*. As the name suggests, it describes how the maximum of the density function deviates from the expected value.

Functions of stochastic variables The next important concept are functions in stochastic variables. If namely $f : \mathbb{R} \to \mathbb{R}$ is real valued function and τ is a random variable, then we can plug in τ in f. What comes out is another

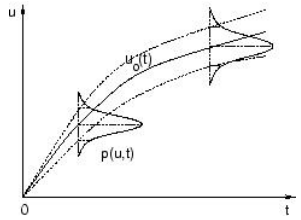

Figure 4: A stochastic process $u(t, \theta)$ with 3 scenarios, one of them – its expectation $u_0(t)$

random variable. We denote this new variable by $f(\tau)$. We do not need the density function of $f(\tau)$ if we know the distribution of τ, e.g.

$$E[f(\tau)] = \int_{\mathbb{R}} f(\tau) d\rho(\tau). \tag{3}$$

Caution! In general for non-linear f $E[f(\tau)] \neq f(E[\tau])$. This means that when dealing with a non-linear model even symmetric deviations in the input will lead to a different expected value in the output.

Stochastic processes Consider a process $x = g(t)$, where t and x represent time and state respectively. If x is now given by a random variable at each time point t we call it *stochastic process*. We write in this case $x = x(t, \theta)$. One can also see such a process as a function of time and chance.

Differential equations with random input parameters Now we arrive at the final destination - differential equations with random input parameters. The solutions are again well-defined as solutions of initial value problems, but they are now stochastic processes.

Returning to the LC-circuit, we may obtain the following system:

$$\dot{U} = -\frac{I}{C}, \qquad C \sim N[C_0, C_1] \tag{4}$$

$$\dot{I} = \frac{U - RI}{L}, \qquad R \sim N[R_0, R_1] \tag{5}$$

Here and later $N[A, B]$ implies a Gaussian distribution with expectation A and variance B.

Our goal is now to construct an algorithm for numerical solution of equations of this type. There are two reasonable demands on such an algorithm.

1. The costs for the solution of the stochastic model should lie in the same range as for the deterministic problem.

2. The result should still mirror the physical properties.

3 Separation of space and time

After a theoretical preparation we will see how to decompose a stochastic process in a time component and a chance component. The idea of separating the components of the problem is quite new for stochastic processes. But when dealing with with non-stationary partial differential equations one of the tools is to split up the model into space and time.

Assumptions For simplicity we assume that only one of the parameters is stochastic and the rest is fixed. Moreover we assume the problem to be well-posed, and so the solution depends continuously on the input parameters. Our last assumption is that the input parameter is Gaussian.

Consider the ODE

$$\dot{y} = F(y, t, \xi) \tag{6}$$

where ξ is a random Gausssian variable with $E[\xi] = 0$ and $Var[\xi] = 1$. If the input parameter a has expectation a_0 and variance a_1^2, we can write it as

$$a = a_0 + a_1\xi. \tag{7}$$

Continuous dependence means that we can write the solution of(6) as $y = y(t, \xi)$. This is a special stochastic process, where y is differentiable in t and continuous in ξ.

Vector spaces Consider the vector space Ξ of random variables with a finite variance which are functions of ξ. It is isomorphic to the space L^2_ρ of square-integrable functions with respect to the weighting function ρ - the density of ξ. In our case is

$$\rho(\theta) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}) \tag{8}$$

In the space Ξ choose the subspace Ξ_0 of variables with expectation 0. Now Ξ_0 is a Hilbert space with the inner product

$$< \xi_1, \xi_2 >= E[\xi_1\xi_2]. \tag{9}$$

Orthogonal polynomials From here we go the standard way of numerical mathematics - we choose an orthogonal basis in this space and represent the problem in this basis. Without any information about the problem we choose the orthogonal polynomials, which are in this case Hermite polynomials.

The Hermite-polynomials are

$$H_0(\xi) = 1, \qquad H_1(\xi) = \xi, \tag{10}$$

$$H_2(\xi) = \xi^2 - 1, \qquad H_3(\xi) = \xi^3 - 3\xi, \ldots \tag{11}$$

In general they are given by a recurrence relation

$$H_0(\xi) = 1; \qquad H_1(\xi) = \xi; \qquad H_{k+1}(\xi) = \xi H_k(\xi) - kH_{k-1}(\xi). \tag{12}$$

When working in Ξ_0 we have to start with H_1 due to $E[H_i(\xi) = 0]$.

The crucial property of Hermite polynomials is orthogonality. For $i \neq j$ holds

$$< H_i(\xi), H_j(\xi) >= E[H_i(\xi)H_j(\xi)] = \int_{\mathbb{R}} H_i(\xi)H_j(\xi)\frac{1}{\sqrt{2\pi}}\exp(-\frac{\xi^2}{2})d\xi = 0. \tag{13}$$

For this reason we can expect superconvergence when approximating with them.

Hermite decomposition Finally we obtain the desired tool. Consider a stochastic variable $x = x(\xi) \in \Xi$, and assume x to be a continuous function. Then $x - E[x] = \tilde{x} \in \Xi_0$. We can write x as

$$x(\xi) = E[x] + \tilde{x} = E[x] + \sum_{i=1}^{\infty} c_i H_i(\xi), \quad c_i \in \mathbb{R}. \tag{14}$$

Subtracting $E[x]$ and calculating the scalar product with $H_k(\xi)$ on both sides we obtain immediately a formula for the coefficients due to orthogonality:

$$\begin{aligned} < \tilde{x}(\xi), H_k(\xi) >=& < \sum_{i=1}^{\infty} c_i H_i(\xi), H_k(\xi) > \\ &= \sum_{i=1}^{\infty} c_i < H_i(\xi), H_k(\xi) >= c_k < H_k(\xi), H_k(\xi) > \end{aligned} \tag{15}$$

$$c_k = \frac{E[x(\xi)H_k(\xi)]}{E[H_k(\xi)H_k(\xi)]}. \tag{16}$$

It makes sense to define $c_0 = E[x]$ and write $x(\xi) = \sum_{i=0}^{\infty} c_i H_i(\xi)$. The formula (16) holds true for $k = 0$ too.

From the coefficients it is rather easy to compute the moments. The expected value is per definition the first coefficient; the variance is

$$\begin{aligned} Var[x] = Var[\tilde{x}] = E[\tilde{x}^2] = E[\left(\sum_{i=1}^{\infty} c_i H_i(\xi)\right)^2] = E[\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} c_i c_j H_i(\xi)H_j(\xi)] \\ = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} c_i c_j \; E[H_i(\xi)H_j(\xi)] = \sum_{i=1}^{\infty} c_i^2 E[H_i(\xi)H_i(\xi)]. \end{aligned} \tag{17}$$

Further moments can be calculated in a similar manner. Note that $E[H_i(\xi)H_i(\xi)]$ are fixed numbers (equal $i!$).

Consider now $y(t, \xi)$ as above. At each fixed time point t we can make the decomposition (14-16), and hence everywhere:

$$y(t, \xi) = \sum_{i=0}^{\infty} u_i(t) H_i(\xi) \tag{18}$$

where $u_i(t)$ are deterministic functions of time. They are given by

$$u_k(t) = \frac{E[y(t,\xi)H_k(\xi)]}{E[H_k(\xi)H_k(\xi)]}. \tag{19}$$

This decomposition is called *Hermite decomposition* and is the central result of this section. It has been used for stochastic finite elements and, recently for stochastic ODE's.

Application to differential equations We return to the differential equation

$$\dot{y} = F(y,t,\xi) \tag{20}$$

As we have learned, we can write $y(t,\xi) = \sum_{i=0}^{\infty} u_i(t)H_i(\xi)$. Since we cannot calculate infinitely many coefficients, we have to truncate the series:

$$y(t,\xi) = \sum_{i=0}^{P} u_i(t)H_i(\xi). \tag{21}$$

Apply to (20):

$$\sum_{i=0}^{P} \dot{u}_i H_i(\xi) = f(t, \sum_{i=0}^{P} u_i(t)H_i(\xi), \xi). \tag{22}$$

Taking on both sides the scalar product with $H_k(\xi)$, $k = 0, \ldots, P$ leads to a system with $P+1$ unknowns and the same number of differential equations. The left side becomes

$$< \sum_{i=0}^{P} \dot{u}_i H_i(\xi), H_k(\xi) >= \sum_{i=0}^{P} \dot{u}_i < H_i(\xi), H_k(\xi) >= \dot{u}_k < H_k(\xi), H_k(\xi) > . \tag{23}$$

On the right side we cannot simplify without knowing the structure of F. Now we obtain the following system:

$$E[H_k(\xi)H_k(\xi)]\dot{u}_k = E[F(y,t,\xi)H_k(\xi)], \quad k = 0, \ldots P \tag{24}$$

This is a purely deterministic system of equations. It seems to be expensive because of the integrals we have to calculate on the right hand side. But usually the structure of F allows us to simplify those calculation. Even if not, we can use the same quadrature rule to all $P+1$ integrals (coming from the $P+1$ equations) and the costs remain low in terms of F-evaluations.

We may expect superconvergence since the Galerkin condition is satisfied: the truncation error is orthogonal to the approximation space (can be checked easily).

4 Example

We consider again the LC-circuit

$$\dot{U} = -\frac{I}{C} \qquad\qquad \dot{I} = \frac{U - RI}{L}$$

where $C = C(\xi)$, e.g. $C = C_0 + C_1\xi$. The approach is

$$U(t,\theta) \approx \sum_{i=0}^{P} u_i(t)H_i(\xi), \qquad\qquad I(t,\theta) \approx \sum_{i=0}^{P} v_i(t)H_i(\xi). \tag{25}$$

Plug in (24) yields for U

$$\begin{aligned} < H_k(\xi), H_k(\xi) > \dot{u}_k &= E[-H_k(\xi), \frac{1}{C(\xi)} \sum_{i=0}^{P} v_i(t)H_i(\xi)] \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} -H_k(\xi) \frac{1}{C(\xi)} \sum_{i=0}^{P} v_i(t)H_i(\xi)\, e^{-\frac{\xi^2}{2}} d\xi \end{aligned} \tag{26}$$

and similarly for I

$$\begin{aligned} &< H_k(\xi), H_k(\xi) > \dot{v}_k = \\ &\qquad \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} H_k(\xi) \frac{\sum_{i=0}^{P} u_i(t)H_i(\xi) - R\sum_{i=0}^{P} v_i(t)H_i(\xi)}{L}\, e^{-\frac{\xi^2}{2}} d\xi \end{aligned} \tag{27}$$

Since k goes from 0 to P we obtain $2(P+1)$ differential equations for $2(P+1)$ unknown coefficients $u_0, v_0, u_1, v_1, \ldots, u_P, v_P$. The initial values are $u_0(0) = U_0,\ v_0(0) = I_0$ and $u_k(0) = v_k(0) = 0\ k = 1, \ldots, P$.

5 Results

Recall that the angular frequency depends on the capacity, but the decay constant does not. We compare the solution deterministic model with the expectation of the solution of the stochastic model with a random Gaussian capacity. Since the decay does not depend on C we expect a change in frequency with the same amplitude. What we should observe is a "phase shift" at later time points. On the Figure 5 we can see on the left the total behaviour in the deterministic and stochastic case; on the left we zoom to see that the amplitudes agree while the frequencies differ. The diagrams are not smooth due to the plotting properties and not because of a rough solution.

The solution was obtained with $P = 6$ and hardly differs from the solution with $P = 5$. The computational time was approximately 5 minutes while the deterministic problem takes less then 1 minute.

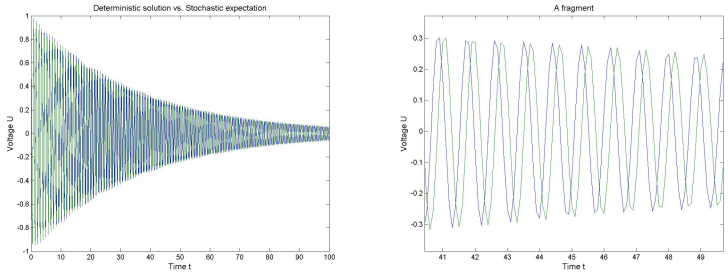

Figure 5: The deterministic (blue) and stochastic expectation (green) of the voltage in an LC-circuit (left) and a fragment showing phase difference (right). The capacity is stochastic

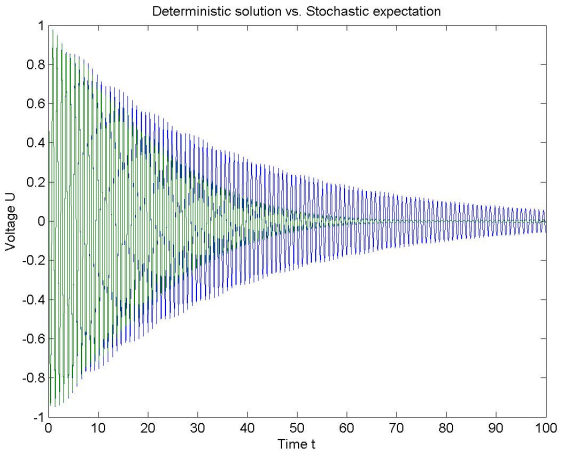

Figure 6: The deterministic (blue) and stochastic expectation (green) of the voltage in an LC-circuit (left). The induction is stochastic

Random Induction If we assume Induction to be stochastic then we expect an impact on the decay behaviour too. We can observe it on the Figure 6. At this point we can emphasize once again the importance of stochastic models. Suppose we want to maintain some certain voltage by repeating impulses after some time T, and calculate this time T taking the exact values. At the diagram we can see, that if the value of the induction is not certain, then the voltage will go below the limit before the calculated time point T.

Comparison with Monte-Carlo The standard tool for problems with random input parameters is Monte-Carlo which is well known for its slow convergence. On the Figure 7 we can see the deterministic solution (green) and the

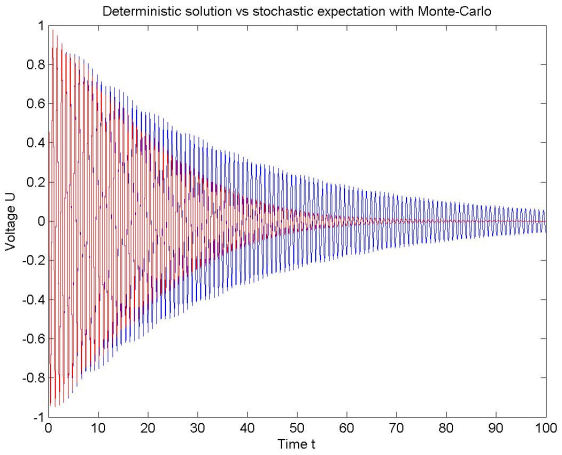

Figure 7: The deterministic (blue) and stochastic expectation calculated with Monte-Carlo (red) of the voltage in an LC-circuit (left). The capacity is stochastic

Monte-Carlo solution (red) using 20000 samples. The computation time is in the range of two hours, the result shows a clearly stronger decay and contradicts to the physical independence of the decay on capacity. Hence the convergence is not reached.

6 Summary and outline

Summary We have seen that deterministic models are always accompanied by stochastic ones, which arise whenever we plug in random variables instead of fixed values. And often those stochastic models are more appropriate to describe for instance the outcome of repeated experiments. We have specialized on models with ordinary differential equations and, using some functional analy-

sis we have constructed an efficient algorithm, which obviously conserves some physical properties and converges fast.

About the work The author worked at the topic during his work placement at Siemens AG. He had to analyze this new approach to the problem. Many examples he has studied may not be published in this paper as well as some new insights.

References

[1] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer Verlag, New York, 1991.

[2] Dongbin Xiu and George Em Karniadakis. The wiener-askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 2:619–644, 2002.