Numerical tricks for studying large rare fluctuations

R Mannella[1]
[1] Dipartimento di Fisica
Università di Pisa and INFM
Piazza Torricelli 2, Pisa ITALY
mannella@difi.unipi.it

September 7, 1998

Abstract

I present some tricks useful in the simulation of stochastic differential equations, with particular attention to the simulation of rare fluctuations, a topic of interest in the light of recent theoretical work on optimal paths. The "best algorithm" and some problems connected to the treatment of the boundaries will be discussed.

1 THE PROBLEM

Although algorithmic elegance always requires that the integrator used to step a differential equation should be optimised, to be able to approach the problem of the integration of a stochastic differential equation (SDE) in the presence of large rare fluctuations we need integration schemes which, with a "macroscopic" time step, faithfully reproduce the activation energies even after a very long integration run. Furthermore, given that typically the large rare fluctuation ought to be stopped at some boundary, we need algorithms to optimally end the integration when the boundary is hit.

1.1 Definitions

A stochastic differential equation has the generic form

$$\dot{x}_i = f_i(\vec{x}) + g_i(\vec{x})\xi(t) \tag{1}$$

$$\langle \xi(t) \rangle = 0 \ \ \langle \xi(t)\xi(s) \rangle = \delta(t-s), \tag{2}$$

and I will assume that the stochastic process ξ is Gaussian, that only one stochastic forcing is present (extensions are possible at the cost of some heavy algebra), and that the integral calculus used is Stratonovich. In the following, I call h the integration time step (more details and a full references list can be found in [1])

A simple minded approach to solve Eq. 1 is to formally integrate it, then to use a Taylor expansion around the point $t = 0$, to find recursively the various contributions.

1.2 Some algebra for a 1D problem

My 1D equation has the form

$$\dot{x} = f(x) + g(x)\xi(t) \tag{3}$$

A formal integration yields

$$x(h) - x(0) = \int_0^h (f(x(t)) + g(x(t))\xi(t))\, dt \tag{4}$$

Let me define $f_0 \equiv f(x(0))$, $g'_t = \frac{\partial g(x(t))}{\partial x(t)}$ and so on. By Taylor expansion I mean $f_t = f_0 + (x(t) - x(0))f'_0 + \ldots$ etc.. The simple minded lowest order in h seems to be

$$x(h)-x(0) = \int_0^h (f_0 + g_0\xi(t))\, dt = hf_0 + g_0 \int_0^h \xi(t)dt. \tag{5}$$

The stochastic integral

$$Z_1(h) = \int_0^h \xi(t)dt$$

is a linear combination of gaussian variables, hence a gaussian variable:

$$\langle Z_1 \rangle = 0 \;\; \langle Z_1^2(h) \rangle = h \;\; Z_1(h) = \sqrt{h}Y_1$$

where (and also in the following), the variables Y_i are gaussian deviates with zero average and standard deviation one. Here, however, we have a problem: really $Z_1(h) \sim \sqrt{h}$. This means that the "simple minded" lowest order is not correct. It turns out that the correct lowest order in h is

$$x(h) - x(0) = g_0 Z_1(h) + f_0 h + \frac{1}{2} g'_0 g_0 Z_1^2$$

Let me introduce other stochastic integrals which appear in the Taylor expansion at higher orders:

$$Z_2(h) = \int_0^h Z_1(s)ds = h^{3/2} \left\{ \frac{Y_1}{2} + \frac{Y_2}{2\sqrt{3}} \right\}$$

$$Z_3(h) = \int_0^h Z_1^2(s)ds \approx \frac{h^2}{3} \left\{ {Y_1}^2 + Y_3 + \frac{1}{2} \right\}$$

For additive noise ($g(x) = \sqrt{2D}$), including terms up to h^2 the algorithm reads

$$x(h) = x(0) + \sqrt{2D}Z_1(h) + f_0 h + \sqrt{2D}Z_2(h)f'_0 + DZ_3(h)f''_0 + \frac{h^2}{2} f'_0 f_0 \tag{6}$$

I will call this the "Full algorithm".
Other possible integration schemes which can be found in the literature are:
Euler scheme: Eq. 6, keeping only the first three terms on the r.h.s..
"Exact propagator": solve exactly $\dot{x} = f(x)$ and then add $Z_1(h)$ to take into account the noise.
Heun scheme: Use the following integrator:

$$\begin{aligned} x_1 &= x(0) + \sqrt{2D}Z_1(h) + f_0 h \\ x(h) &= x(0) + \sqrt{2D}Z_1(h) + \frac{h}{2}(f_0 + f(x_1)) \end{aligned}$$

2 WHICH IS THE BEST ALGORITHM?

The problem of the "best algorithm" for a SDE has two aspects:
Deterministic accuracy: one can disregard the stochastic component, and study only the deterministic integration scheme, using standard techniques. This gives some indications of the "deterministic accuracy" of the scheme considered.
Stochastic behaviour: in this case, one can check the Taylor expansion to "judge" the short times dynamics; and derive equilibrium properties from the propagator to check the large times properties

(there are other possibilities: for instance, one can consider which scheme is the closest to the bona fide trajectory under some measure). Focussing on the large times dynamics, the idea is to start from a generic form of the integrator, $x(t) = x(0) + F(x,t)$, find the equilibrium distribution it generates and compare it to the real one. It follows, writing the propagator, that

$$P(t+h) - P(t) = \sum_{n=1}^{\infty} \sum_{x_i} \frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_n} K_{1\ldots n} P(t)$$

where $P(t)$ is the probability distribution generated in the simulations and

$$K_{1\ldots n} \equiv (-1)^n \frac{1}{n!} \langle F_1 \ldots F_n \rangle_{noise}$$

In general, for systems in detailed balance,

$$P(x,\infty)_{sim} = P(x,\infty)_{true} \times \exp \sum_{n=1}^{\infty} h^n S_n / D$$

where all S_n would be zero if the algorithm were exact.

2.1 Deterministic behaviour

The analysis is carried out like for an ordinary differential equation. The error associated with a given integration scheme is easily evaluated; we find the following:
Euler scheme: accurate up to $O(h)$.
Exact propagator: no numerical error associated with this algorithm (by definition).
Heun scheme: accurate up to $O(h^2)$.
Full algorithm: accurate up to $O(h^2)$.

2.2 Stochastic behaviour

As mentioned, we can consider the short and the large time dynamics. Given that the short time dynamics is studied using a Taylor expansion, it is obvious that the best algorithm at short times is the "Full algorithm". For the equilibrium properties, i.e. the large times dynamics, I focus on the system

$$\dot{x} = -V'(x) + \sqrt{2D}\xi(t)$$

which has the true equilibrium distribution

$$P(x,\infty)_{true} = N \exp\{-V(x)/D\}$$

and the actual equilibrium distribution

$$P(x,\infty)_{sim} = N' \exp\{(-V(x) + hS(h,x))/D\}.$$

Carrying out the necessary algebra, it is straightforward to find for the function $S(x)$ (the "correction" which appears in the exponent and modifies the true equilibrium distribution)
Euler scheme: $S(h,x) = V'^2/4 - DV''/2$
Exact propagator: $S(h,x) = V'^2/2 - DV''/2$
Heun scheme: $S(h,x) = O(h)$
Full algorithm: $S(h,x) = O(h)$

It is clear that the "Full algorithm" and the Heun scheme are the algorithms which most faithfully reproduce the equilibrium distribution. It is also interesting to note that the "Exact propagator" does not do better than the Euler scheme: to improve an algorithm, in other words, it is necessary to take into account that the stochastic force does contribute to higher orders in the integration time step.

3 THE $\sqrt{h}$ PROBLEM AND BOUNDARIES

Regardless of the integration scheme used to carry out the integration, there is an intrinsic problem with a SDE, due to the discreteness of the integration and the sampling; this is a problem present even with an ideal integrator. Fig. 1 shows that a stochastic trajectory appears very different over different time scales (i.e., h's). The problem is most acute when the quantity of interest is, for instance, the Mean First Passage Time (MFPT) to a given boundary, like the dashed line in Fig. 1: the "decimated" trajectory (the trajectory one would obtain integrating with a larger h) simply misses the transition. The cure is simple: take the limit $h \to 0$ in the simulations. . . .

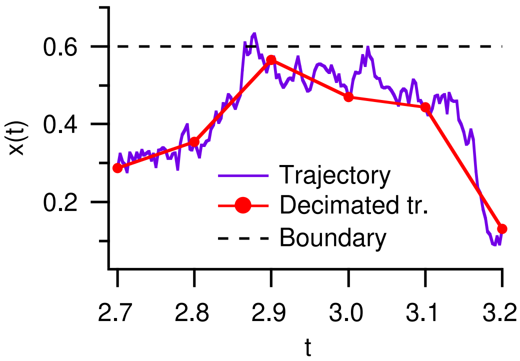

Fig. 1: Comparison between trajectories done with different time steps

In practice, this is done generating an auxiliary stochastic process which evaluates the probability that the particle hit the boundaries within the sampled points.

3.1 Free diffusion with absorbing boundaries

A test system is given by a bunch of particles injected at the origin, which can freely move in 1D, until they reach the boundaries, located at $\pm L$, where they are absorbed.

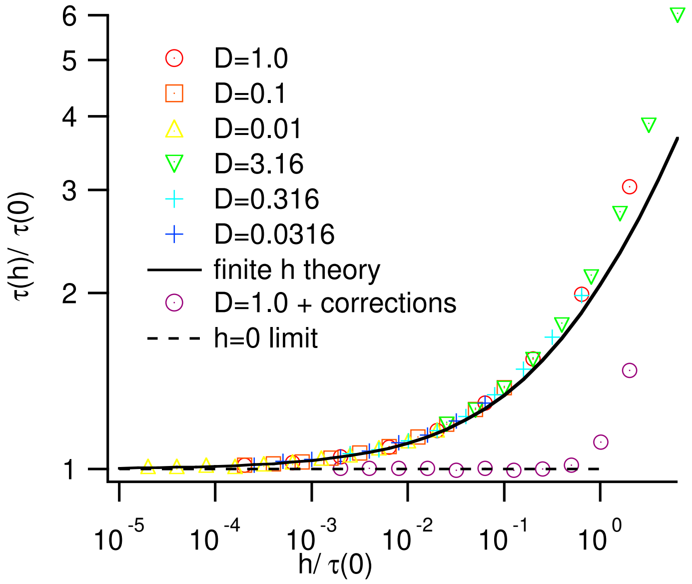

Figure 2: MFPT for a particle freely diffusing to a boundary

The system is defined by the equation

$$\dot{x} = \sqrt{2D}\xi(t), \tag{7}$$

where ξ is a gaussian process with zero average and standard deviation one. For $h = 0$, the MFPT to the boundary would be given by

$$\tau(0) = \frac{L^2}{2D},$$

which, in the presence of a finite h, becomes [2]

$$\tau(h)/\tau(0) = 1 + \sqrt{32/9\pi}\sqrt{h/\tau(0)}. \tag{8}$$

Figure 2 shows that indeed the "simple-minded" simulated MFPT's (symbols) follow Eq. 8 (solid line).

As soon as I introduce the correction in the integrating algorithm, the simulated points (lower circles) fall on the curve given by Eq. 8 (dashed line) with $h = 0$. Let me stress that the integration of this dynamical system is exact, due to the structure of Eq. 7.

3.2 MFPT in a bistable system

Given that the contribution to the MFPT which goes like $\sqrt{h}$ is only caused by the finiteness in the sampling of the stochastic trajectory, we expect that a similar contribution will show up virtually in any MFPT calculation, i.e. in any situation when we need to stop the trajectory when it reaches some boundary. I plotted in Fig. 3 the MFPT to escape from $x = -1$ to $x = 0$ in the system

$$\dot{x} = x - x^3 + \sqrt{2D}\xi(t). \tag{9}$$

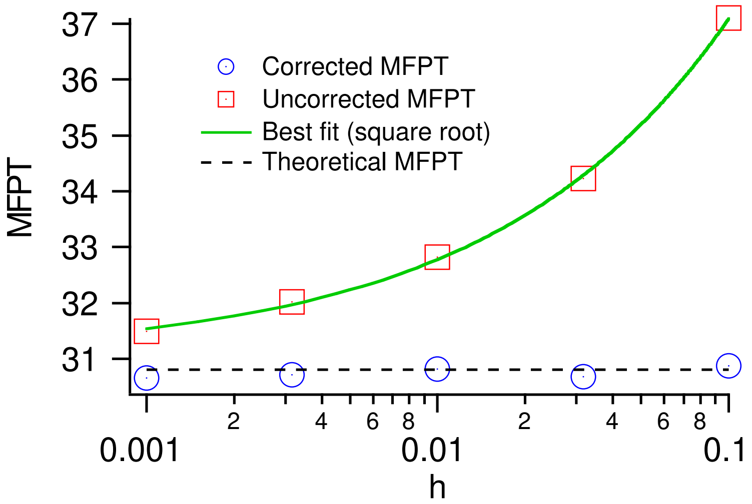

Figure 3: MFPT in the system of Eq. 9

From the figure, it is clear (squares) that the MFPT's computed without any correction in the SDE show a square root dependence on the integration time step. However, when I introduce the additional stochastic process to "simulate" the limit of $h \to 0$, the simulated points (circles) show no dependence on the integration time step. The SDE of Eq. 9 was integrated using the Heung algorithm, described in the previous section. The statistical error associated with the finite number of trajectories considered is order of the symbol dimensions.

4 CONCLUSIONS

I have shown that it is possible to derive integrators for SDE which are able to reproduce the equilibrium properties of a dynamical system to a high accuracy in the integration time step: it is clear, hence, that such integrators will be ideal to study long times dynamics of phenomena like a large rare fluctuation. I have also shown that it is possible to simulate properties of a zero integration time step process, which implies that we are able to determine with high accuracy when a stochastic trajectory reaches a given threshold.

References

[1] Mannella R., Numerical Integration of stochastic differential equations, in *Supercomputation in nonlinear and disordered systems*, Vasquez L, Tirado F and Martin I eds (World Scientific, 1997) pp. 100-129 and references therein.

[2] Strittmatter W, Numerical Simulation of the Mean First Passage Time, *University Freiburg preprint*, THEP 87/12.