Finite sampling interval effects in Kramers–Moyal analysis

Steven J. Lade

Nonlinear Physics Centre, Research School of Physics and Engineering, The Australian National University, ACT 0200, Australia

A R T I C L E   I N F O

Article history:
Received 27 May 2009
Received in revised form 11 August 2009
Accepted 12 August 2009
Available online 15 August 2009
Communicated by C.R. Doering

PACS:
05.10.Gg
05.45.Tp
87.15.hj

Keywords:
Kramers–Moyal coefficients
Fokker–Planck equation
Finite sampling interval
Adjoint operator
Molecular motors
Tethered diffusion

A B S T R A C T

Large sampling intervals can affect reconstruction of Kramers–Moyal coefficients from data. A new method, which is direct, non-stochastic and exact up to numerical accuracy, is developed to estimate these finite time effects. The method is applied numerically to biologically inspired examples. Exact finite time effects are also described analytically for two special cases. The approach developed will permit better evaluation of Langevin or Fokker–Planck based models from data with large sampling intervals. It can also be used to predict the sampling intervals for which finite time effects become significant.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Characterisation of a stochastic process $X(t)$ by computing the Kramers–Moyal coefficients

$$D^{(n)}(x_0,t) = \lim_{\tau \to 0} \frac{1}{n! \tau^n} \left[ \langle x(t+\tau) - x_0 \rangle^n \bigg| x(t) = x_0 \right].$$

from time series data has been recently proposed and developed [1]. The first two Kramers–Moyal coefficients are known as the drift and diffusion coefficients, respectively. They provide an intuitive description of the process, being analogous to effective position-dependent force and noise, or 'temperature', respectively [2]. If the process can be generated by a Langevin (or equivalently, Fokker–Planck) equation, they provide a complete characterisation; furthermore, all higher coefficients will be zero [3]. Unlike many other methods of time series analysis, the Kramers–Moyal approach recovers coefficients with arbitrary nonlinearity, up to the resolution of the bin size chosen in the reconstruction.

The Kramers–Moyal approach has previously been applied to neuroscience [1], cardiology [4], traffic engineering [2], finance [5,6] and turbulence [7]. These applications frequently assume homogeneity, in the sense that the Kramers–Moyal coefficients are time-invariant, so that the ensemble average in Eq. (1) can be replaced by a time average. Homogeneity will also be assumed throughout this article.

The definition (1) requires an infinitesimally small $\tau$, but the smallest $\tau$ available in an experiment is the sampling interval. If the sampling interval is not 'small enough', estimates of the Kramers–Moyal coefficients with this finite sampling interval may not be accurate. (Sometimes, even if shorter sampling intervals are available their use may not be desirable as they may uncover short-time correlations and other dynamics [8].) These finite sampling interval effects may be abbreviated to ‘finite time effects’ throughout this Letter.

Renner et al. [8] numerically investigated this dependence of the Kramers–Moyal coefficients on the sampling interval, Gradišek et al. [9] presented analytical expressions, and Sura and Barsugli [10] approximated error terms, but Ragwitz and Kantz [11] were the first to correct for finite sampling interval. To do so they calculated a partial [12] expansion of the Kramers–Moyal coefficients to first order in the sampling interval (second order in the conditional moments). An exchange with Friedrich et al. ensued [12,13], with Friedrich et al. providing an infinite series expansion for the finite time correction.

E-mail address: steven.lade@anu.edu.au.
These investigations focused on correcting experimentally estimated Kramers–Moyal coefficients for finite time effects. Experimental results and theoretical models could also be compared by introducing finite time effects into the theoretically predicted Kramers–Moyal coefficients. A trivial way of doing so is directly simulating the corresponding Langevin equations with a small time step, then undersampling to calculate the finite time coefficients. Kleinhans et al.\textsuperscript{[14,15]} used such an approach combined with optimisation procedures to estimate model parameters. Such a method is computationally intensive, requiring an average over many stochastic trajectories.

This Letter presents a new method of predicting finite sampling interval effects on Kramers–Moyal coefficients. It is both exact (up to numerical accuracy), unlike the first set of approaches above, and non-stochastic, unlike those of Kleinhans et al. By correcting the theoretical predictions for finite time effects, like Kleinhans et al., the method allows for the comparison of theoretical predictions and experimental data.

The rapidly growing collections of data available with recent advances in quantitative biology beckons use of Kramers–Moyal analysis, for biological systems are often well-modelled by an overdamped and Brownian environment. For example, one phase in the ‘walking’ of naturally occurring molecular motors such as myosin-V and kinesin is believed to be a tethered diffusion state, where one ‘head’ of the bipolar motor is unbound and searching for the next binding site\textsuperscript{[16–18]}. This system will be used to illustrate the results and procedures obtained in this Letter. The unbound head can be modelled as undergoing diffusion on a sphere with centre on the neck juncture and with radius equal to the length of the tether. This case will be referred to as ‘free diffusion on a sphere’. However, the neck juncture itself is not fixed, for its tether is flexible about its preferred orientation\textsuperscript{[19]}. The motion of the juncture will be modelled as diffusion on a sphere with a potential: ‘biased diffusion on a sphere’. Combining these models, a more precise treatment of the unbound head would have it undergoing free diffusion on a sphere about a point which is itself undergoing biased diffusion on a sphere: ‘compound diffusion’. In all cases it is assumed the experiment can only measure in one linear dimension, inclined at angle $\theta_0$ from the direction of the potential minimum (where there is a potential). These models are summarised in Fig. 1.

Fig. 1. Schematic of the biologically-inspired examples used in this Letter: (a) free diffusion on a sphere, or ‘tethered diffusion’; (b) biased diffusion on a sphere; and (c) compound diffusion on a sphere.

In Section 2 Friedrich et al.’s infinite series formula is reviewed, which then leads to the main result of this Letter. Section 3 calculates the finite time effects for analytical special cases, while Section 4 provides numerical examples. These examples are chosen for their relevance to possible biophysical uses of Kramers–Moyal analysis. Points of a general nature are made in Section 5, and concluding remarks in Section 6.

2. Derivation

Assume that the transition probability density $P(x,t|x',t')$ of a stochastic process $x(t)$ obeys a Fokker–Planck equation,

$$\frac{\partial P(x,t|x',t')}{\partial t} = \hat{L}(x)P(x,t|x',t'),$$

where the Fokker–Planck operator is

$$\hat{L}(x) = -\frac{\partial}{\partial x} D^{(1)}(x) + \frac{\partial^2}{\partial x^2} D^{(2)}(x).$$

Denote the Kramers–Moyal coefficients (1) calculated with finite sampling interval $\tau$ from this process by

$$D^{(n)}_{\tau}(x_0) = \frac{1}{n!\tau} \int [x(t+\tau) - x_0]^n |x(t) = x_0| \, dx,$$

This can also be written as

$$D^{(n)}_{\tau}(x_0) = \frac{1}{n!\tau} \int_{-\infty}^{\infty} (x - x_0)^n P(x, t_0 + \tau | x_0, t_0) \, dx,$$

which since homogeneity is assumed is independent of $t_0$. Substituting the formal solution to the Fokker–Planck equation (2)

$$P(x, t + \tau | x_0, t) = e^{L(t)\tau} \delta(x - x_0),$$

results in

$$D^{(n)}_{\tau}(x_0) = \frac{1}{n!\tau} \int_{-\infty}^{\infty} (x - x_0)^n e^{L(t)\tau} \delta(x - x_0) \, dx$$

$$= \frac{1}{n!\tau} e^{L(t)\tau} (x - x_0)^n |_{x=x_0}.$$

Here

$$\hat{L}^\dagger(\tau) = D^{(1)}(x) \frac{\partial}{\partial x} + D^{(2)} \frac{\partial^2}{\partial x^2}$$

is the adjoint of the operator $\hat{L}$ over the inner product $(g,h) = \int_{-\infty}^{\infty} g(x) h(x) \, dx$. Friedrich et al.\textsuperscript{[12]} expand the symbolic exponential and obtain the finite-order correction of Ragwitz and Kantz\textsuperscript{[11]}. Eq. (6) has an alternative interpretation, by a route Zwanzig\textsuperscript{[20]} calls the ‘Heisenberg approach’ in analogy to the complementary formulations of quantum mechanics. This interpretation is that the solution of the partial differential equation

$$\frac{\partial W(x,t)}{\partial t} = \hat{L}^\dagger(\tau) W(x,t)$$

at $(x_0, \tau)$, with initial condition $W(x,0) = (x - x_0)^n$, gives $D^{(n)}_{\tau}(x_0)$ (up to a factor $n!\tau$). Eq. (7) will be referred to as the adjoint Fokker–Planck equation; note it is different to, and should not be confused with, the Kolmogorov backward equation\textsuperscript{[3]}.

Therefore, knowing the continuous-time $D^{(1)}(x)$ and $D^{(2)}(x)$, the effect of finite sampling interval can be predicted exactly (up to numerical accuracy) without direct simulation of the Langevin equations. This is the main result of this article.

To reiterate, the procedure proposed for predicting the effects of a finite sampling interval $\tau$ is to:

(i) Predict the continuous-time Kramers–Moyal coefficients $D^{(n)}(x)$, analytically or numerically, from a theoretical model;
(ii) Solve Eq. (7) with initial condition \( W(x, 0) = (x - x_0)^2 \) up to time \( \tau = \tau' \);
(iii) Report \( W(x_0, \tau') / n! \tau'^n \) as the solution for \( D^{(n)}_r(x_0) \);
(iv) Repeat steps (ii) and (iii) for all the required points \( x_0 \) in the domain of \( D^{(n)}_r(x_0) \); then
(v) Repeat steps (ii)–(iv) for all the required orders \( n \) of the Kramers–Moyal coefficients (usually \( n = 1 \) and 2 for the drift and diffusion coefficients, respectively).

One other obvious non-stochastic approach is to compute Eq. (5) by solving the Fokker–Planck equation (2), but this involves a Dirac-delta initial condition. For some special cases, this initial condition could be treated analytically. Indeed, this is the method by which Ragwitz and Kantz [11] obtained their approximate corrections. In general, however, where numerical methods are required, the Dirac-delta initial condition would present significant difficulties.

3. Analytic special cases

In some cases the adjoint Fokker–Planck equation (7) can be solved analytically. One of the simplest scenarios is a linear drift \( D^{(1)}(x) = -kx \), corresponding to a quadratic potential well, and a constant diffusion \( D^{(2)}(x) = D \). This is the Ornstein–Uhlenbeck process in one dimension [3]. With initial condition \( W(x, 0) = x - x_0 \), the solution is \( W(x, \tau) = x e^{-k \tau} - x_0 \), so the finite time drift coefficient measured at finite sampling interval \( \tau \) is

\[
D^{(1)}_r(\tau) = -\frac{x}{\tau} (1 - e^{-k \tau}).
\]

With initial condition \( W(x, 0) = (x - x_0)^2 \), the solution is \( W(x, \tau) = (x e^{-k \tau} - x_0)^2 + \frac{D}{k} (1 - e^{-2k \tau}) \). Therefore the finite time diffusion coefficient is

\[
D^{(2)}_r(\tau) = \frac{1}{2\tau} \left[ x^2 (1 - e^{-2k \tau})^2 + \frac{D}{k} (1 - e^{-2k \tau}) \right].
\]

These solutions can also be obtained by substituting the transition probability density for the Ornstein–Uhlenbeck process, which is known in closed form [3], directly into Eq. (4).

As \( \tau \) increases, the finite time drift coefficient for the 1-D Ornstein–Uhlenbeck process remains linear but decreases in gradient, while a quadratic component appears in the diffusion. This causes the diffusion estimate to be larger at the edges of a potential well than at the centre, as also noted and discussed by Ragwitz and Kantz [11]. Gottschall and Peinke [21] recently obtained expansions of Eqs. (8) and (9) to first order in sampling interval, observing the same quadratic component in the diffusion.

Free diffusion on the surface of a sphere of radius \( r \), as sketched in Fig. 1(a), can be represented in polar co-ordinates \((\theta, \phi)\) by the Langevin equations [22]

\[
d\theta = D \cot \theta \, dt + \sqrt{2D} \, dw_\theta,
\]
\[
d\phi = \frac{\sqrt{2D}}{\sin \theta} \, dw_\phi,
\]

written in the Itô convention. Projected onto a rectangular co-ordinate \( x = r \cos \theta \), these reduce to the Langevin equation [23]

\[
dx = -2D x dt - \sqrt{2D (r^2 - x^2)} \, dw,
\]

with initial condition \( W(x, 0) = x - x_0 \), the solution to the adjoint Fokker–Planck equation (7) is \( W(x, \tau) = x e^{-2D \tau} - x_0 \), so the finite time drift coefficient is

\[
D^{(1)}_r(\tau) = -\frac{x}{\tau} (1 - e^{-2D \tau}).
\]

With initial condition \( W(x, 0) = (x - x_0)^2 \), the solution is \( W(x, \tau) = x^2 e^{-6D \tau} - 2x_0 x e^{-2D \tau} + x_0^2 - r^2 (e^{-6D \tau} - 1)/3 \). Therefore the finite time diffusion coefficient is

\[
D^{(2)}_r(\tau) = \frac{1}{2\tau} \left[ x^2 (1 + e^{-6D \tau} - 2e^{-2D \tau}) + \frac{r^2}{3} (1 - e^{-6D \tau}) \right].
\]

For free diffusion on a sphere, the gradient of the drift coefficient (11) again simply decreases with \( \tau \). Notice that the diffusion (12), however, undergoes a much more qualitatively significant change: it changes from convex to concave as \( \tau \) increases. This change occurs at \( \tau = \ln(1/2 + \sqrt{5}/2) / 2D \approx 1/4D \).

As a general rule of thumb, if the sampling interval is of order \( 1/k \) (for Ornstein–Uhlenbeck) or 1/4D (free diffusion on a sphere), or larger, one should be concerned about finite sampling interval effects. Also, a concave diffusion is an indicator there may be finite time effects.

In both examples the original \( D^{(1)}(x) \) and \( D^{(2)}(x) \) are recovered in the limit \( \tau \to 0 \).

4. Numerics

In general, to obtain the finite time Kramers–Moyal coefficients the solutions to Eq. (7) must be computed numerically. Two examples of such computations follow, using the procedure outlined in Section 2.

Biased diffusion on a sphere [Fig. 1(b)] can be modelled by the Langevin equations, in polar co-ordinates \((\theta, \phi)\), by [22]

\[
d\theta = (-\theta_0 U + D \cot \theta) \, dt + \sqrt{2D} \, dw_\theta,
\]
\[
d\phi = -\frac{\theta_0 U}{\sin^2 \theta} \, dt + \sqrt{2D} \, dw_\phi.
\]

A harmonic potential \( U = k \theta^2 / 2 \) will be used; for \( k = 0 \) the free diffusion on a sphere of the previous section is recovered. The system was projected onto one-dimensional (continuous-time) Kramers–Moyal coefficients by the method of Lade [23]. This dimension was the rectangular co-ordinate at angle \( \theta_0 \) from the preferred angle of orientation \( \theta = 0 \), that is, \( x = \cos \theta_0 \cos \theta + \sin \theta_0 \sin \theta \cos \phi \).

The finite time \( D^{(1)}(x) \) and \( D^{(2)}(x) \) were then estimated for a variety of \( \tau \) with the procedure outlined in Section 2. The adjoint Fokker–Planck equation was numerically solved with a simple forward-time centred-space scheme and extrapolated boundary conditions. For comparison, the drift and diffusion coefficients were also computed, as per Eq. (3), from direct simulation of the Langevin equations (13). The results are shown in Fig. 2.

The same decrease in gradient of the drift as in the Ornstein–Uhlenbeck process occurs. Since the drift is roughly displacement divided by sampling interval \( \tau \), and since the displacement of the projection \( x \) between measurements is limited (by the radius of the sphere), it makes sense that the estimated drift decreases with increasing \( \tau \). Unlike the Ornstein–Uhlenbeck process, there is some curvature due to geometrical effects. The zero crossing point reflects the position of the potential minimum at \( x = \cos \theta_0 \). The diffusion undergoes a loss of features and convergence to an upright quadratic shape as in the Ornstein–Uhlenbeck process. As observed by Ragwitz and Kantz [11], this shape is because at large \( \tau \), nonzero drift can lead to overestimation of diffusion, and the drift is larger at large position \( x \). In both drift and diffusion curves, there are errors in the direct numerical estimates due to the singularities in the Langevin equation (13) at \( \theta = 0 \) and \( \pi \).

As a second example, compound diffusion on a sphere [Fig. 1(c)] was simulated, where a point undergoes free diffusion on a sphere [10] with respect to a point that is itself undergoing biased diffusion on a sphere (13). Results are in Fig. 3.
and Farahpour et al. [24] and Ghasemi et al. [6], and was used by them in fitting the results of their analyses. If one assumes their process was actually Ornstein–Uhlenbeck, using their sampling intervals and the analytic solutions (8)–(9) on each case, and find a good match with their numerical drift and diffusion curves. Note that, given \( \tau \), this problem is overdetermined: the functional forms leave three simultaneous equations from which to determine \( k \) and \( D \). Thus the processes investigated by those authors (specifically, the logarithmic increments) are well modelled in continuous time by the above simple Ornstein–Uhlenbeck process.

The drift and diffusion coefficients observed in some cascade analyses of time series [5,25,26] also display patterns similar to the above, with the scale size standing in for sampling interval. For example, in the simple Euler–Maruyama [27] algorithm, \( D_t \) by construction provides an unbiased estimate for \( x(t + \tau) \) from \( x(t) \).

It was assumed that the evolution of the time series can be fully described by the Fokker–Planck equation (2) and the drift and diffusion coefficients. This is valid if the process is Markovian. Projections of a larger-dimensional system, such as those used in Section 4, are in general not Markov, even when the original system is [20]. However, it has been shown that the projections of Section 4 are approximately Markov [23], and the excellent agreement with direct numerical simulations in that section confirms that the method worked successfully. Also, the method could easily be extended to other evolution operators \( \tilde{L} \) than the Fokker–Planck operator, in the event that this operator is not appropriate.

The numerical method outlined above, preferably with more sophisticated methods of numerical solution, could be inverted iteratively to find theoretical model parameters from experimentally measured finite time coefficients, as in the stochastic approach of Kleinhans et al. [14,15].

Given a known model for a system, the approach proposed in this Letter can be used to predict the effect of finite experimental sampling intervals on the estimation of Kramers–Moyal

5. Discussion

The characteristic quadratic shape obtained in Section 3 for the diffusion coefficient at large sampling intervals is clearly observable in Farahpour et al. [24] and Ghasemi et al. [6], and was used by them in fitting the results of their analyses. If one assumes their process was actually Ornstein–Uhlenbeck, using their sampling intervals and the analytic solutions (8)–(9) one can solve for \( k \) and \( D \) in each case, and find a good match with their numerical drift and diffusion curves. Note that, given \( \tau \), this problem is overdetermined: the functional forms leave three simultaneous equations from which to determine \( k \) and \( D \). Thus the processes investigated by those authors (specifically, the logarithmic increments) are well modelled in continuous time by the above simple Ornstein–Uhlenbeck process.

The drift and diffusion coefficients observed in some cascade analyses of time series [5,25,26] also display patterns similar to the above, with the scale size standing in for sampling interval. For example, in the simple Euler–Maruyama [27] algorithm, \( D_t \) by construction provides an unbiased estimate for \( x(t + \tau) \) from \( x(t) \).

It was assumed that the evolution of the time series can be fully described by the Fokker–Planck equation (2) and the drift and diffusion coefficients. This is valid if the process is Markovian. Projections of a larger-dimensional system, such as those used in Section 4, are in general not Markov, even when the original system is [20]. However, it has been shown that the projections of Section 4 are approximately Markov [23], and the excellent agreement with direct numerical simulations in that section confirms that the method worked successfully. Also, the method could easily be extended to other evolution operators \( \tilde{L} \) than the Fokker–Planck operator, in the event that this operator is not appropriate.

The numerical method outlined above, preferably with more sophisticated methods of numerical solution, could be inverted iteratively to find theoretical model parameters from experimentally measured finite time coefficients, as in the stochastic approach of Kleinhans et al. [14,15].

Given a known model for a system, the approach proposed in this Letter can be used to predict the effect of finite experimental sampling intervals on the estimation of Kramers–Moyal...
coefficients, without direct simulation of the stochastic Langevin equations. Experimental observations of the Kramers–Moyal coefficients could then be used to support or reject the model.

Given a time series, a quick test to see whether finite time effects may occur could be to check whether the sampling interval approaches any characteristic periods of the system. This would occur if the differences in the time series over a single time step are ‘large’ compared to the full extent of the time series’ fluctuations. To check whether the Kramers–Moyal coefficients change with an undersampling is another simple test.

6. Conclusions

An alternative interpretation of an existing formula (6) was proposed, which permits exact, non-stochastic prediction of finite sampling interval effects on the estimation of Kramers–Moyal coefficients. Special analytical cases were presented, which showed general features of finite sampling interval effects. The approach was then implemented numerically, for examples of particular relevance to biophysics. The results were in good agreement with direct stochastic simulation. Previously published Kramers–Moyal analyses showed features possibly explicable as finite time effects, in particular, concave diffusion curves. The method proposed may permit better evaluation of Langevin or Fokker–Planck based models with data that has large sampling intervals, or to predict the sampling intervals for which finite time effects should become significant.

References

[1] R. Friedrich, S. Siegert, J. Peinke, S. Lück, M. Siefert, M. Lindemann, J. Raethjen, G. Deuschl, G. Pfister, Phys. Lett. A 271 (2000) 217.

[2] S. Kriso, J. Peinke, R. Friedrich, P. Wagner, Phys. Lett. A 299 (2–3) (2002) 287.

[3] H. Risken, The Fokker–Planck Equation: Methods of Solution and Applications, Springer-Verlag, Berlin, 1984.

[4] T. Kuusela, Phys. Rev. E 69 (3) (2004) 031916.

[5] R. Friedrich, J. Peinke, C. Renner, Phys. Rev. Lett. 84 (22) (2000) 5224.

[6] F. Ghasemi, M. Sahimi, J. Peinke, R. Friedrich, G.R. Jafari, M.R.R. Tabar, Phys. Rev. E 75 (6) (2007) 066102(R).

[7] R. Friedrich, J. Peinke, Phys. Rev. Lett. 78 (5) (1997) 863.

[8] C. Renner, J. Peinke, R. Friedrich, J. Fluid Mech. 433 (2001) 383.

[9] J. Gradiešek, S. Siegert, R. Friedrich, I. Grabec, Phys. Rev. E 62 (3) (2000) 3146.

[10] P. Sura, J. Barsugli, Phys. Lett. A 305 (5) (2002) 304.

[11] M. Ragwitz, H. Kantz, Phys. Rev. Lett. 87 (25) (2001) 254501.

[12] R. Friedrich, C. Renner, M. Siefert, J. Peinke, Phys. Rev. Lett. 89 (14) (2002) 149401.

[13] M. Ragwitz, H. Kantz, Phys. Rev. Lett. 89 (14) (2002) 149402.

[14] D. Kleinhans, R. Friedrich, A. Nawroth, J. Peinke, Phys. Lett. A 346 (1–3) (2005) 42.

[15] D. Kleinhans, R. Friedrich, Phys. Lett. A 368 (2007) 194.

[16] S. Burgess, M. Walker, F. Wang, J.R. Sellers, H.D. White, P.J. Knight, J. Trinick, J. Cell Biol. 159 (6) (2002) 983.

[17] A.R. Dunn, J.A. Spudich, Nature Struct. Mol. Biol. 14 (2007) 246.

[18] K. Shiroguchi, K. Kinoshita Jr., Science 316 (5828) (2007) 1208.

[19] A. Vilfan, Biophys. J. 88 (6) (2005) 3792.

[20] R. Zwanzig, Nonequilibrium Statistical Mechanics, Oxford University Press, New York, 2001.

[21] J. Gottschall, J. Peinke, New J. Phys. 10 (8) (2008) 083034.

[22] M. Raible, A. Engel, Appl. Organometal. Chem. 18 (2004) 536.

[23] S.J. Lade, Geometric and projection effects in Kramers–Moyal analysis, Phys. Rev. E, in press.

[24] F. Farahpour, Z. Eskandari, A. Bahraminasab, G. Jafari, F. Ghasemi, M. Sahimi, M. Reza Rahimi Tabar, Physica A 385 (2) (2007) 601.

[25] S. Kimiagar, G.R. Jafari, M.R.R. Tabar, J. Stat. Mech. (2008) 02010.

[26] G.R. Jafari, S.M. Fazeli, F. Ghasemi, S.M. Vaez Allaei, M. Reza Rahimi Tabar, A. Iradjizad, G. Kavei, Phys. Rev. Lett. 91 (22) (2003) 226101.

[27] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer-Verlag, Berlin, 1992.