Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
C. R. Acad. Sci. Paris, Ser. I 338 (2004) 481–486
COMPTES RENDUS MATHEMATIQUE

Probability Theory

Geometrization of Monte-Carlo numerical analysis of an elliptic operator: strong approximation

Ana Bela Cruzeiro [a], Paul Malliavin [b], Anton Thalmaier [c]

[a] *Grupo de Física-Matemática UL and Dep. Matemática IST, Av. Rovisco Pais, 1049-001 Lisboa, Portugal*
[b] *10, rue Saint Louis en l'Isle, 75004 Paris, France*
[c] *Département de mathématiques, Université de Poitiers, téléport 2, BP 30179, 86962 Futuroscope Chasseneuil, France*

Received 16 December 2003; accepted 12 January 2004

Presented by Paul Malliavin

Abstract

A one-step scheme is constructed, which, as the Milstein scheme, has the strong approximation property of order 1; in contrast to the Milstein scheme, our scheme does not involve the simulation of iterated Itô integrals of second order. ***To cite this article: A.B. Cruzeiro et al., C. R. Acad. Sci. Paris, Ser. I 338 (2004).***

Résumé

Monte-Carlo géométrique pour un opérateur elliptique : approximation numérique forte. On propose un schéma à un pas, qui, comme le schéma de Milstein, possède la propriété d'approximation forte à l'ordre 1 ; contrairement au schéma de Milstein, notre schéma ne nécessite pas la simulation d'intégrales itérées de Itô du second degré. ***Pour citer cet article : A.B. Cruzeiro et al., C. R. Acad. Sci. Paris, Ser. I 338 (2004).***

1. Introduction

Numerical integration of SDE consists, an SDE being fixed, in producing a discretization scheme leading to a pathwise approximation of its solution. In this work we discuss the related problem of how to establish a Monte-Carlo simulation to a given real strictly elliptic second order operator $\mathcal{L}$, which leads to good numerical approximations of the fundamental solution to the corresponding heat operator.

Of course many SDE can be associated to $\mathcal{L}$, each choice corresponding to a parametrization by the Wiener space of the Stroock–Varadhan solution of the martingale problem associated to $\mathcal{L}$. For applications to finance all these parametrizations are equivalent. We shall prove that there exists an optimal parametrization for which the one-step Milstein scheme does not involve the computation of iterated stochastic integrals of second order. For our

E-mail addresses: abcruz@math.ist.utl.pt (A.B. Cruzeiro), sli@ccr.jussieu.fr (P. Malliavin), thalmaier@math.univ-poitiers.fr (A. Thalmaier).

1631-073X/$ – see front matter © 2004 Published by Elsevier SAS on behalf of Académie des sciences.
doi:10.1016/j.crma.2004.01.007

search of an optimal parametrization we have to describe the possible parametrizations of the diffusion associated to $\mathcal{L}$; it is sufficient to solve this problem for the Euclidean Laplacian on $\mathbb{R}^d$; in this case it is equivalent to replace the standard Brownian motion W on $\mathbb{R}^d$ by an *orthogonal transform* $\widetilde{W}$ with an Itô differential of the type

$$\mathrm{d}\widetilde{W}_k = \sum_{j=1}^{d} \big[\Omega_W(t)\big]_j^k \,\mathrm{d}W_j(t), \tag{1}$$

where $t \mapsto \Omega_W(t)$ is an adapted process taking values in the group $\mathrm{O}(d)$ of orthogonal matrices. We denote by $\mathrm{O}(\mathcal{W})$ the family of all such orthogonal transforms which is isomorphic to $\mathbb{P}(\mathrm{O}(d))$, the path space on $\mathrm{O}(d)$. Pointwise multiplication defines a group structure on $\mathbb{P}(\mathrm{O}(d)) \simeq \mathrm{O}(\mathcal{W})$.

Given on $\mathbb{R}^d$ the data of $d+1$ smooth vector fields $A_0, A_1, \ldots, A_d$, we consider the Itô SDE

$$\mathrm{d}\xi_W(t) = A_0\big(\xi_W(t)\big)\,\mathrm{d}t + \sum_{k=1}^{d} A_k\big(\xi_W(t)\big)\,\mathrm{d}W_k(t), \quad \xi_W(0) = \xi_0. \tag{2}$$

Throughout this Note we assume ellipticity, that is, for any $\xi \in \mathbb{R}^d$ the vectors $A_1(\xi), \ldots, A_d(\xi)$ constitute a basis of $\mathbb{R}^d$; the components of a vector field U in this basis are denoted $\langle U, A_k\rangle_\xi$ which gives the decomposition $U(\xi) = \sum_{k=1}^d \langle U, A_k\rangle_\xi\, A_k(\xi)$. By *change of parametrization* we mean the substitution of W by $\widetilde{W}$ in (2); we then get an Itô process in $\widetilde{W}$. This change of parametrization does not change the infinitesimal generator associated to (2) which has the form $\mathcal{L} = \frac{1}{2}\sum_{k,\alpha,\beta} A_k^\alpha A_k^\beta D_\alpha D_\beta + \sum_\alpha A_0^\alpha D_\alpha$ where $D_\alpha = \partial/\partial\xi^\alpha$. The group $\mathrm{O}(\mathcal{W})$ operates on the set of elliptic SDE on $\mathbb{R}^d$ and the orbits of this action are classified by the corresponding elliptic operators $\mathcal{L}$.

2. Definition of the scheme $\mathcal{S}$

Denote by $t_\varepsilon := \varepsilon \times$ integer part of t/ε; we define our scheme by

$$\begin{aligned} Z_{W^\varepsilon}(t) - Z_{W^\varepsilon}(t_\varepsilon) &= A_0\big(Z_{W^\varepsilon}(t_\varepsilon)\big)(t - t_\varepsilon) + \sum_k A_k\big(Z_{W^\varepsilon}(t_\varepsilon)\big)\big(W_k(t) - W_k(t_\varepsilon)\big) \\ &\quad + \frac{1}{2}\sum_{k,s} (\partial_{A_k} A_s)\big(Z_{W^\varepsilon}(t_\varepsilon)\big)\big\{\big(W_k(t) - W_k(t_\varepsilon)\big)\big(W_s(t) - W_s(t_\varepsilon)\big) - \varepsilon\eta_k^s\big\} \\ &\quad + \frac{1}{2}\sum_{k,s,i} A_i\big(Z_{W^\varepsilon}(t_\varepsilon)\big)\big\langle [A_s, A_i], A_k\big\rangle_{Z_{W^\varepsilon}(t_\varepsilon)}\big\{\big(W_k(t) - W_k(t_\varepsilon)\big)\big(W_s(t) - W_s(t_\varepsilon)\big) - \varepsilon\eta_k^s\big\}, \end{aligned} \tag{3}$$

where W is standard Brownian motion on $\mathbb{R}^d$, and η_k^s the Kronecker symbol defined by $\eta_k^s = 1$ if $k = s$ and zero otherwise. Denote by $\mathbb{P}(\mathbb{R}^d)$ *the path space* on $\mathbb{R}^d$, that is the Banach space of continuous maps from $[0, T]$ into $\mathbb{R}^d$, endowed with the sup norm: $\|p_1 - p_2\|_\infty = \sup_{t\in[0,T]} |p_1(t) - p_2(t)|_{\mathbb{R}^d}$. Fixing $\xi_0 \in \mathbb{R}^d$, let $\mathbb{P}_{\xi_0}(\mathbb{R}^d)$ be the subspace of paths starting from ξ_0. Given Borel measures $\rho_1,\ \rho_2$ on $\mathbb{P}(\mathbb{R}^d)$, denote by $\mathcal{M}(\rho_1, \rho_2)$ the set of measurable maps $\Psi : \mathbb{P}(\mathbb{R}^d) \to \mathbb{P}(\mathbb{R}^d)$ such that $\Psi_*\rho_1 = \rho_2$; the *Monge transport norm* (see [8,5]) is defined as

$$d_{\mathcal{M}}(\rho_1, \rho_2) := \Big[\inf_{\Psi\in\mathcal{M}(\rho_1,\rho_2)} \int \big\|\Psi(p)\big\|_\infty^2 \rho_1(\mathrm{d}p)\Big]^{1/2}.$$

Theorem 2.1. *Assume ellipticity and assume the vector fields A_k along with their first three derivatives to be bounded; fix $\xi_0 \in \mathbb{R}^d$ and let $\rho_{\mathcal{L}}$ be the measure on $P_{\xi_0}(\mathbb{R}^d)$ defined by the solution of the Stroock–Varadhan martingale problem* [7] *for the elliptic operator $\mathcal{L}$; let $\rho_{\mathcal{S}}$ be the measure obtained by the scheme $\mathcal{S}$ with initial value $Z_{W^\varepsilon}(0) = \xi_0$. Then*

$$\limsup_{\varepsilon\to 0} \frac{1}{\varepsilon} d_{\mathcal{M}}(\rho_{\mathcal{L}}, \rho_{\mathcal{S}}) = c < \infty. \tag{4}$$

Remark. The proof of Theorem 2.1 will provide an explicit transport functional Ψ_0 which puts the statement in a constructive setting; the constant c is effective.

3. The Milstein scheme

The Milstein scheme for SDE (2) (cf., for instance, [6], formula (0.23) or [2], p. 345; see also [3]) is based on the following stochastic Taylor expansion of A_k along the diffusion trajectory: $A_k(\xi_W(t)) = A_k(\xi_W(t_\varepsilon)) + \sum_j (\partial_{A_j} A_k)(\xi_W(t_\varepsilon))(W_j(t) - W_j(t_\varepsilon)) + \mathrm{O}(\varepsilon)$, which leads to

$$\xi_{W^\varepsilon}(t) - \xi_{W^\varepsilon}(t_\varepsilon) = \sum_k A_k\big(\xi_{W^\varepsilon}(t_\varepsilon)\big)\big(W_k(t) - W_k(t_\varepsilon)\big) + (t - t_\varepsilon)A_0\big(\xi_{W^\varepsilon}(t_\varepsilon)\big) + \sum_{i,k} (\partial_{A_i} A_k)\big(\xi_{W^\varepsilon}(t_\varepsilon)\big) \int_{t_\varepsilon}^{t} \big(W_i(s) - W_i(t_\varepsilon)\big)\,\mathrm{d}W_k(s);$$

the computation of $\int_{t_\varepsilon}^{t}(W_i(s) - W_i(t_\varepsilon))\,\mathrm{d}W_k(s)$ gives the Milstein scheme

$$\xi_{W^\varepsilon}(t) - \xi_{W^\varepsilon}(t_\varepsilon) = \sum_k A_k\big(\xi_{W^\varepsilon}(t_\varepsilon)\big)\big(W_k(t) - W_k(t_\varepsilon)\big) + (t - t_\varepsilon)A_0\big(\xi_{W^\varepsilon}(t_\varepsilon)\big) + \frac{1}{2}\sum_{i,k} (\partial_{A_i} A_k)\big(\xi_{W^\varepsilon}(t_\varepsilon)\big)\Big(\big(W_i(t) - W_i(t_\varepsilon)\big)\big(W_k(t) - W_k(t_\varepsilon)\big) - \varepsilon\eta_k^i\Big) + R,$$

where $R = \sum_{i<k}[A_i, A_k](\xi_{W^\varepsilon}(t_\varepsilon)) \int_{t_\varepsilon}^{t}(W_i(s) - W_i(t_\varepsilon))\,\mathrm{d}W_k(s) - (W_k(s) - W_k(t_\varepsilon))\,\mathrm{d}W_i(s)$.

It is well known that the Milstein scheme has the following strong approximation property:

$$\mathbb{E}\Big[\sup_{t\in[0,1]} \big\|\xi_W(t) - \xi_{W^\varepsilon}(t)\big\|^2\Big] = \mathrm{O}\big(\varepsilon^2\big). \tag{5}$$

The numerical difficulty related to the Milstein scheme is how to achieve a fast simulation of R. The purpose of this work is to show that by a change of parametrization this simulation can be avoided.

4. Horizontal parametrization

Given d independent vector fields $A_1, \ldots, A_d$ on $\mathbb{R}^d$, we take the vectors $A_1(\xi), \ldots, A_d(\xi)$ as basis at the point ξ; the functions $\beta_{k,s}^i$, called *structural functions*, are defined (see [1]) by:

$$\beta_{k,\ell}^i(\xi) = \big\langle [A_k, A_\ell], A_i\big\rangle_\xi, \quad [A_k, A_\ell](\xi) = \sum_i \beta_{k,\ell}^i(\xi) A_i(\xi).$$

The structural functions are antisymmetric with respect to the two lower indices. Consider the *connection functions*, defined from the structural functions by

$$\Gamma_{k,s}^i = \frac{1}{2}\big(\beta_{k,s}^i - \beta_{s,i}^k + \beta_{i,k}^s\big). \tag{6}$$

Let Γ_k be the $d \times d$ matrix obtained by fixing the index k in the three indices functions $\Gamma_{k,*}^*$. Then, by means of the antisymmetry of $\beta_{*,*}^*$ in the two lower indices, Γ_k is an antisymmetric matrix:

$$2\big(\Gamma_{k,s}^i + \Gamma_{k,i}^s\big) = \beta_{k,s}^i - \beta_{s,i}^k + \beta_{i,k}^s + \beta_{k,i}^s - \beta_{i,s}^k + \beta_{s,k}^i = 0.$$

The matrix Γ_k operates on the coordinate vectors of the basis $A_s(\xi)$ via $\Gamma_k(A_s) = \sum_i \Gamma_{k,s}^i A_i$. This gives $\Gamma_k(A_s) - \Gamma_s(A_k) = [A_k, A_s]$: the ith component of the l.h.s. is $\frac{1}{2}(\beta_{k,s}^i - \beta_{s,i}^k + \beta_{i,k}^s - \beta_{s,k}^i + \beta_{k,i}^s - \beta_{i,s}^k) = \beta_{k,s}^i$. Let $\mathbb{M} = \mathbb{R}^d \times \mathcal{E}_d$ where $\mathcal{E}_d$ is the vector space of $d \times d$ matrices. Define on $\mathbb{M}$ vector fields $\tilde{A}_k,\ k = 1, \ldots, d$, as follows:

$$\tilde{A}_k(\xi,e)=\left(\sum_\ell e_k^\ell A_\ell(\xi),\mathcal{N}_k(\xi,e)\right),\quad [\mathcal{N}_k]_r^s(\xi,e)=-\sum_{\ell,\ell'} e_k^\ell e_r^{\ell'}\Gamma_{\ell,\ell'}^s(\xi),\quad \xi\in\mathbb{R}^d,\ e\in\mathcal{E}_d. \tag{7}$$

Denoting for a vector Z on $\mathbb{M}$ by Z^H its projection on $\mathbb{R}^d$, we have:

Proposition 4.1. *The vector fields $\tilde{A}_k$ satisfy the relation $[\tilde{A}_k,\tilde{A}_s]^{\mathrm{H}}=0$.*

Proof. The horizontal component $[\partial_{\tilde{A}_k}\tilde{A}_s]^{\mathrm{H}}$ is given by

$$\begin{aligned}
[\partial_{\tilde{A}_k}\tilde{A}_s]^{\mathrm{H}} &= \sum_i \tilde{A}_k^i\,\partial_i\tilde{A}_s^{\mathrm{H}}+\sum_q\left(\sum_{\alpha,\beta}[\mathcal{N}_k]_\beta^\alpha\,\partial_{\binom{\alpha}{\beta}}e_s^q\right)A_q\\
&=\sum_i\left(\sum_\ell e_k^\ell A_\ell^i\right)\left(\sum_{\ell'} e_s^{\ell'}\partial_i A_{\ell'}\right)-\sum_{\alpha,\beta}\left(\sum_{\ell,\ell'} e_k^\ell e_\beta^{\ell'}\,\Gamma_{\ell,\ell'}^\alpha\right)\left(\sum_q \eta_\alpha^q\eta_\beta^s A_q\right)\\
&=\sum_i\left(\sum_\ell e_k^\ell A_\ell^i\right)\left(\sum_{\ell'} e_s^{\ell'}\partial_i A_{\ell'}\right)-\sum_{\alpha,\ell,\ell'} e_k^\ell\,e_s^{\ell'}\,\Gamma_{\ell,\ell'}^\alpha A_\alpha,
\end{aligned}$$

using the fact that $\partial_{\binom{\alpha}{\beta}}e_s^q=\eta_\alpha^q\eta_\beta^s$. We finally get $[\partial_{\tilde{A}_k}\tilde{A}_s]^{\mathrm{H}}=\sum_{\ell,\ell'}e_k^\ell e_s^{\ell'}[\partial_{A_\ell}A_{\ell'}-\sum_\alpha\Gamma_{\ell,\ell'}^\alpha A_\alpha]$. Therefore the horizontal component of the commutator is

$$[\tilde{A}_k,\tilde{A}_s]^{\mathrm{H}}=\sum_{\ell,\ell'}e_k^\ell e_s^{\ell'}[A_\ell,A_{\ell'}]-\sum_{\alpha,\ell,\ell'}e_k^\ell e_s^{\ell'}\big(\Gamma_{\ell,\ell'}^\alpha-\Gamma_{\ell',\ell}^\alpha\big)A_\alpha$$

which vanishes since $\Gamma_\ell(A_{\ell'})-\Gamma_{\ell'}(A_\ell)=[A_\ell,A_{\ell'}]$. □

Denote by e^{T} the transposed of the matrix e and let $\tilde{A}_0(\xi,e)=(A_0(\xi),-\frac{1}{2}eJ)$, $J:=\sum_{k=1}^d\mathcal{N}_k^{\mathrm{T}}\mathcal{N}_k$. Consider the following Itô SDE on the vector space $\mathbb{M}$:

$$\mathrm{d}m_W=\sum_k\tilde{A}_k(m_W)\,\mathrm{d}W_k+\tilde{A}_0(m_W)\,\mathrm{d}t,\quad m_W(0)=(\xi_0,\mathrm{Id}). \tag{8}$$

Proposition 4.2. *Denote $m_W(t)=(\xi_W(t),e_W(t))$, then $e_W(t)$ is an orthogonal matrix for $t\geqslant 0$, and*

$$f\big(\xi_W(t)\big)-\int_0^t(\mathcal{L}f)\big(\xi_W(s)\big)\,\mathrm{d}s\ \textit{is a local martingale, for any } f\in C^2\big(\mathbb{R}^d\big). \tag{9}$$

Proof. We compute the stochastic differential of $e^{\mathrm{T}}e$:

$$\begin{aligned}
\mathrm{d}\big[e^{\mathrm{T}}e\big]_\ell^{\ell'}=\sum_k d\big(e_k^\ell e_k^{\ell'}\big)&=-\sum_{m,k,p}\left(\sum_u e_k^\ell e_m^p e_k^u\Gamma_{p,u}^{\ell'}+\sum_v e_k^{\ell'}e_m^p e_k^v\,\Gamma_{p,v}^\ell\right)\mathrm{d}W_m\\
&\quad+\sum_k\left(e_k^\ell\,(\tilde{A}_0)_k^{\ell'}+e_k^\ell\,(\tilde{A}_0)_k^{\ell'}+\sum_{m,p,q,p',q'}e_m^p e_k^{p'}\Gamma_{p,p'}^{\ell'}e_m^q e_k^{q'}\Gamma_{q,q'}^\ell\right)\mathrm{d}t,
\end{aligned}$$

where the last term of the drift comes from the Itô contraction $\sum_k \mathrm{d}e_k^\ell*\mathrm{d}e_k^{\ell'}=\sum_k[\mathcal{N}_k^{\mathrm{T}}\mathcal{N}_k]_\ell^{\ell'}\,\mathrm{d}t=J_\ell^{\ell'}\,\mathrm{d}t$. The first two terms of the drift are computed by using the definition of $\tilde{A}_0$: $\sum_k[e_k^\ell(\tilde{A}_0)_k^{\ell'}+e_k^{\ell'}(\tilde{A}_0)_k^\ell]=-\frac{1}{2}[(e^{\mathrm{T}}eJ)_\ell^{\ell'}+(e^{\mathrm{T}}eJ)_{\ell'}^\ell]$. Write $e^{\mathrm{T}}e=\mathrm{Id}+\sigma$, then the drift takes the form $-(\sigma J+J\sigma)/2$. We compute the coefficient of $\mathrm{d}W_m$:

$$-\sum_{k,p}\left(\sum_u e_k^\ell e_m^p e_k^u \Gamma_{p,u}^{\ell'} + \sum_v e_k^{\ell'} e_m^p e_k^v \Gamma_{p,v}^\ell\right) = -\sum_p e_m^p\left(\sum_u \left[e^{\mathrm{T}}e\right]_\ell^u \Gamma_{p,u}^{\ell'} + \sum_v \left[e^{\mathrm{T}}e\right]_{\ell'}^v \Gamma_{p,v}^\ell\right).$$

Using the antisymmetry $\Gamma_{p,\ell}^{\ell'} = -\Gamma_{p,\ell'}^{\ell}$ we obtain

$$\mathrm{d}\sigma_{\ell'}^\ell = -\sum_m \mathrm{d}W_m \sum_p e_m^p\left(\sum_u \sigma_\ell^u \Gamma_{p,u}^{\ell'} + \sum_v \sigma_{\ell'}^v \Gamma_{p,v}^\ell\right) - \frac{1}{2}\,[\sigma J + J\sigma]_{\ell'}^\ell\,\mathrm{d}t. \tag{10}$$

Eq. (10), together with Eq. (8), gives an SDE with local Lipschitz coefficients for the triple $(\tilde{\xi}, e, \sigma)$; by uniqueness of the solution, as $\sigma(0) = 0$, we deduce $\sigma(t) = 0$ for all $t \geqslant 0$. □

In terms of the new $\mathbb{R}^d$-valued Brownian motion $\widetilde{W}$ defined by $\mathrm{d}\widetilde{W}_k(t) := \sum_\ell [e_W(t)]_\ell^k\,\mathrm{d}W_\ell$, we have

$$\mathrm{d}\tilde{\xi}_W = \sum_k A_k\big(\tilde{\xi}_W(t)\big)\,\mathrm{d}\widetilde{W}_k(t) + A_0\big(\tilde{\xi}_W(t)\big)\,\mathrm{d}t. \tag{11}$$

5. Reconstruction of the scheme $\mathcal{S}$

We want to prove that our scheme $\mathcal{S}$ is *essentially* the projection of the Milstein scheme $(\tilde{\xi}_{W^\varepsilon}, e_{W^\varepsilon})$ for the solution $m_W = (\tilde{\xi}_W, e_W)$ of the SDE (8). In order to write the first component $\tilde{\xi}_{W^\varepsilon}$ we have to compute the horizontal part of $\partial_{\tilde{A}_k}\tilde{A}_j$, which has been done in the proof to Proposition 4.1: we get

$$\begin{aligned}
&\tilde{\xi}_{W^\varepsilon}(t) - \tilde{\xi}_{W^\varepsilon}(t_\varepsilon)\\
&\quad= A_0\big(\tilde{\xi}_{W^\varepsilon}(t_\varepsilon)\big)(t - t_\varepsilon) + \sum_{k,\ell}\big[e_{W^\varepsilon}(t_\varepsilon)\big]_k^\ell A_\ell\big(\tilde{\xi}_{W^\varepsilon}(t_\varepsilon)\big)\Delta(W_k)\\
&\qquad+ \frac{1}{2}\sum_{k,j}\left\{\sum_{\ell,\ell'}\big[e_{W^\varepsilon}(t_\varepsilon)\big]_k^\ell\big[e_{W^\varepsilon}(t_\varepsilon)\big]_j^{\ell'}\left(\partial_{A_\ell}A_{\ell'} - \sum_i \Gamma_{\ell,\ell'}^i A_i\right)\big(\tilde{\xi}_{W^\varepsilon}(t_\varepsilon)\big)\right\}\big(\Delta(W_k)\Delta(W_j) - \varepsilon\eta_k^j\big),
\end{aligned}$$

where $\Delta(W_k) = W_k(t) - W_k(t_\varepsilon)$. By (5)

$$\mathbb{E}\Big[\sup_{t\in[0,1]}\big\|e_W(t) - e_{W^\varepsilon}(t)\big\|^2\Big] \leqslant c\varepsilon^2, \qquad \mathbb{E}\Big[\sup_{t\in[0,1]}\big\|\tilde{\xi}_W(t) - \tilde{\xi}_{W^\varepsilon}(t)\big\|^2\Big] \leqslant c\,\varepsilon^2. \tag{12}$$

Consider the new process $\xi_W^\sharp$ defined by

$$\begin{aligned}
\xi_W^\sharp(t) - \xi_W^\sharp(t_\varepsilon) &= A_0\big(\xi_W^\sharp(t_\varepsilon)\big)(t - t_\varepsilon) + \sum_{k,\ell}\big[e_W(t_\varepsilon)\big]_k^\ell A_\ell\big(\xi_W^\sharp(t_\varepsilon)\big)\Delta(W_k)\\
&\quad+ \frac{1}{2}\sum_{k,j}\left\{\sum_{\ell,\ell'}\big[e_W(t_\varepsilon)\big]_k^\ell\big[e_W(t_\varepsilon)\big]_j^{\ell'}\left(\partial_{A_\ell}A_{\ell'} - \sum_i \Gamma_{\ell,\ell'}^i A_i\right)\big(\xi_{W^\varepsilon}^\sharp(t_\varepsilon)\big)\right\}\big(\Delta(W_k)\Delta(W_j) - \varepsilon\eta_k^j\big).
\end{aligned}$$

Lemma 5.1. *The process $\xi_W^\sharp$ has the same law as the process Z_{W^ε} defined in* (3).

Proof. By Proposition 4.2, $\widehat{W}_\ell(t) - \widehat{W}_\ell(t_\varepsilon) := \sum_k [e_W(t_\varepsilon)]_k^\ell(W_k(t) - W_k(t_\varepsilon))$ are the increments of an $\mathbb{R}^d$-valued Brownian motion $\widehat{W}$; we get

$$\begin{aligned}
\xi_W^\sharp(t) - \xi_W^\sharp(t_\varepsilon) &= A_0\big(\xi_W^\sharp(t_\varepsilon)\big)(t - t_\varepsilon) + \sum_k A_k\big(\xi_W^\sharp(t_\varepsilon)\big)\big(\widehat{W}_k(t) - \widehat{W}_k(t_\varepsilon)\big)\\
&\quad+ \frac{1}{2}\sum_{k,s}\left(\partial_{A_k}A_s - \sum_i \Gamma_{k,s}^i A_i\right)\big(\xi_W^\sharp(t_\varepsilon)\big)\Big(\big(\widehat{W}_k(t) - \widehat{W}_k(t_\varepsilon)\big)\big(\widehat{W}_s(t) - \widehat{W}_s(t_\varepsilon)\big) - \varepsilon\eta_k^s\Big).
\end{aligned}$$

By Eq. (6), we have $2\sum_i \Gamma^i_{k,s} A_i = [A_k, A_s] + \sum_i (\langle [A_i, A_s], A_k\rangle A_i + \langle [A_i, A_k], A_s\rangle A_i)$, where the first term is antisymmetric in k, s and does not contribute; the remaining sum is symmetric in k, s. Thus we get $-\sum_{i,k,s} \Gamma^i_{k,s} A_i \Delta(\widehat{W}_k)\Delta(\widehat{W}_s) = \sum_{i,k,s} A_i \langle [A_s, A_i], A_k\rangle \Delta(\widehat{W}_k)\Delta(\widehat{W}_s)$ which proves Lemma 5.1. □

Lemma 5.2. *We have* $\mathbb{E}[\sup_{t\in[0,1]} \|\xi^\sharp_W(t) - \tilde{\xi}_{W^\varepsilon}(t)\|^2_{\mathbb{R}^d}] \leqslant c\varepsilon^2$.

Proof. The following method of introducing a parameter λ and differentiating with respect to λ, is continuously used in [4]. For $\lambda \in [0,1]$, let $e^\lambda := \lambda e_W + (1-\lambda) e_{W^\varepsilon}$ and define the process ξ^λ_W by

$$\xi^\lambda_W(t) - \xi^\lambda_W(t_\varepsilon) = A_0\big(\xi^\lambda_W(t_\varepsilon)\big)(t - t_\varepsilon) + \sum_{k,\ell}\big[e^\lambda_W(t_\varepsilon)\big]^\ell_k A_\ell\big(\xi^\lambda_W(t_\varepsilon)\big)\Delta(W_k)$$
$$+ \frac{1}{2}\sum_{k,j}\left\{\sum_{\ell,\ell'}\big[e^\lambda_W(t_\varepsilon)\big]^\ell_k\big[e^\lambda_W(t_\varepsilon)\big]^{\ell'}_j\left(\partial_{A_\ell}A_{\ell'} - \sum_i \Gamma^i_{\ell,\ell'} A_i\right)\big(\xi^\lambda_{W^\varepsilon}(t_\varepsilon)\big)\right\}\big(\Delta(W_k)\Delta(W_j) - (t - t_\varepsilon)\eta^j_k\big).$$

Let $u^\lambda_W := \partial\xi/\partial\lambda$; then $\xi^\sharp_W(t) - \tilde{\xi}_{W^\varepsilon}(t) = \int_0^1 u^\lambda_W(t)\,\mathrm{d}\lambda$. Denote by A'_k, $(\partial_{A_\ell}A_{\ell'})'$, $(\Gamma^i_{\ell,\ell'}A_i)'$ the matrices obtained by differentiating A_k, $\partial_{A_\ell}A_{\ell'}$, $\Gamma^i_{\ell,\ell'}A_i$ with respect to ξ, and consider the delayed matrix SDE

$$d J_{t\leftarrow t_0} = \Bigg[A'_0\big(\xi^\lambda_W(t_\varepsilon)\big)\,\mathrm{d}t + \sum_{k,\ell}\big[e^\lambda_W(t_\varepsilon)\big]^\ell_k A'_\ell\big(\xi^\lambda_W(t_\varepsilon)\big)\,\mathrm{d}W_k + \frac{1}{2}\sum_{k,j}\Bigg\{\sum_{\ell,\ell'}\big[e^\lambda_W(t_\varepsilon)\big]^\ell_k\big[e^\lambda_W(t_\varepsilon)\big]^{\ell'}_j$$
$$\times\left((\partial_{A_\ell}A_{\ell'})' - \sum_i\big(\Gamma^i_{\ell,\ell'}A_i\big)'\right)\big(\xi^\lambda_{W^\varepsilon}(t_\varepsilon)\big)\Bigg\}\big(\Delta(W_k)\,\mathrm{d}W_j(t) + \Delta(W_j)\,\mathrm{d}W_k(t) - \eta^j_k\,\mathrm{d}t\big)\Bigg]J_{t\leftarrow t_0}$$

with initial condition $J_{t_0\leftarrow t_0} = \mathrm{Id}$. Then

$$u^\lambda_W(T) = J_{T\leftarrow t_0}\int_0^T J^{-1}_{t\leftarrow t_0}\Bigg[\sum_{k,\ell}\big[e_W(t_\varepsilon) - e_{W^\varepsilon}(t_\varepsilon)\big]^\ell_k A_\ell\big(\xi^\lambda_W(t_\varepsilon)\big)\,\mathrm{d}W_k(t) + \sum_{k,j}\Bigg\{\sum_{\ell,\ell'}\big[e_W(t_\varepsilon) - e_{W^\varepsilon}(t_\varepsilon)\big]^\ell_k$$
$$\times\big[e^\lambda_W(t_\varepsilon)\big]^{\ell'}_j\left(\partial_{A_\ell}A_{\ell'} - \sum_i \Gamma^i_{\ell,\ell'}A_i\right)\big(\xi^\lambda_{W^\varepsilon}(t_\varepsilon)\big)\Bigg\}\big(\Delta(W_k)\,\mathrm{d}W_j(t) + \Delta(W_j)\,\mathrm{d}W_k(t) - \eta^j_k\,\mathrm{d}t\big)\Bigg]$$

which along with (12) proves Lemma 5.2. □

References

[1] A.-B. Cruzeiro, P. Malliavin, Renormalized differential geometry on path space: structural equation, curvature, J. Funct. Anal. 139 (1996) 119–181.

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

[3] P. Malliavin, Paramétrix trajectorielle pour un opérateur hypoelliptique et repère mobile stochastique, C. R. Acad. Sci. Paris, Sér. A-B 281 (1975) A241–A244.

[4] P. Malliavin, A. Thalmaier, Numerical error for SDE: asymptotic expansion and hyperdistributions, C.R. Math. Acad. Sci. Paris, Sér. I 336 (2003) 851–856.

[5] F. Malrieu, Convergence to equilibrium for granular media equations and their Euler schemes, Ann. Appl. Probab. 13 (2003) 540–560.

[6] G.N. Milstein, Numerical Integration of Stochastic Differential Equations, in: Math. Appl., vol. 313, Kluwer Academic, Dordrecht, 1995. Translated and revised from the 1988 Russian original.

[7] D.W. Stroock, S.R.S. Varadhan, Multidimensional Diffusion Processes, in: Grundlehren Math. Wiss., vol. 233, Springer-Verlag, Berlin, 1979.

[8] C. Villani, Topics in Optimal Transportation, in: Grad. Stud. Math., vol. 58, Americal Mathematical Society, Providence, RI, 2003.