Cubature Methods and Applications

Dan Crisan

Imperial College London

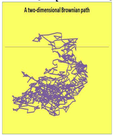
CEMRACS 2017
Numerical methods for stochastic models:
control, uncertainty quantification, mean-field
July 17 - August 25
CIRM, Marseille

- Feynman-Kac representations - a common platform
  - Theoretical Analysis of solution of PDEs
  - Approximation of Feynman-Kac representations
    - Wiener measure approximation
    - Computational effort
- Example 1: Semilinear PDEs (joint work with K. Manolarakis)
  - The Feynman-Kac representation
  - Functional discretization
  - Theoretical results
  - Numerical Implementation
- Example 2: linear parabolic SPDEs (joint work with S. Ortiz-Latorre)
  - The Feynman-Kac representation
  - Functional discretization
  - Theoretical results
  - Numerical Implementation
- Example 3: McKean-Vlasov PDEs (joint work with E. McMurray)
  - The Feynman-Kac representation
  - Functional discretization
  - Theoretical results
  - Numerical Implementation
- Final remarks

Common feature of many PDEs: their solutions can be represented as integrals of certain nonlinear functionals with respect to the Wiener measure.

Feynman-Kac formula

$$u(t,x) = E[\Lambda_{t,x}(W)] = \int_{\omega \in C([0,\infty),\mathbb{R}^d)} \Lambda_{t,x}(\omega) dP_W(\omega)$$

Microscopic level	Macroscopic level		Timeline
Brownian motion $W = \{W_t, t \geq 0\}$ 	Heat equation	$\begin{cases} \partial_t u_t = \frac{1}{2}\Delta u_t \\ u_0 = \Phi \end{cases}$	Feynman 1948 Kac 1949
	Zakai equation	$du_t = Lu_t + hu_t dY_t$	Duncan, Mortensen, Zakai 1970
	McKean-Vlasov PDEs	$\partial_t u_t = \sum_{i,j=1}^d a_{ij}(u_t)\partial_i\partial_j u_t + \sum_{i=1}^d b_i(u_t)\partial_i u_t + c(u_t)u_t$	Gärtner 1988
	Semilinear PDEs	$\begin{cases} \partial_t u_t = Lu_t + f(t,x,u_t,\nabla u_t) \\ u_0 = \Phi \end{cases}$	Pardoux & Peng 1990, 1992
	Fully Nonlinear PDEs	$F(t,x,u_t,\nabla u_t,\Delta u_t) = 0$	Soner, Touzi & Victoir 2007
	3 − d incompressible Navier − Stokes equation	$\begin{cases} \partial_t u_t + (u_t \cdot \nabla)u_t - \nu\Delta u_t + \nabla p = 0 \\ \nabla \cdot u_t = 0 \end{cases}$	Constantin & Iyer 2008
	K-S equation	$du_t = Lu_t + u_t(\bar{h})(dY_t - u_t(h)dt)$	Crisan & Xiong 2009
	viscous Burgers equation	$\partial_t u_t + u_t\partial_x u_t - \nu\partial_x^2 u_t = 0$	Novikov & Iyer 2010

Smoothness of u_t $\equiv$ Smoothness of $\Lambda_{t,x}$ in Malliavin sense

Ex: Heat equation

$$\begin{cases} \partial_t u_t & = \frac{1}{2}\Delta u_t \\ u_0 & = \Phi \end{cases} \qquad u_t(x) = E[\Phi(x + W_t)] = \int \Phi(y) \frac{1}{\sqrt{2\pi t}} e^{\frac{-(y-x)^2}{2t}} dy$$

Via an integration by parts formula one can prove that

$$u_t(x)' = \frac{1}{t} E[\phi(x + W_t)W_t] \quad \Rightarrow |u_t(x)'| \leq \frac{E[|W_t|]}{t} \|\phi\|_\infty = \sqrt{\frac{2}{\pi}} \frac{\|\phi\|_\infty}{\sqrt{t}}.$$

Remarks:

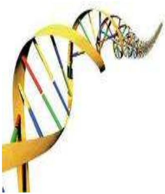
- Fundamental progress (though not complete) for F-K formulae for linear PDEs - notably through the Kusuoka-Stroock programme : Kusuoka & Stroock [1985,1987, 2003], further developed in DC & Ghazali [2007], DC, Manolarakis, Nee [2013], DC & Ottobre [2016].
- Some progress for F-K formulae for non-linear PDEs
  - Semilinear equations : DC-Delarue [2012]
  - McKean-Vlasov equations: DC & McMurray [2017]

$$u(t,x) = E[\Lambda_{t,x}(W)] = \int_{\omega \in C([0,\infty),\mathbb{R}^d)} \Lambda_{t,x}(\omega) dP_W(\omega)$$

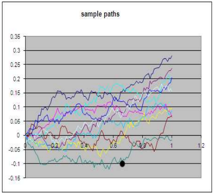
A three-step scheme:

- replace P_W with $P_{\tilde{W}} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\omega_i}$ - $\tilde{W}$ approximates the signature of W
- approximate $\Lambda_{t,x}$ with an explicit/simple version $\tilde{\Lambda}_{t,x}$
- control the computational effort (use the TBBA)

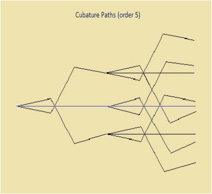
$$u(t,x) \simeq \frac{1}{n}\sum_{i=1}^{n} \tilde{\Lambda}_{t,x}(\omega_i)$$



Full DNA



Typical Paths



Representative Paths

Truncated DNA

Chen's iterated integrals expansion - signature of a path

Let $T\left(\mathbb{R}^d\right) = \bigoplus_{i=0}^{\infty}\left(\mathbb{R}^d\right)^{\otimes i}$ and $T^{(m)}\left(\mathbb{R}^d\right) = \bigoplus_{i=0}^{m}(\mathbb{R}^d)^{\otimes i}$ be the tensor algebra of all non-commutative polynomials over $\mathbb{R}^d$ and, respectively the tensor algebra of all non-commutative polynomials of degree less than $m+1$. For a path $\omega \in C_{bv}([0,\infty),\mathbb{R}^d)$ define its signature $S_t(\omega)$ and, respectively, its truncated signature $S_t^m(\omega)$ to be the corresponding Chen's iterated integrals expansion:

$$S : C_{bv}([0,\infty),\mathbb{R}^d) \to T\left(\mathbb{R}^d\right) \qquad S_t(\omega) = \sum_{k=0}^{\infty}\int_{0<t_1\ldots t_k<t} d\omega_{t_1}\otimes\ldots\otimes d\omega_{t_k}$$

$$S^m : C_{bv}([0,\infty),\mathbb{R}^d) \to T^{(m)}\left(\mathbb{R}^d\right) \qquad S_t^m(\omega) = \sum_{k=0}^{m}\int_{0<t_1\ldots t_k<t} d\omega_{t_1}\otimes\ldots\otimes d\omega_{t_k}.$$

The (random) signature and, respectively, the truncated signature of the Brownian motion are

$$S_t(W) = \sum_{k=0}^{\infty}\int_{0<t_1\ldots t_k<t} dW_{t_1}\otimes\ldots\otimes dW_{t_k}, \quad S_t^m(W) = \sum_{k=0}^{m}\int_{0<t_1\ldots t_k<t} dW_{t_1}\otimes\ldots\otimes dW_{t_k}.$$

- $E\left[S_t(W)\right]$ uniquely identifies the Wiener measure P_W.
- If $\tilde{W}$ is another process such that $E\left[S_t^m(W)\right] = E\left[S_t^m(\tilde{W})\right]$, then

$$E[\Lambda_{t,x}(W)] \simeq E[\Lambda_{t,x}(\tilde{W})] \quad \text{high order approximation of} \quad u(t,x).$$

  See DC and Ghazali [2007] for conditions.
- Several choices for $\tilde{W}$: Kusuoka [2001,2004], Kusuoka and Ninomiya [2004], Lyons and Victoir. [2004], Ninomiya and Victoir [2004], etc.

Theorem (Lyons & Victoir (2004))

For any $t > 0$, there exists paths $\omega_1, \ldots, \omega_N \in C^0_{0,bv}([0,t];\mathbb{R}^d)$ and $\lambda_1, \lambda_2, \ldots, \lambda_N$ ($\sum_{i=1}^N \lambda_i = 1$), such that if $P(\tilde{W} = \omega_i) = \lambda_i$ then

$$E\left[S_t^m(W)\right] = E\left[S_t^m(\tilde{W})\right].$$

If the above is true, we call $\mathcal{L}_{\tilde{W}} = \sum_{i=1}^N \lambda_i \delta_{\omega_i}$ the cubature measure and denote it by $\mathbb{Q}_t^m$.

For example, if we want to approximate $\mathbb{E}[\alpha(X_t)]$, where X is the the solution of the following SDE

$$dX_t = V_0(X_t)dt + \sum_{i=1}^{d} V_i(X_t) \circ dW_t^i$$

Then X can be expressed as $X = \Psi_t(W)$ giving a representation of the form $\mathbb{E}[\Lambda_t(W)]$. Choose X^j to be the solution of the following ODE

$$dX_t^j = V_0(X_t^j)dt + \sum_{i=1}^{d} V_i(X_t^j)d\omega_t^{j,i}$$

In this case:

$$\mathbb{E}[\Lambda_t(\tilde{W})] = \mathbb{E}_{\mathbb{Q}^m}[\alpha(X_t)] = \sum_{i=1}^{N} \lambda_i g(X_t^j)$$

and

$$|\mathbb{E}[\Lambda_t(W)] - \mathbb{E}[\Lambda_t(\tilde{W})]| \leq C\delta^{\frac{m-1}{2}}.$$

Cubature of order 3: For d=1, we can use 2 paths with equal weights $\lambda_j = \frac{1}{2}$ defined as

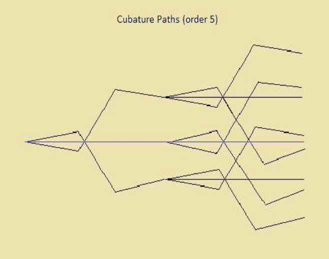
$$\omega_t^j = tz^i,$$

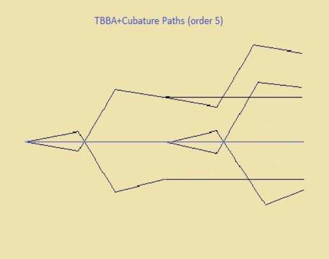
where $z^i \in \{-1, 1\}$. For $d \geq 2$, we can use $\left[\dfrac{d(d+2)}{6}\right] + 1$ linear paths with equal weights λ_j.

Cubature of order 5: For d=1, we can use 3 paths, ω, $-\omega$ and 0 with ω is defined as

$$\omega(t) = \begin{cases} \frac{\sqrt{3}}{2}\left(4-\sqrt{22}\right)t & t \in \left[0, \frac{1}{3}\right] \\ \frac{\sqrt{3}}{6}\left(4-\sqrt{22}\right) + \sqrt{3}\left(-1+\sqrt{22}\right)\left(t-\frac{1}{3}\right) & t \in \left[\frac{1}{3}, \frac{2}{3}\right] \\ \frac{\sqrt{3}}{6}\left(2+\sqrt{22}\right) + \frac{\sqrt{3}}{2}\left(4-\sqrt{22}\right)\left(t-\frac{2}{3}\right) & t \in \left[\frac{2}{3}, 1\right] \end{cases}$$

• An additional procedure is used to control the computational effort.
• The measure $\mathbb{Q}^m$ is replaced by a measure $\tilde{\mathbb{Q}}^{m,N}$ with support of size N by using a *tree based branching algorithm* (TBBA).
• The TBBA was introduced in DC & Lyons (2002) as the optimal stratified sampling procedure in the context of the filtering problem. The method has a wider applicability: it is applicable to any method that uses branching trees.
• By merging the TBBA with the cubature method we keep the number of particles on the support of the intermediate measures *constant*.





• Assume that we constructed $\mathbb{Q}^m = \sum_{j=1}^{M} \lambda_j \delta_{\gamma_j}$ with M paths and we want to reduce the number to at most N paths.
• We replace $\mathbb{Q}^m$ with a random measure $\tilde{\mathbb{Q}}^{m,N}$ such that $\mathrm{supp}(\tilde{\mathbb{Q}}^{m,N}) \subseteq \mathrm{supp}(\mathbb{Q}^m)$ and that the size on its support is at most N. For an arbitrary path $\gamma \in \mathrm{supp}(\mathbb{Q}^m)$, we will have

$$\hat{\mathbb{Q}}_k^m(\gamma) = \begin{cases} \frac{\lfloor N\mathbb{Q}^m(\gamma)\rfloor}{N} & \text{with probability } 1 - \{N\mathbb{Q}^m(\gamma)\} \\ \frac{\lfloor N\mathbb{Q}^m(\gamma)\rfloor + 1}{N} & \text{with probability } \{N\mathbb{Q}^m(\gamma)\} \end{cases} . \tag{1}$$

• In addition, $\tilde{\mathbb{Q}}^{m,N}$ is constructed so that it is a (random) probability measure, i.e.,

$$\sum_{\gamma \in \mathrm{supp}\mathbb{Q}^m} \hat{\mathbb{Q}}_k^m(\gamma) = 1. \tag{2}$$

• The mass allocated to each path $\gamma \in \mathrm{supp}(\mathbb{Q}^m)$ is either 0 or an integer multiple of $1/N \Rightarrow$ the support of any realization of $\tilde{\mathbb{Q}}^{m,N}$ has size at most N.
• The maximum number of paths is achieved when $\tilde{\mathbb{Q}}^{m,N}$ puts mass $1/N$ on N of the M paths. This is the basis of the control of the computational effort.

Let $u \in C^{1,2}([0,T] \times \mathbb{R}^m)$ be the solution of the final value Cauchy problem

$$\begin{cases} (\partial_t + L)u = -f\left(t, x, u, V_1 u, ..., V_d u\right)(x)\right), & t \in [0,T),\ x \in \mathbb{R}^m \\ u(T,x) = \Phi(x), & x \in \mathbb{R}^m \end{cases}, \tag{3}$$

where

- L is the second order differential operator

$$L = V_0 + \frac{1}{2}\sum_{i=1}^{d} V_i^2. \tag{4}$$

- In (4) V_i, $i = 0, 1..., d$ are the first differential operators associated to $V_i = (V_i^j)_{j=1}^d$.

$$V_i = \sum_{j=1}^{m} V_i^j \partial_{x_i}$$

- Particular case

$$\begin{cases} (\partial_t + \Delta)u = -f\left(t, x, u, \nabla u\right)(x)\right), & t \in [0,T),\ x \in \mathbb{R}^m \\ u(T,x) = \Phi(x), & x \in \mathbb{R}^m \end{cases}, \tag{5}$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be complete probability space endowed with a filtration that satisfies the usual conditions $\{\mathcal{F}_t\}_{t\geq 0}$. Let W be an $\{\mathcal{F}_t\}$-adapted Brownian motion and $(X, Y, Z) = \{(X_t, Y_t, Z_t), t \in [0, T]\}$ be the solution of the (decoupled) system:

Forward-Backward SDE

$$X_t = X_0 + \int_0^t V_0(X_s)ds + \sum_{i=1}^d \int_0^t V_i(X_s) \circ dW_s^i, \qquad \text{forward c.} \tag{6}$$

$$Y_t = \Phi(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \sum_{i=1}^d \int_t^T Z_s^i dW_s^i, \qquad \text{backward c.} \tag{7}$$

- X m-dimensional, Y 1-dimensional, Z d-dimensional
- $V_i : \mathbb{R}^m \to \mathbb{R}^m$ smooth vector fields $V_i \in \mathbb{C}_b^\infty(\mathbb{R}^m, \mathbb{R}^m)$, $i = 0, 1, ..., d$
- The stochastic integral in (6) is a *Stratonovitch* type integral
- $\Phi(X_T)$ called *the final condition*
- $f : [0, T] \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ Lipschitz, called "the driver".

Theorem (Pardoux & Peng (1990))

There exists a unique $\mathcal{F}_t$-adapted solution (X, Y, Z) of the system (6)+(7).

Theorem (Peng 1991,1992, Pardoux & Peng 1992)

Under additional smoothness assumptions on the coefficients of the FBDSE, the unique solution of the Cauchy problem (5) *admits the following* **Feynman-Kac representation**

$$u(t,x) = Y_t^{t,x} = \mathbb{E}\left[\Phi(X^{t,x}(T)) + \int_t^T f(s, X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})ds\right], \tag{8}$$

where $(X^{t,x}, Y^{t,x}, Z^{t,x})$ *is the 'stochastic flow' associated FBSDE* (6)+(7)

$$X_s^{t,x} = x + \int_t^s V_0(X_u^{t,x})du + \sum_{i=1}^d \int_t^s V_i(X_u^{t,x}) \circ dW_u^i, \quad s \in [t,T], \tag{9}$$

$$Y_s^{t,x} = \Phi(X_T^{t,x}) + \int_s^T f(u, X_u^{t,x}, Y_u^{t,x}, Z_u^{t,x})du - \sum_{i=1}^d \int_s^T (Z_u^{t,x})^i dW_u^i. \tag{10}$$

Moreover $(Z_s^{t,x})^i = V_i u(s, X_s^{t,x})$ *for* $s \in [t,T]$.

We deduce from the Feynman-Kac representation (8) and the formula for Z that:

$$u(t,x) = \mathbb{E}\left[\Phi(X_T) + \int_t^T f(s, X_s, u(s,X_s), (V_1 u, \ldots, V_d u)(s, X_s))ds\right]$$

Hence there exists $\Lambda_t : C_{\mathbb{R}^d}[t,T] \to \mathbb{R}$ such that

$$u(t,x) = \mathbb{E}\left[\Lambda_t\left(X^{t,x}_{[t,T]}\right)\right].$$

Let $\pi := \{0 = t_0 < t_1 < \ldots < t_n = T\}$ be a partition of $[0,T]$, with $\delta = t_{i+1} - t_i$, $\Delta W_{i+1} = W_{t_{i+1}} - W_{t_i}$.
Define $R_n g = g$ and $S_n g = \nabla g V$ or $S_n g = 0$ if g not Lipschitz.
For $i = 0, \ldots, n-1$, we define the operators R^i, S^i :

$$\begin{aligned} R_i g(x) &= \mathbb{E}\left[R_{i+1} g(X^{t_i,x}_{t_{i+1}})\right] + \delta f(t_i, x, R_i g(x), S_i g(x)) \\ S_i g(x) &= \frac{1}{\delta}\mathbb{E}\left[R_{i+1} g(X^{t_i,x}_{t_{i+1}})\Delta W_{i+1}\right] \end{aligned}$$

Let:

$$u^{\delta}(0,x) = \mathbb{E}\left[\tilde{\Lambda}(X_{t_0}, X_{t_1}, ..., X_{t_n})\right] = R_0\Phi(x).$$

Theorem (Bouchard & Touzi 2004, Gobet & Labart 2007, DC & Manolarakis 2010)

For f Lipschitz

$$\sup_{x\in\mathbb{R}^d} |u^{\delta}(0,x) - u(0,x)| \le C\sqrt{\delta}.$$

For f smooth

$$\sup_{x\in\mathbb{R}^d} |u^{\delta}(0,x) - u(0,x)| \le C\delta.$$

Other discretizations methods are possible: Zhang 2005, DC & Manolarakis 2010, Chassagneux & DC 2013

Next we use the cubature measure $\mathbb{Q}^m$ of degree $m = 3, 5, 7, ...$ coupled with the TBBA procedure to produce a measure $\tilde{\mathbb{Q}}^{m,N}$ with support of size N . For every $i = 0, \ldots, n-1$ define the operators $\hat{R}^i$, $\hat{S}^i$

$$\begin{aligned} \tilde{R}_i g(x) &= \mathbb{E}_{\tilde{\mathbb{Q}}^{m,N}}\left[\tilde{R}_{i+1} g(X_{t_{i+1}}^{t_i,x})\right] + \delta f(t_i, x, \tilde{R}_i g(x), \tilde{S}_i g(x)) \\ \tilde{S}_i g(x) &= \frac{1}{\delta}\mathbb{E}_{\tilde{\mathbb{Q}}^{m,N}}\left[\tilde{R}_{i+1} g(X_{t_{i+1}}^{t_i,x})\Delta W_{i+1}\right] \end{aligned}$$

Let $\tilde{u}^\delta(0, x) = \mathbb{E}_{\tilde{\mathbb{Q}}^{m,N}}\left[\tilde{\Lambda}(X_{t_0}, X_{t_1}, ..., X_{t_n})\right] = \tilde{R}_0\Phi(x).$

Theorem (D.C. & Manolarakis, 2011)

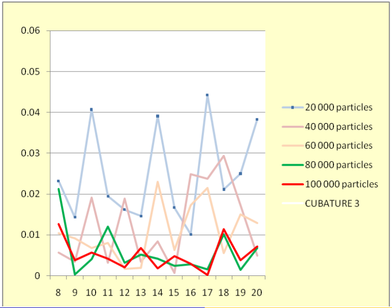
$$\sup_{x\in\mathbb{R}^d} \mathbb{E}[|\tilde{u}^\delta(0, x) - u(0, x)|^2] \leq C\left(\delta + \delta^{\frac{m-1}{2}} + \frac{1}{N}\right).$$

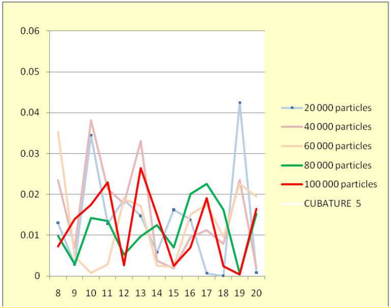
Let $u \in C^{1,2}([0,T] \times \mathbb{R}^4)$ be the solution of the final value Cauchy problem

$$\begin{cases} (\partial_t + L)u = -f\left(t, x, u, \Delta u\right), & t \in [0,T),\ x \in \mathbb{R}^4 \\ u(T,x) = \Phi(x), & x \in \mathbb{R} \end{cases},$$

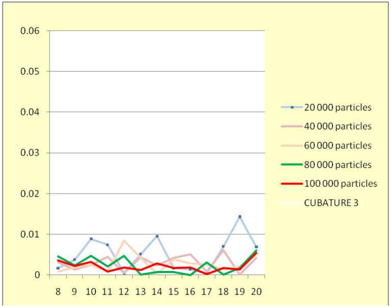
where

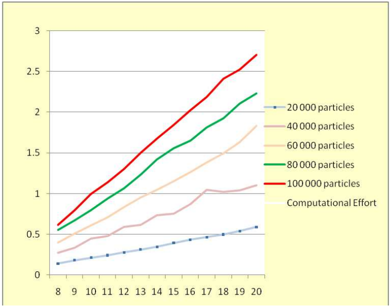
- $L\varphi(x) = \sum_{i=1}^{4} \mu_i x_i \frac{d\varphi}{dx_i} + \frac{\sigma_i^2 x_i^2}{2} \frac{d^2\varphi}{dx_i^2} \quad x \in \mathbb{R}.$
- $f(t,x,y,z) = -ry - \sum_{i=1}^{4}(\mu_i - r)z_i,\ x \in \mathbb{R}.$
- $\Phi(x) = (x_1 x_2 x_3 x_4 - k)_+.$





The randomness can be reduced by averaging over independent copies





$(\Omega, \mathcal{F}, P)$ probability space,
W m-dimensional standard Brownian motion
$\rho = \{\rho_t, t \geq 0\}$ $\mathcal{M}_F\left(\mathbb{R}^d\right)$-valued stochastic process.

$$d\rho_t(\varphi) = \rho_t(A\varphi)dt + \sum_{k=1}^{m} \rho_t(\gamma_k \varphi)dW_t^k. \quad (11)$$

where

$$\begin{aligned} A\varphi(x) &= \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\,\partial_i\partial_j\varphi(x) + \sum_{i=1}^{d} b_i(x)\,\partial_i\varphi(x) \\ a &= \sigma\sigma^\top \end{aligned}$$

Particular case:

$$d\rho_t(z) = \Delta\rho_t(z)dt + \sum_{k=1}^{m} \gamma_k \rho_t(z)dW_t^k.$$

$(\Omega, \mathcal{F}, \tilde{P})$ probability space
$V = (V_t^i)_{i=1}^p, t \geq 0\}$, $U = \{\left(U_t^i\right)_{i=1}^m, t \geq 0\}$ independent Brownian motions

$$\begin{aligned} X_t &= X_0 + \int_0^t b(X_s)\, ds + \int_0^t \sigma(X_s)\, dV_s \\ W_t &= \int_0^t \gamma(X_s)\, ds + U_t, \end{aligned}$$

The filtering problem: Find the conditional law of the *signal* X_t given $\mathcal{W}_t = \sigma(W_s, s \in [0,t])$, i.e.,

$$\vartheta_t(\varphi) = \tilde{\mathbb{E}}[\varphi(X_t)|\mathcal{W}_t], \qquad t \geq 0, \qquad \varphi \in \mathcal{B}(\mathbb{R}^d).$$

Theorem

$$\vartheta_t = \frac{\rho_t}{\rho_t(\mathbf{1})},$$

where $\rho_t(\mathbf{1}) = \int_{\mathbb{R}^d} 1 \rho_t(dx) = \rho_t(\mathbb{R}^d)$.

The process W becomes a Brownian motion via a change of measure (Girsanov's theorem)

$$\left.\frac{dP}{d\tilde{P}}\right|_{\mathcal{F}_t} = Z_t \triangleq \exp\left(-\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right) dU_s^k - \frac{1}{2}\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right)^2 ds\right), t \geq 0.$$

Under P, W is a Brownian motion and X satisfies:

$$X_t = X_0 + \int_0^t b\left(X_s\right) ds + \int_0^t \sigma\left(X_s\right) dV_s.$$

The Feynman-Kac formula

$$\rho_t\left(\varphi\right) = E\left[\varphi(X_t) \exp\left(\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right) dW_s^k - \frac{1}{2}\int_0^t \sum_{k=1}^m \gamma_k\left(X_s\right)^2 ds\right)\middle| \mathcal{W}_t\right] \tag{12}$$

- $\rho_t(\varphi)$ is the expected value of a functional of X which depends W
- to approximate numerically $\rho_t(\varphi)$ we need to
  - approximate the functional with a simpler version
  - approximate the law of the process X
  - control the computational effort

Consider the equidistant partition $\{\frac{it}{n}\}_{i=0}^{n}$ and let g_i be the noise dependent functions

$$g_i = \sum_{j=1}^{m} (\gamma^j (W^j_{\frac{it}{n}} - W^j_{\frac{(i-1)t}{n}}) - \frac{t}{2n} (\gamma^j)^2).$$

Define the operators

$$\begin{aligned} R^n_s \varphi(x) &= \mathbb{E}[\varphi(X_s(x))] \\ R^i_s \varphi(x) &= \mathbb{E}[\varphi(X_s(x)) \exp(g_i(X_s(x))) | \mathcal{W}_t] \end{aligned}$$

for $i = 0, .., n-1$. Let ρ^n_t be the approximate measure

$$\begin{aligned} \rho^n_t(\varphi) &= \mathbb{E}[\varphi(X_t) Z^n_t(X, W) | \mathcal{W}_t] = \mathbb{E}\left[R^0_{\frac{t}{n}} R^1_{\frac{t}{n}} ... R^n_{\frac{t}{n}} \varphi(X_0) \middle| \mathcal{W}_t \right] \\ Z^n_t(X, W) &= \prod_{i=1}^{n} \exp\left(\sum_{k=1}^{m} \left(\gamma_k(X_s)(W_{\frac{it}{n}} - W_{\frac{(i-1)t}{n}}) - \frac{t}{2n} \gamma_k(X_s)^2 \right) \right) \end{aligned}$$

Finally define $\vartheta^n_t(\varphi)$ by the formula $\vartheta^n_t(\varphi) = \dfrac{\rho^n_t(\varphi)}{\rho^n_t(\mathbf{1})}$.

M. All moments of X_0 are finite. The functions b, σ are Lipschitz.
FLp. $\mathbb{E}\left[Z_t(X,W)^p\right] < \infty$ and $\sup_n \mathbb{E}\left[Z_t^n(X,W)^p\right] < \infty$ for some $p > 2$.

Condition **FLp** holds true if γ is bounded. If γ is unbounded, but it has linear growth, then the condition is satisfied if X has exponential moments uniformly bounded on $[0, t]$.

Theorem

*Assume that conditions **M** and **FLp** hold true and $\gamma_k,\ k = 1, ..., m$ are Lipschitz. Then, if φ has polynomial growth, there exists a constant $c = c\,(\varphi, t)$ independent of n such that*

$$\mathbb{E}\left[\left|\rho_t^n\varphi - \rho_t\varphi\right|^2\right] \leq \frac{c}{n}.$$

Moreover, if $\sup_n \mathbb{E}[(\vartheta_t^n(\varphi))^2] < \infty$, then

$$\mathbb{E}\left[\left|\vartheta_t^n\varphi - \vartheta_t\varphi\right|\right] \leq \frac{c}{\sqrt{n}},$$

where, again, $c = c\,(\varphi, t)$ is a constant independent of n.

AP. Let $(P_s)_{s\geq 0}$ be the semigroup associated to the Markov process X. We will assume that, for any Lipschitz continuous function $\psi:\mathbb{R}^d\to\mathbb{R}$, $P_s\psi$ is twice differentiable for any $s\in[0,t]$. Moreover, if

$$P_{a,b}\psi \triangleq P_a\psi - P_b\psi, a, b \in [0,t]\,,$$

we will assume that there exists a constant $c_7 = c_7\,(t)$ independent of a and b such that

$$\sup_{x\in\mathbb{R}^d} |P_{a,b}\psi\,(x)| \quad \leq \quad ck_\psi\left(\sqrt{a}-\sqrt{b}\right) \tag{13}$$

$$\sup_{x\in\mathbb{R}^d} |\partial_i P_{a,b}\psi| \quad \leq \quad \frac{c}{b}k_\psi\,(a-b)\,, \quad i=1,...,d, \tag{14}$$

where k_ψ is the Lipschitz constant of ψ.

Inequalities (13) and (14) are satisfied if, for example, $f,\sigma=\left(\sigma^i\right)_{i=1}^d\in C_b^\infty(\mathbb{R}^d)$ and the vector fields $\left(\sigma^i\right)_{i=1}^d$ satisfy the Hörmander condition.

Theorem

*Assume that conditions **M, FLp** and **AP** hold true. Assume also that the functions φ and γ_i $i = 1, .., m$ are Lipschitz. Then there exists $N > 0$ such that for all $n > N$ and $\varepsilon \in (0, 1)$, there exists a constant $c = c(\varphi, t, N, \varepsilon)$ independent of the partition such that*

$$\mathbb{E}\left[\left|\rho_t^n\varphi - \rho_t\varphi\right|^2\right] \leq \frac{c}{n^{2-\varepsilon}}.$$

Moreover, if $\sup_n \mathbb{E}[(\vartheta_t^n(\varphi))^2] < \infty$, then for all $n > N$ and $\varepsilon \in (0, 1)$, there exists a constant $c = c(\varphi, t, N, \varepsilon)$ independent of the partition such that

$$\mathbb{E}\left[\left|\vartheta_t^n\varphi - \vartheta_t\varphi\right|\right] \leq \frac{c}{n^{1-\varepsilon}}.$$

Theorem

*Assume that conditions **M, FLp** are satisfied and that the functions $\gamma_i \in C_b^2(\mathbb{R}^d)$ for $i = 1, .., m$. Then, if φ has polynomial growth, there exists a constant $c = c(\varphi, t)$ independent of n such that*

$$\mathbb{E}\left[|\rho_t^n \varphi - \rho_t \varphi|^2\right] \leq \frac{c}{n^2}. \tag{15}$$

Moreover, if $\sup_n \mathbb{E}[(\vartheta_t^n(\varphi))^2] < \infty$,

$$\mathbb{E}\left[|\vartheta_t^n \varphi - \vartheta_t \varphi|\right] \leq \frac{c}{n},$$

where, again, $c = c(\varphi, t)$ is a constant independent of n.

Introduce the operators

$$\begin{aligned}\bar{R}_s^n\varphi(x) &= \mathbb{E}_{\mathbb{Q}^m}[\varphi(X_s^x)]\\ \bar{R}_s^i\varphi(x) &= \mathbb{E}_{\mathbb{Q}^m}[\varphi(X_s^x)\exp(g_i(X_s^x))]\end{aligned}$$

Recall that $\rho_t^n(\varphi) = \mathbb{E}\left[R^0_{\frac{t}{n}} R^1_{\frac{t}{n}} \ldots R^n_{\frac{t}{n}}\varphi(X_0)\middle|\mathcal{W}_t\right]$ with corresponding approximation

$$\bar{\rho}_t^n(\varphi) = \mathbb{E}_{\mathbb{Q}^m}\left[\bar{R}^0_{\frac{t}{n}} \bar{R}^1_{\frac{t}{n}} \ldots \bar{R}^n_{\frac{t}{n}}\varphi(X_0)\middle|\mathcal{W}_t\right].$$

Theorem

For all $\varphi \in C_b^{m+2}(\mathbb{R}^d)$ and $p \geq 1$,

$$||\rho_t^n(\varphi) - \bar{\rho}_t^n(\varphi)||_p \leq \frac{c}{n^{(m-1)/2}}\sum_{i=1}^{m+2}||\nabla^i\varphi||_\infty.$$

where $c = c(t, m, p)$ is independent of n.

A third procedure is used to control the computational effort. The measure $\mathbb{Q}^m$ is replaced by a measure $\tilde{\mathbb{Q}}^{m,N}$ with support of size N by using a tree based branching algorithm (TBBA). Recall that

$$\begin{aligned} \tilde{R}^n_s\varphi(x) &= \mathbb{E}_{\mathbb{Q}^{m,N}}[\varphi(X^x_s)], \qquad \tilde{R}^i_s\varphi(x) = \mathbb{E}_{\mathbb{Q}^{m,N}}[\varphi(X^x_s)\exp(g_i(X^x_s))], \\ \tilde{\rho}^n_t(\varphi) &= \mathbb{E}_{\mathbb{Q}^{m,N}}\left[\tilde{R}^0_{\frac{t}{n}}\tilde{R}^1_{\frac{t}{n}}\ldots\tilde{R}^n_{\frac{t}{n}}\varphi(X_0)\middle|\mathcal{W}_t\right] \end{aligned}$$

Theorem

$$\mathbb{E}[|\tilde{\rho}^n_t(\varphi) - \bar{\rho}^n_t(\varphi)|^2] \leq \frac{C}{N}.$$

Corollary

$$\mathbb{E}[|\tilde{\rho}^n_t(\varphi) - \rho_t(\varphi)|^2] \leq C\left(\frac{1}{n^\alpha} + \frac{1}{n^{\frac{m-1}{2}}} + \frac{1}{N}\right).$$

Consider the 1-dimensional model:

$$d\rho_t(\varphi) = \rho_t(A\varphi)dt + \rho_t(\gamma\varphi)dW_t.$$

where

$$\begin{aligned} A\varphi(x) &= \frac{\sigma^2}{2}\frac{d^2\varphi}{dx^2}(x) + \mu\sigma \tanh\left(\frac{\mu x}{\sigma}\right)\frac{d\varphi}{dx}(x) \\ \gamma(x) &= h_1 x + h_2 \end{aligned}$$

Then

$$\rho_t \simeq w^+\mathcal{N}(A_t^+/(2B_t), 1/(2B_t)) + w^-\mathcal{N}(A_t^-/(2B_t), 1/(2B_t)),$$

where

$$\begin{aligned} w_t^\pm &\triangleq \exp\left((A_t^\pm)^2/(4B_t)\right)/\left(\exp\left((A_t^+)^2/(4B_t)\right) + \exp\left((A_t^-)^2/(4B_t)\right)\right) \\ A_t^\pm &\triangleq \pm\frac{\mu}{\sigma} + h_1\Psi_t + \frac{h_2 + h_1 x_0}{\sigma \sinh(h_1\sigma t)} - \frac{h_2}{\sigma}\coth(h_1\sigma t), \\ B_t &\triangleq \frac{h_1}{2\sigma}\coth(h_1\sigma t), \\ \Psi_t &\triangleq \int_0^t \frac{\sinh(h_1\sigma s)}{\sinh(h_1\sigma t)} dW_s, \end{aligned}$$

Consider

$$\partial_t u_t(x) = \sum_{i=1}^{d} V^i(x, u_t(\varphi_i)))^2 u_t(x) + V^0(x, u_t(\varphi_0)))u_t(x), \quad u_t = f$$

Then

$$u(0,x) = \mathbb{E}\left[f(X_T^x)\right],$$

where X is a solution of a McKean-Vlasov SDE with smooth scalar interaction

$$X_t^x = x + \int_0^t V_0(X_s^x, \mathbb{E}[\varphi_0(X_s^x)])\, ds + \sum_{i=1}^{d} \int_0^t V_i(X_s^x, \mathbb{E}[\varphi_i(X_s^x)]) \circ dB_s^i, \qquad (16)$$

- scalar interaction := the dependence on the solution through integrals against scalar functions,

$$\mathbb{E}[\varphi_0(X_s^x)], \quad \mathbb{E}[\varphi_i(X_s^x)], \quad i = 1, ..., d.$$

- $\varphi_i \in \mathcal{C}_b^\infty(\mathbb{R}^N; \mathbb{R})$, $V_i \in \mathcal{C}_b^\infty(\mathbb{R}^{N+1}; \mathbb{R}^N)$.
- $B = \left(B^1, \ldots, B^d\right)$ d-dim standard Brownian motion.
- The process X gives a probabilistic representation of a nonlinear PDE.

$V_0, \ldots, V_d : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$ are vector fields $V_i(\cdot, x')$ on $\mathbb{R}^N$ parametrised by the second variable, $x' \in \mathbb{R}$, with the Lie Bracket between any two given by

$$[V_i, V_j](x, x') = \partial_x V_j(x, x') V_i(x, x') - \partial_x V_i(x, x') V_j(x, x'),$$

where $\partial V_i(x, x') := (\partial_{x_l} V_i^k(x, x'))_{1 \le k, l \le N}$ is the Jacobian matrix of V_i and similarly for $\partial_x V_j$. Then, for $\alpha \in \bigcup_{k \ge 1} \{1, \ldots, N\}^k$ and $i \in \{1, \ldots, N\}$, we define inductively

$$V_{[i]} := V_i, \qquad V_{[\alpha * i]} := [V_i, V_{[\alpha]}] \quad .$$

(A1): Uniform strong Hörmander condition: There exist $\delta > 0$ and $m \in \mathbb{N}$ such that for all $\xi \in \mathbb{R}^N$,

$$\inf_{(x,x') \in \mathbb{R}^N \times \mathbb{R}} \sum_{\alpha \in \cup_{k=1}^m \{1, \ldots, N\}^k} \langle V_{[\alpha]}(x, x'), \xi \rangle^2 \ge \delta \, |\xi|^2$$

(A2): Smoothness of coefficients:

$$\varphi_i \in \mathcal{C}_b^\infty(\mathbb{R}^N; \mathbb{R}), \quad V_i \in \mathcal{C}_b^\infty(\mathbb{R}^N \times \mathbb{R}; \mathbb{R}^N) \quad i = 0, \ldots, d$$

Two algorithms

Partition $[0, T]$ into $\{0 = t_0 < t_1 < \ldots < t_n = T\}$

$$X_t^x = x + \int_0^t V_0(X_s^x, \mathbb{E}[\varphi_0(X_s^x)])\, ds + \sum_{i=1}^d \int_0^t V_i(X_s^x, \mathbb{E}[\varphi_i(X_s^x)]) \circ dB_s^i,$$

Taylor method TM

• Introduced by Chaudru de Raynal & Garcia-Trillos [2015].
• Over the interval $[t_j, t_{j+1}]$ replace $\mathbb{E}\varphi_i(X_t^x)$ appearing in the coefficients with the cubature approximation of the Taylor expansion of the path $t \mapsto \mathbb{E}\varphi_i(X_t^x)$ around t_j up to some order, q.

Lagrange interpolation method LIM

• Over the interval $[t_j, t_{j+1}]$ replace $\mathbb{E}\varphi_i(X_t^x)$ with a Lagrange polynomial which interpolates the cubature approximation of $\mathbb{E}\varphi_i(X_t^x)$ at the previous r points in the time partition.

$\mathcal{E}(T, x, l, \Pi_n)$ - the global error. Parameters:

- r the maximal no of interpolation points required to approximate $\mathbb{E}\varphi_i(X_t^{0,y})$
- q the truncation of the Taylor expansion of $\mathbb{E}\varphi_i(X_t^{0,y})$
- n no. of partition points.
- l the cubature order.

Theorem (DC, McMurray, 2016)

Let $f \in \mathcal{C}_b^\infty(\mathbb{R}^N; \mathbb{R})$. Then, the error for the Taylor method satisfies

$$\sup_{x \in \mathbb{R}^N} |\mathcal{E}(T, x, l, \Pi_n)| \leq C \sum_{j=0}^{n-1} (t_{j+1} - t_j)^{A(q,l)},$$

where $A(q, l) := (q + 2) \wedge (l + 1)/2$. The error in the Lagrange interpolation method is

$$\sup_{x \in \mathbb{R}^N} |\mathcal{E}(T, x, l, \Pi_n)| \leq C \sum_{j=0}^{n-1} \left\{ \frac{(t_{j+1} - t_j)}{((j + 1) \wedge r)!} \prod_{k=0}^{j \wedge (r-1)} (t_{j+1} - t_{j-k}) + (t_{j+1} - t_j)^{(l+1)/2} \right\}.$$

Parameters:

- γ controls the choice of the Kusuoka partition Π_n^γ
- m level required for the uniform strong Hörmander condition

Theorem (DC, McMurray, 2016)

Suppose f is only Lipschitz continuous. Assuming that we use the Kusuoka partition Π_n^γ with $\gamma > l-1$, then the error in the TM satisfies

$$m = 1: \qquad \sup_{x\in\mathbb{R}^N} |\mathcal{E}(T,x,l,\Pi_n^\gamma)| \leq C\,n^{-B(q,l)-1/2}, \tag{17}$$

$$m \geq 2: \qquad \sup_{x\in\mathbb{R}^N} |\mathcal{E}(T,x,l,\Pi_n^\gamma)| \leq C\,n^{-B(q,l)}, \tag{18}$$

where $B(q,l) = (q+\frac{1}{2}) \wedge \frac{l-2}{2}$. Assuming that we use the modified Kusuoka partition $\Pi_n^{\gamma,r}$ with $\gamma \in (l-1,l)$, then the error in the LIM satisfies

$$m = 1: \qquad \sup_{x\in\mathbb{R}^N} \left|\mathcal{E}(T,x,l,\Pi_n^{\gamma,r})\right| \leq C\,n^{-D(r,l)-1/2}\,(1-r/n)^{-l/2}, \tag{19}$$

$$m \geq 2: \qquad \sup_{x\in\mathbb{R}^N} \left|\mathcal{E}(T,x,l,\Pi_n^{\gamma,r})\right| \leq C\,n^{-D(r,l)}\,(1-r/n)^{-l/2}, \tag{20}$$

where $D(r,l) = (r-\frac{3}{2}) \wedge \frac{l-2}{2}$.

Example 1

$N = d = 1$:

$$X_t^{0,x} = x + \int_0^t \mathbb{E}\left[X_s^{0,x}\right] ds + B_t, \quad X_t^x = xe^t + B_t, \quad \mathbb{E}X_t^{0,x} = xe^t$$

Terminal function $f(x) = x^+ := \max\{x, 0\}$ and, by integrating the Gaussian density, we can compute

$$\mathbb{E}(X_t^{0,x})^+ = \sqrt{t}\phi\left(\frac{xe^t}{\sqrt{t}}\right) + xe^t\left(1 - \Phi\left(-\frac{xe^t}{\sqrt{t}}\right)\right),$$

where ϕ and Φ are the density and cumulative distribution function, respectively, of a standard Gaussian random variable.

- The Taylor approximation of order q is easy to compute: $\mathcal{T}_t^q\left(\mathbb{E}X_t^{0,x}\right) = \sum_{k=0}^q \frac{x}{k!} t^k$.
- Cubature formula of degree 5
- A fourth order adaptive Runge-Kutta scheme to solve the ODEs.
- To achieve order 2 convergence with a cubature formula of degree 5 choose $q \geq 1$ and $\gamma \in (4, 5)$ with Taylor method and $r \geq 3$ in the Lagrange interpolation method.

- Parameters $(x, T, \gamma, q, r) = (0.5, 10, 4.5, 2, 3)$.

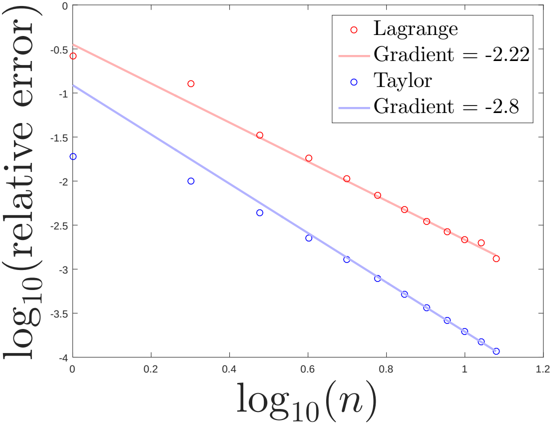


Figure: log-log error plot comparison between the LIM and TM. The gradient of each solid line is given by a linear regression on the last 5 points.

- Both methods achieve the expected quadratic convergence rate.
- TM performs better than LIM.

Example 2

- coefficients are not uniformly elliptic and $N = d = 2$.
- $X_t^{0,x} = (X_t^1, X_t^2)$

$$\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \int_0^t \begin{pmatrix} \left[2 + \sin\left(\mathbb{E}\, X_s^2\right)\right] \\ X_t^1 \end{pmatrix} \circ dB_s^1 + \int_0^t \begin{pmatrix} X_t^2 \\ X_t^1 \end{pmatrix} \circ dB_s^2,$$

where the coefficients are

$$V_0 \equiv 0,\ \ V_1(x_1, x_2, x') = \begin{pmatrix} 2 + \sin(x') \\ x_1 \end{pmatrix},\ \ V_2(x_1, x_2, x') = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix},\ \ \varphi_1(x_1, x_2) = x_2,$$

for all $(x_1, x_2, x') \in \mathbb{R}^3$.

- At $x_1 = 0$ the coefficients degenerate.
- $$V_{[(1,2)]}(x_1, x_2, x') = \begin{pmatrix} x_1 \\ 2 + \sin(x') - x_1 \end{pmatrix}.$$

- Since x_1 and $2 + \sin(x') - x_1$ cannot both be zero at the same time, we see that V_1, V_2 and $V_{[(1,2)]}$ span $\mathbb{R}^2$. The coefficients satisfy the uniform strong Hörmander condition, for $m = 2$.

- For $m = 2$, with a cubature formula of degree 5 ($N_{Cub} = 13$), we expect to achieve a convergence rate of $3/2$. To do so, we have to choose $\gamma \in (4, 5)$ and $r > 7/2$.
- Parameters $(x_1, x_2, T, \gamma, r) = (1, 0.5, 1, 4.5, 4)$ and $f(x) = x^+$.
- No explicit solution, we compare the cubature approximation to a Monte Carlo approximation with Euler-Maruyama discretisation.

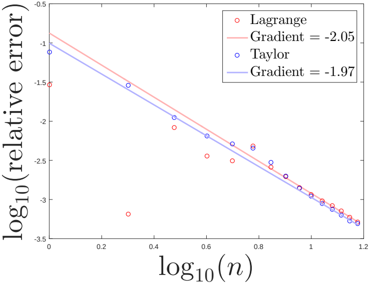


Figure: The gradient of each line is given by a linear regression on the last 5 points.

The performance of the algorithms is similar. Empirically we observe second order convergence, whereas a rate of $3/2$ is predicted.

- The solution of a variety of deterministic and stochastic PDEs can be analyzed and/or approximated through their corresponding Feynman-Kac representations.
- The common methodology contains three steps: functional discretization, Wiener measure approximation (cubature method) and computational effort reduction (TBBA).
- The cubature method is essentially *deterministic*. The diffusion approximation uses a set of ordinary differential equations to approximate the distribution of the solution of the SDE.
- The (exponentially) increase in the computational effort is controlled by the TBBA (a random method).
- A first step towards a unified theory of approximations