

Improved stabilized multilevel Monte Carlo method for stiff stochastic differential equations

Assyr Abdulle and Adrian Blumenthal

ANMC, Section de Mathématiques, École Polytechnique Fédérale de Lausanne, Switzerland, {assyr.abdulle, adrian.blumenthal}@epfl.ch*

Abstract. An improved stabilized multilevel Monte Carlo (MLMC) method is introduced for stiff stochastic differential equations in the mean square sense. Using S-ROCK2 with weak order 2 on the finest time grid and S-ROCK1 (weak order 1) on the other levels reduces the bias while preserving all the stability features of the stabilized MLMC approach. Numerical experiments illustrate the theoretical findings.

1 Introduction

Estimating the expectation of a functional depending on a stochastic process is essential in many applications ranging from biology, chemistry, physics to economics [9,12,8,13]. A popular approach for such problems is the use of classical and improved Monte Carlo (MC) techniques, in particular the multilevel Monte Carlo (MLMC) method using Euler-Maruyama (EM) [10] as numerical integrator [7]. An *explicit* stabilized multilevel Monte Carlo method has been proved to be useful and efficient for stiff problems in a mean square sense and specially attractive for problems of large dimensions (as it avoids solving nonlinear systems typically arising with implicit methods) [2]. Here we present an improved version of the stabilized MLMC method by using a higher weak order scheme on the finest time grid.

We consider the Itô stochastic differential equation (SDE)

$$dX(t) = f(X(t))dt + \sum_{r=1}^m g^r(X(t))dW_r(t), \quad 0 \leq t \leq T, \quad X(0) = X_0, \quad (1)$$

where $X(t) \in \mathbb{R}^d$ is a random variable, $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ the drift function, $g^r : \mathbb{R}^d \rightarrow \mathbb{R}^d$ the diffusion functions and $W_r(t)$ independent one-dimensional Wiener processes (with $r = 1, 2, \dots, m$). Further, we take into account a numerical approximation of the solution of (1) by using a discrete map $X_{n+1} = \Psi(X_n, h, \xi_n)$, where $\Psi(\cdot, h, \xi_n) : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $X_n \in \mathbb{R}^d$ for $n \in \mathbb{N}$, h a time stepsize and ξ_n some random vector. Let us now consider any

* This work is partially supported by the Swiss National Foundation Grant 200021_140692.

$\tau_n = nh \in [0, T]$ for h sufficiently small. The numerical approximation is said to be of strong order of convergence θ_s if $\max_{0 \leq n \leq T/h} \mathbb{E}|X_n - X(\tau_n)| \leq Ch^{\theta_s}$ for a constant C (independent of h). It is said to be of weak order θ_w if for any function $\phi \in C_P^{2(\gamma+1)}(\mathbb{R}^d, \mathbb{R})$ (space of $2(\gamma+1)$ times continuously differentiable functions with all partial derivatives of polynomial growth) it exists a constant C (independent of h) such that $|\mathbb{E}[\phi(X_n)] - \mathbb{E}[\phi(X(\tau_n))]| \leq Ch^{\theta_w}$.

The stability of a numerical method is another important issue for computations. A stochastic process $(X(t))_{t \geq 0}$ is said to be mean square stable if $\mathbb{E}[X(t)^2]$ tends to zero as t goes to infinity. The scalar linear SDE $dX(t) = \lambda X(t)dt + \mu X(t)dW(t)$, $X(0) = 1$, with $\lambda \in \mathbb{C}$ and $\mu \in \mathbb{C}$ is commonly used to figure as test problem [9]. The stability domain of the exact solution is given by $\mathcal{S}_{exact} := \{(\lambda, \mu) \in \mathbb{C}^2 \mid \Re\{\lambda\} + \frac{1}{2}|\mu|^2 < 0\}$. Similarly a numerical method is mean square stable if $\mathbb{E}[X_n^2]$ tends to zero as n goes to infinity. The stability domain of the EM method (that has strong order 1/2 and weak order 1) is specified by $\mathcal{S}_{EM} = \{(p, q) \in \mathbb{C}^2 \mid |1+p|^2 + q^2 < 1\}$ with $(p, q) = (h\lambda, \sqrt{h}|\mu|)$ (see [9]).

Stabilized stochastic methods were introduced in [3,4] with the aim of improving the stability behavior of the EM method, while staying explicit. These methods are s -stage *explicit* methods with fixed order based on

- a deterministic stabilization procedure;
- a finishing stochastic procedure to achieve the desired accuracy.

We will consider the s -stage S-ROCK1 method [4] defined by (for $s \geq 2$)

$$\begin{aligned} K_0 &= X_n, \quad K_1 = X_n + h\frac{\omega_1}{\omega_0}f(K_0), \\ K_i &= 2h\omega_1\frac{T_{i-1}(\omega_0)}{T_i(\omega_0)}f(K_{i-1}) + 2\omega_0\frac{T_{i-1}(\omega_0)}{T_i(\omega_0)}K_{i-1} - \frac{T_{i-2}(\omega_0)}{T_i(\omega_0)}K_{i-2}, \\ K_s &= 2h\omega_1\frac{T_{s-1}(\omega_0)}{T_s(\omega_0)}f(K_{s-1}) + 2\omega_0\frac{T_{s-1}(\omega_0)}{T_s(\omega_0)}K_{s-1} - \frac{T_{s-2}(\omega_0)}{T_s(\omega_0)}K_{s-2} \\ &\quad + \sum_{r=1}^m g^r(K_{s-1})\Delta W_{n+1,r}, \end{aligned} \tag{2}$$

with $i = 2, 3, \dots, s-1$ and where $\omega_0 = 1 + \frac{\eta}{s^2}$ with η a damping parameter, $\omega_1 = \frac{T_s(\omega_0)}{T_s(\omega_0)}$ with $(T_i(x))_{i \geq 0}$ the orthogonal Chebyshev polynomials, $\Delta W_{n+1,r} \sim \mathcal{N}(0, h)$ and $X_{n+1} = K_s$. Here the first $s-1$ represent the stabilization procedure and the last stage a finishing procedure to achieve strong order 1/2 and weak order 1 [3,4]. We will also consider the S-ROCK2 method introduced in [1]. Similar to the S-ROCK1 this scheme uses a stabilization procedure (in this case ROCK2 [5]) on the first $s-2$ stages and then a finishing procedure on the last two stages to obtain a weak order of 2 and a strong order of 1/2. While remaining explicit the two S-ROCK methods have an extended stability domain. In fact, defining $\mathcal{S}_{SDE,a} = \{(p, q) \in [-a, 0] \times \mathbb{R} \mid |q| \leq \sqrt{-2p}\}$ a ‘‘portion’’ of \mathcal{S}_{exact} and $a^* = \sup\{a > 0 \mid \mathcal{S}_{SDE,a} \subset \mathcal{S}_{num}\}$, we get for S-ROCK1 and S-ROCK2

$a^* \approx c_{SR1}s^2$ and $a^* \approx c_{SR2}(s+2)^2$, respectively. The parameters c_{SR1} and c_{SR2} quickly reach a value independent of the stage number that can be estimated numerically as 0.33 (S-ROCK1) and 0.42 (S-ROCK2) [3,4,1].

2 Multilevel Monte Carlo method for stiff SDEs

We are interested in estimating $E := \mathbb{E}[\phi(X(T))]$, the expectation of some Lipschitz continuous functional $\phi: \mathbb{R}^d \rightarrow \mathbb{R}$ depending on the stochastic process $(X(t))_{t \in [0, T]}$ specified through (1). The classical Monte Carlo method uses Euler-Maruyama [10] as numerical integrator with a fixed time stepsize to approximate the stochastic process and sample averages to estimate the expectation. This approach is easy to implement, but computationally expensive. In fact defining the computational cost (or complexity) by the number of function evaluations, one can show that to achieve a mean square accuracy of $\mathcal{O}(\varepsilon^2)$ a computational cost of $\mathcal{O}(\varepsilon^{-3})$ is required (with $\varepsilon > 0$) (see e.g. [8]).

One way to improve the performance of standard Monte Carlo techniques is to use the multilevel Monte Carlo method, which is based on hierarchical sampling. In this approach Monte Carlo is applied to a sequence of nested time stepsizes. Simultaneously the number of samples is balanced according to the stepsize. Combining many samples of computationally cheap approximations (the ones using a large stepsize) with a few samples of computationally expensive approximations (the ones based on a fine time grid) reduces the cost significantly to $\mathcal{O}(\varepsilon^{-2}(\log(\varepsilon))^2)$ while maintaining the same mean square accuracy of $\mathcal{O}(\varepsilon^2)$ [7].

For stiff problems, using the standard MLMC approach with EM, there is some time stepsize restriction due to stability issues. Let $k \geq 2$ be some positive integer indicating the refinement factor. We consider the time stepsizes

$$h_\ell = \frac{T}{M_\ell}, \quad \ell = 0, 1, \dots, L, \quad (3)$$

with L the total number of levels and $M_\ell = k^\ell$ denoting the number of time steps at level ℓ . Suppose there is some stability constraint given by $k^{-\ell_{EM}}\rho \leq 1$ with ℓ_{EM} corresponding to the largest possible stepsize such that the EM scheme is mean square stable and ρ a given stiffness parameter. Furthermore suppose that a root mean square accuracy of $\varepsilon = k^{-L}$ is desired. One has to distinguish between two cases:

1. $\ell_{EM} > L$.
MLMC cannot be applied only classical MC with $h = T/k^{\ell_{EM}}$ which results in a computational cost of $\mathcal{O}(\varepsilon_{MC}^{-3})$ and a precision $\mathcal{O}(\varepsilon_{MC}^2)$, where $\varepsilon_{MC} = k^{-\ell_{EM}}$.
2. $0 < \ell_{EM} \leq L$.
MLMC can only be applied to the levels $\ell_{EM}, \ell_{EM} + 1, \dots, L$. The resulting computational complexity is $\mathcal{O}(\varepsilon^{-2}((\log(\varepsilon))^2 + \varepsilon^{-\ell_{EM}/L}))$.

In [2] a stabilized multilevel Monte Carlo method is introduced which uses as numerical integrator S-ROCK1 (2), an explicit Runge-Kutta method based on orthogonal Chebyshev polynomials. The stability constraint of this scheme is given by $\frac{k^{-\ell} \rho}{c_{SR1} s_\ell^2} \leq 1$, where s_ℓ is the number of stages at level ℓ and c_{SR1} a positive constant. Due to the extended stability domain of S-ROCK1 all levels are accessible in the stabilized MLMC approach. A computational cost of $\mathcal{O}\left(\varepsilon^{-2} (\log(\varepsilon))^2 \left(1 + \frac{\sqrt{p}}{|\log(\varepsilon)|}\right)\right)$ is necessary to attain a mean square precision of $\mathcal{O}(\varepsilon^2)$. For stiff problems as well as for nonstiff problems with no small noise, this means a significant improvement over the standard MLMC approach (which uses EM) as it is shown in [2].

3 Improved stabilized multilevel Monte Carlo method for stiff SDEs

In this section we describe how the stabilized multilevel Monte Carlo method can further be improved. As mentioned in the introduction, the EM method as well as the S-ROCK1 method are both of weak order 1 and strong order 1/2. The idea is to use a numerical integrator of higher weak order for the finest time grid (see [6]), in our case S-ROCK2 [1] with weak order 2, which leads to a reduction of the bias. In fact due to the telescopic sum representation of the multilevel estimator, only the estimator based on the smallest time stepsize (which uses S-ROCK2) appears in the bias. A smaller bias yields a reduction of the total number of levels, and thus a reduced computational cost, without decreasing the accuracy. Note that in the following we focus on problems that are either stiff or nonstiff but with significant noise. Problems with no stability issues can be treated in a similar way.

Recall the sequence of nested stepsizes (3). For $\ell = 0, 1, \dots, L-1$ we denote by ϕ_ℓ the approximation of $\phi(X(T))$ using S-ROCK1 with time stepsize h_ℓ . The approximation of $\phi(X(T))$ using S-ROCK2 on the finest time grid which is based on h_L is indicated by ϕ_L . The improved stabilized multilevel Monte Carlo estimator is defined by

$$\tilde{E} := \sum_{\ell=0}^L \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} \left(\phi_\ell^{(i)} - \phi_{\ell-1}^{(i)} \right) \quad \text{with } \phi_{-1} \equiv 0, \quad (4)$$

a sum of sample averages over N_ℓ independent and identically distributed samples. Note that $\phi_\ell^{(i)}$ and $\phi_{\ell-1}^{(i)}$ are based on the same Wiener path. The accuracy of the estimator \tilde{E} can be measured, e.g., by the mean square error (see e.g. [8]), which can be split into bias and variance as follows:

$$\text{MSE}(\tilde{E}) = \mathbb{E} \left[(\tilde{E} - E)^2 \right] = \text{Var}(\tilde{E}) + \left(\text{bias}(\tilde{E}) \right)^2.$$

Using the properties of the expectation we obtain

$$\mathbb{E} \left[\tilde{E} \right] = \sum_{\ell=0}^L (\mathbb{E} [\phi_\ell] - \mathbb{E} [\phi_{\ell-1}]) = \mathbb{E} \left[\sum_{\ell=0}^L (\phi_\ell - \phi_{\ell-1}) \right] = \mathbb{E} [\phi_L].$$

Hence the bias satisfies

$$\text{bias} \left(\tilde{E} \right) = \mathbb{E} \left[\tilde{E} \right] - E = \mathbb{E} [\phi_L] - E = \mathcal{O} (k^{-2L}) \quad (5)$$

since the S-ROCK2 method, on which ϕ_L is based, is of weak order 2. Furthermore, for the variance we use the Cauchy-Schwarz inequality to obtain

$$\text{Var} (\phi_\ell - \phi_{\ell-1}) \leq \left(\text{Var} (\phi_\ell - E)^{1/2} + \text{Var} (\phi_{\ell-1} - E)^{1/2} \right)^2.$$

Both numerical integrators, S-ROCK1 and S-ROCK2, are of strong order 1/2 and ϕ is Lipschitz continuous by assumption. Thus

$$\text{Var} (\phi_\ell - E) \leq \mathbb{E} \left[(\phi_\ell - E)^2 \right] \leq \mathbb{E} \left[(\phi (X_{M_\ell}) - \phi (X(T)))^2 \right] \leq Ck^{-\ell}$$

and therefore

$$\text{Var} \left(\tilde{E} \right) = \sum_{\ell=0}^L \frac{\text{Var} (\phi_\ell - \phi_{\ell-1})}{N_\ell} \leq C \left(\sum_{\ell=0}^{L-1} \frac{k^{-\ell}}{N_\ell} + \frac{k^{-L}}{N_L} \right), \quad (6)$$

where C is a positive constant.

Assume now a mean square precision of $\text{MSE} \left(\tilde{E} \right) = \mathcal{O} (\varepsilon^2)$ is desired for some $\varepsilon > 0$. Taking into account (5) we obtain a total number of levels $L = -\frac{1}{2} \frac{\log(\varepsilon)}{\log(k)}$ (or equivalently $\varepsilon = k^{-2L}$). Inspired by (6) the number of simulations per level ℓ is set to $N_\ell = k^{-\ell} k^{4L} (L-1)$ for $\ell = 0, 1, \dots, L-1$ and $N_L = k^{-L} k^{4L}$, which yields $\text{Var} \left(\tilde{E} \right) \leq Ck^{-4L} \left(2 + \frac{1}{L-1} \right) = \mathcal{O} (\varepsilon^2)$.

As mentioned above the stability constraint of S-ROCK1 is given by $\frac{k^{-\ell} \rho}{c_{SR1} s_\ell^2} \leq 1$. In a similar way we can define a stability criterion for S-ROCK2 $\frac{k^{-L} \rho}{c_{SR2} (s_L+2)^2} \leq 1$ with $s_L \geq 2$ and with c_{SR2} as defined above.

Complexity. The computational cost of \tilde{E} and \hat{E} , the estimator of the stabilized MLMC of [2], can both be estimated as

$$\text{Cost} \left(\tilde{E} \right) = \text{Cost} \left(\hat{E} \right) = \mathcal{O} \left(\varepsilon^{-2} (\log(\varepsilon))^2 \left(1 + \frac{\sqrt{\rho}}{|\log(\varepsilon)|} \right) \right),$$

however with a smaller constant prefactor for \tilde{E} allowing for a cost reduction.

Indeed, we have $\text{Cost}(\tilde{E})$

$$\begin{aligned}
&= \sum_{\ell=0}^{L-1} N_\ell M_\ell (s_\ell + m) + N_L M_L (s_L + 8 + 2m) \\
&= \sum_{\ell=0}^{L-1} k^{4L} (L-1) \left(\sqrt{\frac{\rho}{c_{SR1}}} k^{-\ell/2} + m \right) + k^{4L} \left(\sqrt{\frac{\rho}{c_{SR2}}} k^{-L/2} + 6 + 2m \right) \\
&= k^{4L} (L-1) \left(\sqrt{\frac{\rho}{c_{SR1}}} \frac{\sqrt{k} - k^{-L/2+1/2}}{\sqrt{k-1}} + mL \right) + k^{4L} \left(\sqrt{\frac{\rho}{c_{SR2}}} k^{-L/2} + 6 + 2m \right) \\
&= \varepsilon^{-2} \left(\frac{1}{2} \frac{\log(\varepsilon)}{\log(k)} \right)^2 (\zeta_1 - \alpha_1) ,
\end{aligned}$$

where $\zeta_1 = m \frac{L-1}{L} + \frac{1}{L} \left(\frac{\sqrt{k}}{\sqrt{k-1}} \right) \sqrt{\frac{\rho}{c_{SR1}}}$ and $\alpha_1 = d_1 \frac{\varepsilon^{1/4} \sqrt{\rho}}{L} + d_2 \frac{(\sqrt{\rho} - d_3)}{L^2}$ with d_1, d_2, d_3 some positive constants. In comparison we have

$$\text{Cost}(\hat{E}) = \varepsilon^{-2} \left(\frac{\log(\varepsilon)}{\log(k)} \right)^2 (\zeta_2 - \alpha_2) ,$$

where $\zeta_2 = m \frac{L+1/2}{L} + \frac{1}{2L} \left(\frac{\sqrt{k}}{\sqrt{k-1}} \right) \sqrt{\frac{\rho}{c_{SR1}}}$ and $\alpha_2 = d_4 \frac{\varepsilon^{1/2} \sqrt{\rho}}{L}$ with d_4 a positive constant. Hence, asymptotically we notice for the improved stabilized MLMC method roughly a reduction of the computational cost by a factor between 0.25 (nonstiff problems but significant noise) and 0.5 (stiff problems).

4 Numerical experiments

In this section we investigate a two-dimensional nonlinear noncommutative SDE inspired by the one-dimensional population dynamics model (see [11])

$$\left\{ \begin{aligned} d \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} &= \begin{pmatrix} \alpha(X_2(t) - 1) - \lambda_1 X_1(t)(1 - X_1(t)) \\ -\lambda_2 X_2(t)(1 - X_2(t)) \end{pmatrix} dt \\ &+ \begin{pmatrix} -\mu_1 X_1(t)(1 - X_1(t)) \\ -\mu_2 X_2(t)(1 - X_2(t)) \end{pmatrix} dW_1(t) \\ &+ \begin{pmatrix} -\mu_2(1 - X_1(t)) \\ 0 \end{pmatrix} dW_2(t), \quad 0 \leq t \leq T, \end{aligned} \right.$$

with initial condition $(X_1(0), X_2(0)) = (0.95, 0.95)$ and where $(W_1(t))_{t \in [0,1]}$ and $(W_2(t))_{t \in [0,1]}$ are two independent Wiener processes. We consider two different scenarios. First a stiff problem with drift term $\lambda_1 \in \{-1, -100, -10000\}$ and noise term $\mu_1 = \sqrt{|\lambda_1|}$. And then a nonstiff problem with no small noise by fixing $\lambda_1 = -1$ and varying $\mu_1 = \sqrt{-2\lambda_1 - \delta}$ with $\delta \in \{10^{-1}, 10^{-2}, 10^{-4}\}$.

In addition we pick $\alpha = 2$, $\lambda_2 = -1$, $\mu_2 = 0.5$, $k = 2$. As root mean square accuracy we choose k^{-2L} with $L \in \{1, 2, \dots, 5\}$. Stability is guaranteed by assessing the second moment at the time end point. In Figure 1 we compare the number of function evaluations (by counting the drift and diffusion evaluations) of the improved stabilized (using S-ROCK1 and S-ROCK2), the stabilized (using S-ROCK1) and the standard (using EM) MLMC method. As expected the improved stabilized approach yields a cost reduction over the other two methods. In Table 1 we present a selection of the results in numbers.

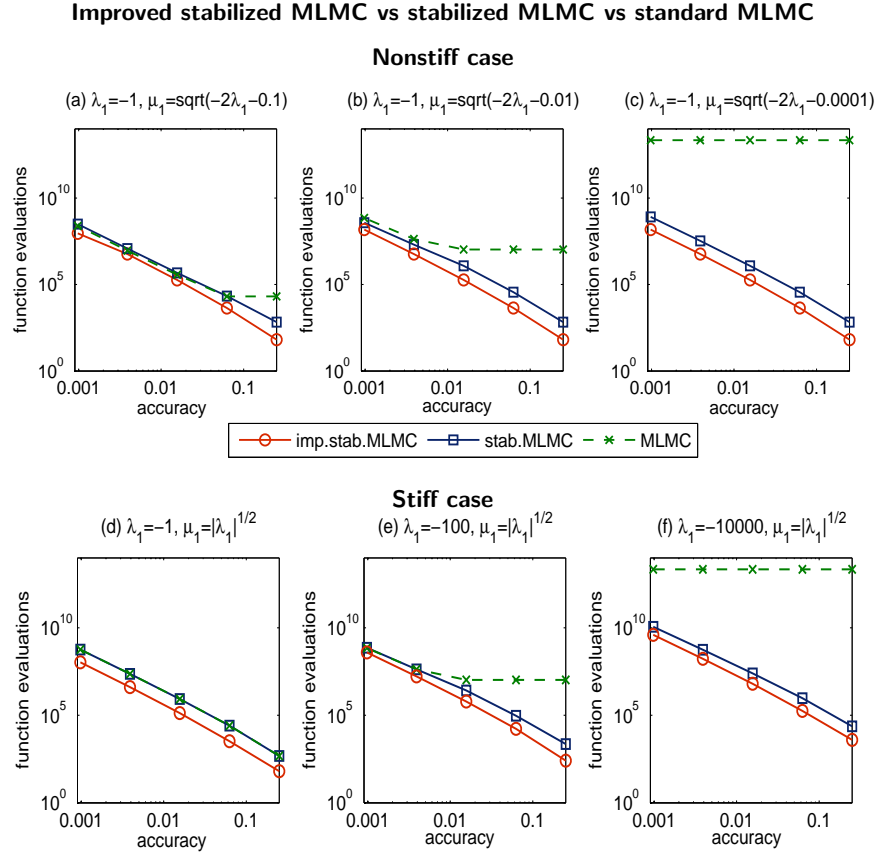


Fig. 1. Function evaluations against root mean square accuracy comparing the improved stabilized MLMC method using S-ROCK1 and S-ROCK2 with the stabilized (S-ROCK1) and the standard (EM) MLMC method.

Table 1. Number of function evaluations of the improved stabilized MLMC (using S-ROCK1 and S-ROCK2), the stabilized MLMC (using S-ROCK1) and standard MLMC (using EM) for different values of the root mean square error. As parameters we take $\lambda_1 = -1, \mu_1 = \sqrt{-2\lambda_1 - 0.01}$ (b) and $\lambda_1 = -100, \mu_1 = \sqrt{|\lambda_1|}$ (e).

precision	2^{-2}	2^{-4}	2^{-6}	2^{-8}	2^{-10}
(b) imp.stab.MLMC	64	4352	184320	5.70×10^6	14.99×10^7
stab.MLMC	672	35840	1204224	19.92×10^6	37.12×10^7
MLMC	10.49×10^6	10.49×10^6	10.49×10^6	42.27×10^6	70.25×10^7
(e) imp.stab.MLMC	256	16896	614400	16.91×10^6	39.53×10^7
stab.MLMC	2272	95232	2629632	42.73×10^6	73.61×10^7
MLMC	10.49×10^6	10.49×10^6	10.49×10^6	42.27×10^6	70.25×10^7

References

1. A. ABDULLE, G. VILMART, AND K. ZYGALAKIS, *Weak second order explicit stabilized methods for stiff stochastic differential equations*, SIAM J. Sci. Comput. **35**:4 (2013), A1792–A1814.
2. A. ABDULLE AND A. BLUMENTHAL, *Stabilized multilevel Monte Carlo method for stiff stochastic differential equations*, J. Comput. Phys. **251** (2013), 445–460.
3. A. ABDULLE AND S. CIRILLI, *Stabilized methods for stiff stochastic systems*, C. R. Math. Acad. Sci. Paris **345**:10 (2007), 593–598.
4. A. ABDULLE AND T. LI, *S-ROCK methods for stiff Itô SDEs*, Commun. Math. Sci. **6**:4 (2008), 845–868.
5. A. ABDULLE AND A. MEDOVIKOV, *Second order Chebyshev methods based on orthogonal polynomials*, Numer. Math. **90**:1 (2001), 1–18.
6. K. DEBRABANT AND A. RÖSSLER, *On the Acceleration of the Multi-Level Monte Carlo Method*, ArXiv preprint **1301.7650** (2013).
7. M. GILES, *Multilevel Monte Carlo path simulation*, Operations Research **56**:3 (2008), 607–617.
8. P. GLASSERMAN, *Monte Carlo Methods in Financial Engineering*, Applications of Mathematics (New York), vol. 53, Springer-Verlag, New York, 2004.
9. P. KLOEDEN AND E. PLATEN, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin and New York, 1992.
10. G. MARUYAMA, *Continuous Markov processes and stochastic equations*, Rend. Circ. Mat. Palermo **4** (1955), 48–90.
11. J. MURRAY, *Mathematical Biology I: An Introduction*, Springer, Seattle, 2002.
12. H. C. ÖTTINGER, *Stochastic processes in polymeric fluids*, Springer-Verlag, Berlin, 1996, Tools and examples for developing simulation algorithms.
13. P. RUÉ, J. VILLÀ-FREIXA, AND K. BURRAGE, *Simulation methods with extended stability for stiff biochemical Kinetics*, BMC Systems Biology **4**:110 (2010), 1–13.