ÖAW Austrian Academy of Sciences

Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

JKU JOHANNES KEPLER UNIVERSITY LINZ

A Fully Parallelizable Space-Time Multilevel Monte Carlo Method for Stochastic Differential Equations with Additive Noise

Martin Neumüller Andreas Thalhammer

DK-Report No. 2017-06 09 2017

A–4040 LINZ, ALTENBERGERSTRASSE 69, AUSTRIA

Supported by

Austrian Science Fund (FWF) Upper Austria

FWF Der Wissenschaftsfonds.

LAND OBERÖSTERREICH

DK sponsors:

- **Johannes Kepler University Linz** (JKU)
- **Austrian Science Fund** (FWF)
- **Upper Austria**

A FULLY PARALLELIZABLE SPACE-TIME MULTILEVEL MONTE CARLO METHOD FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH ADDITIVE NOISE*

MARTIN NEUMÜLLER† AND ANDREAS THALHAMMER‡

Abstract. In this work a combination of parallelizable space-time multigrid methods for deterministic parabolic partial differential equations with multilevel Monte Carlo methods for stochastic differential equations with additive noise is developed. Instead of applying the backward Euler–Maruyama scheme sequentially for every time step, the basic idea for the considered space-time method is to solve a large linear system at once, for which a parallelizable multigrid algorithm is constructed that inherits the space-time hierarchy of the multilevel Monte Carlo method. Overall this results in a fully parallelizable algorithm with respect to space, time and probability. As model problems for the numerical testing of the proposed method serve in finite dimensions the Ornstein-Uhlenbeck process and in infinite dimensions the stochastic heat equation in 2 and 3 space dimensions.

Key words. Monte Carlo estimators, space-time multigrid, stochastic differential equations, stochastic partial differential equations, parallel-in-time algorithms

AMS subject classifications. 65C05, 60H35, 65Y05, 60H10, 60H15

1. Introduction. In many fields of science, stochastic differential equations (SDEs) are commonly used to model time-dependent problems that are perturbed under random influences. Since the solution to SDEs can rarely be computed analytically, the numerical analysis of stochastic ordinary differential equations (SODEs) and stochastic partial differential equations (SPDEs) has gained a lot of attention over the last decades. In this work we focus on the approximation of $\mathbb{E}[\varphi(X(T))]$, where $X(T)$ denotes the (mild or strong) solution of a linear SDE with additive noise evaluated at a terminal time $T > 0$.

A standard way to approximate $\mathbb{E}[\varphi(X(T))]$ is to use Monte Carlo estimators that are based on independent realizations of the quantity of interest $\varphi(\hat{X}(T))$, where $\hat{X}(t), t \in [0,T]$, denotes a numerical approximation of the solution $X(t), t \in [0,T]$. The corresponding approximation error of such Monte Carlo simulation methods decomposes into the systematic error induced by the weak convergence error of the considered numerical scheme and into the Monte Carlo error, see e.g. [16]. Since the Monte Carlo error converges in the root mean-square error only with $\mathcal{O}(M^{-1/2})$, where M denotes the number of independent realizations used in the Monte Carlo estimator, see [4], one needs in order to obtain a certain accuracy $\varepsilon > 0$ in total $M = \mathcal{O}(\varepsilon^2)$ numerical realizations. This is rather computationally expensive, especially, if one considers the numerical approximation of SPDEs or high-dimensional SODEs, where the simulation of a single trajectory as such is already a computationally demanding task.

In order to improve the efficiency of standard Monte Carlo estimators, the multilevel Monte Carlo (MLMC) method can be applied to reduce the computational

*Acknowledgement: The authors want to thank Evelyn Buckwar and Ulrich Langer for their support and for fruitful discussions.

Funding: The second author work was funded by the Austrian Science Fund (FWF): W1214-N15, project DK14.

†Johannes Kepler University, Institute of Computational Mathematics, Altenberger Straße 69, A-4040 Linz, Austria (neumueller@numa.uni-linz.ac.at).

‡Johannes Kepler University, Doctoral Program "Computational Mathematics" and Institute for Stochastics, Altenberger Straße 69, A-4040 Linz, Austria (andreas.thalhammer@jku.at).

complexity of approximating $\mathbb{E}[\varphi(X(T))]$, see [11] for an extensive overview of a large range of applications of the MLMC method. The basic idea is to use instead of a numerical approximation with respect to a single refinement level, a whole collection of approximations on different refinement levels with increasing accuracy and computational cost, see e.g. in [10, 11] for details on the MLMC method for SODEs and in [3, 4, 13] for details on its applications to SPDEs.

High performance computers have a large number of cores in the present days, which allow to solve larger and larger problems. Since (multilevel) Monte Carlo estimators consist of a linear combination of independent samples of the considered random variable, Monte Carlo methods are known to be easily parallelizable. A further parallelization of numerical methods for the approximation of solution trajectories can be exploited in order to reduce the overall computation time. This could be e.g. achieved by using (semi-)implicit time stepping schemes to SPDEs (such as the backward Euler–Maruyama scheme) in combination with a spatial discretization (by e.g. using a standard finite element method), where efficiently parallelizable methods are applied in every time step sequentially to solve the arising linear systems in parallel. Another possibility to add parallelization is given by the time direction. In fact parallelization in time has become a very interesting topic of research in the past years, see [8] for a historical overview of different parallel-in-time methods. Space-time multigrid methods have shown to be very efficient and robust methods [12, 6, 9] for solving parabolic problems, which will be considered in this work. First results on time-parallel algorithms for stochastic ordinary differential equations using Parareal can be found in [2].

The idea in this work is to combine the MLMC method with space-time multigrid methods by using the space-time hierarchy coming from the MLMC estimator also for solving the arising linear systems. This work is an extension of [17], where the stochastic heat equation is only in the one-dimensional case considered.

The outline of this article is as follows: In Section 2, we introduce two model problems (the Ornstein-Uhlenbeck process and the stochastic heat equation in various space dimensions) together with discretization techniques for these model problems with respect to space and time. Afterwards, we consider the multilevel Monte Carlo (MLMC) method for approximating the expectation in Section 3 and we discuss parallelizable space-time multigrid methods based on the inherited space-time hierarchy of the MLMC estimator in Section 4. We present in Section 5 numerical experiments for the model problems that illustrate the parallelization and convergence properties of the proposed numerical method. Finally, we summarize the obtained results in Section 6.

2. Model problems. Let $T > 0$ and let $(\Omega, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathcal{F}, \mathbb{P})$ be a complete probability space. At first, we consider a one-dimensional model problem given by the stochastic ordinary differential equation (SODE)

$$\begin{aligned} \mathrm{d}u(t) + \lambda u(t)\,\mathrm{d}t &= \sigma\,\mathrm{d}\beta(t) \qquad \text{for } t \in (0,T], \\ u(0) &= u_0, \end{aligned} \tag{1}$$

where $\lambda \in \mathbb{R}_0^+, \sigma, u_0 \in \mathbb{R}$ and $\beta = (\beta(t), t \in [0,T])$ is a standard Brownian motion. The solution of this SODE is a special *Ornstein-Uhlenbeck* process defined by

$$u(t) = u_0 e^{-\lambda t} + \sigma \int_0^t e^{-\lambda(t-s)}\,\mathrm{d}\beta(s), \qquad t \in [0,T]. \tag{2}$$

As second model problem we consider the *stochastic heat equation* on a bounded and convex domain $D \subset \mathbb{R}^d, d = 1, 2, 3$, with homogeneous Dirichlet boundary conditions. If $d = 2$ then D is assumed to be polygonal and if $d = 3$, then the domain D is polyhedral.

We rewrite the stochastic partial differential equation (SPDE) as a stochastic evolution equation on the Hilbert space $H = L^2(D)$

$$\begin{aligned} \mathrm{d}U(t) &= \mathcal{A}U(t)\,\mathrm{d}t + G\,\mathrm{d}W(t) \qquad \text{for } t \in (0,T], \\ U(0) &= U_0 \in \mathcal{D}(\mathcal{A}), \end{aligned} \tag{3}$$

where $G \in L(U;H)$ and $\mathcal{A}v = \Delta_x v$ for $v \in \mathcal{D}(\mathcal{A})$. It is a well-known result that the operator $\mathcal{A}$ is the generator of an analytic semigroup $(S(t), t \in [0,T])$. Subsequently, we denote by $(e_j, j \in \mathbb{N})$ the set of eigenfunctions of the differential operator $\mathcal{A}$, which forms an orthonormal basis of H.

Let U be a separable Hilbert space. Then we assume that $W = (W(t), t \in [0,T])$ is a U-valued Q-Wiener process with a linear, positive definite, symmetric, trace class covariance operator Q. By results from e.g. [5, 15], W can be represented as

$$W(t) = \sum_{j=1}^{\infty} \sqrt{\mu_j} f_j \beta_j(t), \tag{4}$$

where $(\mu_j, j \in \mathbb{N})$ denotes the set of eigenvalues of Q with corresponding eigenfunctions $(f_j, j \in \mathbb{N})$ satisfying $Qf_j = \mu_j f_j$ and $(\beta_j, j \in \mathbb{N})$ is a sequence of independent standard Brownian motions.

Then, by [5], there exists a unique, square-integrable mild solution to SPDE (3)

$$U(t) = S(t)U_0 + \int_0^t S(t-s)G\,\mathrm{d}W(s) \qquad \text{for } t \in [0,T]. \tag{5}$$

2.1. Discretization of model problems. In this section, we present fully discrete schemes for approximating the solution processes from Eq. (2) and Eq. (5). For this we fix an equidistant partition Θ_K of the time interval $[0,T]$ given by

$$\Theta_K = \{0 = t_0 < t_1 < \cdots < t_K = T\},$$

where $t_j = j\Delta t$ for $0 \le j \le K$ with time step size $\Delta t = T/K$.

Discretization of the Ornstein–Uhlenbeck process. For the approximation of the strong solution of SODE (1), we consider the backward Euler–Maruyama scheme given by

$$\begin{aligned} (1 + \lambda\Delta t)\mathbf{u}_j &= \mathbf{u}_{j-1} + \sigma\Delta\beta^j, \qquad \text{for } 1 \le j \le K, \\ \mathbf{u}_0 &= u_0, \end{aligned} \tag{6}$$

where the Brownian increments are given by $\Delta\beta^j = \beta(t_j) - \beta(t_{j-1})$. By rewriting the recursion (6) in a matrix-vector representation we obtain the following linear system

$$\begin{pmatrix} (1+\lambda\Delta t) & & & \\ -1 & (1+\lambda\Delta t) & & \\ & \ddots & \ddots & \\ & & -1 & (1+\lambda\Delta t) \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_K \end{pmatrix} = \begin{pmatrix} \sigma\Delta\beta^1 + \mathbf{u}_0 \\ \sigma\Delta\beta^2 \\ \vdots \\ \sigma\Delta\beta^K \end{pmatrix}, \tag{7}$$

that is abbreviated in this article by

$$\mathcal{L}_\tau \mathbf{u} = \mathbf{f}(\omega),$$

where we use the ω-dependency in $\mathbf{f}(\omega)$ to indicate that the right hand side is a random vector.

Discretization of the stochastic heat equation. For approximating the mild solution (5) of the stochastic heat equation (3) we consider a standard Galerkin finite element (FE) discretization based on a regular family $(\mathcal{T}_h, h \in (0,1])$ of triangulations of $D = (0,1)^d, d = 2,3$, with maximal mesh size h. Hence we want to find a fully discrete approximation U_h^j of the mild solution $U(t_j), t_j \in \Theta_K$, where U_h^j attains values in a finite-dimensional subspace $V_h \subset \mathcal{D}((-\mathcal{A})^{1/2}) \subset H$ given by the space of globally continuous and on $\mathcal{T}_h$ piecewise linear functions. By using the nodal basis functions $(\phi_i, 1 \leq i \leq N_h)$, where $N_h \in \mathbb{N}$ denotes the dimension of V_h, the fully discrete approximation scheme based on Galerkin finite elements in space and on the backward Euler–Maruyama scheme in time is given by (see e.g. [3])

$$(M_h + \Delta t K_h)\mathbf{U}_j = M_h \mathbf{U}_{j-1} + \Delta \mathbf{W}^j \qquad \text{for } 1 \leq j \leq K, \tag{8}$$

where $\Delta \mathbf{W}^j$ denotes the vector representation of the FE approximation of the Q-Wiener increments $G\Delta W^j(\mathbf{x}) = GW(t_j, \mathbf{x}) - GW(t_{j-1}, \mathbf{x}), \mathbf{x} \in D$, and

$$U_h^j = \sum_{i=1}^{N_h} \mathbf{U}_j[i]\phi_i,$$

where $j = 0, \ldots, K$ and $\mathbf{U}_j[i]$ denotes the ith component of the vector $\mathbf{U}_j \in \mathbb{R}^{N_h}$. Here, we denote by M_h the standard *mass* matrix and K_h the standard *stiffness* matrix given by

$$M_h[i,j] := \int_D \phi_j(\mathbf{x})\phi_i(\mathbf{x})\,\mathrm{d}\mathbf{x},$$
$$K_h[i,j] := \int_D \nabla\phi_j(\mathbf{x}) \cdot \nabla\phi_i(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

for $i, j = 1, \ldots, N_h$. Finally, by rewriting the numerical scheme (8) in a matrix-vector formulation we obtain the large linear system

$$\begin{pmatrix} B_h & & & \\ -M_h & B_h & & \\ & \ddots & \ddots & \\ & & -M_h & B_h \end{pmatrix} \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \vdots \\ \mathbf{U}_K \end{pmatrix} = \begin{pmatrix} \Delta\mathbf{W}^1 + M_h\mathbf{U}_0 \\ \Delta\mathbf{W}^2 \\ \vdots \\ \Delta\mathbf{W}^K \end{pmatrix}, \tag{9}$$

where $B_h = M_h + \Delta t K_h$. Subsequently we abbreviate the linear system (9) by

$$\mathcal{L}_{h,\tau}\mathbf{U} = \mathbf{F}(\omega).$$

Simulation of Q-Wiener increments. By truncating the Karhunen–Loéve expansion (4) of the Q-Wiener process one can simulate the Q-Wiener increments $(\Delta W^i, i = 1, \ldots, K)$ by defining for $J \in \mathbb{N}$

$$\Delta^J W^i(\mathbf{x}) = \sum_{j=1}^{J} \sqrt{\mu_j} f_j(\mathbf{x}) \Delta\beta_j^i, \qquad \text{for } \mathbf{x} \in D \tag{10}$$

where $\Delta\beta_j^i = \beta_j(t_i) - \beta_j(t_{i-1})$. In order to preserve the convergence properties of the backward Euler scheme (8), the truncation parameter J has to be in general coupled to the dimension of the FE space V_h and to the decay rate of the eigenvalues of Q, see e.g. [3, 15].

In general the evaluation of the sum in Eq. (10) is computationally expensive - especially for high-dimensional spaces V_h. However, specific structural properties of the eigenvalues $(f_j, j \in \mathbb{N})$ can be exploited to speed up the simulation of the right hand side $\mathbf{F}(\omega)$. For the choice of f_j being the eigenvalues of the Laplace operator endowed with periodic boundary conditions on $D = (0,1)^d, d = 2, 3$, only a single fast Fourier transform (FFT) is needed to get two independent realizations of the right hand side $\mathbf{F}(\omega)$ on an equidistant grid, see [14, 15]. This has also been implemented in Section 5 for the numerical studies of the stochastic heat equation in higher dimensions (i.e., $D \subset \mathbb{R}^d, d = 2, 3$), where for the FFT computations the C++-library $FFTW3$, [7], has been used in the numerical experiments. Furthermore, we want to emphasize that the evaluation of the FFT can also be performed in parallel.

3. Multilevel Monte Carlo methods. The goal is to approximate $\mathbb{E}[\varphi(u(T))]$ or $\mathbb{E}[\varphi(U(T))]$ for a sufficiently smooth mapping $\varphi : H \to B$, where B is a separable Hilbert space, by using suitable estimators. For $Y \in L^2(\Omega; B)$ a common way to approximate $\mathbb{E}[Y]$ is to use a standard *Monte Carlo* (MC) estimator defined by

$$E_M[Y] := \frac{1}{M} \sum_{i=1}^{M} Y^{(i)},$$

where $(Y^{(i)}, i = 1, \dots, M)$ are independent realizations of Y. Here, $L^2(\Omega; B)$ denotes the space of strongly measurable random variables Y that satisfy

$$\|Y\|^2_{L^2(\Omega;B)} := \mathbb{E}[\|Y\|_B^2] < \infty.$$

Due to the rather slow convergence of the MC estimator of order $M^{-1/2}$ in the $L^2(\Omega; B)$-sense, see e.g. [4], the efficient *multilevel Monte Carlo* (MLMC) estimator has been proposed in [10]. For its definition we consider a sequence $(Y_\ell, \ell \in \mathbb{N}_0)$ of approximations of the random variable $Y \in L^2(\Omega; B)$ based on different refinement levels $\ell \in \mathbb{N}_0$ with increasing accuracy and also with increasing computational cost. The MLMC estimator is then given by

$$E^L[Y_L] := \sum_{\ell=0}^{L} E_{M_\ell}[Y_\ell - Y_{\ell-1}],$$

where $Y_{-1} = 0$. The $L^2(\Omega; B)$-error of the MLMC estimator satisfies (see [13])

$$\|\mathbb{E}[Y] - E^L[Y_L]\|_{L^2(\Omega;B)} \le \|\mathbb{E}[Y - Y_L]\|_B + \left(\sum_{\ell=0}^{L} M_\ell^{-1} \operatorname{Var}[Y_\ell - Y_{\ell-1}] \right)^{1/2} \tag{11}$$

where $\operatorname{Var}[Y] = \mathbb{E}[\|Y - \mathbb{E}[Y]\|_B^2]$ for $Y \in L^2(\Omega; B)$.

3.1. Parameter selection for model problems. We now discuss how to choose the number of samples $(M_\ell, \ell \in \mathbb{N}_0)$ and the refinement parameters h_ℓ and Δt_ℓ in order to guarantee the convergence of the MLMC estimator.

Ornstein-Uhlenbeck process. Let u be given in Eq. (2) and for $\ell \in \mathbb{N}_0$ let $\mathbf{u}_{K_\ell}$ be the numerical approximation of $u(T)$ based on the backward Euler–Maruyama scheme (6) with respect to the partition Θ_{K_ℓ} with time step size Δt_ℓ. Furthermore, let $\varphi \in C_b^2(\mathbb{R}, \mathbb{R})$, i.e., $\varphi : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable with bounded first and second derivatives. Due to the additive noise structure of SDE (1) we obtain by results from [16] that

$$\begin{aligned} |\mathbb{E}[\varphi(u(T)) - \varphi(\mathbf{u}_{K_L})]| &\leq C\Delta t_L, \\ \mathrm{Var}[\varphi(\mathbf{u}_{K_\ell}) - \varphi(\mathbf{u}_{K_{\ell-1}})]^{1/2} &\leq C\Delta t_\ell. \end{aligned}$$

Thus, by similar arguments as in [13, 17], if we choose for any $\varepsilon, C_M > 0$,

$$\begin{aligned} M_0 &= \lceil C_M \Delta t_L^{-2} \rceil, \\ M_\ell &= \lceil C_M \Delta t_\ell^2 \Delta t_L^{-2} \ell^{1+\varepsilon} \rceil \quad \text{for } \ell = 1, \dots, L, \end{aligned} \tag{12}$$

then

$$\mathbb{E}[\varphi(u(T))] - E^L[\varphi(\mathbf{u}_{K_L})]\|_{L^2(\Omega;\mathbb{R})} = \mathcal{O}(\Delta t_L).$$

Stochastic heat equation. Let U be given in Eq. (5) and for $\ell \in \mathbb{N}_0$ let $U_{h_\ell}^{K_\ell}$ be an approximation of $U(T)$ based on the FE backward Euler–Maruyama scheme (8) with respect to the partition Θ_{K_ℓ} and the FE space V_{h_ℓ}. Furthermore, let $\varphi \in \mathcal{C}_b^2(H, B)$, i.e., $\varphi : H \to B$ is twice Fréchet differentiable with bounded first and second Fréchet derivatives. Then by using the results from [1], we get by choosing $\Delta t_\ell = h_\ell^2$ for any $\gamma \in [0, 1)$

$$\begin{aligned} \|\mathbb{E}[\varphi(U(T)) - \varphi(U_{h_L}^{K_L})]\|_B &\leq C h_L^{2\gamma}, \\ \mathrm{Var}[\varphi(U_{h_\ell}^{K_\ell}) - \varphi(U_{h_{\ell-1}}^{K_{\ell-1}})] &\leq C h_\ell^{2\gamma}. \end{aligned}$$

Thus, by [13, 17], if we choose $\Delta t_\ell = h_\ell^2$ and for any $\varepsilon, C_M > 0$,

$$\begin{aligned} M_0 &= \lceil C_M h_L^{-2\gamma} \rceil, \\ M_\ell &= \lceil C_M h_\ell^{2\gamma} h_L^{-2\gamma} \ell^{1+\varepsilon} \rceil \quad \text{for } \ell = 1, \dots, L, \end{aligned} \tag{13}$$

then

$$\|\mathbb{E}[\varphi(U(T))] - E^L[\varphi(U_{h_L}^{K_L})]\|_{L^2(\Omega;B)} = \mathcal{O}(h_L^\gamma).$$

4. Space-time multigrid methods. We combine the MLMC method with a space-time multigrid method by using the hierarchy obtained from the MLMC estimator (see Section 3.1) also for the solution process for the arising linear systems. In detail we will apply the space-time multigrid method presented in [9] to solve the linear system (7) and (9) at once. This will allow parallelization in time direction and for the linear system (9) it is also possible to add parallelization with respect to the space dimension. Overall by re-using the hierarchy from the MLMC method for the solution process this will result in a method which is fully parallel with respect to space, time and probability. The key ingridient is a parallel smoothing iteration which is used in the space-time multigrid method. Here we will use an (inexact) damped block Jacobi smoother, see also [9], i.e. for the problem (7) we use

$$\mathbf{u}^{(n+1)} = \mathbf{u}^{(n)} + \alpha \mathcal{D}_\tau^{-1} \left[\mathbf{f}(\omega) - \mathcal{L}_\tau \mathbf{u}^{(n)} \right] \qquad \text{for } n = 0, 1, \dots,$$

with the diagonal matrix $\mathcal{D}_\tau := \mathrm{diag}(1 + \lambda\Delta t)$. Whereas, for the problem (9) we use the smoothing iteration

$$\mathbf{U}^{(n+1)} = \mathbf{U}^{(n)} + \alpha\mathcal{D}_{h,\tau}^{-1}\left[\mathbf{F}(\omega) - \mathcal{L}_{h,\tau}\mathbf{U}^{(n)}\right] \qquad \text{for } n = 0, 1, \ldots,$$

with the block diagonal matrix $\mathcal{D}_{h,\tau} := \mathrm{diag}(B_h)$. To speed up the application of the smoothing procedure we replace the exact inverse of $\mathcal{D}_{h,\tau}$ by applying one iteration of a multigrid V-cycle with respect to the matrix B_h. For the relaxation parameter we choose $\alpha = \frac{1}{2}$. For the restriction and prolongation operators we will use the same operators as for the MLMC method. We refer the reader to [9] for more details, where this space-time multigrid method has been analyzed in detail. Overall by choosing $\Delta t \approx h^2$ we obtain a fully parallel and robust solver which is independent of the number of time steps K and the time step size Δt.

5. Numerical experiments. In this section we present numerical experiments for the Ornstein–Uhlenbeck process (Section 5.1) and for the stochastic heat equation on $D = (0, 1)^d, d = 2, 3$, (Section 5.2). All results were computed on the RADON1 cluster (a distributed memory cluster with 1088 CPU Cores, 8.7TB Memory; consisting of 64 compute nodes each with two 8-core Intel Haswell processors "Xeon E5-2630v3", 2.4Ghz and 128 GB of memory).

5.1. Ornstein-Uhlenbeck process. We consider the SODE (1) with $\lambda = 1, \sigma = 1, T = 1$ and $u_0 = 1$. We want to approximate $\mathbb{E}[\varphi_j(u(T))], j = 1, 2$, where $\varphi_1(x) = x$ and $\varphi_2(x) = x^2$. For both test cases the solution can be computed analytically, where

$$\mathbb{E}[\varphi_1(u(T))] = \mathbb{E}[u(T)] = e^{-T}$$

and

$$\mathbb{E}[\varphi_2(u(T))] = \mathbb{E}[u^2(T)] = \frac{1 + e^{-2T}}{2}.$$

For the numerical approximation we consider the backward Euler–Maruyama scheme from Eq. (6) in the matrix-vector representation $\mathcal{L}_\tau \mathbf{u} = \mathbf{f}(\omega)$, which is solved by the time multigrid method described in Section 4. For the approximation of the expectation we consider a multilevel Monte Carlo estimator based on the sample size selection from Eq. (12) with $\varepsilon = \frac{1}{2}$ and $C_M = 10$.

In Table 1, we consider the root mean-square errors of the MLMC method with respect to $\varphi_j, j = 1, 2$, where $\|\mathbb{E}[\varphi_1(u(T))] - E^L[\varphi_1(\mathbf{u}_{K_L})]\|_{L^2(\Omega;\mathbb{R})}$ is approximated by a standard Monte Carlo estimator given by

$$\text{RMS-err}_1(L) = \left(\frac{1}{M}\sum_{i=1}^{M}\left|\mathbb{E}[u(T)] - E^L[\mathbf{u}_{K_L}]^{(i)}\right|^2\right)^{1/2},$$

and $\|\mathbb{E}[\varphi_2(u(T))] - E^L[\varphi_2(\mathbf{u}_{K_L})]\|_{L^2(\Omega;\mathbb{R})}$ is approximated by

$$\text{RMS-err}_2(L) = \left(\frac{1}{M}\sum_{i=1}^{M}\left|\mathbb{E}[u^2(T)] - E^L[\mathbf{u}_{K_L}^2]^{(i)}\right|^2\right)^{1/2}.$$

Here $(E^L[\varphi_j(\mathbf{u}_{K_L})]^{(i)}, 1 \le i \le M)$ are independent realizations of the MLMC estimator $E^L[\varphi_j(\mathbf{u}_{K_L})]$. For the Monte Carlo simulations we choose $M = 100$ in the

numerical experiments from Table 1. The experimental order of convergence (EOC) of the root mean-square error is computed by

$$\mathrm{EOC}_i(L) = \frac{\log(\text{RMS-err}_i(L)) - \log(\text{RMS-err}_i(L-1))}{\log(\Delta t_L) - \log(\Delta t_{L-1})}, \quad i = 1, 2.$$

In Table 1 we observe for both text function $\varphi_i, i = 1, 2$, the right convergence behaviour as predicted by the theory.

TABLE 1
Numerical test for SODE (1) (Ornstein-Uhlenbeck process) - convergence.

L	K_L	M_0	M_L	RMS-err$_1(L)$	EOC$_1(L)$	RMS-err$_2(L)$	EOC$_2(L)$
0	1	10	10	2.03312E-1	-	1.93142E-1	-
1	2	40	20	1.16614E-1	0.80	1.10059E-1	0.81
2	4	160	50	6.15999E-2	0.92	5.87172E-2	0.91
3	8	640	80	2.99257E-2	1.04	3.06077E-2	0.94
4	16	2560	110	1.52328E-2	0.97	1.70489E-2	0.84
5	32	10240	140	7.75317E-3	0.97	8.50822E-3	1.00
6	64	40960	180	3.71647E-3	1.06	4.33203E-3	0.97
7	128	163840	220	1.81249E-3	1.04	2.27805E-3	0.93
8	256	655360	270	9.65028E-4	0.91	1.26140E-3	0.85
9	512	2621440	310	4.48929E-4	1.10	6.26638E-4	1.01
10	1024	10485760	360	1.96168E-4	1.19	3.42444E-4	0.87

5.2. Stochastic heat equation. For the stochastic heat equation (3) we consider the domain $D = (0, 1)^d, d = 2, 3$, and the initial value

$$U_0(\mathbf{x}) = \prod_{i=1}^{d} \sin(\pi x_i) \qquad \text{for } \mathbf{x} \in D.$$

By choosing $T = 0.1$ and $\varphi(v) = v$ for all $v \in L^2(D)$, we are interested in approximating

$$\mathbb{E}[U(T, \mathbf{x})] = \exp(-d\pi^2 T) \prod_{i=1}^{d} \sin(\pi x_i), \qquad \text{for } \mathbf{x} \in D.$$

For the numerical approximation in space and time, we consider the FE Euler–Maruyama scheme from Eq. (8) on an equidistant spatial triangulation in the matrix-vector formulation $\mathcal{L}_{h,\tau}\mathbf{U} = \mathbf{F}(\omega)$, which is again solved by the space-time multigrid method described in Section 4. For the approximation of the expectation we consider the MLMC method based on the sample size selection (13) with $\varepsilon = \frac{1}{2}$ and $C_M = 10$.

In numerical experiments the root mean-square error $\|\mathbb{E}[U(T)] - E^L[U_{h_L}^{K_L}]\|_{L^2(\Omega;B)}$ is approximated by a standard Monte Carlo estimator, i.e., we consider

$$\text{RMS-err}(L) = \left(\frac{1}{M} \sum_{i=1}^{M} \left\| \mathbb{E}[U(T)] - E^L[U_{h_L}^{K_L}]^{(i)} \right\|_{L^2(D)}^2 \right)^{1/2},$$

where $(E^L[U_{h_L}^{K_L}]^{(i)}, 1 \le i \le M)$ are independent realizations of the estimator $E^L[U_{h_L}^{K_L}]$ and

$$\|\mathbb{E}[U(T)] - E^L[U_{h_L}^{K_L}]^{(i)}\|_{L^2(D)}^2 = \int_D \left| \mathbb{E}[U(T, \mathbf{x})] - E^L[U_{h_L}^{K_L}(\mathbf{x})]^{(i)} \right|^2 \mathrm{d}\mathbf{x}.$$

The experimental order of convergence (EOC) of the root mean-square error is now computed by

$$\mathrm{EOC}(L) = \frac{\log(\text{RMS-err}(L)) - \log(\text{RMS-err}(L-1))}{\log(h_L) - \log(h_{L-1})}.$$

In this section we consider the stochastic heat equation (3) in 2D and 3D, i.e., let $D = (0,1)^d, d = 2, 3$. In order to apply the FFT-based simulation techniques of the Q-Wiener increments described in Section 2.1 we consider here the specific choice of eigenvalues

$$\mu_{\mathbf{k}} = \exp(-\mathrm{i}a|\mathbf{k}|^2),$$

where $\mathbf{k} = (k_1, \ldots, k_d)^T \in \mathbb{Z}^d$, $a = 2$ and i denotes the imaginary unit, see [15] for details. The corresponding eigenfunctions $f_{\mathbf{k}}$ are given by

$$f_{\mathbf{k}}(\mathbf{x}) = \exp(2\pi\mathrm{i}\langle \mathbf{k}, \mathbf{x}\rangle) \qquad \text{for } \mathbf{x} \in D.$$

For approximating realizations of the Q-Wiener increments we consider the following truncation of Eq. (4): For $\mathcal{J} = (J_1, \ldots, J_d)^T \in \mathbb{N}^d$ define

$$\Delta^{\mathcal{J}} W^i(\mathbf{x}) = \sum_{\substack{\mathbf{k}\in\mathbb{Z}^d,\\ -J_1/2<|k_1|\leq J_1/2,\\ \cdots,\\ -J_d/2<|k_d|\leq J_d/2}} \sqrt{\mu_{\mathbf{k}}} f_{\mathbf{k}}(\mathbf{x}) \Delta\widetilde{\beta}^i_{\mathbf{k}}, \qquad \text{for } \mathbf{x} \in D.$$

Here the increments $\Delta\widetilde{\beta}^i_{\mathbf{k}} = \Delta\beta^{i,\mathrm{real}}_{\mathbf{k}} + \mathrm{i}\Delta\beta^{i,\mathrm{imag}}_{\mathbf{k}}$ are complex, where the real and imaginary parts consist of independent, real-valued Brownian increments $\Delta\beta^{i,\mathrm{real}}_{\mathbf{k}}, \Delta\beta^{i,\mathrm{imag}}_{\mathbf{k}}$, see [15]. In the numerical experiments below we choose $J_i = N_h, i = 1, 2, 3$.

In Tables 2 and 4 we use $M = 100$ realizations of the MLMC estimator to approximate the $L^2(\Omega; H)$-error by a Monte Carlo simulation for the 2D and 3D stochastic heat equation. The experimental outcome illustrates for both test cases that the predicted convergence order $\gamma \approx 1$ is attained.

In Tables 3 and 5 we measure the computation times for one run of the MLMC estimator, where we consider different distributions of 512 cores with respect to solving the linear system $\mathcal{L}_{h,\tau}\mathbf{U} = \mathbf{F}(\omega)$ in parallel or for parallelizing the involved Monte Carlo estimators. In the considered numerical experiments we obtain that a balanced distribution of cores for the parallelization in time and for the parallelization of the Monte Carlo estimators yields the best computation time. For instance for the numerical experiments in the case $d = 2$ at level $L = 6$ the best setting is given by 128 cores for the time parallelization and only 4 cores for the Monte Carlo simulations (844.61 sec.). Overall, this results in an approximately 19-times faster solving time than for a straightforward parallelization of the MLMC estimator (16278.1 sec.).

6. Conclusions. In this work we combined a space-time multigrid technique for simulating trajectories of the solution process of the Ornstein–Uhlenbeck process and the stochastic heat equation on $D = (0,1)^d, d = 2, 3$, with a multilevel Monte Carlo estimator to approximate $\mathbb{E}[\varphi(u(T))]$ and $\mathbb{E}[\varphi(U(T))]$, respectively. This combination leads to a fully parallelizable method with respect to space, time and probability.

Special emphasis was laid on constructing the MLMC estimator in such a way that the root mean-square errors converge with order 1 (with respect to the time step size Δt_L for the Ornstein–Uhlenbeck process and with respect to the spatial refinement

TABLE 2
Numerical test for SPDE (3) (stochastic heat equation in 2D) – convergence.

L	K_L	N_{h_L}	DOF (ST)	M_0	M_L	RMS-err(L)	EOC(L)
1	4	1	4	40	20	3.72225E-02	-
2	16	9	144	160	30	1.85333E-02	1.01
3	64	49	3136	640	60	9.49312E-03	0.97
4	256	225	57600	2560	90	4.65968E-03	1.03
5	1024	961	984064	10240	120	2.10172E-03	1.15
6	4096	3969	16257024	40960	150	1.03803E-03	1.02

TABLE 3
Numerical test for SPDE (3) (stochastic heat equation in 2D) – computation time with respect to different distributions of 512 *cores (in sec).*

	cores time / cores Monte Carlo				
L	1 / 512	2 / 256	4 / 128	8 / 64	16 / 32
1	0.07	0.08	0.05	0.07	0.07
2	0.17	0.05	0.08	0.13	0.06
3	0.40	0.22	0.14	**0.12**	0.13
4	6.92	3.59	1.94	1.84	**1.56**
5	115.53	69.83	35.30	34.94	35.31
6	16278.10	7993.41	6310.75	3081.63	1575.11
	cores time / cores Monte Carlo				
L	32 / 16	64 / 8	128 / 4	256 / 2	512 / 1
1	**0.00**	0.03	0.03	0.01	0.06
2	0.05	**0.03**	0.09	0.13	0.17
3	0.14	0.22	0.42	0.70	1.50
4	1.69	2.07	2.81	5.01	10.11
5	**34.44**	34.68	39.93	50.37	103.75
6	1154.21	953.59	**844.61**	978.95	1168.93

parameter h_L for the stochastic heat equation). For this the number of numerical realizations $(M_\ell, \ell = 0, \dots, L)$ for the MLMC estimator could be precomputed. In all considered numerical experiments the convergence results could be numerically observed.

For the stochastic heat equation, the robustness of the proposed space-time multigrid solver is guaranteed by [9], since we were using the space-time hierarchy of the MLMC estimator also for the backward Euler–Maruyama approximation of the numerical trajectories, where the space-time coupling satisfies $\Delta t_\ell / h_\ell^2 \approx 1$ on all considered levels $\ell = 1, \dots, L$.

Furthermore, we investigated the optimal distribution of 512 cores with respect to solving the linear system of the space-time formulation and the parallelization of the involved Monte Carlo estimators. In the presented numerical experiments on the stochastic heat equation it turned out that a balanced distribution of cores between parallelization in time and parallelization of the MLMC estimator provides the best computation times.

REFERENCES

TABLE 4
Numerical test for SPDE (3) (stochastic heat equation in 3D) – convergence.

L	K_L	N_{h_L}	DOF (ST)	M_0	M_L	RMS-err(L)	EOC(L)
1	4	1	4	40	20	1.47372E-02	-
2	16	27	432	160	30	1.23371E-02	0.26
3	64	343	21952	640	60	7.48574E-03	0.72
4	256	3375	864000	2560	90	3.62278E-03	1.05
5	1024	29791	30505984	10240	120	1.76545E-03	1.04

TABLE 5
Numerical test for SPDE (3) (stochastic heat equation in 3D) – computation time with respect to different distributions of 512 *cores (in sec).*

	cores time / cores Monte Carlo				
L	1 / 512	2 / 256	4 / 128	8 / 64	16 / 32
1	0.25	0.17	**0.02**	0.05	0.03
2	0.22	0.21	0.14	0.21	0.06
3	2.66	1.36	0.75	**0.36**	0.42
4	133.55	63.82	33.58	32.89	**25.12**
5	5167.79	2886.78	1571.55	1597.29	1501.86
	cores time / cores Monte Carlo				
L	32 / 16	64 / 8	128 / 4	256 / 2	512 / 1
1	0.03	0.03	0.03	0.05	0.05
2	**0.04**	0.07	0.11	0.42	0.37
3	0.56	0.91	1.36	3.57	5.84
4	25.28	27.92	33.79	51.16	99.39
5	1525.51	**1491.09**	1582.12	1671.24	1935.34

[1] A. ANDERSSON, R. KRUSE, AND S. LARSSON, *Duality in refined Sobolev-Malliavin spaces and weak approximation of SPDE*, Stoch. Partial Differ. Equ. Anal. Comput., 4 (2016), pp. 113–149, https://doi.org/10.1007/s40072-015-0065-7, http://dx.doi.org/10.1007/s40072-015-0065-7.

[2] G. BAL, *Parallelization in time of (stochastic) ordinary differential equations*, (2017), https://galton.uchicago.edu/~guillaumebal/PAPERS/paralleltime.pdf.

[3] A. BARTH AND A. LANG, *Simulation of stochastic partial differential equations using finite element methods*, Stochastics An International Journal of Probability and Stochastic Processes, 84 (2012), pp. 217–231.

[4] A. BARTH, A. LANG, AND C. SCHWAB, *Multilevel Monte Carlo method for parabolic stochastic partial differential equations*, BIT, 53 (2013), pp. 3–27.

[5] G. DA PRATO AND J. ZABCZYK, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, 44. Cambridge University Press, Cambridge, 1992.

[6] R. D. FALGOUT, S. FRIEDHOFF, T. V. KOLEV, S. P. MACLACHLAN, AND J. B. SCHRODER, *Parallel time integration with multigrid*, SIAM J. Sci. Comput., 36 (2014), pp. C635–C661.

[7] M. FRIGO AND S. G. JOHNSON, *The design and implementation of FFTW3*, Proceedings of the IEEE, 93 (2005), pp. 216–231. Special issue on "Program Generation, Optimization, and Platform Adaptation".

[8] M. GANDER, *50 years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition, T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, eds., Springer-Verlag, 2015, pp. 69–114, http://www.unige.ch/~gander/Preprints/50YearsTimeParallel.pdf.

[9] M. J. GANDER AND M. NEUMÜLLER, *Analysis of a new space-time parallel multigrid algorithm for parabolic problems*, SIAM J. Sci. Comput., 38 (2016), pp. A2173–A2208, https://doi.org/10.1137/15M1046605, http://dx.doi.org/10.1137/15M1046605.

[10] M. B. GILES, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617,

https://doi.org/10.1287/opre.1070.0496.

[11] M. B. Giles, *Multilevel Monte Carlo methods*, Acta Numer., 24 (2015), pp. 259–328, https://doi.org/10.1017/S096249291500001X.

[12] G. Horton and S. Vandewalle, *A space-time multigrid method for parabolic partial differential equations*, SIAM J. Sci. Comput., 16 (1995), pp. 848–864.

[13] A. Lang, *A note on the importance of weak convergence rates for SPDE approximations in multilevel Monte Carlo schemes*, in Monte Carlo and quasi-Monte Carlo methods, vol. 163 of Springer Proc. Math. Stat., Springer, [Cham], 2016, pp. 489–505, https://doi.org/10.1007/978-3-319-33507-0_25, http://dx.doi.org/10.1007/978-3-319-33507-0_25.

[14] A. Lang and J. Potthoff, *Fast simulation of Gaussian random fields*, Monte Carlo Methods Appl., 17 (2011), pp. 195–214, https://doi.org/10.1515/MCMA.2011.009, http://dx.doi.org/10.1515/MCMA.2011.009.

[15] G. J. Lord, C. E. Powell, and T. Shardlow, *An Introduction to Computational Stochastic PDEs*, Cambridge Texts in Applied Mathematics. Cambridge University Press, New York, 2014.

[16] G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Scientific Computation. Springer-Verlag, Berlin, 2004.

[17] M. Neumüller and A. Thalhammer, *Combining Space-Time Multigrid Techniques with Multilevel Monte Carlo Methods for SDEs*, DK Report 2017-04, Doctoral Program in Computational Mathematics, Johannes Kepler University Linz, Austria, Linz, 2017, (submitted for DD24 proceedings), https://www.dk-compmath.jku.at/publications/dk-reports/2017-05-07/view.

Technical Reports of the Doctoral Program "Computational Mathematics"

2017

2017-01 E. Buckwar, A. Thalhammer: *Importance Sampling Techniques for Stochastic Partial Differential Equations* January 2017. Eds.: U. Langer, R. Ramlau

2017-02 C. Hofer, I. Toulopoulos: *Discontinuous Galerkin Isogeometric Analysis for parametrizations with overlapping regions* June 2017. Eds.: U. Langer, V. Pillwein

2017-03 C. Hofer, S. Takacs: *Inexact Dual-Primal Isogeometric Tearing and Interconnecting Methods* June 2017. Eds.: B. Jüttler, V. Pillwein

2017-04 M. Neumüller, A. Thalhammer: *Combining Space-Time Multigrid Techniques with Multilevel Monte Carlo Methods for SDEs* June 2017. Eds.: U. Langer, E. Buckwar

2017-05 C. Hofer, U. Langer, M. Neumüller: *Time-Multipatch Discontinuous Galerkin Space-Time Isogeometric Analysis of Parabolic Evolution Problems* August 2017. Eds.: V. Pillwein, B. Jüttler

2017-06 M. Neumüller, A. Thalhammer: *A Fully Parallelizable Space-Time Multilevel Monte Carlo Method for Stochastic Differential Equations with Additive Noise* September 2017. Eds.: U. Langer, E. Buckwar

2016

2016-01 P. Gangl, U. Langer: *A Local Mesh Modification Strategy for Interface Problems with Application to Shape and Topology Optimization* November 2016. Eds.: B. Jüttler, R. Ramlau

2016-02 C. Hofer: *Parallelization of Continuous and Discontinuous Galerkin Dual-Primal Isogeometric Tearing and Interconnecting Methods* November 2016. Eds.: U. Langer, W. Zulehner

2016-03 C. Hofer: *Analysis of Discontinuous Galerkin Dual-Primal Isogeometric Tearing and Interconnecting Methods* November 2016. Eds.: U. Langer, B. Jüttler

2016-04 A. Seiler, B. Jüttler: *Adaptive Numerical Quadrature for the Isogeometric Discontinuous Galerkin Method* November 2016. Eds.: U. Langer, J. Schicho

2016-05 S. Hubmer, A. Neubauer, R. Ramlau, H. U. Voss: *On the Parameter Estimation Problem of Magnetic Resonance Advection Imaging* December 2016. Eds.: B. Jüttler, U. Langer

2016-06 S. Hubmer, R. Ramlau: *Convergence Analysis of a Two-Point Gradient Method for Nonlinear Ill-Posed Problems* December 2016. Eds.: B. Jüttler, U. Langer

2015

2015-01 G. Grasegger, A. Lastra, J. Rafael Sendra, F. Winkler: *A Solution Method for Autonomous First-Order Algebraic Partial Differential Equations in Several Variables* January 2015. Eds.: U. Langer, J. Schicho

2015-02 P. Gangl, U. Langer, A. Laurain, H. Meftahi, K. Sturm: *Shape Optimization of an Electric Motor Subject to Nonlinear Magnetostatics* January 2015. Eds.: B. Jüttler, R. Ramlau

2015-03 C. Fürst, G. Landsmann: *Computation of Dimension in Filtered Free Modules by Gröbner Reduction* May 2015. Eds.: P. Paule, F. Winkler

2015-04 A. Mantzaflaris, H. Rahkooy, Z. Zafeirakopoulos: *Efficient Computation of Multiplicity and Directional Multiplicity of an Isolated Point* July 2015. Eds.: B. Buchberger, J. Schicho

2015-05 P. Gangl, S. Amstutz, U. Langer: *Topology Optimization of Electric Motor Using Topological Derivative for Nonlinear Magnetostatics* July 2015. Eds.: B. Jüttler, R. Ramlau

2015-06 A. Maletzky: *Exploring Reduction Ring Theory in Theorema* August 2015. Eds.: B. Buchberger, W. Schreiner

The complete list since 2009 can be found at
`https://www.dk-compmath.jku.at/publications/`

Doctoral Program
"Computational Mathematics"

Director:

Prof. Dr. Peter Paule
Research Institute for Symbolic Computation

Deputy Director:

Prof. Dr. Bert Jüttler
Institute of Applied Geometry

Address:

Johannes Kepler University Linz
Doctoral Program "Computational Mathematics"
Altenbergerstr. 69
A-4040 Linz
Austria
Tel.: ++43 732-2468-6840

E-Mail:

office@dk-compmath.jku.at

Homepage:

http://www.dk-compmath.jku.at

Submissions to the DK-Report Series are sent to two members of the Editorial Board who communicate their decision to the Managing Editor.