Convergence of the Micro-Macro Parareal Method for a Linear Scale-Separated Ornstein-Uhlenbeck SDE

Ignace Bossuyt, Stefan Vandewalle, and Giovanni Samaey

1 Model problem and motivation

In this work, we consider a two-dimensional slow-fast Ornstein-Uhlenbeck (OU) stochastic differential equation (SDE) [9], modelling the coupled evolution of a slowly evolving variable $x \in \mathbb{R}$ and a variable $y \in \mathbb{R}$ that quickly reaches its equilibrium distribution:

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma/\epsilon & \zeta/\epsilon \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} dt + \sigma \begin{bmatrix} 1 & 0 \\ 0 & 1/\sqrt{\epsilon} \end{bmatrix} dW. \tag{1}$$

where $dW \in \mathbb{R}^2$ is a two-dimensional Brownian motion and $\epsilon \in \mathbb{R}$ is a (small) time scale separation parameter $\epsilon \ll 1$. The initial condition has a distribution with mean $\begin{bmatrix} m_{x,0} & m_{y,0} \end{bmatrix}$ and covariance matrix $\begin{bmatrix} \Sigma_{x,0} & \Sigma_{xy,0} \\ \Sigma_{xy,0} & \Sigma_{y,0} \end{bmatrix}$, and time $t \in [0, T]$.

Model problem (1) mimics the general situation where x is a low-dimensional quantity of interest whose evolution is influenced by a quickly evolving, high-dimensional variable y, all described by SDEs. The joint probability density of x and y obeys a Fokker-Planck equation (see, e.g. [3]). Instead of directly solving this partial differential equation using classical deterministic techniques, which suffer from the curse of dimensionality, the corresponding SDE can be solved using a Monte Carlo method. In this paper, our aim is to obtain insight in the convergence of a parallel-in-time (PinT) method applied to the low-dimensional linear OU model problem (1). In our method, the fine propagator of the SDE is based on a high-dimensional slow-fast microscopic model; the coarse propagator is based on a model-reduced version of the latter, that captures the low-dimensional, effective dynamics at the slow time scales. This problem allows for an analytic treatment, if

Ignace Bossuyt, Stefan Vandewalle, Giovanni Samaey
Department of Computer Science, KU Leuven, Celestijnenlaan 200A, 3001 Leuven, Belgium,
e-mail: ignace.bossuyt1@kuleuven.be, stefan.vandewalle@kuleuven.be,
giovanni.samaey@kuleuven.be

the quantities of interest are the mean and the (co)variance of x and y. We expect that this convergence analysis can be useful as a stepping stone for analysing PinT methods for higher-dimensional (nonlinear) SDEs.

1.1 Derivation of a reduced model

The averaging technique from [7, chapter 10, see, e.g., Remark 10.2] allows to define the reduced dynamics variable X, that approximates the slow variable x in (1). This technique exploits time-scale separation in order to integrate out the fast variable with respect to $\rho^\infty(y|x)$, the invariant distribution of the fast variable y conditioned on a fixed slow variable x.

The reduced model reads as follows (Λ_Σ and Σ_Σ are defined implicitly):

$$dX = A(X)dt + S(X)dW \tag{2}$$

with

$$A(X) = \int_{\mathcal{Y}} a(X,y)\rho^\infty(y|X)dy = \Lambda_\Sigma X := \left(\alpha - \frac{\beta\gamma}{\zeta}\right) X$$

$$S(X)S(X)^T = \int_{\mathcal{Y}} s(X,y)s(X,y)^T \rho^\infty(y|X)dy = \Sigma_\Sigma := \sigma,$$

where $\mathcal{Y}$ denotes the domain of y. It can be shown that for the OU system (1), the conditional distribution $\rho^\infty(y|x) = \mathcal{N}\left(\frac{\gamma x}{\zeta}, \frac{\sigma^2}{2\zeta}\right)$ (see [7, Example 6.19]).

The reduced model (2), while it is only an approximation to the slow dynamics, offers two computational advantages w.r.t. the full, scale-separated system (1): (i) it contains fewer degrees of freedom, and (ii) it is less stiff with a computational cost that is independent of ϵ. As ϵ approaches zero, the multiscale model (1) gets more stiff, while the (cheaper) reduced model becomes a more accurate approximation.

1.2 Moment system for the Ornstein-Uhlenbeck process

The evolution of mean and variance of a linear SDE can be described exactly using the moment models from [1]. Thus, for the linear Ornstein-Uhlenbeck SDE model problem, we can use these linear ODEs instead of using a Monte Carlo simulation.

Moments for reduced model. The evolution of the mean of X in (2) is given by

$$\frac{dm_X}{dt} = \left(\alpha - \frac{\beta\gamma}{\zeta}\right) m_X. \tag{3}$$

The evolution of the variance of the reduced system is given by the ODE

$$\frac{d\Sigma_X}{dt} = \Lambda_\Sigma \Sigma_X + \Sigma_\Sigma^2 = 2\left(\alpha - \frac{\beta\gamma}{\zeta}\right)\Sigma_X + \sigma^2. \tag{4}$$

Moments for multiscale model. The evolution of the mean of the multiscale SDE (1) is described by the following linear ODE:

$$\frac{d}{dt}\begin{bmatrix} m_x \\ m_y \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma/\epsilon & \zeta/\epsilon \end{bmatrix} \begin{bmatrix} m_x \\ m_y \end{bmatrix}. \tag{5}$$

The evolution of the covariance of (1) is given by the linear ODE $\dot{\Sigma} = B_\Sigma \Sigma + b_\Sigma$:

$$\frac{d}{dt}\begin{bmatrix} \Sigma_x \\ \Sigma_{xy} \\ \Sigma_y \end{bmatrix} = \left[\begin{array}{c|cc} 2\alpha & 2\beta & 0 \\ \hline \gamma/\epsilon & \alpha + \zeta/\epsilon & \beta \\ 0 & 2\gamma/\epsilon & 2\zeta/\epsilon \end{array}\right] \begin{bmatrix} \Sigma_x \\ \Sigma_{xy} \\ \Sigma_y \end{bmatrix} + \begin{bmatrix} \sigma^2 \\ 0 \\ \sigma^2/\epsilon \end{bmatrix} \tag{6}$$

where we define $\Sigma_q = \begin{bmatrix} \Sigma_{xy} & \Sigma_y \end{bmatrix}^T$, and where the blocks of B_Σ are named as $B_\Sigma = \left[\begin{array}{c|c} 2\alpha & p_\Sigma^T \\ \hline q_\Sigma/\epsilon & -A_\Sigma/\epsilon \end{array}\right]$, where $A_\Sigma = -\begin{bmatrix} \alpha + \zeta/\epsilon & \beta \\ 2\gamma/\epsilon & 2\zeta/\epsilon \end{bmatrix}$. To ensure stability of the fast dynamics, we assume that the parameters in (1) are chosen such that the real part of the eigenvalues of the matrix A_Σ are all positive $\mu_{\Sigma,i} \geq \mu_- > 0$. This condition is satisfied for instance for any $\alpha, \beta \in \mathbb{R}$ if ζ and γ are sufficiently small.

2 The Micro-Macro Parareal algorithm

The Micro-Macro Parareal (mM-Parareal) for scale-separated ODEs [5] and for SDEs [4], is a generalisation of the Parareal algorithm [6]. It combines two levels of description: (i) the micro variable u, with corresponding fine propagator $\mathcal{F}$, and (ii) the macro variable ρ, which is lower-dimensional, with coarse propagator C. These levels are related through coupling operators: the restriction operator $\mathcal{R}$ extracts macro information from a micro state, the lifting operator $\mathcal{L}$ produces a micro state that is consistent with a given macro state, and finally the matching operator $\mathcal{M}$ produces a micro state that is consistent with a given macro state, based on prior information of the micro state. Examples of these operators are given in the sequel. The mM-Parareal algorithm iterate at iteration k and time step n is given next. For $k = 0$ (initialization), we have

$$\rho_{n+1}^0 = C(\rho_n^0) \qquad u_{n+1}^0 = \mathcal{L}(\rho_{n+1}^0), \tag{7}$$

and for $k \geq 1$,

$$\begin{aligned} \rho_{n+1}^{k+1} &= C(\rho_n^{k+1}) + \mathcal{R}(\mathcal{F}(u_n^k)) - C(\rho_n^k) \\ u_{n+1}^{k+1} &= \mathcal{M}(\rho_{n+1}^{k+1}, \mathcal{F}(u_n^k)). \end{aligned} \tag{8}$$

If the coupling operators are chosen such that $\mathcal{M}(\mathcal{R}u, u) = u$, then at each iteration it holds that $\rho_n^k = \mathcal{R}u_n^k$. Classical Parareal [6] corresponds to the case $\mathcal{R} = \mathcal{L} = \mathcal{M} = \mathcal{I}$.

Convergence of Micro-Macro Parareal for linear scale-separated ODEs. In [5], the convergence of mM-Parareal for a linear scale-separated ODE is studied.

We briefly review the main ingredients of the theory, because we will use them further on to study the convergence for our model problem (1).

The test system in [5], modelling the coupled evolution of a slow variable $r \in \mathbb{R}$ and a fast variable $v \in \mathbb{R}^p$, $p \geq 1$, has the following structure:

$$\begin{bmatrix} \dot{r} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} a & p^T \\ q/\epsilon & -A/\epsilon \end{bmatrix} \begin{bmatrix} r \\ v \end{bmatrix} \tag{9}$$

where $A \in \imath \times \imath$ has positive eigenvalues: the fast component v is dissipative. The model for the approximate slow variable U, and the parameter Λ, are defined as follows:

$$\dot{U} = \Lambda U = \left(a + p^T A^{-1} q\right) U, \tag{10}$$

with $U(0) = U_0 = r_0$. In [5, equations (2.8), (2.13) and (2.14)] the following properties of the multiscale system (9) and its reduced model (10) are proven (the subscript $\cdot_0$ denotes the initial condition):

$$\sup_{t \in [0,T]} |r(t) - U_0 \exp(\Lambda t)| \leq C\epsilon(|r_0| + \|v_0 - A^{-1} q r_0\|), \tag{11}$$

$$\sup_{t \in [0,T]} |r(t)| \leq C(|r_0| + \epsilon\|v_0\|), \tag{12}$$

$$\sup_{t \in [t_{\mathrm{BL}},T]} \|v(t)\| \leq C(|r_0| + \epsilon\|v_0\|), \tag{13}$$

where the constant C only depends only on A, p, q, a and T (see (9)).

Using the properties (11)–(13), in [5], the convergence of mM-Parareal for the linear test problem (9) with coarse model (10) is analysed, using the restriction operator $\mathcal{R}([r, v]^T) = r$ (with $\mathcal{R}^{\perp}([r, v]^T) = v$), the lifting operator $\mathcal{L}(U) = [U, A^{-1} q U]^T$ and the matching operator $\mathcal{M}(U, u) = [U, \mathcal{R}^{\perp} u]^T$. We now present two minor extensions to existing Micro-Macro Parareal convergence lemmas for later use.

Lemma 1 (Convergence of mM-Parareal for nonhomogenous linear ODEs) *The mM-Parareal solution of the system $\dot{u} = Au + b$ equals the mM-Parareal solution of the system $\dot{v} = Av$, with $v = u = A^{-1} b$, if $v(0)$ is chosen $v(0) = u(0) - A^{-1} b$, with A and b constant. Assume that the (numerical) fine propagator satisfies the following property when it is applied on a linear system: $\mathcal{F}(u) = (I + A_{\mathcal{F}}) u + B_{\mathcal{F}}$ with $B_{\mathcal{F}} = 0$ for the homogeneous system. (This assumption is not restrictive, e.g., it is satisfied by any Runge Kutta method.) Futher assume that $\mathcal{M}(\rho, u) - \mathcal{M}(\sigma, v) = \mathcal{M}(\rho - \sigma, u - v)$ and that the coarse propagator is linear. Then, the mM-Parareal iterates satisfy*

$$u_n^k = v_n^k + A^{-1} B \tag{14}$$

The proof of Lemma 1 can be constructed by induction on n.

Lemma 2 (Convergence of mM-Parareal without lifting in the zeroth iteration for linear scale-separated ODEs)

Using trivial lifting, that is $\mathcal{L}(X) = [X, v_0]$*, and using mM-Parareal, defined in* (7)–(8)*, with the specific choice of operators* (9)–(10)*, let* $E^k_{\Sigma_X,n} = u^k_n - \mathcal{R}u_n$ *be the macro error and* $e^k_n = u^k_n - u_n$ *be the micro error. Then, there exists* $\epsilon_0 \in (0,1)$*, that only depends on* α*,* p*,* q*,* A *and* T*, such that, for all* $\epsilon < \epsilon_0$ *and all* $\Delta t > t^{BL}_\epsilon$*, there exists a constant* C_k*, independent of* ϵ*, such that for all* $k \geq 0$*:*

$$\sup_{0\leq n\leq N} |E^k_n| \leq C_k \epsilon^{1+\lfloor (k+1)/2 \rfloor}, \tag{15}$$

$$\sup_{0\leq n\leq N} \|e^k_n\| \leq C_k \epsilon^{\lceil (k+1)/2 \rceil}. \tag{16}$$

The proof of Lemma 2 closely follows [5, proof of Theorem 13].

3 Convergence of Micro-Macro Parareal for model problem: theoretical analysis

The model problem is the multiscale Ornstein-Uhlenbeck process (1). We define the micro variable, describing the first two moments of its solution as $u^k_n = [m_x\ m_y\ \Sigma_x\ \Sigma_{xy}\ \Sigma_y]$. The macro variable is defined as $\rho^k_n = [m_x\ \Sigma_x]$.

For the fine propagator $\mathcal{F}$, we use the SDE (1), which we model via its moment models (5) and (6). The coarse propagator $\mathcal{C}$ simulates the reduced system (2), or equivalently the scalar ODEs (6) and (4). The restriction operator is defined as $\mathcal{R}\left([m_x\ m_y\ \Sigma_x\ \Sigma_y\ \Sigma_{xy}]\right) = [m_x\ \Sigma_x]$, the lifting operator as: $\mathcal{L}\left([M_X\ S_X]^T\right) = [M_X\ m_{y,0}\ S_X\ \Sigma_{y,0}\ \Sigma_{xy,0}]^T$, and the matching operator as $\mathcal{M}\left([M_X\ S_X]^T, [m_x\ m_q\ \Sigma_x\ \Sigma_y\ \Sigma_{xy}]^T\right) = [M_X\ m_y\ S_X\ \Sigma_y\ \Sigma_{xy}]^T$. The lifting operator thus initializes the moments of the fast variable to its initial value.

Convergence of first moment. The moment equations (5) and (3), describing the evolution of the first moment obey the structure of the multiscale system (9), and therefore we can, after using Lemma 1, apply Lemma 2.

Converence of covariance. The evolution of the multiscale covariance (6) does not satisfy the same property as the model in equation (9) because (i) the submatrix A_Σ contains the parameter ϵ, and (ii) the reduced model is not defined using (10). Next we will prove that, although the models (6) and (9) are different, they both satisfy some key theoretical properties that were used in [5].

Lemma 3 (An equivalent of (11) **for model** (6) **instead of model** (9)**)** *For system* (6) *and its reduced model* (4)*, it holds true that*

$$\sup_{t\in[0,T]} |\Sigma_x(t) - \Sigma_{x,0}\exp(\Lambda_\Sigma t)| \leq C\epsilon(|\Sigma_{x,0}| + \|\Sigma_{y,0} - A_\Sigma^{-1} q_\Sigma \Sigma_{x,0}\|). \tag{17}$$

Proof From (11), (6) and (19), we have

$$\sup_{t\in[0,T]} |\Sigma_x(t) - \Sigma_{x,0}\exp(\lambda_\Sigma t)| \le C\epsilon(|\Sigma_{x,0}| + \|\Sigma_{z,0}\|). \tag{18}$$

If we define

$$\lambda_\Sigma = 2\alpha + p_\Sigma^T A_\Sigma^{-1} q_\Sigma, \tag{19}$$

we can interpret the averaged model (4) as a limit of the reduced model (10) for the system (6): $\Lambda_\Sigma = \lim_{\epsilon\to 0} \lambda_\Sigma = 2\alpha - 2\frac{\beta\gamma}{\zeta}$. Now we define $\Delta\Lambda_\Sigma = \Lambda_\Sigma - \lambda_\Sigma$ (see (4) and (19)) and we observe that $\Delta\Lambda_\Sigma = \mathcal{O}(\epsilon)$. It then holds that $\exp(\Lambda_\Sigma t) = \exp((\lambda_\Sigma + \Delta\Lambda_\Sigma)t) = \exp(\lambda_\Sigma t)\,[1 + \mathcal{O}(\epsilon)]$. From the triangle inequality and the inequality (11) we have that

$$\begin{aligned}\sup_{t\in[0,T]} |\Sigma_x(t) - \Sigma_{x,0}\exp(\Lambda_\Sigma t)| &\le \sup_{t\in[0,T]} |\Sigma_x(t) - \Sigma_{x,0}\exp(\lambda_\Sigma t)(1+\mathcal{O}(\epsilon))| \\ &\le C\epsilon(|\Sigma_{x,0}| + \|\Sigma_{z,0}\|) + |\Sigma_{x,0}||\mathcal{O}(\epsilon)| \\ &\le K\epsilon(|\Sigma_{x,0}| + \|\Sigma_{z,0}\|),\end{aligned} \tag{20}$$

where $K > C$. This proves equation (17). □

Lemma 4 (An equivalent of (12) **and** (13) **for model** (6) **instead of model** (9)**)** *Assuming that that the eigenvalues $\mu_{\Sigma,i}$ of the matrix A_Σ (see* (6)*) are all positive, the properties in equation* (12) *and* (13) *hold true for the system* (6)*:*

$$\begin{aligned}\sup_{t\in[0,T]} |\Sigma_x(t)| &\le C(|\Sigma_{x,0}| + \epsilon\|\Sigma_{q,0}\|), \\ \sup_{t\in[t_{BL},T]} \|\Sigma_q(t)\| &\le C(|\Sigma_{x,0}| + \epsilon\|\Sigma_{q,0}\|).\end{aligned} \tag{21}$$

Proof The proof is similar to [5, Proof of Corollary 3]. In [5], the assumption that the eigenvalues of A_Σ are all positive is important. The structure of λ_Σ (or Λ_Σ) does not further influence the proof. □

The preceding lemmas allow us to formulate our main result.

Lemma 5 (Convergence of mM-Parareal for evolution of covariance) *Consider mM-Parareal, defined in* (7)–(8)*, with fine and coarse propagators the full system* (6) *and the reduced system* (4)*, respectively. Let $E^k_{\Sigma_X,n} = \rho^k_n - \mathcal{R}u_n$ be the macro error and $e^k_n = u^k_n - u_n$ be the micro error. Then there exists $\epsilon_0 \in (0,1)$, that only depends on α, p_Σ, q_Σ, A_Σ and T, such that, for all $\epsilon < \epsilon_0$ and all $\Delta t > t^{BL}_\epsilon$, there exists a constant C_k, independent of ϵ, such that for all $k \ge 0$:*

$$\sup_{0\le n\le N} |E^k_n| \le C_k \epsilon^{1+\lfloor (k+1)/2 \rfloor} \tag{22}$$

$$\sup_{0\le n\le N} \|e^k_n\| \le C_k \epsilon^{\lceil (k+1)/2 \rceil} \tag{23}$$

Proof Using Lemmas 1, 2, 3, and 4 the proof follows from [5, Proof of Theorem 2]. □

4 Numerical experiments

The test parameters for the numerical experiments are chosen to be:

$$\begin{bmatrix} \alpha & \beta \\ \gamma/\epsilon & \zeta/\epsilon \end{bmatrix} = \begin{bmatrix} -1. & -1. \\ 0.1/\epsilon & -1./\epsilon \end{bmatrix}, \qquad \sigma = 0.5 \tag{24}$$

The time interval is chosen as $[0, 10]$, the number of time intervals $N = 10$, and the initial value $\left[m_{x,0}\ m_{q,0}\ \Sigma_{x,0}\ \Sigma_{q,0}\ \Sigma_{xq,0}\right]^T = \left[100\ 100\ 100\ 0\ 0\right]^T$. In the experiments, which are shown in Figure 1, it is seen that the micro and macro errors in the mean follow the behaviour given by Lemma 2; those in the variance follow the behaviour as given by Lemma 5. Observe that mM-Parareal converges faster for computationally more expensive models (with small ϵ).

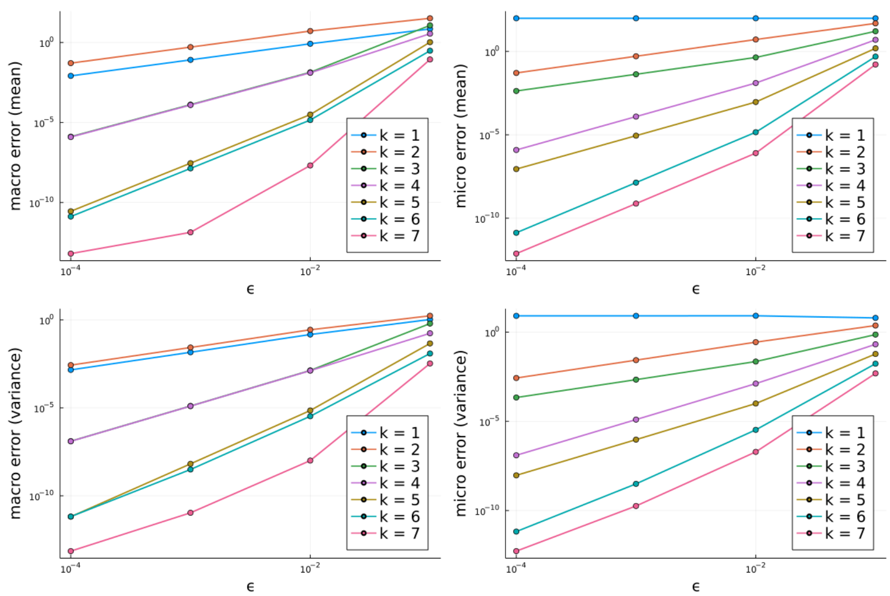

Fig. 1 Error as function of time-scale separation parameter ϵ. We used ∞-norm over time (only considering coarse discretisation points) and the 2-norm for the micro error. **Top left**: macro error on mean, **Top right**: micro error on mean, **Bottom left**: macro error on variance, **Bottom right**; micro error on variance. We used a numerical solver to discretise the moment equations (3)–(6) with a very stringent tolerance, so that the effect of numerical discretisation errors can be neglected.

5 Discussion and conclusion

Summary. We presented a convergence analysis of the Micro-Macro Parareal algorithm on scale-separated Ornstein-Uhlenbeck SDEs. We analysed its convergence behaviour w.r.t. the time scale separation parameter ϵ, using moment models. The convergence of the first moment is closely related to the analysis in [5]. For the covariance we presented some extensions to this theory.

Limitations. While the analysis using moment models quantifies the error on the mean and variance of the SDE solution, we cannot say anything about other quantities of interest, such as higher moments of the SDE solutions.

Also, by using the moment model (an ODE that we solved using very stringent tolerances), we exclusively looked at the model error, neglecting the discretisation errors and statistical errors (in e.g. Monte Carlo simulations) that arise in the discretisation of an SDE.

Open questions. It remains to be studied how the analysis generalises to higher dimensions, for instance when the slow variable is multi-dimensional. Also, an extension of the convergence analysis could cover nonlinear SDEs, or linear SDEs for which there is a coupling between mean and variance in the moment model ODEs. Another open problem is an analysis of convergence of the method w.r.t. the iteration number, in contrast to convergence w.r.t. the parameter ϵ. This would be more useful in practice.

Software. The code that is used for the numerical experiments, is available[1]. We used the Julia language [2] and the DifferentialEquations.jl package [8].

References

1. Arnold, L. *Stochastic differential equations: theory and applications*. Wiley, New York (1974).
2. Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. Julia: A Fresh Approach to Numerical Computing. *SIAM Review* **59**(1), 65–98 (2017).
3. Gardiner, C. W. *Handbook of stochastic methods for physics, chemistry, and the natural sciences*. No. v. 13 in Springer series in synergetics. Springer-Verlag, Berlin ; New York (1983).
4. Legoll, F., Lelièvre, T., Myerscough, K., and Samaey, G. Parareal computation of stochastic differential equations with time-scale separation: a numerical convergence study. *Comput. Vis. Sci.* **23**(1-4), Paper No. 9, 18 (2020).
5. Legoll, F., Lelièvre, T., and Samaey, G. A micro-macro parareal algorithm: application to singularly perturbed ordinary differential equations. *SIAM J. Sci. Comput.* **35**(4), A1951–A1986 (2013).
6. Lions, J.-L., Maday, Y., and Turinici, G. Résolution d'EDP par un schéma en temps "pararéel". *C. R. Acad. Sci. Paris Sér. I Math.* **332**(7), 661–668 (2001).
7. Pavliotis, G. A. and Stuart, A. M. *Multiscale methods: Averaging and homogenization*, *Texts in Applied Mathematics*, vol. 53. Springer, New York (2008).
8. Rackauckas, C. and Nie, Q. DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. *Journal of Open Research Software* **5**(1), 15 (2017).
9. Uhlenbeck, G. E. and Ornstein, L. S. On the Theory of the Brownian Motion. *Physical Review* **36**(5), 823–841 (1930).

[1] https://gitlab.kuleuven.be/numa/public/mm-parareal-convergence-sde