Coupling heterogeneous multiscale FEM with Runge-Kutta methods for parabolic homogenization problems: a fully discrete space-time analysis

Assyr Abdulle, Gilles Vilmart

Coupling heterogeneous multiscale FEM with Runge-Kutta methods for parabolic homogenization problems: a fully discrete space-time analysis

Assyr Abdulle[1] and Gilles Vilmart[1,2]

September 7, 2011

Abstract

Numerical methods for parabolic homogenization problems combining finite element methods (FEMs) in space with Runge-Kutta methods in time are proposed. The space discretization is based on the coupling of macro and micro finite element methods following the framework of the Heterogeneous Multiscale Method (HMM). We present a fully-discrete analysis in both space and time. Our analysis relies on new (optimal) error bounds in the norms $L^2(H^1)$, $\mathcal{C}^0(L^2)$, and $\mathcal{C}^0(H^1)$ for the fully discrete analysis in space. These bounds can then be used to derive fully discrete space-time error estimates for a variety of Runge-Kutta methods, including implicit methods (e.g., Radau methods) and explicit stabilized method (e.g., Chebyshev methods). Numerical experiments confirm our theoretical convergence rates and illustrate the performance of the methods.

Keywords: multiple scales, fully discrete, numerical homogenization, finite elements, Runge-Kutta methods, Chebychev methods, parabolic problems

AMS subject classification (2010): 65N30, 65M60 (74Q10, 35K15).

1 Introduction

A wide range of applications such as thermal diffusion in composite materials or water infiltration in porous soil are modeled by parabolic homogenization problems. In such problems, microscopic heterogeneities are assumed to occur at a much smaller length scale than the scale of interest, typically at a macroscopic level. Classical numerical methods for such problems need grid resolution to resolve the finest scale and become therefore quickly inefficient due to the very large systems arising with such discretization. Yet, when there is a separation of scales between the microscopic length scale and the macroscopic scale, numerical methods based on macro and micro solvers can be efficient.

Numerous numerical methods based on macro and micro solvers for the solution of partial differential equations (PDEs) with multiple scales have been developed the past few years. For parabolic problems we mention [6, 26, 38, 41] (we note that nonlinear and stochastic problems are considered in [38] and [26]). In this paper we consider the framework of the heterogeneous multiscale method (HMM) introduced in [22]. Finite difference and finite

[1]Mathematics Section, École Polytechnique Fédérale de Lausanne, Station 8, CH-1015 Lausanne, Switzerland, Assyr.Abdulle@epfl.ch

[2]new address: École Normale Supérieure de Cachan, Antenne de Bretagne, av. Robert Schuman, F-35170 Bruz, France, Gilles.Vilmart@bretagne.ens-cachan.fr

element methods for parabolic homogenization problems in the HMM framework have been developed in [6] and [38], respectively.

From a computational perspective, it is important to balance time and space discretization to obtain a desired accuracy while minimizing the computational cost. Yet for multiscale problems it is important to further balance the discretization effort at the macro and micro levels involved in the spatial discretization. This has first been addressed in [2] for the HMM applied to linear elliptic problems (see also [3] for a review), in [28] for nonlinear elliptic problems of monotone type, and in [11] for nonlinear elliptic problems of nonmonotone type. However none of the existing work for parabolic homogenization problems [6, 26, 41, 38] discuss a fully discrete discretization in space.

Furthermore, while implicit solvers are usually chosen to integrate parabolic problems in time, it is sometimes (in particular for large systems) desirable to avoid the use of linear algebra for solving their implicit part. It is known that the (severe) CFL constraint leading to (unacceptable) stepsize restriction for classical explicit solvers can be overcome by using stabilized (Chebyshev) methods [7, 1, 31, 32]. While remaining fully explicit and straightforward to implement, Chebyshev methods free the explicit schemes from stepsize restriction, provided an appropriate selection of the number of internal stages of the methods. As the analysis presented in [6, 26, 38, 41] proceeds directly with a discrete scheme in space and time, it is tailored to the specific numerical integrator used in time (usually the implicit Euler scheme) and cannot be straightforwardly generalized to classes of implicit methods or to Chebyshev methods.

The aim of this paper is to give a fully discrete analysis in both space and time for a variety of time integration methods including the family of (implicit) Radau type methods as well as the family of (explicit) Chebyshev methods. Our analysis proceeds in two steps. First, we derive new (optimal) error bounds in the $L^2(H^1)$, $\mathcal{C}^0(L^2)$, and $\mathcal{C}^0(H^1)$ for the fully discrete method in space. These error bounds, involving macro and micro meshsizes, allow to set the optimal rate of the micro mesh refinement in function of the macro mesh in order to have the most accurate output with the smallest computational effort. Then, fully discrete space-time error estimates are obtained in a second step for several classes of time integration methods as described previously. The fully discrete error bounds rely on analytic semigroups techniques in a Hilbert space framework, following [37]. The fully discrete bounds in space, derived in the first part of the paper, are instrumental for such an analysis. Results of extensive numerical simulations are discussed for both periodic and nonperiodic (e.g. random) coefficients, linear and nonlinear problems. While our analysis focuses on linear problems, numerical tests indicate that the convergence rates derived in this paper are also valid for a class of nonlinear problems.

The paper is organized as follows: In section 2 we introduce the homogenization problem for parabolic problems and recall some analytical background. In section 3 we introduce the FEM for the multiscale parabolic equation. In section 4 we present the analysis of the method. Numerical experiments for various type of oscillating coefficients and for both linear and nonlinear problems are given in section 5.

Notations. In what follows we consider the usual Sobolev space $W^{s,p}(\Omega)$. For $p = 2$ we use the notation $H^s(\Omega)$ and $H^1_0(\Omega)$, and we denote $W^1_{per}(Y) = \{v \in H^1_{per}(Y);\ \int_Y v dx = 0\}$, where $H^s_{per}(Y)$ is defined as the closure of $\mathcal{C}^\infty_{per}(Y)$ (the subset of $\mathcal{C}^\infty(\mathbb{R}^d)$ of periodic functions in the unit cube $Y = (0,1)^d$) for the H^s norm. For a domain $D \subset \Omega$, $|D|$ denotes the measure of D. We will sometimes use the notation $\partial_t, \partial_{tt}, \ldots$ or alternatively ∂_t^k for $\frac{\partial}{\partial t}, \frac{\partial^2}{\partial t^2}, \ldots$. For a Banach space X with norm $\|\cdot\|_X$, we denote by $L^p(0,T;X) = L^p(X)$, $1 \le p < \infty$ the

space of functions $t \to v(t)$ which are L^p on $(0,T)$ with values in X. Equipped with the norm $\|v\|_{L^p(0,T;X)} = \left(\int_0^T \|v(t)\|_X^p dt\right)^{1/p}$, the space $L^p(X)$ is a Banach space. In addition, we denote $\mathcal{C}^0([0,T],X)$ the Banach space of continuous functions $t \to v(t)$ with values in X, equipped with the norm $\|v\|_{\mathcal{C}^0([0,T],X)} = \sup_{t\in[0,T]} \|v(t)\|_X$.

2 Model Problem

We consider a bounded, convex and polyhedral domain Ω in $\mathbb{R}^d$. Let $T > 0$. The class of problems considered in this paper are the following parabolic problems

$$\begin{aligned} &\partial_t u_\varepsilon - \nabla \cdot (a^\varepsilon \nabla u_\varepsilon) = f \ \text{ in } \Omega \times (0,T) \\ &u_\varepsilon = 0 \ \text{ on } (0,T) \times \partial\Omega \\ &u_\varepsilon(x,0) = g(x) \ \text{ in } \ \Omega, \end{aligned} \tag{1}$$

where $a^\varepsilon(x,t)$ satisfies $a^\varepsilon \in (L^\infty(\Omega \times (0,T)))^{d\times d}$ and is uniformly elliptic and bounded, i.e.,

$$\begin{aligned} &\exists \lambda, \Lambda > 0 \text{ such that } \lambda|\xi|^2 \le a^\varepsilon(x,t)\xi\cdot\xi, \ \ \|a^\varepsilon(x,t)\xi\| \le \Lambda\|\xi\|, \\ &\forall \xi \in \mathbb{R}^d \text{ and a.e. } x \in \Omega, t \in (0,T), \forall \varepsilon > 0. \end{aligned} \tag{2}$$

The homogeneous Dirichlet boundary conditions are taken here for simplicity and our analysis applies to other types of boundary conditions (e.g. Neumann, mixte, etc.). Here and in what follows, ε represents a small scale in the problem that characterizes the multiscale nature of the tensor $a^\varepsilon(x)$. We make the following assumption on the parameters of the problem:

$$f \in L^2(0,T;L^2(\Omega)), \quad g \in L^2(\Omega), \tag{3}$$

and we consider the following Banach space

$$E = \{v; \ v \in L^2(0,T;H_0^1(\Omega)), \ \partial_t v \in L^2(0,T;H^{-1}(\Omega)\} \tag{4}$$

with the norm

$$\|v\|_E = \|v\|_{L^2(0,T;H^1(\Omega))} + \|\partial_t v\|_{L^2(0,T;H^{-1}(\Omega))}. \tag{5}$$

Under the assumption (3), the weak form of (1) has a unique solution $u_\varepsilon \in E$ (see for example [36]). As we have the continuous inclusion $E \subset \mathcal{C}^0([0,T];L^2(\Omega))$ (see [27, Sect. 5.9.2] or [36, Chap. 1, Thm. 3.1]), the solution u_ε satisfies $u_\varepsilon \in \mathcal{C}^0([0,T];L^2(\Omega))$. Moreover the following a priori estimate holds

$$\|u_\varepsilon\|_E + \|u_\varepsilon\|_{\mathcal{C}^0([0,T];L^2(\Omega))} \le C(\|f\|_{L^2(0,T;L^2(\Omega))} + \|g\|_{L^2(\Omega)}). \tag{6}$$

Consider now the family of problems (1) (indexed by ε) and the family of corresponding solutions $\{u_\varepsilon\}$. The estimate (6) together with standard compactness results ensure the existence of a subsequence of $\{u_\varepsilon\}$ (still denoted by ε) such that

$$u_\varepsilon \rightharpoonup u_0 \text{ weakly in } L^2(0,T;H_0^1(\Omega)), \ \partial_t u_\varepsilon \rightharpoonup \partial_t u_0 \text{ weakly in } L^2(0,T;H^{-1}(\Omega)). \tag{7}$$

The next task is to find a limiting equation for the solution u_0. Asymptotic expansions can be used to find a candidate for such a limiting equation. To show that the solution of this latter equation is the limit (in some sense) of the oscillating family of functions $\{u_\varepsilon\}$, one uses usually the notion of G or H convergence (see [42, 19, 39]), the former being restricted

to symmetric tensors. It is then possible to show that there exist a subsequence of $\{u_\varepsilon\}$ (still denoted by ε) satisfying (7) and such that u_0 is the solution of a so-called homogenized parabolic problem

$$\begin{aligned} &\partial_t u_0 - \nabla \cdot \left(a^0 \nabla u_0\right) = f \ \text{ in } \Omega \times (0,T) \\ &u_0 = 0 \ \text{ on } (0,T) \times \partial\Omega \\ &u_0(x,0) = g(x) \ \text{ in } \ \Omega, \end{aligned} \tag{8}$$

where the so-called homogenized tensor a^0 again satisfies (2), possibly with different constants. We refer to [12, 15, 20] for details. Let us note that if $a^\varepsilon(x,t)$ has more structure as for example if $a^\varepsilon(x,t) = a(x,x/\varepsilon,t) = a(x,y,t)$ is Y-periodic in y then one can show that the whole sequence $\{u_\varepsilon\}$ weakly converges in the sense described above.

The discretization of the problem (1) with FEM is a well understood problem. Taking a finite dimensional subspace $S(\Omega,\mathcal{T}_h)$ of the Banach space (4), we search for a piecewise polynomial solution $u^h(t) \in S(\Omega,\mathcal{T}_h)$ of the variational equation corresponding to (1) in $S(\Omega,\mathcal{T}_h)$. However, the major issue is that usual convergence rates can only be obtained if the spatial meshsize h satisfies $h < \varepsilon$. For multiscale problems with order of magnitude of discrepancies between the scale of interest (for which we would like to set the spatial grid) and ε, the restriction $h < \varepsilon$ can be prohibitive in terms of degrees of freedom of the computational procedure if not impossible to realize.

The idea of the multiscale method described in the next section is to rely on two grids, and in turn on two FE methods. A macroscopic method relying on a macroscopic mesh $H > \varepsilon$ which does not discretize the fine scale and a microscopic mesh (defined on sampling domains within the macroscopic mesh) which discretizes the smallest scale. Proper averaging of the microscopic FEM on the sampling domains allows to recover macroscopic (averaged) data related to the homogenized problem whose coefficients are unknown beforehand.

3 Multiscale method for parabolic homogenization problems

In this section we describe the numerical method under study. It is based on the finite element heterogeneous multiscale methods, introduced an analyzed in [22, 2, 23] (see [3, 4] for a review). In the HMM context, two approaches for the (spatial) numerical homogenization of parabolic problems have been proposed. The method in [6] is based on finite difference discretization techniques while the method in [38] is based on finite element discretization techniques.

We first describe the spatial discretization in Section 3.1 based on the HMM framework, using macro and micro FEM. We then discuss the time-discretization in Section 3.2

3.1 Spatial discretization

The numerical method is based on a macro and a micro FEM. We denote by $\mathcal{T}_H$ a family of (macro) partitions[1] of Ω in simplicial or quadrilateral elements K. The diameter of an element $K \in \mathcal{T}_H$ is denoted by H_K and we set $H = \max_{K\in\mathcal{T}_H} H_K$. We then consider a macro FE space

$$S_0^\ell(\Omega,\mathcal{T}_H) = \{v^H \in H_0^1(\Omega);\ v^H|_K \in \mathcal{R}^\ell(K),\ \forall K \in \mathcal{T}_H\}, \tag{9}$$

[1] By macro partition we mean that $H \gg \varepsilon$ is allowed.

where $\mathcal{R}^\ell(K)$ is the space $\mathcal{P}^\ell(K)$ of polynomials on K of total degree at most ℓ if K is a simplicial FE, or the space $\mathcal{Q}^\ell(K)$ of polynomials on K of degree at most ℓ in each variables if K is a rectangular FE. Crucial to the FE-HMM is the definition of a quadrature formula on the elements K. We thus consider within each macro element $K \in \mathcal{T}_H$ for $j = 1, \ldots, J$,

- quadrature points $x_{K_j} \in K$,
- quadrature weights ω_{K_j}.

Quadrature formula. A proper choice must ensure the coercivity of the bilinear form (16). Furthermore, minimizing the number of quadrature points is crucial as each quadrature point involves the solution of a boundary value problem (the micro problem). We consider $\hat{K}$ the reference element and for every element of the triangulation the mapping F_K (a $\mathcal{C}^1$-diffeomorphism) such that $K = F_K(\hat{K})$. For every K we consider the quadrature formula $\{x_{K_j}, \omega_{K_j}\}_{j=1}^J$ and we suppose that the quadrature weights and integration points are given by $x_{K_j} = F_K(\hat{x}_j),\ \omega_{K_j} = \hat{\omega}_j |\det(\partial F_K)|, j = 1, \ldots, J$, where $\{\hat{x}_j, \hat{\omega}_j\}_{j=1}^J$ is a quadrature formula on $\hat{K}$. We will make the following assumption on the quadrature formula

(**Q1**) $\hat{\omega}_j > 0,\ j = 1, \ldots, J, \quad \sum_{j=1}^J \hat{\omega}_j |\nabla \hat{p}(\hat{x}_j)|^2 \geq \hat{\lambda} \|\nabla \hat{p}\|^2_{L^2(\hat{K})},\ \forall \hat{p}(\hat{x}) \in \mathcal{R}^\ell(\hat{K}),$ with $\hat{\lambda} > 0$;

(**Q2**) $\int_{\hat{K}} \hat{p}(\hat{x}) d\hat{x} = \sum_{j \in J} \hat{\omega}_j \hat{p}(\hat{x}_j),\ \forall \hat{p}(\hat{x}) \in \mathcal{R}^\sigma(\hat{K})$, where $\sigma = \max(2\ell - 2, \ell)$ if $\hat{K}$ is a simplicial FE, or $\sigma = \max(2\ell - 1, \ell + 1)$ if $\hat{K}$ is a rectangular FE.

These requirements on the quadrature formula ensure that the optimal convergence rates for elliptic FEM hold when using numerical integration [17].

Based on the quadrature points, we define sampling domains

$$K_{\delta_j} = x_{K_j} + \delta I,\ I = (-1/2, 1/2)^d \quad (\delta \geq \varepsilon), \tag{10}$$

the "domains" of the micro FE method. We consider a (micro) partition $\mathcal{T}_h$ of each sampling domain K_{δ_j} in simplicial or quadrilateral elements Q of diameter h_Q and denote $h = \max_{Q \in \mathcal{T}_h} h_Q$. The sampling domains K_{δ_j} are typically of size ε, that is $|K_{\delta_j}| = \mathcal{O}(\varepsilon^d)$, and hence $h < \varepsilon \leq \delta$ holds for the micro mesh size. For this partition we define a micro FE space

$$S^q(K_{\delta_j}, \mathcal{T}_h) = \{z^h \in W(K_{\delta_j});\ z^h|_T, \in \mathcal{R}^q(Q),\ Q \in \mathcal{T}_h\}, \tag{11}$$

where $W(K_{\delta_j})$ is a certain Sobolev space. Various spaces can be chosen for the micro numerical method, and the choice of the particular space has important consequences in the numerical accuracy of the method as we will see below (this choice sets the coupling condition between macro and micro solvers). We consider here

$$W(K_{\delta_j}) = W^1_{per}(K_{\delta_j}) = \{z \in H^1_{per}(K_{\delta_j});\ \int_{K_{\delta_j}} z dx = 0\}, \tag{12}$$

for a periodic coupling or

$$W(K_{\delta_j}) = H^1_0(K_{\delta_j}) \tag{13}$$

for a coupling through Dirichlet boundary conditions. Essential for the definition of the multiscale method below is the definition of the following microfunctions. Let $w^H \in S^\ell_0(\Omega, \mathcal{T}_H)$ and consider its linearization

$$w^H_{lin} = w^H(x_{K_j}) + (x - x_{K_j}) \cdot \nabla w^H(x_{K_j})$$

at the integration point x_{K_j}. Associated to w^H_{lin} in the sampling domain K_{δ_j} we define a microfunction $w^{h,t}_{K_j}$, depending on t, such that $(w^{h,t}_{K_j} - w^H_{lin}) \in S^q(K_{\delta_j}, \mathcal{T}_h)$ and

$$\int_{K_{\delta_j}} a^\varepsilon(x,t)\nabla w^{h,t}_{K_j} \cdot \nabla z^h dx = 0 \quad \forall z^h \in S^q(K_{\delta_j}, \mathcal{T}_h). \tag{14}$$

Multiscale method. Find $u^H \in [0,T] \times S^\ell_0(\Omega, \mathcal{T}_H) \to \mathbb{R}$ such that

$$\begin{aligned} &(\partial_t u^H, v^H) + B_H(t; u^H, v^H) = (f(t), v^H) \ \ \forall v^H \in S^\ell_0(\Omega, \mathcal{T}_H) \\ &u^H = 0 \ \text{ on } \partial\Omega \times (0,T) \\ &u^H(x,0) = u^H_0, \end{aligned} \tag{15}$$

where

$$B_H(t; u^H, v^H) := \sum_{K \in \mathcal{T}_H} \sum_{j=1}^{J} \frac{\omega_{K_j}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x,t)\nabla u^{h,t}_{K_j} \cdot \nabla v^{h,t}_{K_j} dx, \tag{16}$$

and $u^H_0 \in S^\ell_0(\Omega, \mathcal{T}_H)$ is chosen to approximate the exact initial condition g (see Remark 4.6 below). Here, $u^{h,t}_{K_j}, v^{h,t}_{K_j}$ are the solution of the microproblems (14) constrained by u^H_{lin}, v^H_{lin}, respectively. We emphasize that the above numerical method aims at capturing the homogenized solution of (8). A numerical corrector can be defined extending the known micro solutions u^h (defined in the sampling domains K_{δ_j}) on the whole macro element K. With the help of the numerical corrector, an approximation of the fine scale solution u^ε can be obtained (see [23],[3, Chap. 3.3.2]). For nonlinear monotone elliptic problems, the convergence of such reconstruction procedure has been proved in [24, 25] for the Multiscale Finite Element method (MsFEM) in the stochastic case, and in [28] for both the MsFEM and HMM in the general case of an arbitrary G-converging sequence.

3.2 Time discretization

Let U^H be the column vector of the coefficients of u^H in the basis $\{\phi_j\}_{j=1}^M$ of $S^\ell_0(\Omega, \mathcal{T}_H)$, and consider the stiffness matrices

$$K(t) = (B_H(t; \phi_i, \phi_j))_{i,j=1}^M \qquad M = ((\phi_i, \phi_j))_{i,j=1}^M.$$

This permits to rewrite (15) in the form of an ordinary differential equation

$$MU'^H(t) + K(t)U^H(t) = F^H(t), \qquad U(0) = U_0,$$

where $F^H(t)$ corresponds to the source term and is defined in the usual manner.

With these notations, (15) can be written as

$$U'^H(t) = F(t, U^H(t)), \qquad U(0) = U_0, \tag{17}$$

where $F(t, U^H) = -M^{-1}K(t)U^H + M^{-1}F^H(t)$.

The fully discrete space time method relies on the discretization of (17) by a (s-stage) Runge-Kutta method

$$U_{n+1} = U_n + \Delta t_n \sum_{j=1}^{s} b_j S_{nj}, \quad U_{ni} = U_n + \Delta t_n \sum_{j=1}^{s} \gamma_{ij} S_{nj}, \tag{18}$$

$$S_{ni} = F(t_n + c_i \Delta t_n, U_{ni}), \quad i = 1 \dots s. \tag{19}$$

where γ_{ij}, b_j, c_i with $i, j = 1 \dots s$ are the given coefficients of the method (with $\sum_{j=1}^{s} \gamma_{ij} = c_i$) and $t_n = \sum_{i=0}^{n} \Delta t_n$. We sometimes use the notations

$$\Gamma = (\gamma_{ij})_{i,j=1}^{s}, \quad b = (b_1, \dots, b_s)^T, \quad c = (c_1, \dots, c_s)^T = \Gamma \mathbf{1}, \quad \mathbf{1} = (1, \dots, 1)^T.$$

We recall that the order of the method is r if the following error holds after one step with the same initial conditions for both the exact and the numerical solution

$$U_1 - U(t_1) = \mathcal{O}\left(\Delta t^{r+1}\right), \text{ for } \Delta t \to 0,$$

for all sufficiently differentiable initial value systems of differential equations.

Two classes of time integrator will be considered in the analysis.

Implicit methods. We consider a subclass of Runge-Kutta methods which are of order r with stage order (the accuracy of the internal stages) $r - 1$, and which are strongly $A(\theta)$-stable with $0 \leq \theta \leq \pi/2$. This latter condition means that $I - z\Gamma$ is a nonsingular matrix in the sector $|\arg(-z)| \leq \theta$ and the stability function [2] $R(z) = 1 + zb^T(I - z\Gamma)^{-1}\mathbf{1}$ satisfies $|R(z)| < 1$ in $|\arg(-z)| \leq \theta$ (we refer to [31, Sect. IV.3,IV.15] for details on the stability concepts described here).

Notice that all s-stage Radau Runge-Kutta methods satisfy the above assumptions (with $\theta \geq \pi/2$) [31]. In particular, for $s = 1$, we retrieve the implicit Euler method

$$(M + \Delta t K(t_{n+1}))U_{n+1}^H = MU_n^H + F_H(t_{n+1}). \tag{20}$$

Our analysis for implicit methods covers variable time step methods, provided that the stepsize sequence $\{\Delta t_n\}_{0 \leq n \leq N-1}$ with $\Delta t_n = t_{n+1} - t_n > 0$ and $t_N = T$ satisfies for $C, c > 0$

$$\sum_{n=0}^{N-1} |\Delta t_{n+1}/\Delta t_n - 1| \leq C, \tag{21}$$

$$c\Delta t \leq \Delta t_n \leq \Delta t \qquad \text{for all } 0 \leq n \leq N-1. \tag{22}$$

Remark 3.1 *The condition* (22) *may appear restrictive. However, a finite subdivision of the interval* $[0, T]$ *into subintervals can be considered and* (22) *is required only on each of the subintervals (see [37, Sect. 5]). This permits to use stepsizes of different scales.*

Chebyshev methods. Chebyshev methods are a subclass of explicit Runge-Kutta methods with extended stability domains along the negative real axis. Their favorable stability properties originate from the fact that is is possible to construct s-stage methods with stability functions $R_s(z) = 1 + zb^T(I - z\Gamma)^{-1}\mathbf{1}$ satisfying

$$|R_s(x)| \leq 1 \quad \text{for } x \in [-L_s, 0] \tag{23}$$

with $L_s = Cs^2$, where the constant C depends on the order of the method. Such methods have been constructed for order up to $r = 4$ [7, 1, 32, 35]. As eigenvalue of a parabolic problem lie on the negative real axis, Chebyshev methods are suitable for such problems. For first order methods, we have

$$R_s(x) = T_s(1 + x/s^2), \tag{24}$$

where $T_s(\cdot)$ denotes the Chebyshev polynomial of degree s and $L_s = 2s^2$. The corresponding Runge-Kutta method can be efficiently implemented by using the induction relation of the

[2] We recall that the stability function of a Runge-Kutta method is the rational function $R(\Delta t\lambda) = R(z)$ obtained after applying the method over one step Δt to the scalar problem $dy/dt = \lambda y,\ y(0) = 1,\ \lambda \in \mathbb{C}$.

Chebyshev polynomials. Higher order methods based on induction relations can also be built [7, 1, 32]. We briefly explain why it can be advantageous to use such methods. Let ρ_H be the spectrum of the discretized parabolic problem (depending on the macro spatial meshsize H) and assume that the stepsize to achieve the desired accuracy is Δt. Integrating the parabolic problem with an Euler method requires to take a stepsize δt satisfying the CFL constraints $\delta t \leq 2/\rho_H$ and we thus need $\Delta t/\delta t \geq (\Delta t \rho_H)/2$ function evaluations (a measure of the numerical work). With a Chebyshev method of the same order with stability function given above, one can first choose a stepsize Δt and then choose an s-stage method such that $\Delta t \rho_H \leq 2s^2$. This sets the number of stages $s = \sqrt{(\Delta t \rho_H)/2}$, which represents the square root of the numerical work needed by the explicit Euler method (notice that for Chebyshev methods, there is one new function evaluation per stage). Chebyshev methods are usually used in a "damped" form, where the stability function satisfies the strong stability condition

$$\sup_{z \in [-L_s, -\gamma],\, s \geq 1} |R_s(z)| < 1, \quad \text{for all } \gamma > 0. \tag{25}$$

For instance by changing the function (24) to

$$R_s(z) = \frac{T_s(\omega_0 + \omega_1 z)}{T_s(\omega_0)}, \quad \text{with} \quad \omega_0 = 1 + \frac{\eta}{s^2}, \quad \omega_1 = \frac{T_s(\omega_0)}{T_s'(\omega_0)}. \tag{26}$$

The parameter η is called the damping parameter. Choosing any fixed $\eta > 0$ ensures that $R_s(z)$ satisfies (25). It also ensures that a strip around the negative real axis is enclosed in the stability domain $\mathcal{S} := \{z \in \mathbb{C}; |R_s(z)| \leq 1\}$. The growth on the negative real axis shrinks a bit but remains quadratic [32],[31, Chap. IV.2]. For the analysis, we shall also need that the stability functions are bounded in a neighbourhood of zero uniformly with respect to s, precisely, there exist $\gamma > 0$ and $C > 0$ such that

$$|R_s(z)| \leq C \text{ for all } |z| \leq \gamma \text{ and all } s. \tag{27}$$

This can be easily checked for the Chebyshev methods with stability functions (24), (26).

3.3 Preliminaries

In order to perform the analysis of the FE-HMM, it is convenient to introduce the following auxiliary bilinear form

$$B(t; v, w) = \int_\Omega a^0(x,t) \nabla v(x) \cdot \nabla w(x) dx, \quad \forall v, w \in H_0^1(\Omega), \tag{28}$$

where $a^0(x,t)$ is the homogenized tensor of (8). Consider also the associated bilinear form for standard FEM with numerical quadrature,

$$B_{0,H}(t; v^H, w^H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^{J} \omega_{K_j} a^0(x_{K_j}, t) \nabla v^H(x_{K_j}) \cdot \nabla w^H(x_{K_j}), \tag{29}$$

for all $v^H, w^H \in S_0^\ell(\Omega, \mathcal{T}_H)$. Of course, $a^0(x,t)$ is usually unknown, otherwise there is no need for a multiscale method. In order to define FEM with numerical quadrature (as in the above bilinear form), some regularity on the tensor $a^0(x,t)$ is needed. We suppose that

(**H1**) $\quad a^0_{ij}, \partial_t a^0_{ij} \in \mathcal{C}^0([0,T] \times \overline{K})$ for all $K \in \mathcal{T}_H$ for all $i, j = 1, \ldots, d$.

The following construction of a numerically homogenized tensor is usefull (see [4, 5] for details). Let $\mathbf{e}_i,\ i = 1,\dots,d$, denote the canonical basis of $\mathbb{R}^d$. For each $\mathbf{e}_i$ and each $t \in [0,T]$, we consider the following elliptic problem

$$\int_{K_{\delta_j}} a^\varepsilon(x,t)\nabla\psi^{i,h,t}_{K_j}(x)\cdot\nabla z^h(x)dx = -\int_{K_{\delta_j}} a^\varepsilon(x,t)\mathbf{e}_i\cdot\nabla z^h(x)dx,\ \forall z^h \in S^q(K_{\delta_j},\mathcal{T}_h), \quad (30)$$

where $S^q(K_{\delta_j},\mathcal{T}_h)$ is defined in (11) with either periodic or Dirichlet boundary conditions. We also consider the problem

$$\int_{K_{\delta_j}} a^\varepsilon(x,t)\nabla\psi^{i,t}_{K_j}(x)\cdot\nabla z(x)dx = -\int_{K_{\delta_j}} a^\varepsilon(x,t)\mathbf{e}_i\cdot\nabla z(x)dx,\ \forall z \in W(K_{\delta_j}), \quad (31)$$

where the Sobolev space $W(K_{\delta_j})$ is defined in (12) or (13). We then define two tensors

$$a^0_K(x_{K_j},t) := \frac{1}{|K_{\delta_j}|}\int_{K_{\delta_j}} a^\varepsilon(x,t)\left(I + J^T_{\psi^{h,t}_{K_j}(x)}\right)dx, \quad (32)$$

where $J_{\psi^{h,t}_{K_j}(x)}$ is a $d\times d$ matrix with entries $\left(J_{\psi^{h,t}_{K_j}(x)}\right)_{i\ell} = (\partial\psi^{i,h,t}_{K_j})/(\partial x_\ell)$ and

$$\bar{a}^0_K(x_{K_j},t) := \frac{1}{|K_{\delta_j}|}\int_{K_{\delta_j}} a^\varepsilon(x,t)\left(I + J^T_{\psi^{t}_{K_j}(x)}\right)dx, \quad (33)$$

where $J_{\psi^{t}_{K_j}(x)}$ is a $d\times d$ matrix with entries $\left(J_{\psi^{t}_{K_j}(x)}\right)_{ik} = (\partial\psi^{i,t}_{K_j})/(\partial x_k)$.

Using the numerically homogenized tensors (32) or and the results of [5] (see also Lemmas 11 and 12 of [4]) we obtain the following reformulation of the bilinear form $B_H(\cdot,\cdot)$ of (16) which will be useful for the analysis.

Lemma 3.2 *The bilinear form $B_H(\cdot,\cdot)$ defined in* (16) *can be written as*

$$B_H(t;v^H,w^H) = \sum_{K\in\mathcal{T}_H}\sum_{j=1}^{J}\omega_{K_j}a^0_K(x_{K_j},t)\nabla v^H(x_{K_j})\cdot\nabla w^H(x_{K_j}). \quad (34)$$

Using (33) we can also define a bilinear form useful for the subsequent analysis

$$\bar{B}_H(t;v^H,w^H) = \sum_{K\in\mathcal{T}_H}\sum_{j=1}^{J}\omega_{K_j}\bar{a}^0_K(x_{K_j},t)\nabla v^H(x_{K_j})\cdot\nabla w^H(x_{K_j}). \quad (35)$$

Solving the parabolic problem (15) with the bilinear form $\bar{B}_H$ amounts to neglecting the micro errors, as the micro functions in (31) are exact.

Coercivity and boundedness of the bilinear forms. We shall also consider the bilinear forms $B'(t;v,w)$, $B'_H(t;v^H,w^H)$ defined as the time differentiations of (28), (34), respectively,

$$\begin{aligned} B'(t;v,w) &= \int_\Omega \left(\partial_t a^0(x,t)\right)\nabla v(x)\cdot\nabla w(x)dx, \\ B'_H(t;v^H,w^H) &= \sum_{K\in\mathcal{T}_H}\sum_{j=1}^{J}\omega_{K_j}\left(\partial_t a^0_K(x_{K_j},t)\right)\nabla v^H(x_{K_j})\cdot\nabla w^H(x_{K_j}). \end{aligned}$$

In view of (2) and (**H1**), we have for some positive constants γ_1, γ_2,

$$\begin{aligned} &\gamma_1 \|v\|^2_{H^1(\Omega)} \leq B(v,v), \\ &|B(t;v,w)| + |B'(t;v,w)| \leq \gamma_2 \|v\|_{H^1(\Omega)} \|w\|_{H^1(\Omega)}, \end{aligned} \tag{36}$$

for all $v, w \in H^1_0(\Omega)$ and all t. Assuming that the quadrature formula satisfies (**Q1**) and that (2) and (**H1**) hold, then (29) is coercive and bounded (see [16]). It can also be shown that the bilinear forms $B_H(\cdot,\cdot)$ defined in (16) is (uniformly w.r. to ε) coercive and bounded (see [2],[23],[3, Sect. 3.3.1] for a proof). i.e.,

$$\gamma_1 \|v^H\|^2_{H^1(\Omega)} \leq B_H(t; v^H, v^H), \tag{37}$$

$$|B_H(t; v^H, w^H)| + |B'_H(t; v^H, w^H)| \leq \gamma_2 \|v^H\|_{H^1(\Omega)} \|w^H\|_{H^1(\Omega)}, \tag{38}$$

for all $v^H, w^H \in S^\ell_0(\Omega, \mathcal{T}_H)$ and all t. The same bounds (37),(38) can be shown for the bilinear form $\bar{B}_H(t; v^H, w^H)$. No spatial structure assumption (as e.g. periodicity, random homogeneity) on the small scale tensor a^ε is required to prove coercivity and boundedness (only (2) and (**Q1**) are needed).

We quantify the discrepancy between the bilinear forms $B_{0,H}(t;\cdot,\cdot)$ defined in (29) and $B_H(t;\cdot,\cdot)$ defined in (16). This will account for the error done at the microscale as well as the so-called modeling error, the error induced by artificial micro boundary conditions or non-optimal sampling of the micro structure. Consider the quantity

$$\begin{aligned} r_{HMM} \quad := \quad & \sup_{K \in \mathcal{T}_H, x_{K_j} \in K, t \in [0,T]} \|a^0(x_{K_j},t) - a^0_K(x_{K_j},t)\|_F \\ + \quad & \sup_{K \in \mathcal{T}_H, x_{K_j} \in K, t \in [0,T]} \|\partial_t a^0(x_{K_j},t) - \partial_t a^0_K(x_{K_j},t)\|_F, \end{aligned} \tag{39}$$

where $\|\cdot\|_F$ denotes the Frobenius norm[3]. Following the strategy developed in [5, 3, 4] for the error analysis, we can further decompose r_{HMM} into micro and modeling error terms as

$$\begin{aligned} r_{HMM} \quad = \quad & \underbrace{\sum_{k=0}^{1} \sup_{K \in \mathcal{T}_H, x_{K_j} \in K, t \in [0,T]} \|\partial_t^k a^0(x_{K_j},t) - \partial_t^k \bar{a}^0_K(x_{K_j},t)\|_F}_{r_{MOD}} \\ + \quad & \underbrace{\sum_{k=0}^{1} \sup_{K \in \mathcal{T}_H, x_{K_j} \in K, t \in [0,T]} \|\partial_t^k \bar{a}^0_K(x_{K_j},t) - \partial_t^k a^0_K(x_{K_j},t)\|_F}_{r_{MIC}}, \end{aligned} \tag{40}$$

where we have used the tensor (33).

Lemma 3.3 *Let $B_{0,H}(t;\cdot,\cdot)$ and $B_H(t;\cdot,\cdot)$ be the bilinear forms defined in* (29) *and* (16)*, respectively. Then we have*

$$\begin{aligned} &|B_{0,H}(t; v^H, w^H) - B_H(t; v^H, w^H)| + |B'_{0,H}(t; v^H, w^H) - B'_H(t; v^H, w^H)| \\ &\leq C r_{HMM} \|v^H\|_{H^1(\Omega)} \|w^H\|_{H^1(\Omega)} \end{aligned}$$

[3]The Frobenius norm of a matrix M is defined as $\|M\|_F = \sqrt{\text{trace}(M^T M)}$.

Proof. By using the continuity of the bilinear forms $B_{0,H}(\cdot,\cdot)$ and $B_H(\cdot,\cdot)$, we have using the Cauchy-Schwarz inequality,

$$
\begin{aligned}
&|B_{0,H}(t;v^H,w^H) - B_H(t;v^H,w^H)| \\
&\le \sum_{K\in\mathcal{T}_H} |\sum_{j=1}^{J} \omega_{K_j}(a^0(x_{K_j},t) - a^0_K(x_{K_j},t))\nabla v^H(x_{K_j})\nabla w^H(x_j)dx| \\
&\le \sup_{K\in\mathcal{T}_H, x_{K_j}\in K, t\in[0,T]} \|a^0(x_{K_j},t) - a^0_K(x_{K_j},t)\|_F \sqrt{\sum_{K\in\mathcal{T}_H}\sum_{j=1}^{J}\omega_{K_j}|\nabla v^H(x_{K_j})|^2} \\
&\cdot\sqrt{\sum_{K\in\mathcal{T}_H}\sum_{j=1}^{J}\omega_{K_j}|\nabla w^H(x_{K_j})|^2} \\
&\le C \sup_{K\in\mathcal{T}_H, x_{K_j}\in K, t\in[0,T]} \|a^0(x_{K_j},t) - a^0_K(x_{K_j},t)\|_F \|v^H\|_{H^1(\Omega)}\|w^H\|_{H^1(\Omega)}.
\end{aligned}
$$

We proceed similarly for $B'_{0,H}(v^H,w^H) - B'_H(v^H,w^H)$. □

The modeling and the micro error can be traced in the following lemma, whose proof follows the arguments of Lemma 3.3.

Lemma 3.4 *Let $B_{0,H}(t;\cdot,\cdot), B_H(t;\cdot,\cdot)$ and $\bar{B}_H(t;\cdot,\cdot)$ be the bilinear forms defined in* (29), (16), *and* (35), *respectively. Then we have for all $v^H, w^H \in S^\ell_0(\Omega,\mathcal{T}_H)$,*

$$
\begin{aligned}
&|B_{0,H}(t;v^H,w^H) - \bar{B}_H(t;v^H,w^H)| + |B'_{0,H}(t;v^H,w^H) - \bar{B}'_H(t;v^H,w^H)| \\
&\le C r_{MOD}\|v^H\|_{H^1(\Omega)}\|w^H\|_{H^1(\Omega)}, \\
&|\bar{B}_H(t;v^H,w^H) - B_H(t;v^H,w^H)| + |\bar{B}'_H(t;v^H,w^H) - B'_H(t;v^H,w^H)| \\
&\le C r_{MIC}\|v^H\|_{H^1(\Omega)}\|w^H\|_{H^1(\Omega)}.
\end{aligned}
$$

Standard estimates for bilinear forms with numerical quadrature. Consider the usual nodal interpolant $\mathcal{I}_H : C^0(\overline{\Omega}) \to S^\ell_0(\Omega,\mathcal{T}_H)$ onto the FE space $S^\ell_0(\Omega,\mathcal{T}_H)$ defined in (9). The following estimates are based on the Bramble-Hilbert lemma and have first been derived in [17, Thm. 4 and Thm. 5]. They will often be used in our analysis. Assuming (**Q2**) and the regularity assumptions of Theorem 4.1 (see next section), we have for all $v^H, w^H \in S^\ell_0(\Omega,\mathcal{T}_H)$ (where $\mu = 0$ or 1),

$$
\begin{aligned}
|B(t;v^H,w^H) - B_{0,H}(t;v^H,w^H)| &\le CH\|v^H\|_{H^1(\Omega)}\|w^H\|_{H^1(\Omega)}, &(41)\\
|B(t;\mathcal{I}_H u_0,w^H) - B_{0,H}(t;\mathcal{I}_H u_0,w^H)| &\le CH^\ell\|u_0(t)\|_{W^{\ell+1,p}(\Omega)}\|w^H\|_{H^1(\Omega)}, &(42)\\
|B(t;\mathcal{I}_H u_0,w^H) - B_{0,H}(t;\mathcal{I}_H u_0,w^H)| &\le CH^{\ell+\mu}\|u_0(t)\|_{W^{\ell+1,p}(\Omega)}\Big(\sum_{K\in\mathcal{T}_H}\|w^H\|^2_{H^2(K)}\Big)^{1/2}. &(43)
\end{aligned}
$$

An α-accretive operator. For the time-discretization analysis, we introduce for each time t the linear operator $A_H(t) : S^\ell_0(\Omega,\mathcal{T}_H) \to S^\ell_0(\Omega,\mathcal{T}_H)$ defined as

$$
(-A_H(t)v^H, w^H) = B_H(t;v^H,w^H), \quad \text{for all } v^H, w^H \in S^\ell_0(\Omega,\mathcal{T}_H), \tag{44}
$$

where B_H is the bilinear form defined in (16). Consider the sector in the complex plane

$$
S_\alpha = \{\rho e^{i\theta} \ ; \ \rho \ge 0, |\theta| \le \alpha\}.
$$

The operator A_H can be extended straightforwardly to a complex Hilbert space based on $S_0^\ell(\Omega, \mathcal{T}_H)$ equipped with the complex scalar product $(u, v) = \int_\Omega u(x)\overline{v}(x)dx$ which is an extension of the usual L^2 scalar product. It can be shown that $-A_H$ is a so-called α-accretive operator[4]: there exist $0 \leq \alpha \leq \pi/2$ and $C > 0$ such that for all $z \notin S_\alpha$, the operator $zI + A_H(t)$ is an isomorphism and

$$\|(zI + A_H(t))^{-1}\|_{L^2(\Omega)\to L^2(\Omega)} \leq \frac{1}{d(z, S_\alpha)} \qquad \text{for all } z \notin S_\alpha, \tag{45}$$

where $d(z, S_\alpha)$ is the distance between z and S_α. In general one can show $\alpha \leq \arccos(\gamma_1/\gamma_2)$ using (37),(38). In the case of a symmetric tensor, all the eigenvalues of A_H are real and located on the negative real axis of the complex plane, and one has simply $\alpha = 0$. The proof of (45) is omitted as this is a classical result for the time discretization of parabolic PDEs. More details can be found for instance in [18].

4 Main results

In this section we shall present our main results. We first analyze the spatial discretization errors, and then we focus on the time discretization errors.

4.1 Fully discrete results in space

Theorem 4.1 *Consider u_0, u^H the solutions of (8), (15), respectively. Let $\mu = 0$ or 1, $\ell \geq 1$ and $2 \leq p \leq \infty$ such that $\ell > d/p$. Assume (**Q1**),(**Q2**),(**H1**),(36),(37) and*

$$\begin{aligned} &u_0, \partial_t u_0 \in L^2(0, T; W^{\ell+1,p}(\Omega)), \\ &a^0_{ij}, \partial_t a^0_{ij} \in L^\infty(0, T; W^{\ell+\mu,\infty}(\Omega)), \quad \forall i, j = 1 \dots d. \end{aligned}$$

Then we have the $L^2(H^1)$ and $\mathcal{C}^0(L^2)$ estimates

$$\|u_0 - u^H\|_{L^2([0,T];H^1(\Omega))} \leq C(H^\ell + r_{HMM} + \|g - u_0^H\|_{L^2(\Omega)}), \tag{46}$$

$$\|u_0 - u^H\|_{\mathcal{C}^0([0,T];L^2(\Omega))} \leq C(H^{\ell+1} + r_{HMM} + \|g - u_0^H\|_{L^2(\Omega)}), \quad \text{if } \mu = 1. \tag{47}$$

If in addition, the tensor is symmetric, then we have the $\mathcal{C}^0(H^1)$ estimate

$$\|u_0 - u^H\|_{\mathcal{C}^0([0,T];H^1(\Omega))} \leq C(H^\ell + r_{HMM} + \|g - u_0^H\|_{H^1(\Omega)}). \tag{48}$$

The constants C are independent of H, r_{HMM}.

The first term in the right-hand side of the above estimates quantifies the error of the macro solver. It shows that the proposed multiscale FEM gives optimal (macroscopic) convergence rates in the $\mathcal{C}^0(L^2)$ and $L^2(H^1)$ norms (and $\mathcal{C}^0(H^1)$ for symmetric tensors) of the fully discrete FE-HMM (15). We emphasize that the above error estimates have been derived without specific assumptions on the oscillation of the multiscale tensor. We recall that the additional term r_{HMM} defined in (39), that appears in the right-hand side of (46) or (47), encodes the so-called modeling and micro error, i.e., the error due to a possible mismatch of the averaging procedure in the FE-HMM, the boundary conditions and size of the sampling domains as well as the discretization error done of the micro FEMs.

[4]Equivalently, $+A_H$ is called an α-dissipative operator.

To quantify further the term r_{HMM} we need some regularity and growth assumption (in terms of ε) of the solution of the microproblems (31). Motivated by the case of periodic tensors (e.g. the chain rule applied to $a^\varepsilon = a(x, x/\varepsilon, t)$) we consider the following regularity assumption on the solution of problem (31)

(**H2**) $|\psi^{i,t}_{K_j}|_{H^{q+1}(K_{\delta_j})} + |\partial_t \psi^{i,t}_{K_j}|_{H^{q+1}(K_{\delta_j})} \leq C\ \varepsilon^{-q}\sqrt{|K_{\delta_j}|}$, where C is independent of ε, the time t, the quadrature points x_{K_j}, and the domain K_{δ_j}. We also suppose that the map $t \to a^\varepsilon(\cdot, t) \in (L^\infty(\Omega))^{d\times d}$ is C^1 and $|\partial_t a^\varepsilon_{ij}(t,\cdot)|_{L^\infty(\Omega)} \leq C$, for all $t \in (0, T)$ and all $\varepsilon > 0$. We make the same assumptions on the solution of the modified problem (31) where the tensor a^ε is replaced by $a^{\varepsilon T}$ (the adjoint problem).

Remark 4.2 *When Dirichlet boundary conditions* (13) *are imposed in* (30)*, the assumption (***H2***) can be easily satisfied (without any further knowledge about the structure of the oscillating tensor a^ε) for $q = 1$ as $|\psi^{i,t}_{K_j}|_{H^2(K_{\delta_j})} \leq C\varepsilon^{-1}\sqrt{|K_{\delta_j}|}$ follows from classical H^2 regularity results ([34, Chap. 2.6]), provided that $|a^\varepsilon_{ij}(\cdot, t)|_{W^{1,\infty}(\Omega)} \leq C\varepsilon^{-1}$ for $i, j = 1, \ldots, d$. Then, following the proof of [11, Lemma 4.12], $|\partial_t \psi^{i,t}_{K_j}|_{H^2(K_{\delta_j})} \leq C\varepsilon^{-1}\sqrt{|K_{\delta_j}|}$ holds, provided $|\partial_t a^\varepsilon_{ij}(\cdot, t)|_{W^{1,\infty}(\Omega)} \leq C\varepsilon^{-1}$. For periodic boundary conditions* (12) *in* (30)*, (***H2***) can be established for any given q, provided $a^\varepsilon = a(x, x/\varepsilon, t) = a(x, y, t)$ is Y-periodic in y, $\delta/\varepsilon \in \mathbb{N}$, and a^ε is sufficiently smooth, by following classical regularity results for periodic problems (see [13, Chap. 3]).* [5]

We then have the following theorem.

Theorem 4.3 *Consider u_0, u^H the solutions of* (8)*,* (15)*, respectively. In addition to the assumptions of Theorem 4.1, assume that (***H2***) hold. Then we have*

$$\|u_0 - u^H\|_{L^2([0,T];H^1(\Omega))} \leq C(H^\ell + \left(\frac{h}{\varepsilon}\right)^{2q} + r_{MOD} + \|g - u_0^H\|_{L^2(\Omega)}), \tag{49}$$

$$\|u_0 - u^H\|_{\mathcal{C}^0([0,T];L^2(\Omega))} \leq C(H^{\ell+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + r_{MOD} + \|g - u_0^H\|_{L^2(\Omega)}), \text{ if } \mu = 1. \tag{50}$$

If in addition, the tensor is symmetric, then

$$\|u_0 - u^H\|_{\mathcal{C}^0([0,T];H^1(\Omega))} \leq C(H^\ell + \left(\frac{h}{\varepsilon}\right)^{2q} + r_{MOD} + \|g - u_0^H\|_{H^1(\Omega)}).$$

The constants C are independent of $H, h, r_{MOD}, \varepsilon$.

The first term as is Theorem 4.1 quantifies the error coming from the macro solver. The second term quantities the error coming from the micro solver – when discretizing the microproblems by a FEM – transmitted to macroscale. This term does not appear in the analysis given in [38], where the microsolutions u^h, v^h in (16) where supposed to be exact. The additional analysis of the micro error allows to derive a macro-micro refinement strategy.
We emphasize that the remaining term r_{MOD} defined in (40) does not depend on the macro and micro mesh sizes H and h. In particular, any result concerning the approximation of the homogenized tensor with artificial micro boundary conditions or modified cell problems

[5] We also note that $\partial_t^k a^\varepsilon_{ij}|_K \in W^{1,\infty}(K)\ \forall K \in \mathcal{T}_H$ and $|\partial_t^k a^\varepsilon_{ij}|_{W^{1,\infty}(K)} \leq C\varepsilon^{-1}$ with $k = 0$ and 1 are sufficient, if the macro mesh is aligned with the (possible) discontinuities of a^ε (see [5] for details).

(e.g. [14],[23],[29],[30],[44]) could be used in our analysis. Here, we consider a class of non uniformly periodic tensors of the form

(**H3**) $a^\varepsilon = a(x, x/\varepsilon, t) = a(x, y, t)$ Y-periodic in y, where we set $Y = (0,1)^d$.

We then have the following theorem.

Theorem 4.4 *Consider u_0, u^H the solutions of (8), (15), respectively. In addition to the assumptions of Theorem 4.1, assume (**H2**) and (**H3**). Assume also that $\psi^{i,t}_{K_{\delta_j}}$ is the solution of the cell problem (30) in the space $W^1_{per}(K_{\delta_j})$, that $\varepsilon/\delta \in \mathbb{N}$, and that the tensor $a(x, x/\varepsilon, t)$ is collocated at the quadrature points $a(x_{K_j}, x/\varepsilon, t)$ in the FE-HMM macro bilinear form (16) and in the micro problems (14). Then we have*

$$\begin{aligned} \|u_0 - u^H\|_{L^2([0,T];H^1(\Omega))} &\leq C(H^\ell + \left(\frac{h}{\varepsilon}\right)^{2q} + \|g - u_0^H\|_{L^2(\Omega)}), \\ \|u_0 - u^H\|_{\mathcal{C}^0([0,T];L^2(\Omega))} &\leq C(H^{\ell+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + \|g - u_0^H\|_{L^2(\Omega)}), \quad \text{if } \mu = 1. \end{aligned} \tag{51}$$

If in addition, the tensor is symmetric, then

$$\|u_0 - u^H\|_{\mathcal{C}^0([0,T];H^1(\Omega))} \leq C(H^\ell + \left(\frac{h}{\varepsilon}\right)^{2q} + \|g - u_0^H\|_{H^1(\Omega)}).$$

The constants C are independent of H, h, ε.

Remark 4.5 *If the tensor $a(x, x/\varepsilon, t)$ is not collocated at the slow variable in the above theorem, we get for the modeling error (see [11, Appendix],[3, Prop. 14])*

$$r_{MOD} \leq C\ \delta.$$

If the solution of the cell problem (30) in $H^1_0(K_{\delta_j})$, a resonance error contributes to r_{MOD}. For a tensor independent of time, the results in [23] can be readily used in the framework developed in this paper for the analysis of parabolic problems and we have

$$r_{MOD} \leq C(\delta + \frac{\varepsilon}{\delta}).$$

This results could be extended for time-dependent tensor by following [23] and [11, Appendix].

4.2 Fully discrete estimates in space and time

In this section, we explain how fully discrete estimates in both space and time can be derived. We focus on the one hand on implicit time discretizations (Runge-Kutta methods) with variable timesteps, and on the other hand on stabilized explicit time discretizations (Chebyshev methods). We assume that the numerical initial condition u_0^H of the FE-HMM in (15) is chosen to approximate the exact initial condition g as

$$\|u_0^H - g\|_{L^2(\Omega)} \leq C(H^{\ell+1} + r_{HMM}), \tag{52}$$

$$\|u_0^H - g\|_{H^1(\Omega)} \leq C(H^\ell + r_{HMM}). \tag{53}$$

Remark 4.6 *There are several natural choices for the initial condition u_0^H to satisfy (52)-(53). For instance, one can take $u_0^H = \Pi_H g$, the L^2 projection of g on $S_0^\ell(\Omega, \mathcal{T}_H)$, defined as*

$$(\Pi_H g - g, z^H) = 0, \quad \forall z^H \in S_0^\ell(\Omega, \mathcal{T}_H), \tag{54}$$

and then (52)-(53) hold without the r_{HMM} terms[6]. One can also consider the elliptic projec-

[6]Notice that the regularity assumed on $u_0, \partial_t u_0$ in Theorem 4.1 implies $u_0(0) = g \in W^{\ell+1,p}(\Omega)$.

tion $u_0^H = P_H g$ *with respect to the bilinear forms* B *in* (28) *and* B_H *in* (16),

$$B_H(0; P_H g, z^H) = B(0; g, z^H), \qquad \forall z^H \in S_0^\ell(\Omega, \mathcal{T}_H), \tag{55}$$

and (52)-(53) *hold (see Corollary 5.2 in Sect. 5.1 below where we notice* $w^H(0) = \Pi_H g$*).*

The first theorem treats the case of implicit methods and is obtained by combining our fully discrete error estimates in space (Theorem 4.4) with the results of [37].

Theorem 4.7 *Consider* u_0 *the exact solution of* (8) *and* u_n^H *the numerical solution of a Runge-Kutta method for* (17)*, with variable timesteps* $\{\Delta t_n\}$ *satisfying* (21)-(22)*. Given an integer* $r \geq 1$*, assume that the Kunge-Kutta method has order* r *when applied to ordinary differential equations, that it has stage order* $r-1$*, and that it is strongly* $A(\theta)$*-stable with* $\alpha < \theta$ *where* α *is the angle in* (45) *of accretivity of* $-A_H$*. Assume the hypotheses of Theorem 4.4 with* $\mu = 1$*. Assume further* (52)*,*

$$f \in H^r(0,T; L^2(\Omega)), \qquad a^\varepsilon \in \mathcal{C}^r([0,T], L^\infty(\Omega)^{d\times d}) \text{ with } \|\partial_t^k a^\varepsilon\|_{(L^\infty(\Omega))^{d\times d}} \leq C, \ k = 1 \dots r,$$

and

$$\|\partial_t^r u^H(0)\|_{L^2(\Omega)} \leq C, \tag{56}$$

where u^H *is the solution of* (15)*. Then, we have the* $\mathcal{C}^0(L^2)$ *estimate*

$$\max_{0\leq n\leq N} \|u_n^H - u_0(t_n)\|_{L^2(\Omega)} \leq C\left(H^{\ell+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + \Delta t^r\right). \tag{57}$$

Assuming in addition (53) *and that* a^ε *is symmetric, then we have the* $L^2(H^1)$ *estimate*

$$\sum_{n=0}^{N-1} \Delta t_n \|u_n^H - u_0(t_n)\|_{H^1(\Omega)}^2 \leq C\left(H^\ell + \left(\frac{h}{\varepsilon}\right)^{2q} + \Delta t^r\right)^2. \tag{58}$$

All the above constants C *are independent of* $H, h, \varepsilon, \Delta t$*.*

The assumption (56) can be satisfied in dimension $d = \dim \Omega \leq 3$ as proved in Proposition 5.3 of Sect. 5.2 below. We mention that for $r = 1$ the symmetry assumption on the tensor can be removed for (58) (see Sect. 5.2). The next theorem treats the case of Chebyshev methods where we focus for simplicity on the case where tensor a^ε is symmetric and time-independent. Recall from Sect. 3.2 that it is essential when considering Chebychev methods that the eigenvalues of the differential operator of the problem remain close to the negative real axis. This is automatically the case when the tensor is symmetric.

Theorem 4.8 *Consider* u_0 *the exact solution of* (8) *and* u_n^H *the numerical solution of a Chebyshev method for* (17)*, applied with a constant timestep* $\Delta t = T/N$*, and with stability functions* $\{R_s(z)\}_{s\geq 1}$*. Assume that the tensor* a^ε *is symmetric and time-independent, and that* $f = 0$*. Assume* (52) *and the hypotheses of Theorem 4.4 with* $\mu = 1$*. Given* $r \geq 1$*, assume that the order of the Chebyshev method is* r*, precisely,*

$$\lim_{z\to 0} \left|\frac{e^z - R_s(z)}{z^{r+1}}\right| < \infty \quad \text{ for all } s \geq 1. \tag{59}$$

In addition to (27)*, assume the strong stability condition* (25) *holds with the number of stages* s *chosen such that* $\rho\Delta t \leq L_s$*, where* ρ *is the spectral radius of the operator* A_H *defined in* (44)*. Then,*

$$\max_{0\leq n\leq N} \|u_n^H - u_0(t_n)\|_{L^2(\Omega)} \leq C\left(H^{\ell+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + \Delta t^r\right). \tag{60}$$

For the sake of brevity of the analysis, we assumed in Theorem 4.8 above that the source term f is zero. Notice that a non-zero time-independent source $f(x)$ could also be considered in the analysis by using a change of variable of the standard form $u_0(x,t) \leftrightarrow u_0(x,t) - \overline{u}_0(x)$ where $\overline{u}_0$ denotes the stationary solution of the problem, to retreive the zero source case (we omit the details). Moreover, in the case where the strong stability condition (25) is not satisfied (for instance if the damping is zero in the Chebyshev method (26)), we can still show the convergence by exploiting the regularity of the initial condition, as illustrated in Theorem 5.6 of Sect. 5.2.1.

Remark 4.9 *For simplicity, we assumed $r_{MOD} = 0$ in Theorems 4.7 and 4.8. If (***H2***) does not hold (similarly to Theorem 4.3), then* (57), (58), *and* (60) *remain valid provided the term r_{MOD} defined in* (40) *is added in the right-hand sides of these estimates.*

5 Analysis

5.1 Fully discrete analysis in space

In this section we perform the analysis of the numerical method (15).

We shall often use the following continuous embedding result [27, Sect. 5.9.2]. For all $v \in H^1(0,T;X)$ where X is a real Banach space, e.g. $X = L^2(\Omega)$ or $X = H^1(\Omega)$, we have $v \in \mathcal{C}^0([0,T];X)$, $v(t) = v(s) + \int_s^t v'(\tau)d\tau$, for all $0 \le s \le t \le T$ and

$$\|v\|_{\mathcal{C}^0([0,T];X)} \le C\|v\|_{H^1(0,T;X)} \tag{61}$$

where C depends only on T and X. The macro FE space $S_0^\ell(\Omega, \mathcal{T}_H)$ equipped with the H^1 norm will be denoted V^ℓ in the following. Let $u_0(t)$ be the solution of (8). Following an idea of Raviart [40] for FEMs with numerical quadrature applied to (single scale) parabolic equations, we define an elliptic projection as follows. For all $t \in (0,T)$, we let $w^H(t) \in S_0^\ell(\Omega, \mathcal{T}_H)$ equipped with the norm of $H^1(\Omega)$ be the elliptic projection of $u_0(t)$ with respect to B and B_H as

$$B_H(t; w^H(t), z^H) = B(t; u_0(t), z^H), \qquad \forall z^H \in S_0^\ell(\Omega, \mathcal{T}_H), \quad t \in (0,T). \tag{62}$$

Using the ellipticity and continuity (37) of B_H and (36), $w^H(t)$ is well defined for all t and $w^H \in L^2(0,T;V^\ell)$. Differentiating (62) with respect to time, we obtain for all $z^H \in S_0^\ell(\Omega, \mathcal{T}_H)$ and almost all $t \in (0,T)$

$$B_H(t; \partial_t w^H(t), z^H) = B'(t; u_0(t), z^H) + B(t; \partial_t u_0(t), z^H) - B'_H(t; w^H(t), z^H). \tag{63}$$

We deduce similarly $\partial_t w^H \in L^2(0,T;V^\ell)$. Using (61) with $X = V^\ell$ we have that w^H is in fact a continuous function of time on $[0,T]$. As a preparation for the derivation of Theorem 4.1 we need the following lemma.

Lemma 5.1 *Let $u_0(t)$ be the solution of* (8) *and $w^H(t)$ be the elliptic projection defined in* (62). *Assume that the hypothesis of Theorem 4.1 hold. Then we have*

$$\|\partial_t^k \left(w^H - u_0\right)\|_{L^2(0,T;H^1(\Omega))} \le C(H^\ell + r_{HMM}), \ k = 0,1, \tag{64}$$

$$\|\partial_t^k \left(w^H - u_0\right)\|_{L^2(0,T;L^2(\Omega))} \le C(H^{\ell+\mu} + r_{HMM}), \ k = 0,1, \ \mu = 0,1. \tag{65}$$

Proof. *Step 1: estimation of* $\|w^H - u_0\|_{L^2(0,T;H^1(\Omega))}$.
Using (62), we get for all $z^H \in S_0^\ell(\Omega, \mathcal{T}_H)$,

$$\begin{aligned} B_H(t; w^H - \mathcal{I}_H u_0, z^H) &= B(t; u_0 - \mathcal{I}_H u_0, z^H) \\ &+ B(t; \mathcal{I}_H u_0, z^H) - B_{0,H}(t; \mathcal{I}_H u_0, z^H) \\ &+ B_{0,H}(t; \mathcal{I}_H u_0, z^H) - B_H(t; \mathcal{I}_H u_0, z^H) \\ &\leq CH^\ell \|u_0(t)\|_{W^{\ell+1,p}(\Omega)} \|z^H\|_{H^1(\Omega)} \\ &+ r_{HMM} \|\mathcal{I}_H u_0(t)\|_{H^1(\Omega)} \|z^H\|_{H^1(\Omega)} \end{aligned}$$

where we used (42) and Lemma 3.3. Using the ellipticity (37) of B_H and taking $z^H = w^H(t) - \mathcal{I}_H u_0$ gives

$$\gamma_1 \|w^H(t) - \mathcal{I}_H u_0(t)\|_{H^1(\Omega)} \leq CH^\ell \|u_0(t)\|_{W^{\ell+1,p}(\Omega)} + r_{HMM} \|u_0(t)\|_{H^2(\Omega)}.$$

Integrating between 0 and T, and using $\|\mathcal{I}_H u_0 - u_0\|_{L^2(0,T;H^1(\Omega))} \leq CH^\ell$ gives

$$\|w^H - u_0\|_{L^2(0,T;H^1(\Omega))} \leq C(H^\ell + r_{HMM}). \tag{66}$$

Step 2: estimation of $\|\partial_t(w^H - u_0)\|_{L^2(0,T;H^1(\Omega))}$.
Using (63), we obtain

$$\begin{aligned} B_H(t; \partial_t w^H - \mathcal{I}_H \partial_t u_0, z^H) &= B(t; \partial_t u_0 - \mathcal{I}_H \partial_t u_0, z^H) + B'(t; u_0 - \mathcal{I}_H u_0, z^H) \\ &+ B(t; \mathcal{I}_H \partial_t u_0, z^H) - B_{0,H}(t; \mathcal{I}_H \partial_t u_0, z^H) \\ &+ B_{0,H}(t; \mathcal{I}_H \partial_t u_0, z^H) - B_H(t; \mathcal{I}_H \partial_t u_0, z^H) \\ &+ B'(t; \mathcal{I}_H u_0, z^H) - B'_{0,H}(t; \mathcal{I}_H u_0, z^H) \\ &+ B'_{0,H}(t; \mathcal{I}_H u_0, z^H) - B'_H(t; \mathcal{I}_H u_0, z^H) \\ &- B'_H(t; w^H - \mathcal{I}_H u_0, z^H) \end{aligned}$$

This yields using (42) and Lemma 3.3,

$$\begin{aligned} B_H(t; \partial_t w^H - \mathcal{I}_H \partial_t u_0, z^H) &\leq CH^\ell (\|u_0(t)\|_{W^{\ell+1,p}(\Omega)} + \|\partial_t u_0(t)\|_{W^{\ell+1,p}(\Omega)}) \|z^H\|_{H^1(\Omega)} \\ &+ r_{HMM} (\|\mathcal{I}_H u_0(t)\|_{H^1(\Omega)} + \|\mathcal{I}_H \partial_t u_0(t)\|_{H^1(\Omega)}) \|z^H\|_{H^1(\Omega)} \\ &+ C \|w^H - \mathcal{I}_H u_0(t)\|_{H^1(\Omega)} \|z^H\|_{H^1(\Omega)} \end{aligned}$$

We take $z^H = \partial_t w^H(t) - \mathcal{I}_H \partial_t u_0(t)$. Using (66) to estimate the last term in above inequality, we obtain the following estimate, similarly to the Step 1,

$$\|\partial_t(w^H - u_0)\|_{L^2(0,T;H^1(\Omega))} \leq C(H^\ell + r_{HMM}). \tag{67}$$

Step 3: estimation of $\|w^H - u_0\|_{L^2(0,T;L^2(\Omega))}$.
We use the classical duality argument of Aubin-Nitsche. Given $v \in L^2(0,T;L^2(\Omega))$, consider for almost every $t \in (0,T)$ the solution $\varphi(t) \in H_0^1(\Omega)$ of the problem

$$B(z, \varphi(t)) = (v(t), z), \quad \forall z \in H_0^1(\Omega). \tag{68}$$

Since $a \in (L^\infty(0,T; W^{1,\infty}(\Omega)))^{d \times d}$, using the convexity of the polyhedral domain Ω, we know that $\varphi \in L^2(0,T; H^2(\Omega))$ and

$$\|\varphi\|_{L^2(0,T;H^2(\Omega))} \leq C \|v\|_{L^2(0,T;L^2(\Omega))}. \tag{69}$$

We take $z = v = w^H(t) - u_0(t)$ in (68). Using (62) we obtain for all $\varphi^H(t) \in S_0^\ell(\Omega, \mathcal{T}_H)$,

$$\begin{aligned}(w^H - u_0, w^H - u_0) &= B(t; w^H - u_0, \varphi - \varphi^H) \\ &+ B(t; w^H - \mathcal{I}_H u_0, \varphi^H) - B_H(t; w^H - \mathcal{I}_H u_0, \varphi^H) \\ &+ B(t; \mathcal{I}_H u_0, \varphi^H) - B_H(t; \mathcal{I}_H u_0, \varphi^H). \end{aligned} \tag{70}$$

We take $\varphi^H = \mathcal{I}_H \varphi(t)$. Using (36), (41) and (43) respectively, we deduce

$$\begin{aligned}(w^H - u_0, w^H - u_0) &\leq C\|\varphi(t)\|_{H^2(\Omega)}\big((H + r_{HMM})\|w^H(t) - u_0(t)\|_{H^1(\Omega)} \\ &+ (H^{\ell+\mu} + r_{HMM})\|u(t)\|_{H^{\ell+1}(\Omega)}\big)\end{aligned}$$

where we used also Lemma 3.3. We deduce with (69) and (66) that

$$\|w^H - u_0\|_{L^2(0,T;L^2(\Omega))} \leq C(H^{\ell+\mu} + r_{HMM}). \tag{71}$$

Step 4: Estimation of $\|\partial_t(w^H - u_0)\|_{L^2(0,T;L^2(\Omega))}$.
We consider again the dual problem (68), with $z = v = \partial_t(w^H - u_0)$

$$\big(\partial_t(w^H - u_0), \partial_t(w^H - u_0)\big) = B(t; \partial_t(w^H - u_0), \varphi), \tag{72}$$

and estimate for $\varphi(t) \in H^2(\Omega)$ and $\varphi^H(t) \in S_0^\ell(\Omega, \mathcal{T}_H)$

$$\begin{aligned}B(t, \partial_t(w^H - u_0), \varphi) &= B(t; \partial_t(w^H - u_0), \varphi - \varphi^H) \\ &+ B(t; \partial_t w^H - \mathcal{I}_H \partial_t u_0, \varphi^H) - B_H(t; \partial_t w^H - \mathcal{I}_H \partial_t u_0, \varphi^H) \\ &+ B(t; \mathcal{I}_H \partial_t u_0, \varphi^H) - B_H(t; \mathcal{I}_H \partial_t u_0, \varphi^H) \\ &+ B'(t; u_0 - w^H, \varphi) \\ &+ B'(t; u_0 - w^H, \varphi^H - \varphi) \\ &+ B'(t; w^H - \mathcal{I}_H u_0, \varphi^H) - B'_H(t; w^H - \mathcal{I}_H u_0, \varphi^H) \\ &+ B'(t; \mathcal{I}_H u_0, \varphi^H) - B'_H(t; \mathcal{I}_H u_0, \varphi^H). \end{aligned} \tag{73}$$

The fourth term in the right-hand side of (73) can be bounded as

$$B'(t; u_0 - w^H, \varphi) = (u_0 - w^H, \mathcal{L}\varphi) \leq C\|u_0(t) - w^H(t)\|_{L^2(\Omega)}\|\varphi(t)\|_{H^2(\Omega)},$$

where we consider the differential operator $\mathcal{L} := -\sum_{i,j=1}^d \frac{\partial}{\partial x_i}\big(\frac{\partial a^0_{ji}}{\partial t}(x,t)\frac{\partial}{\partial x_j}\big)$. All other terms in the right-hand side of (73) can be bounded similarly as those in the right-hand side of (70). From (72), using (73) with $\varphi^H = \mathcal{I}_H \varphi$ we obtain

$$\|\partial_t(w^H(t) - u_0(t))\|_{L^2(\Omega)} \leq C\Big(\|u_0(t) - w^H(t)\|_{L^2(\Omega)} + H^{\ell+\mu} + r_{HMM}\Big)\|\varphi(t)\|_{H^2(\Omega)}.$$

Integrating this estimate between 0 and T and using the estimates (69),(66), (67), (71), we deduce from the Cauchy-Schwarz inequality

$$\|\partial_t(w^H - u_0)\|_{L^2(0,T;L^2(\Omega))} = C(H^{\ell+\mu} + r_{HMM}). \tag{74}$$

This concludes the proof of Lemma 5.1. □

Corollary 5.2 *Under the assumption of Lemma 5.1, we have*

$$\|w^H - u_0\|_{\mathcal{C}^0(0,T;H^1(\Omega))} \leq C(H^\ell + r_{HMM}) \tag{75}$$
$$\|w^H - u_0\|_{\mathcal{C}^0(0,T;L^2(\Omega))} \leq C(H^{\ell+\mu} + r_{HMM}). \tag{76}$$

Proof. The results follow from (61) with $X = H^1(\Omega)$ or $X = L^2(\Omega)$ using (66),(67) or (71),(74). □

We can now prove our first main result.
Proof of Theorem 4.1 In view of the triangle inequality

$$\|u_0 - u^H\| \leq \|w^H - u_0\| + \|u^H - w^H\|$$

and Lemma 5.1 it remains to estimate $\|u^H - w^H\|$ for the $L^2(H^1), \mathcal{C}^0(L^2)$ and the $\mathcal{C}^0(H^1)$ norms.
Step 1: Estimation of $\|u^H - w^H\|_{L^2(0,T;H^1(\Omega))} + \|u^H - w^H\|_{\mathcal{C}^0([0,T];L^2(\Omega))}$.
We set $\xi^H(t) = u^H(t) - w^H(t), t \in [0,T]$. A simple calculation using (8),(15) and (62) gives for all $z^H \in S_0^\ell(\Omega, \mathcal{T}_H)$,

$$(\partial_t \xi^H, z^H) + B_H(t; \xi^H, z^H) = (\partial_t u_0, z^H) - (\partial_t w^H, z^H). \tag{77}$$

Integrating this equality from 0 to t with $z^H = \xi^H$, using the continuity of u^H, w^H with respect to time, the coercivity of $B_H(\cdot,\cdot)$, (36) and the Cauchy-Schwarz and the Young inequalities, we obtain

$$\|\xi^H(t)\|^2_{L^2(\Omega)} + \gamma_1 \int_0^t \|\xi^H(s)\|^2_{H^1(\Omega)} ds \leq \|\xi^H(0)\|^2_{L^2(\Omega)} + \frac{1}{\gamma_1} \int_0^t \|\partial_t u_0 - \partial_t w^H\|^2_{L^2(\Omega)} ds, \tag{78}$$

and thus

$$\|\xi^H\|^2_{L^2(0,T;H^1(\Omega))} \leq C\Big(\|\xi^H(0)\|^2_{L^2(\Omega)} + \|\partial_t u_0 - \partial_t w^H\|^2_{L^2(0,T;L^2(\Omega))}\Big).$$

Using the decomposition $\xi(0) = (u_0 - w^H)(0) + (u_0^H - g)$ and (76) yields

$$\|\xi(0)\|_{L^2(\Omega)} \leq C(H^{\ell+\mu} + r_{HMM}) + \|u_0^H - g\|_{L^2(\Omega)}. \tag{79}$$

By taking the supremum with respect to t in (78), using (79) and (65) we deduce

$$\|u^H - w^H\|_{\mathcal{C}^0([0,T];L^2(\Omega))} + \|u^H - w^H\|_{L^2(0,T;H^1(\Omega))} \leq C(H^{\ell+\mu} + r_{HMM} + \|u_0^H - g\|_{L^2(\Omega)}).$$

Step 2: Estimation of $\|u^H - w^H\|_{\mathcal{C}^0([0,T];H^1(\Omega))}$.
For $\xi^H(t) = u^H(t) - w^H(t), t \in [0,T]$, we set $z^H = \partial_t \xi^H$ in (77). Using the symmetry of the tensor, and integrating from 0 to t, we obtain for $0 \leq t \leq T$

$$\begin{aligned} 2\int_0^t \|\partial_t \xi^H(s)\|^2_{L^2(\Omega)} ds + B_H(t; \xi^H(t), \xi^H(t)) &= B_H(0; \xi^H(0), \xi^H(0)) + \int_0^t B'_H(s; \xi^H, \xi^H) ds \\ &+ 2\int_0^t (\partial_t u_0 - \partial_t w^H, \partial_t \xi^H)\, ds. \end{aligned}$$

Using the coercivity of $B_H(\cdot,\cdot)$, (36) and the Cauchy-Schwarz inequality, and the Young inequality, we obtain, similarly to (78),

$$\begin{aligned}\int_0^t \|\partial_t \xi^H(s)\|^2_{L^2(\Omega)} ds + \gamma_1 \|\xi^H(t)\|^2_{H^1(\Omega)} &\leq \gamma_2 \|\xi^H(0)\|^2_{H^1(\Omega)} + \gamma_2 \int_0^t \|\xi^H(s)\|^2_{H^1(\Omega)} ds \\ &+ \int_0^t \|\partial_t u_0(s) - \partial_t w^H(s)\|^2_{L^2(\Omega)} ds. \end{aligned} \tag{80}$$

Using again $\xi^H(0) = (u_0 - w^H)(0) + (u_0^H - g)$ and (75) gives

$$\|\xi^H(0)\|_{H^1(\Omega)} \leq C(H^\ell + r_{HMM}) + \|u_0^H - g\|_{H^1(\Omega)}. \tag{81}$$

Taking the suppremum with respect to t in (80), and using (81), (64), we deduce

$$\|u^H - w^H\|_{\mathcal{C}^0([0,T];H^1(\Omega))} \leq C(H^\ell + r_{HMM} + \|u_0^H - g\|_{H^1(\Omega)}).$$

This together with (75) concludes the proof of (48). □

Micro and modeling errors.
Proof of Theorem 4.3 In view of (40), we have to estimate

$$r_{MIC} = \sum_{k=0}^{1} \sup_{K\in\mathcal{T}_H, x_{K_j}\in K, t\in[0,T]} \|\partial_t^k \bar{a}_K^0(x_{K_j},t) - \partial_t^k a_K^0(x_{K_j},t)\|_F. \tag{82}$$

For $k=0$, the estimate of the micro error r_{MIC} was first presented in [2], generalized to high order in [3, Lemma 10],[5, Corollary 10] (see also [4]). These results can be generalized for nonsymmetric tensor [21] (see also [11] for a short proof). The time-dependence of the tensor does not change the above proof and we have

$$\sup_{K\in\mathcal{T}_H, x_{K_j}\in K, t\in[0,T]} \|\bar{a}_K^0(x_{K_j},t) - a_K^0(x_{K_j},t)\|_F \leq C\left(\frac{h}{\varepsilon}\right)^{2q}.$$

To estimate the term for $k=1$ we follow [11, Lemma 4.6, Lemma 7.1]. We first observe (see [11, Lemma 4.6]) that

$$\begin{aligned}&(\bar{a}_K^0(x_{K_j},t) - a_K^0(x_{K_j},t))_{mn} \\ &= \frac{-1}{|K_{\delta_j}|}\int_{K_{\delta_j}} a^\varepsilon(x,t)\left(\nabla\psi_{K_j}^{n,t}(x) - \nabla\psi_{K_j}^{n,h,t}(x)\right)\cdot\left(\nabla\overline{\psi}_{K_j}^{m,t}(x) - \nabla\overline{\psi}_{K_j}^{m,h,t}(x)\right)dx, \end{aligned} \tag{83}$$

where $\overline{\psi}_{K_j}^{m,h,t}, \overline{\psi}_{K_j}^{m,t}$ are the solutions of the problems (30), (31), respectively, with $a^\varepsilon(x,t)$ replaced by $a^\varepsilon(x,t)^T$ (the adjoint problem). Differentiating with respect to t and using (2) we obtain

$$\begin{aligned}&|\frac{d}{dt}(\bar{a}_{K_j}^0(x_{K_j},t) - a_{K_j}^0(x_{K_j},t))_{mn}| \\ &\leq \frac{1}{|K_{\delta_j}|}\Big(\|\nabla\psi_{K_j}^{n,t} - \nabla\psi_{K_j}^{n,h,t}\|_{L^2(K_{\delta_j})}\|\nabla\overline{\psi}_{K_j}^{m,t} - \nabla\overline{\psi}_{K_j}^{m,h,t}\|_{L^2(K_{\delta_j})} \\ &+\|\partial_t\nabla\psi_{K_j}^{n,t} - \partial_t\nabla\psi_{K_j}^{n,h,t}\|_{L^2(K_{\delta_j})}\|\nabla\overline{\psi}_{K_j}^{m,t} - \nabla\overline{\psi}_{K_j}^{m,h,t}\|_{L^2(K_{\delta_j})} \\ &+\|\nabla\psi_{K_j}^{n,t} - \nabla\psi_{K_j}^{n,h,t}\|_{L^2(K_{\delta_j})}\|\partial_t\nabla\overline{\psi}_{K_j}^{m,t} - \partial_t\nabla\overline{\psi}_{K_j}^{m,h,t}\|_{L^2(K_{\delta_j})}\Big). \end{aligned}$$

Following the proof of [11, Lemma 7.1] one can show that

$$\partial_t \nabla \psi_{K_j}^{m,t} = \nabla \partial_t \psi_{K_j}^{m,t}, \qquad \partial_t \nabla \psi_{K_j}^{m,h,t} = \nabla \partial_t \psi_{K_j}^{m,h,t},$$

and similarly for $\overline{\psi}_{K_j}^{m,t}, \overline{\psi}_{K_j}^{m,h,t}$. Using hypothesis (**H2**) gives the result. □

Proof of Theorem 4.4 Under hypothesis (**H3**) one can show the identity $a_{K_j}^0(x_{K_j}, t) = \bar{a}_{K_j}^0(x_{K_j}, t)$ similarly as in [9, Appendix A] (see also [4, Theorem 17]) and the result follows. □

5.2 Fully discrete analysis in space and time

In this section, we explain how the time discretization error can be analyzed, to derive an error analysis that is fully discrete in both time and space.

We consider the s-stage Runge-Kutta methods based on the FE-HMM spatial discretization as defined in (18). We assume that the method has order (at least) r, stage order $r-1$ and is strongly $A(\theta)$-stable with $\alpha < \theta \leq \pi/2$ (see Section 3.2). The following proof combines our fully discrete error bounds in space and results in [37, Thm. 3.2 & 5.1].

Proof of Theorem 4.7 We first show the estimate

$$\|\partial_t^r u^H\|_E \leq C(\|f\|_{H^r(0,T;L^2(\Omega))} + \max_{k=0\ldots r} \|\partial_t^k u^H(0)\|_{L^2(\Omega)}). \tag{84}$$

The assumptions on the tensor ensure that the bilinear form $B_H(t;\cdot,\cdot)$ is bounded and elliptic uniformly in t, i.e. (37) holds, and its derivatives up to order r are continuous and bounded uniformly in t. In the case $r = 1$, the estimate (84) is shown in [27, Sect. 7.1]. We notice that the proof of this estimate in [27, Sect. 7.1] remains valid for a multidimensional valued function. Applying this result to the augmented system associated to (15) and taking as unknown the vector function $(u^H, \partial_t u^H, \ldots, \partial_t^r u^H)$, we deduce that (84) holds for a general r. Using [37, Thm. 3.2 & 5.1] (we take the pair $V = (S_0^\ell(\Omega, \mathcal{T}_H), \|\cdot\|_{H_0^1(\Omega)})$ and $H = (S_0^\ell(\Omega, \mathcal{T}_H), \|\cdot\|_{L^2(\Omega)})$) we then have

$$\sum_{n=0}^{N-1} \Delta t_n \|u_n^H - u^H(t_n)\|_{H^1(\Omega)}^2 + \max_{0\leq n\leq N} \|u_n^H - u^H(t_n)\|_{L^2(\Omega)}^2 \leq C(\Delta t^r)^2 \|\partial_t^r u^H(t)\|_E^2, \tag{85}$$

where $\|\cdot\|_E$ is defined in (5) and the constant C is independent of $H, \Delta t$.

The estimates (57)-(58) are then an immediate consequence of the triangle inequality applied to the decomposition $u_n^H - u_0(t_n) = (u_n^H - u^H(t_n)) + (u^H(t_n) - u_0(t_n))$, Theorem 4.4, and (84),(56) and (85). □

We next discuss (56) used in the above theorem, i.e., $\max_{k=0\ldots r}(\|\partial_t^k u^H(0)\|_{L^2(\Omega)}) \leq C$. For the case of order $r = 1$, this assumption is automatically satisfied if one chooses the initial condition $u_0^H = P_H g$ defined in (55) and by assuming $\partial_t u_0 \in L^2(\Omega)$. Indeed, in this case, we have $\partial_t u_0^H = \Pi_H \partial_t u_0$. For arbitrary r, the assumption (56) can be replaced by an assumption on the exact solution u_0 as proved in the following proposition.

Proposition 5.3 *Assume $\ell \geq r$. Assume that $d = \dim \Omega \leq 3$ and the family of macro meshes $\{\mathcal{T}_H\}$ satisfies the inverse assumption*

$$\frac{H}{H_K} \leq C \quad \textit{for all } K \in \mathcal{T}_H \textit{ and all } \mathcal{T}_H. \tag{86}$$

In addition to the hypotheses of Theorem 4.7, assume for all $k = 1 \dots r$,

$$\partial_t^k u_0 \in \mathcal{C}^0([0,T]; W^{\ell+1,p}(\Omega)).$$

*In the case $r > 1$ assume in addition to (**H2**) that $|\partial_t^k \psi^{i,t}_{K_j}|_{H^{q+1}(K_{\delta_j})} \leq C\varepsilon^{-q}$ for all $k < r$, and that the homogenized tensor has the smoothness*

$$\partial_t^k a^0_{ij} \in L^\infty(0,T; W^{\ell,\infty}(\Omega)), \ \forall i,j = 1 \dots d, \quad \textit{for all } k < r.$$

Assume further that the macro and micro mesh sizes satisfy $(h/\varepsilon)^{2q} \leq H^r$. Then (56) is satisfied with a constant C independent of H, h, ε.

Proof. We consider first the case $r = 1$. Subtracting (8) and (15), we obtain

$$\begin{aligned}
(\partial_t u^H - \mathcal{I}_H \partial_t u_0, z^H) &= B(t; u_0 - \mathcal{I}_H u_0, z^H) + B_H(t; \mathcal{I}_H u_0 - u^H, z^H) \\
&+ B(t; \mathcal{I}_H u_0, z^H) - B_{H,0}(t; \mathcal{I}_H u_0, z^H) \\
&+ B_{H,0}(t; \mathcal{I}_H u_0, z^H) - B_H(t; \mathcal{I}_H u_0, z^H) \\
&+ (\partial_t u_0 - \mathcal{I}_H \partial_t u_0, z^H), \quad \forall z^H \in S_0^\ell(\Omega, \mathcal{T}_H).
\end{aligned} \tag{87}$$

Taking $z^H = \partial_t u^H - \mathcal{I}_H \partial_t u_0$, using (42) and Lemma 3.3, we deduce

$$\|\partial_t u^H - \mathcal{I}_H \partial_t u_0\|^2_{L^2(\Omega)} \leq C(H^\ell + r_{HMM}) \|\partial_t u^H - \mathcal{I}_H \partial_t u_0\|_{H^1(\Omega)} \leq CH \|\partial_t u^H - \mathcal{I}_H \partial_t u_0\|_{H^1(\Omega)}.$$

In view of (86), using the inverse estimate $\|z^H\|_{H^1(\Omega)} \leq CH^{-1}\|z^H\|_{L^2(\Omega)}$ [16, Thm. 17.2], we deduce $\|\partial_t u^H - \mathcal{I}_H \partial_t u_0\|_{\mathcal{C}^0([0,T],L^2(\Omega))} \leq C$, which concludes the proof in the case $r = 1$. The proof in the case of a general r is obtained similarly by time differentiating (87), and by proving by induction that for all $k = 1 \dots r$,

$$\|\partial_t^k u^H - \partial_t^k u_0\|_{\mathcal{C}^0([0,T],L^2(\Omega))} \leq CH^{r-k}.$$

This concludes the proof. □

We conclude this section by proving that the symmetry assumption used for (58) in Theorem 4.7 can be removed in certain situations.

Proposition 5.4 *For $r = 1$, and for a general initial condition u_0^H, assuming*

$$\|\partial_t^2 u^H(0)\|_{L^2(\Omega)} \leq C,$$

then (53) and the symmetry assumption on the tensor can be removed for (58) in the proof of the Theorem 4.7.

Proof. Using $\|\partial_t^2 u^H(0)\|_{L^2(\Omega)} \leq C$ and the estimate (84), we deduce $\|\partial_t u^H\|_{\mathcal{C}([0,T],L^2(\Omega))} \leq C$. Then, setting $e(t) = \nabla u^H(t) - \nabla u_0(t)$, and using the inequality

$$\begin{aligned}
\Delta t_n \|e(t_n)\|^2_{L^2(\Omega)} &= \int_{t_n}^{t_{n+1}} \|e(t)\|^2_{L^2(\Omega)} dt - 2\int_{t_n}^{t_{n+1}} (t_{n+1}-t)(e(t), \partial_t e(t)) dt \\
&\leq \|e\|^2_{L^2(t_n,t_{n+1},L^2(\Omega))} + \int_{t_n}^{t_{n+1}} ((t_{n+1}-t)^2 + \|e(t)\|^2_{L^2(\Omega))}) dt \\
&\leq \|e\|^2_{L^2(t_n,t_{n+1},L^2(\Omega))} + (\Delta t_n^3/3 + \|e\|^2_{L^2(t_n,t_{n+1},L^2(\Omega))}),
\end{aligned}$$

we deduce, after summation on n,

$$\sum_{n=0}^{N-1} \Delta t_n \|e(t_n)\|^2_{L^2(\Omega)} \leq C(\|e\|^2_{L^2(0,T,L^2(\Omega))} + \Delta t^2).$$

We conclude the proof using the triangle inequality applied to the decomposition $\nabla u_n^H - \nabla u_0(t_n) = \nabla(u_n^H - u^H(t_n)) + \nabla(u^H(t_n) - u_0(t_n))$, and applying (85) and (47). □

5.2.1 Stabilized explicit methods

The proof of Theorem 4.8 relies on the following lemma, whose proof is inspired by the analysis in [18, Thm. 9] in the context of standard implicit Runge-Kutta methods.

Lemma 5.5 *Assume the hypothesis of Theorem 4.8. Then,*

$$\max_{0\leq n\leq N} \|u_n^H - u^H(t_n)\|_{L^2(\Omega)} \leq C\Delta t^r \|u_0^H\|_{L^2(\Omega)}.$$

Proof. Let $\varphi_{n,s}(z) = e^{nz} - R_s(z)^n$. It is sufficient to show the estimate

$$|\varphi_{n,s}(z)| \leq Cn^{-r} \quad \text{for all } z \in [-L_s, 0], \tag{88}$$

where C is independent of n, s. Indeed, consider the operator A_H in (44). Using the symmetry of A_H, there exists an orthonormal basis where the operator A_H is in diagonal form, which yields,

$$\|\varphi_{n,s}(\Delta t A_H)\|_{L^2(\Omega)\to L^2(\Omega)} = \sup_{z\in sp(A_H)} |\varphi_{n,s}(\Delta t z)|, \tag{89}$$

where $sp(A_H)$ denotes the spectrum of A_H. Using (88), (89), we deduce the estimate

$$\|u_n^H - u^H(t_n)\|_{L^2(\Omega)} = \|\varphi_{n,s}(\Delta t A_H) u_0^H\|_{L^2(\Omega)} \leq Cn^{-r}\|u_0^H\|_{L^2(\Omega)},$$

which concludes the proof. It remains to prove (88). The order condition (59) implies that $(e^z - R_s(z))/z^{r+1}$ is an analytic function in a neighbourhood of zero. Applying the maximum principle and using the bound (27), we deduce

$$\sup_{|z|\leq\gamma} \left|\frac{e^z - R_s(z)}{z^{r+1}}\right| \leq \sup_{|z|=\gamma} \frac{e^\gamma + |R_s(z)|}{\gamma^{r+1}} \leq C_0, \tag{90}$$

where C_0 is independent of s and γ is the constant in (27). For $z \in [-\gamma, 0]$, we notice

$$|R_s(z)| = e^{z/2}|e^{z/2} - e^{-z/2}\varphi_{1,s}(z)| \leq e^{z/2}(1 - |z|/2 + |z|^2/8 + C_0|z|^{r+1}e^{-z/2})$$

where we use (90), which yields

$$|R_s(z)| \leq e^{z/2} \text{ for all } z \in [-\gamma, 0], \tag{91}$$

where γ is chosen small enough. Using the identity

$$\varphi_{n,s}(z) = \varphi_{1,s}(z) \sum_{k=0}^{n-1} e^{kz} R_s(z)^{n-1-k}, \tag{92}$$

we deduce from (90) and (91) the bound

$$|\varphi_{n,s}(z)| \leq C_0 |z|^{r+1} n e^{(n-1)z/2} \leq C_0 \frac{n(2(r+1)|z|)^{r+1}}{(e(n-1)|z|)^{r+1}} \leq \frac{C_0(4(r+1)/e)^{r+1}}{n^r},$$

where we use the estimate $e^{-x} \leq (\frac{r+1}{ex})^{r+1}$ for all $x \geq 0$. For the case $z \in [-L_s, -\gamma]$, let $\rho < 1$ denote the quantity in the left-hand side of (25). We have

$$|\varphi_{n,s}(z)| \leq e^{-n|z|} + \rho^n \leq e^{-n\gamma} + e^{-n(1-\rho)} \leq \frac{(r/e)^r(\gamma^{-r} + (1-\rho)^{-r})}{n^r}$$

where we used twice the estimate $e^{-x} \leq (\frac{r}{ex})^r$. This proves the estimate (88) and concludes the proof. □

Proof of Theorem 4.8. The proof is an immediate consequence of

$$\|u_n^H - u_0(t_n)\|_{L^2(\Omega)} \leq \|u_n^H - u^H(t_n)\|_{L^2(\Omega)} + \|u^H(t_n) - u_0(t_n)\|_{L^2(\Omega)},$$

Lemma 5.5, and (51) in Theorem 4.4. □

We end this section by proving a result analogous to Theorem 4.8 but in the case where the strong stability condition is not satisfied, i.e., when only (23) instead of (25) holds.

Theorem 5.6 *Consider u_0 the exact solution of* (8) *and u_n^H the numerical solution of a Chebyshev method for* (17)*. Assume the hypothesis of Theorem 4.8 hold for $r = 1$, with the exception that the strong stability condition* (25) *is replaced by* (23)*. Then,*

$$\max_{0 \leq n \leq N} \|u_n^H - u_0(t_n)\|_{L^2(\Omega)} \leq C \left(H^{\ell+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + \Delta t \right).$$

The proof of Theorem 5.6 relies on the following lemma which takes advantage of the regularity assumed on the initial condition g. Recall that $g \in H^2(\Omega)$ follows from (61) and the regularity assumed on $u_0, \partial_t u_0$ in Theorem 4.1.

Lemma 5.7 *Assume the hypothesis of Theorem 5.6. Then,*

$$\max_{0 \leq n \leq N} \|u_n^H - u^H(t_n)\|_{L^2(\Omega)} \leq C(\Delta t \|g\|_{H^2(\Omega)} + \|P_H g - u_0^H\|_{L^2(\Omega)}),$$

where $P_H g$ denotes the elliptic projection of g defined in (55)*.*

Proof. We consider first the case where the initial condition is $u_0^H = P_H g$. We first show

$$\left| \frac{e^{nz} - R_s(z)^n}{z} \right| \leq C, \quad \text{for all } z \in [-L_s, 0].$$

Consider the function $\varphi_{n,s}(z) = (e^{nz} - R_s(z)^n)/z$. For $z \in [-\gamma, 0]$, we deduce from (23)-(90) and an identity similar to (92) the estimate

$$|\varphi_{n,s}(z)| \leq C_0 |z| \sum_{k=0}^{n-1} e^z \leq C_0 \frac{\gamma}{1 - e^{-\gamma}}.$$

For $z \in [-L_s, -\gamma]$, from (23), we have $|\varphi_{n,s}(z)| \leq 2/|z| \leq 2/\gamma$. We deduce from

$$u_n^H - u^H(t_n) = \Delta t \varphi_{n,s}(\Delta t A_H) A_H u_0^H$$

and the identity (89) that

$$\|u_n^H - u^H(t_n)\|_{L^2(\Omega)} \leq \Delta t C \|A_H u_0^H\|_{L^2(\Omega)}.$$

The operator P_H in (55) can be decomposed as $P_H = A_H^{-1} \Pi_H A$ where A is the operator defined by $-(Av, w) = B(v, w)$, $\forall v, w \in H^1(\Omega)$ (see the discrete analog (44)). This yields $A_H u_0^H = \Pi_H A g$. Using $\|\Pi_H A g\|_{L^2(\Omega)} \leq \|Ag\|_{L^2(\Omega)} \leq \|g\|_{H^2(\Omega)}$, this concludes the proof of Lemma 5.7 in the case $u_0^H = P_H g$.

For a general initial condition u_0^H, we use the inequality

$$\begin{aligned} \|u_n^H - u^H(t_n)\|_{L^2(\Omega)} &\leq \|(\exp(t_n A_H) - R_s(\Delta t A_H)) P_H u_0\|_{L^2(\Omega)} \\ &+ \|R_s(\Delta t A_H)(P_H u_0 - u_0)\|_{L^2(\Omega)} \\ &\leq C \Delta t \|g\|_{H^2(\Omega)} + \|P_H u_0 - u_0^H\|_{L^2(\Omega)} \end{aligned}$$

where the first term is bounded as previously and the second term is bounded using (23) and the identity (89) applied with R_s instead of $\varphi_{n,s}$. □

Proof of Theorem 5.6. Using the estimate $\|P_H g - g\|_{L^2(\Omega)} \leq C(H^{\ell+1} + r_{HMM})$ which follows from Corollary 5.2 where we notice that $w^H(0) = \Pi_H g$, we deduce

$$\|P_H g - u_0^H\|_{L^2(\Omega)} \leq C(H^{\ell+1} + r_{HMM}) + \|g - u_0^H\|_{L^2(\Omega)}.$$

We conclude the proof by combining Lemma 5.7 and (51) in Theorem 4.4. □

Remark 5.8 *The Theorem 5.6 can be generalized for arbitrary r provided one can bound the second term of the following inequality*

$$\max_{0 \leq n \leq N} \|u_n^H - u^H(t_n)\|_{L^2(\Omega)} \leq C(\Delta t^r \|A^r g\|_{L^2(\Omega)} + \|A_H^{-r} \Pi_H A^r g - u_0^H\|_{L^2(\Omega)}),$$

which is obtained following the lines of the proof of Lemma 5.7. This term vanishes if one chooses the initial condition $u_0^H = A_H^{-r} \Pi_H A^r g$ *which is equivalent to impose* $\partial_t^r u^H(0) = \Pi_H(\partial_t^r u_0(0))$.

6 Numerical experiments

In this section, we show numerically that the spatial estimates of Theorem 4.4 for linear parabolic problems with locally periodic tensors are sharp for $\mathcal{P}^1$ or $\mathcal{Q}^1$-finite elements, and smooth initial data. We also illustrate the time discretization error estimates of Theorem 4.7 (implicit Euler method) and Theorem 5.6 (Chebyshev method). Numerical experiments for a class of non-linear parabolic problems are reported. The case of a non periodic tensor with variable cell sizes is also investigated on a random problem.

6.1 Convergence rates

We recall that for a tensor of the form $a^\varepsilon(x,t) = a(x, x/\varepsilon, t) = a(x,y,t)$ periodic with respect to the y variable and collocated in the slow variable at the quadrature points, the $L^2(H^1)$ and $\mathcal{C}^0(L^2)$ errors satisfy (see Theorem 4.4 with $\ell = q = 1$)

$$\|u^H - u_0\|_{L^2(0,T,H^1(\Omega))} \leq C\left(H + \left(\frac{h}{\varepsilon}\right)^2\right),$$
$$\|u^H - u_0\|_{\mathcal{C}^0([0,T],L^2(\Omega))} \leq C\left(H^2 + \left(\frac{h}{\varepsilon}\right)^2\right),$$

where we have considered periodic boundary conditions on the micro domains and $\delta/\varepsilon \in \mathbb{N}$. As usual, C in the above estimates is independent of H, h, ε. Here, we consider for the numerical initial value, $u_0^H = \Pi_H g$, the L^2-projection (54) on $S_0^\ell(\Omega, \mathcal{T}_H)$.

In the case of a linear or non-linear periodic tensor, to be able to compare our numerical solution with the exact homogenized solution, we consider test problems with tensors which are simple enough so that an analytical formula for the homogenized tensor can be derived. For the case of a random tensor, a finescale reference solution is computed.

The $\mathcal{C}^0(L^2)$ and $L^2(H^1)$ relative errors between the numerical and exact homogenized solutions are approximated by quadrature formulas. We compute[7]

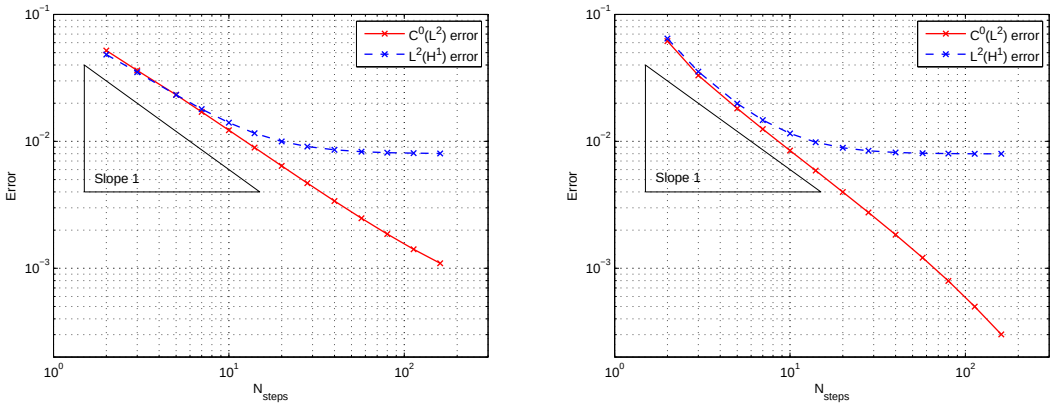
$$e^2_{\mathcal{C}^0(L^2),T} := \|u_0\|^{-2}_{\mathcal{C}^0([0,T];L_2(\Omega))} \max_{n=0\ldots N} \sum_{K\in\mathcal{T}_H} \sum_{\ell=1}^{\mathcal{L}} \omega_{K_\ell} |u_n^H(x_{K_\ell}) - u_0(x_{K_\ell}, t_n)|^2, \tag{93}$$
$$e^2_{L^2(H^1),T} := \|u_0\|^{-2}_{L^2(0,T;H_0^1(\Omega))} \frac{T}{N}{\sum_{n=0}^{N}}' \sum_{K\in\mathcal{T}_H} \sum_{\ell=1}^{\mathcal{L}} \omega_{K_\ell} |\nabla u_n^H(x_{K_\ell}) - \nabla u_0(x_{K_\ell}, t_n)|^2,$$

so that

$$e_{\mathcal{C}^0(L^2),T} \approx \frac{\|u_0 - u^H\|_{\mathcal{C}^0([0,T];L_2(\Omega))}}{\|u_0\|_{\mathcal{C}^0([0,T];L_2(\Omega))}}, \qquad e_{L^2(H^1),T} \approx \frac{\|u_0 - u^H\|_{L^2(0,T;H_0^1(\Omega))}}{\|u_0\|_{L^2(0,T;H_0^1(\Omega))}}.$$

The Gauss quadrature formula with $\mathcal{L} = 4$ nodes is chosen for quadrilateral elements, while the quadrature formula with $\mathcal{L} = 6$ nodes (the vertices, and the middles of the edges) is chosen for triangular elements. Here, the norms $\|u_0\|^2_{\mathcal{C}^0([0,T];L_2(\Omega))}$ and $\|u_0\|^2_{L^2(0,T;H_0^1(\Omega))}$, and also the quantities $u_0(x_{K_\ell}, t_n)$ and $\nabla u_0(x_{K_\ell}, t_n)$ are computed using the analytical formulas for the exact solution $u_0(x,t)$. The prime in ${\sum_{n=0}^{N}}'$ indicates that the first and last terms of the sum are divided by 2 (trapezoidal rule).

[7]The prime in ${\sum_{n=0}^{N}}'$ indicates that the first and last terms of the sum are divided by 2 (trapezoidal rule).



(a) Implicit Euler method.

(b) Chebyshev method of order 1.

Figure 1: Linear problem (94)-(97). $e_{\mathcal{C}^0(L^2),T}$ error (solid lines) and $e_{L^2(H^1),T}$ error (dashed lines) versus the number N_{steps} of timesteps. Fine macro and micro meshes ($N_{macro} = N_{micro} = 128$).

6.2 Linear case with a (non-uniform) periodic tensor

We consider the numerical resolution of multiscale parabolic problems of the form (1), on the domain $\Omega = [0,1]^2$ discretized by a uniform mesh of $N_{macro} \times N_{macro}$ $\mathcal{Q}^1$-quadrilateral elements which corresponds to $M_{macro} = \mathcal{O}(N_{macro}^2)$ degrees of freedom. The micro sampling domains (10) are discretized by a uniform mesh of $M_{micro} = N_{micro} \times N_{micro}$ with $\mathcal{Q}^1$-quadrilateral elements ($M_{micro} = \mathcal{O}(N_{micro}^2)$ DOF). On these quadrilateral elements, we consider the Gauss quadrature formula with $J = 4$ nodes for the macro and micro domains. Notice that similar results are also obtained when considering $\mathcal{P}^1$-simplicial elements, together with the quadrature formula with $J = 1$ node at the barycenter. A detailed description of the practical implementation and discretization of the above numerical method for elliptic and parabolic problems is discussed in [8].

In the linear case with a (non-uniform) periodic tensor, we consider two test problems on the time interval $(0,T)$ where $T = 1$. The first problem reads

$$\begin{aligned} &\partial_t u_\varepsilon - \nabla \cdot (a(x, x/\varepsilon, t)\nabla u_\varepsilon) = f \quad \text{in } \Omega \times (0,T) \\ &u_\varepsilon = 0 \quad \text{on } \partial\Omega \times (0,T) \\ &u_\varepsilon(x,0) = g(x) \quad \text{in } \Omega, \end{aligned} \tag{94}$$

where the tensor is time-dependent, non-symmetric, and depends on the slow variables x and on the fast variable $y = x/\varepsilon$ in both directions,

$$\begin{aligned} a(x, x/\varepsilon, t) &= \nu e^t \begin{pmatrix} (2 + x_2 \sin(\pi x_1))\omega(x_1/\varepsilon) & e^{x_1}\omega(x_2/\varepsilon) \\ (1 - x_1 x_2)\omega(x_1/\varepsilon) & (3 + x_1^2 \sin(\pi x_2))\omega(x_2/\varepsilon) \end{pmatrix}, \quad \nu = 0.1, \\ g(x) &= 16 x_1(1-x_1)x_2(1-x_2), \end{aligned} \tag{95}$$

where we consider the periodic function

$$\omega(y) = 2 + \sin(2\pi y).$$

The homogenized tensor $a^0(x,t)$ for $\varepsilon \to 0$ can by computed analytically with standard formulas and is given by the same formula as in (95) but with the oscillating terms $\omega(x_1/\varepsilon), \omega(x_2/\varepsilon)$

replaced by the constant $\sqrt{3}$ (see for instance [12, Sect. 1.2]). The expression for the source $f(x,t)$ in (94) is computed analytically using Maple (we do not reproduce it here), so that the exact solution of the homogenized problem is

$$u_0(x,t) = g(x)\cos(\pi t x_2/2). \tag{96}$$

For the second homogenization problem, we consider again (94), but with the symmetric and time-independent tensor

$$a(x,x/\varepsilon) = \nu\frac{64}{9\sqrt{17}}(\sin(2\pi x_1/\varepsilon)+\frac{9}{8})(\cos(2\pi x_2/\varepsilon)+\frac{9}{8})I_2, \quad \nu = 0.1. \tag{97}$$

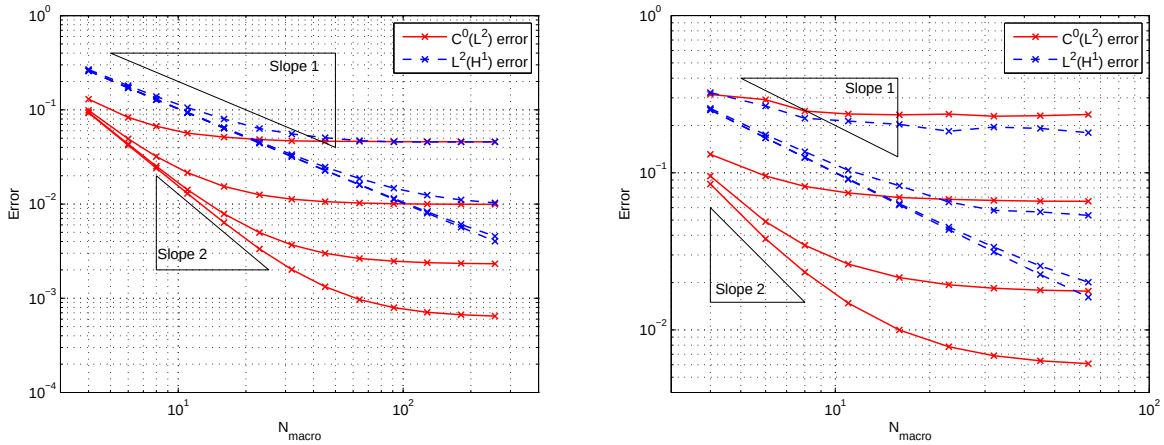
The coefficients for the tensor $a^\varepsilon(x) = a(x,x/\varepsilon)$ are chosen so that the homogenized tensor reduces to $a^0(x) = \nu I_2$ where I_2 is the identity matrix (see e.g. [33, Chap. 1.2]). The source $f(x,t)$ in (94) is again adjusted analytically so that the exact homogenized solution is (96).

6.2.1 Time discretization errors

For problem (94)-(97), we take fine macro and micro mesh sizes with $N_{macro} = N_{micro} = 128$ and we focus on the time discretization error. We consider various constant timesteps of sizes $\Delta t = T/N_{steps}$ and consider respectively the implicit Euler method (20) and the first order Chebyshev method (see (24)). For the latter method, to make the method explicit, we use a standard mass lumping technique and replace the mass matrix by a diagonal matrix. In Figure 1(a), we plot the $L^2(H^1)$ and $\mathcal{C}^0(L^2)$ errors for many different timesteps Δt for the implicit Euler method (20). We observe curves of slopes 1, as predicted by Theorem 4.7 (case $r = 1$). For small timesteps, the $L^2(H^1)$ error becomes constant, this is due to the residual spatial discretization error (which is in our case about two order of magnitude larger than the $\mathcal{C}^0(L^2)$ error). In Figure 1(a), we consider the first order Chebyshev method (24). The spectral radius of the operator A_H can be approximated by $\rho \sim 6.56\cdot 10^3 \sim 4\nu N_{macro}^2$. To guaranty the stability of the method, we thus take $s = \sqrt{2\Delta t\nu}N_{macro} + 1$ stages (strictly speaking we should choose $s = \sqrt{2\Delta t\nu}N_{macro}$; the factor one added to s is a safety factor usually taken for Chebyshev methods [1],[7]). For $N_{steps} = 2,3,5,7,10,14,20,28,40,57,80,113,160$ time steps, this corresponds respectively to $s = 41,34,27,23,19,16,14,12,10,9,7,6,6$ stages. We observe again lines of slopes 1 (recall that for the sake of brevity of the analysis, only the case of a zero source $f = 0$ is considered in the hypothesis of Theorem 4.8). Notice that for the explicit Euler method, we would have the severe stability CFL condition $\Delta t \leq 2/\rho \sim 3\cdot 10^{-4}$.

6.2.2 Space discretization errors

For problem (94)-(95), we perform $N = 400$ time steps of constant size $\Delta t = 2.5\cdot 10^{-3}$ using the implicit Euler method (20). Our numerical tests indicate that for this timestep, the time discretization error is negligible compared to the space discretization error. In Figure 2(a), we plot the $\mathcal{C}^0(L^2)$ and $L^2(H^1)$ errors for the numerical solution compared to the analytical exact solution, versus the number N_{macro} of macro elements in each space dimension. The macro mesh size is therefore $H = 1/N_{macro}$ (recall that the number of macro elements is $M_{macro} = N_{macro}\times N_{macro}$). This is done for many different values of the scaled micro mesh size $\hat{h} = h/\varepsilon = 1/N_{micro}$. When N_{macro} is small, the macro error is dominant. In agreement with Theorem 4.1, it has size $\mathcal{O}(H)$ for the $L^2(H^1)$ error (lines of slope one), whereas its size is $\mathcal{O}(H^2)$ for the $\mathcal{C}^0(L^2)$ error (lines of slope two). When N_{macro} gets larger, the micro error $\mathcal{O}\left((h/\varepsilon)^2\right)$ becomes dominant and independent of N_{micro} (horizontal lines). For the



(a) Linear problem (94)-(95).

(b) Nonlinear problem (98)-(99).

Figure 2: $e_{\mathcal{C}^0(L^2),T}$ error (solid lines) and $e_{L^2(H^1),T}$ error (dashed lines) versus N_{macro}. The lines correspond respectively to $N_{micro} = 4, 8, 16, 32$.

$\mathcal{C}^0(L^2)$ error, we observe that when N_{micro} is multiplied by 2, the micro error is divided by 4, which illustrates that the micro error decreases quadratically as predicted by Theorem 4.4 with $q = \ell = 1$. For the $L^2(H^1)$ error, one can see that the macro error dominates (when the microstructure is reasonably sampled) indicating a smaller influence of the error originating from the micro discretization. This is again in agreement with Theorem 4.4. (Notice that the curves of the $L^2(H^1)$ error are almost identical for $N_{micro} = 16, 32$ in Figure 2(a)). Finally we note that our numerical tests indicate that the $\mathcal{C}^0(H^1)$ errors behave similarly to the $L^2(H^1)$ errors, for both symmetric and non-symmetric tensors.

6.3 Non-linear problem

In this section, we investigate numerically if the error estimates derived in this paper are still valid for a class of non-linear problems. We now consider non-linear multiscale parabolic problems of a form similar to (1), but with a tensor $a(u_\varepsilon(x), x, x/\varepsilon)$ (instead of $a^\varepsilon(x)$) which depends non-linearly on u_ε. To the best of our knowledge, no numerical experiments have been given for HMM type methods applied to nonlinear parabolic PDEs.

We consider for all $u^H \in S^\ell(\Omega, \mathcal{T}_H)$ the bilinear form $B_H(u^H; \cdot, \cdot)$ defined by (16) with tensor $a^\varepsilon(x)$ replaced by $a(u^H(x_{K_j}), x, x/\varepsilon)$. Here, the micro functions $v^h_{K_j}$ (and similarly $w^h_{K_j}$) are the solutions of the micro problems (14) with the modified tensor $a^\varepsilon(u^H(x_{K_j}), x)$. Details on the implementation of such a nonlinear bilinear form can be found in [11] in the context of nonlinear elliptic problems. A similar FE-HMM method has been proposed in [38] along with a semi-discrete analysis (in space). We investigate here the influence of the micro discretization which has not been taken into account in [38]. We consider the following nonlinear test problem:

$$\begin{aligned} &\partial_t u_\varepsilon - \nabla \cdot (a(u_\varepsilon(x), x, x/\varepsilon)\nabla u_\varepsilon) = f \quad \text{in } \Omega \times (0,T) \\ &u_\varepsilon = 0 \quad \text{on } \partial\Omega \times (0,T) \\ &u_\varepsilon(x,0) = g(x) \quad \text{in } \Omega, \end{aligned} \tag{98}$$

where the tensor depends nonlinearly on u_ε as

$$\begin{aligned} a(u,x,x/\varepsilon) &= \nu(1+u^2)\frac{64}{9\sqrt{17}}(\sin(2\pi x_1/\varepsilon)+\frac{9}{8})(\cos(2\pi x_2/\varepsilon)+\frac{9}{8})I_2 \quad \nu=0.1, \\ g(x) &= 16x_1(1-x_1)x_2(1-x_2). \end{aligned} \tag{99}$$

In this situation, the homogenized tensor reduces to $a^0(u,x,y)=\nu(1+u^2)I_2$. The source $f(x,t)$ in (98) is adjusted so that the exact solution is (96).

For the integration in time, we consider the implicit Euler method. We consider a time interval of length $T=0.5$ and we perform $N=20$ time steps of size $\Delta t = 2.5\cdot 10^{-2}$. In Figure 2(b), we observe that the $\mathcal{C}^0(L^2)$ and $L^2(H^1)$ relative errors behave similarly to the linear case. When N_{macro} is small, the macro error is dominant, and we observe lines of slope one for the $L^2(H^1)$ error whereas the slope is two for the $L^2(H^1)$ error. When N_{macro} gets large, the micro error becomes dominant, we observe in the $\mathcal{C}^0(L^2)$ error that it is divided by 4 each time N_{micro} is multiplied by 2, which illustrates that the micro error decreases quadratically as for the linear case. These numerical experiments indicate that our fully discrete estimates may be valid also for nonlinear problem of the type considered (a rigorous proof of this observation remains to be done).

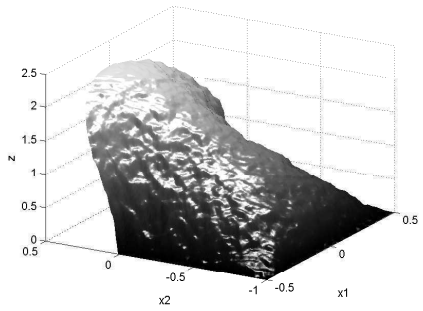
6.4 Linear problem with a random tensor

In practical situations, the tensor may not be periodic, or an analytic expression of the tensor may not be available. This is the case for instance for random models involved for the pressure equation in porous media flow [43]. We refer to [44, 14, 23, 30] for numerical homogenization results in the case of linear problems with stochastic tensors. We consider here the parabolic multiscale problem (1) with a random tensor as considered in [8, Sect. 4.2]. It corresponds to a log-normal stochastic field with mean zero and variance $\sigma=0.01$, generated by the moving ellipse average method, as described in [43, Sect. 4.1]. The role of ε which is not available here is played by the correlation lengths with known values $\varepsilon_{x_1}=0.01$ and $\varepsilon_{x_2}=0.02$. The source term is $f(x,t)=1$ and the initial condition is $g(x)=7(0.5-x_1)(0.5+x_1)(1+x_2)$. Notice that the other initial conditions as e.g. $g(x)=0$ yield similar conclusions in numerical experiments.

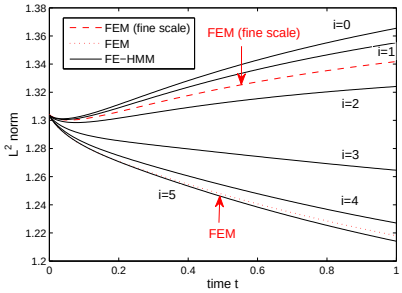
The computational domain Ω consists of a half disk meshed using 576 (macro) triangles, and a rectangle meshed using 784 (macro) quadrilaterals. Notice that for $\mathcal{P}^1$ and $\mathcal{Q}^1$-elements, this mesh corresponds to about $M_{macro}\approx 1100$ degrees of freedom. We consider mixed boundary conditions, with Dirichlet conditions on the three edges of the rectangular, and Neumann conditions on the boundary of the half disk.

Then, for $i=0,1\ldots 5$, we compute the numerical solution for different sampling domain of sizes $\delta_i=\delta_0/2^i$ where $\delta_0=0.12$ is six times larger than the correlation length $\varepsilon_{x_2}\geq\varepsilon_{x_1}$. For each index i, we fix the size of $N_{micro,i}$ so that the size $h=\delta_i/N_{micro,i}\approx 0.94\cdot 10^{-3}$ of the micro elements remains constant for all experiments.

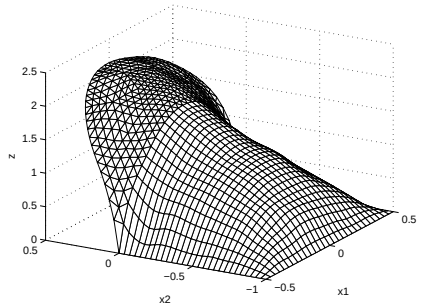
In all experiments, the solution of the parabolic problem is obtained using the implicit Euler method with stepsize $\Delta t=5\times 10^{-3}$ on a time interval of length $T=1$. The shape of the solutions at final time for experiments $i=0$ and $i=4$ are plotted in Figures 3(c) and (d). We compute a reference solution using a standard FEM with simplicial $\mathcal{P}^1$-elements with a fine mesh with about 10^6 degrees of freedom (see the solution at final time in Figure 3(a)). For this method, as the chosen size of the discretization does not resolve the fine scale, we do not expect an accurate output as confirmed by the numerical experiments.



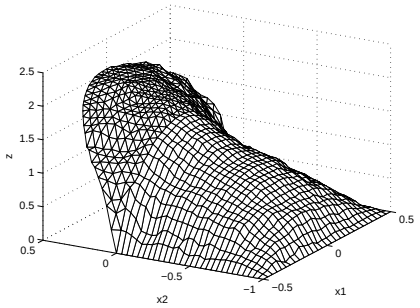
(a) Standard FEM. 10^6 DOF (finescale).



(b) Norm $\|u_n^H\|_{L^2(\Omega)}$ for various sampling domains of sizes $\delta_i = 0.12 \cdot 2^{-i}, N_{micro} = 2^{7-i}$.



(c) FE-HMM with $N_{micro} = 128$, $\delta = 0.12$. 1100 DOF (for macro mesh).



(d) FE-HMM with $N_{micro} = 16$, $\delta = 1.5 \cdot 10^{-2}$. 1100 DOF (for macro mesh).

Figure 3: Random problem of Section 6.4. Solutions at final time $t = 1$.

In Figure 3(b) we plot the L^2 norm of the solution u_n^H along the time interval $[0, T]$ for the different sampling domains with $i = 0, 1, \ldots 5$. We observe that when the sampling domains are sufficiently large compared to the correlation lengths of the random field, the profile of the solution gets closer to the reference fine scale solution. Notice that for finite elements close to the boundaries of Ω, the sampling domains may not be contained in Ω. In this case, since the tensor is given only on Ω, we use for simplicity the arbitrary value $a(x) := a(\text{argmin}_{\overline{x}\in\overline{\Omega}}\|x-\overline{x}\|)$ for $x \notin \Omega$. Here, for evaluating the L^2 norm, we use a quadrature formulas as given in (93).

7 Conclusion

We have presented and analyzed a fully discrete multiscale numerical method in space and time for parabolic homogenization problems. The analysis is valid for several classes of time integration methods (of arbitrary order) including stabilized explicit methods. The analysis has been obtained in two steps. The first step provides new convergence rates in the norms $L^2(H^1)$, $\mathcal{C}^0(L^2)$, and $\mathcal{C}^0(H^1)$ in term of the micro and macro discretization parameters. Such estimates are crucial to determine (a priori) the refinement strategy of the macro and the micro meshes so that an optimal convergence rate with a minimal computational cost can be achieved. Recently, an a posteriori error analysis has been proposed for the FE-HMM for elliptic problems [10]. Fully discrete analysis in space is also instrumental for such an

posteriori error analysis (the micro mesh has appropriately refined in the macro elements marked for refinement) has and the present work could be used to derive such an analysis for parabolic problems. The semi-discrete error bounds also allow, in a second step, to use semigroup techniques in Hilbert space to derive fully discrete error estimates in time and space for various ordinary differential equation solvers. Finally, we have discussed a generalization of the numerical method for nonlinear parabolic problems. It is proved in [11] that the error estimates in terms of the micro and macro meshes obtained for the linear case still hold for a class of nonlinear elliptic problems. Numerical results seem to indicate that this is also the case for a class of nonlinear parabolic problems. The corresponding macro-micro analysis will be reported elsewhere.

References

[1] A. Abdulle, *Fourth order Chebyshev methods with recurrence relation*, SIAM J. Sci. Comput., 23 (2002), 2041-2054.

[2] A. Abdulle, *On a-priori error analysis of the fully discrete Heterogeneous Multiscale FEM*, SIAM Multiscale Model. Simul., 4, no. 2, (2005), 447–459.

[3] A. Abdulle, *The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs,* GAKUTO Int. Ser. Math. Sci. Appl., 31 (2009), 135–184.

[4] A. Abdulle, *A priori and a posteriori analysis for numerical homogenization: a unified framework*, to appear in Higher Education Press, Beijing; World Scientific Publishing Co. Pte. Ltd., Singapore (2010).

[5] A. Abdulle, *Discontinuous Galerkin finite element heterogeneous multiscale method for elliptic problems with multiple scales*, to appear in Math. Comp. (2011).

[6] A. Abdulle and W. E, *Finite difference heterogeneous multiscale method for homogenization problems*, J. Comput. Phys., 191 (2003), 18–39.

[7] A. Abdulle and A.A. Medovikov, *Second order Chebyshev methods based on orthogonal polynomials*, Numer. Math., 90 (2001), 1-18.

[8] A. Abdulle and A. Nonnenmacher, *A short and versatile finite element multiscale code for homogenization problems*, Comput. Methods Appl. Mech. Engrg. 198 (2009), 2839–2859.

[9] A. Abdulle and C. Schwab, *Heterogeneous multiscale FEM for diffusion problems on rough surfaces*, SIAM Multiscale Model. Simul., 3, no. 1 (2005), 195–220.

[10] A. Abdulle and A. Nonnenmacher, *Adaptive finite element heterogeneous multiscale method for homogenization problems*, Comput. Methods Appl. Mech. Engrg. 200 (2011), 2710–2726.

[11] A. Abdulle and G. Vilmart, *Analysis of the finite element heterogeneous multiscale method for nonmonotone elliptic homogenization problems,* preprint. `http://infoscience.epfl.ch/record/163326`

[12] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic analysis for periodic structures*, North Holland, Amsterdam, 1978.

[13] L. Bers, F. John, and M. Schechter, *Partial differential equations*, Lectures in Applied Mathematics, Proceedings of the Summer Seminar, Boulder, CO, 1957.

[14] A. Bourgeat and A. Piatnitski, *Approximations of effective coefficients in stochastic homogenization*, Ann. I. H. Poincaré 40 (2004), 153–165.

[15] S. Brahim-Otsmane, G. A. Francfort, and F. Murat, *Correctors for the homogenization of the wave and heat equations*, J. Math. Pures Appl., 71 (1992), 197–231.

[16] P.G. Ciarlet, *Basic error estimates for elliptic problems*, Handb. Numer. Anal., Vol. 2, North-Holland, Amsterdam (1991), 17–351.

[17] P.G. Ciarlet and P.A. Raviart, *The combined effect of curved boundaries and numerical integration in isoparametric finite element method*, in A. K Aziz (Ed), Math. Foundation of the FEM with Applications to PDE, Academic Press, New York, NY, (1972), 409–474.

[18] M. Crouzeix, *Approximation of parabolic equations*, 2005. Lecture notes available at `http://perso.univ-rennes1.fr/michel.crouzeix/`

[19] E. De Giorgi and S. Spagnolo, *Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine,* Boll. Un. Mat. Ital., 4 (1973), 391–411.

[20] P. Donato and D. Cioranescu, *An introduction to Homogenization*, Oxford University Press, 1999.

[21] R. Du and P. Ming, *Heterogeneous multiscale finite element method with novel numerical integration schemes* Commun. Math. Sci., 8(4) (2010), 863–885.

[22] W. E and B. Engquist, *The Heterogeneous Multi-Scale Methods*, Commun. Math. Sci., 1 (2003), 87–132.

[23] W. E, P. Ming, and P. Zhang, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. of AMS, 18 (2005), 121–156.

[24] Y.R. Efendiev and A. Pankov. *Numerical homogenization of monotone elliptic operators*, Multiscale Model. Simul., 2 (2003), 62–79.

[25] Y.R. Efendiev and A. Pankov. *Numerical homogenization and correctors for nonlinear elliptic equations*, SIAM J. Appl Math., 65 (2004), 43–68.

[26] Y.R. Efendiev, and A. Pankov, *Numerical homogenization of nonlinear random parabolic operators*, SIAM Multiscale Model. Simul., 2 (2004), 237–268.

[27] L.C. Evans, *Partial Differential Equations*, AMS, Providence, RI, 1998.

[28] A. Gloria. *An analytical framework for the numerical homogenization of monotone elliptic operators and quasiconvex energies*, Multiscale Model. Simul., 5 (2006), 996–1043.

[29] A. Gloria. *Reduction of the resonance error - Part 1: Approximation of homogenized coefficients*, To appear in Math. Models Methods Appl. Sci. (2011).

[30] A. Gloria. *Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations*, M2AN Math. Model. Numer. Anal., 46 (2012), 1–38.

[31] E. Hairer and G. Wanner, *Solving ordinary differential equations II. Stiff and differential-algebraic problems.* Second edition, Springer-Verlag, Berlin and New York, 1992.

[32] P.J. van der Houwen and B.P. Sommeijer, *On the internal stage Runge-Kutta methods for large m-values.* Z. Angew. Math. Mech., 60 (1980), 479-485.

[33] V.V. Jikov, S.M. Kozlov, and O.A. Oleinik, *Homogenization of differential operators and integral functionals,* Springer-Verlag, 1994.

[34] O.A. Ladyzhenskaya, *The boundary value problems of mathematical physics*, Applied Mathematical Sciences, 49, Springer-Verlag New York Inc., 1985.

[35] V.I. Lebedev, *How to solve stiff systems of differential equations by explicit methods.* CRC Pres, Boca Raton, pp. 45–80, 1994.

[36] J.-L. Lions and E. Magenes, *Problème aux limites homogènes et applications,* 1, Dunod, Paris, 1968.

[37] C. Lubich and A. Ostermann, *Runge-Kutta approximation of quasi-linear parabolic equations,* Math. Comp., 64 (1995), no. 210, 601-627.

[38] P. Ming and P. Zhang, *Analysis of the heterogeneous multiscale method for parabolic homogenization problems,* Math. Comp., 76 (2007), 153-177.

[39] F. Murat and L. Tartar, *H-convergence,* in "Topics in the mathematical modeling of composite materials", A. Cherkaev and R. Kohn Eds., Birkhäuser, Boston (1997), 21–43.

[40] P.A. Raviart, *The use of numerical integration in finite element methods for solving parabolic equations*, in Miller, J. J. H. (ed.), Topics in Numerical Analysis, Academic Press (1973), 233–264.

[41] G. Samaey, D. Roose, and I. G. Kevrekidis, *The gap-tooth scheme for homogenization problems*, Multiscale Model. Simul., 4 (2005), 278–306.

[42] S. Spagnolo, *Sulla convergenza di soluzioni di equationi paraboliche ed ellittiche*, Ann. Scuola Norm. Sup. Pisa, 22 (1968), 571–597.

[43] T.C. Wallstrom, S. Hou, M.A. Christie, L.J. Durlofsky, and D.H. Sharp, *Accurate scale up of two phase flow using renormalization and nonuniform coarsening*, Comput. Geosci. 3 (1) (1999) 67–87.

[44] V.V. Yurinskiĭ. *Averaging of symmetric diffusion in random medium*, Sibirskii Matematicheskii Zhurnal, 27 (1986), 167–180.