Advances in Computational Mathematics (2019) 45:1291–1327
https://doi.org/10.1007/s10444-018-09657-7

High dimensional finite elements for time-space multiscale parabolic equations

Wee Chin Tan[1] **· Viet Ha Hoang**[1]

Received: 5 January 2018 / Accepted: 18 December 2018 /
Published online: 6 March 2019

Abstract

The paper develops the essentially optimal sparse tensor product finite element method for a parabolic equation in a domain in $\mathbb{R}^d$ which depends on a microscopic scale in space and a microscopic scale in time. We consider the critical self similar case which has the most interesting homogenization limit. We solve the high dimensional time-space multiscale homogenized equation, which provides the solution to the homogenized equation which describes the multiscale equation macroscopically, and the corrector which encodes the microscopic information. For obtaining an approximation within a prescribed accuracy, the method requires an essentially optimal number of degrees of freedom that is essentially equal to that for solving a macroscopic parabolic equation in a domain in $\mathbb{R}^d$. A numerical corrector is deduced from the finite element solution. Numerical examples for one and two dimensional problems confirm the theoretical results. Although the theory is developed for problems with one spatial microscopic scale, we show numerically that the method is capable of solving problems with more than one spatial microscopic scale.

Keywords High dimensional finite elements · Time-space multiscale parabolic equations · Optimal complexity · Numerical corrector

Mathematics Subject Classification (2010) 35B27 · 65M12 · 65M60

1 Introduction

Let $D \subset \mathbb{R}^d$ be a bounded domain where $d = 1, 2$ or 3. Let $T > 0$. Let $Y = (0, 1)^d$ be the unite cube in $\mathbb{R}^d$. We consider a symmetric matrix valued function

Communicated by: Ivan Oseledets

✉ Viet Ha Hoang
vhhoang@ntu.edu.sg

[1] Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore

$a(t, x, \tau, y) \in C([0, T] \times \bar{D} \times [0, 1] \times \bar{Y}; \mathbb{R}^{d\times d}_{sym})$. The function a is periodic with respect to τ and y, with the period being $(0, 1)$ and Y respectively; from now on, we say that it is $(0, 1) \times Y$ periodic with respect to τ and y. We assume that a is uniformly coercive and bounded, i.e, there are positive constants c_1 and c_2 such that for all $\xi, \zeta \in \mathbb{R}^d$

$$c_1|\xi|^2 \le a(t, x, \tau, y)\xi \cdot \xi, \quad a(t, x, \tau, y)\xi \cdot \zeta \le c_1|\xi||\zeta| \tag{1.1}$$

for all $(t, x, \tau, y) \in (0, T) \times D \times (0, 1) \times Y$, where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^d$. Let $\varepsilon > 0$ be a small number that represents the microscopic scale. We consider the time-space multiscale coefficient

$$a^\varepsilon(t, x) = a\left(t, x, \frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right).$$

We denote by $V = H^1_0(D)$ and $H = L^2(D)$. We note that $V \subset H \subset V'$ form a Gelfand triple. We define by $\langle\cdot,\cdot\rangle_H$ the inner product in H, extended to the duality pairing between V and V' by density. Let $T > 0$, $f \in L^2((0, T), V')$ and $g \in H$. We consider the parabolic problem

$$\begin{aligned} \frac{\partial u^\varepsilon}{\partial t} - \nabla \cdot (a^\varepsilon(t, x)\nabla u^\varepsilon) &= f(t, x), \; x \in D, \; t \in (0, T), \\ u^\varepsilon(0, x) &= g, \; x \in D \end{aligned} \tag{1.2}$$

with the Dirichlet boundary condition for $u^\varepsilon(t, \cdot)$. Problem (1.2) has a unique solution $u^\varepsilon \in L^2((0, T), V) \cap H^1((0, T), V')$ which satisfies

$$\|u^\varepsilon\|_{L^2((0,T),V)} + \|u^\varepsilon\|_{H^1((0,T),V')} \le c(\|f\|_{L^2((0,T),V')} + \|g\|_H)$$

where the constant c only depends on the constants c_1, c_2 in (1.1) and T ([31] Chapter 4).

We develop an efficient finite element (FE) method for approximating the solution to the parabolic Eq. 1.2 which depends on microscopic scales in both the temporal and spatial variables. Homogenization of (1.2) is first studied by Benssousan et al. in [5] where the general time scale ε^k is considered. However, the most interesting case is when $k = 2$ where the derivative with respect to the fast time variable plays a role in the limiting equation. When $k < 2$, this variable only plays the role of a parameter; and when $k > 2$, it is averaged out in the homogenization limit. Thus, we only consider the critical case $k = 2$.

A direct numerical discretization to take into account all the space and time scales is prohibitively expensive. For general multiscale problems, there have been extensive efforts in reducing the complexity of approximating the solutions, see, e.g. [1] and [13]. For parabolic problems with microscopic scales in both time and space variables, Efendiev and Pankov [14] employ the ideas of the multiscale finite element method (MsFEM) ([13, 22]) to perform numerical homogenization for quasilinear parabolic equations, where multiscale FE basis functions are employed which are solutions of multiscale local problems. The generalized multiscale finite element method (GMsFEM) ([15]) is used for general parabolic equations with multiple space and time scales in [10]. The Heterogeneous Multiscale Method [1] is employed by Ming and Zhang for parabolic equations that depend on multiple time scale in [27].

Owhadi and Zhang [29] construct a multiscale basis by solving a set of multiscale parabolic problems with a non-homogeneous boundary condition. Though general, the cost of these approaches can be high as microscopic meshes with respect to both time and space have to be used. For general multiscale parabolic problems whose coefficients are independent of time (and therefore are independent of microscopic time scales), we note the work of Chen, E, and Shu [8]; Abdulle and Vilmart [3]; and Abdulle and Huber [2] using heterogeneous multiscale method. Malqvist and Persson [25] use the local orthogonal decomposition technique introduced in [26] for parabolic equations that depend on spatial multiscales.

Restricting the consideration to the case where the coefficient is locally periodic with respect to both the temporal and spatial microscopic scales, we develop an essentially optimal method for finding all the necessary macroscopic and microscopic information. We employ the high dimensional finite element method to solve the multiscale homogenized equation derived from multiscale convergence which was introduced by Nguetseng [28] and developed further by Allaire [4]. The method was initiated by Hoang and Schwab [20] for multiscale elliptic problems and employed for other equations in [9, 33–35]. For parabolic monotone equations that depend only on spatial multiscales, Tan and Hoang apply the method in [30]. The method requires an essentially optimal number of degrees of freedom to approximate the solution of the high dimensional multiscale homogenized equation within a prescribed level of accuracy. It exploits the regularity of the corrector terms with respect to all the slow and the fast variables at the same time. We note that for a one dimensional two scale elliptic problem, assuming that the periodic coefficient a and the forcing f are analytic, Kazeev et al. [23] prove that an approximation for the solutions of two scale elliptic problems can be obtained with an exponential convergence rate with respect to the complexity. It is an interesting problem to study the convergence of the method in [23] when the solution only possesses Sobolev regularity, and to develop it for multiscale parabolic equations such as (1.2).

For Eq. 1.2, as the coefficient depends also on the microscopic time scale, the concept of multiscale convergence of Nguetseng [28] and Allaire [4] needs to be extended. This was first done by Holmbom et al in [21]. Using time-space multiscale convergence, the multiscale homogenized equation is derived. Solving it, we obtain the solution to the homogenized equation which describes the solution to the multiscale Eq. 1.2 macroscopically, and the corrector which encodes the microscopic information. From the FE solution of this equation, we construct a numerical corrector.

The paper is organized as follows. In Section 2, we recall the concept of multiscale convergence in both time and space, we prove several results on time space multiscale limit of a sequence of functions, and use them to derive the multiscale homogenized equation. Numerical approximation of the multiscale homogenized equation is studied in Section 3. We first consider a numerical scheme with general FE spaces and prove the convergence. We then consider the scheme using the full tensor product FE spaces and the sparse tensor product FE spaces for the corrector in Sections 3.2 and 3.3 respectively. Assuming regularity for the solution of the multiscale homogenized equation, we derive FE error estimates in terms of the mesh size. We show that the sparse tensor product FE approximation produces essentially equal level of accuracy as the full tensor product FE approximation, but uses only an essentially optimal number of degrees of freedom. In Section 4, we construct a numerical corrector from the FE solution. In Section 5, we show that the regularity required to get the error estimates for the full and sparse tensor product FE approximations hold under some regularity conditions for the coefficient of the multiscale equation, and the functions f and g in (1.2). In Section 6, we present some numerical examples in one and two dimensions to verify the FE rate of convergence for sparse tensor product FE approximations. Though we only consider the theory for the case with only one microscopic spatial scale, our method is fully capable of solving equations with more than one microscopic spatial scales, e.g., those considered in Holmbom et al. [21]. We show this by solving some examples with one microscopic time scale and two microscopic spatial scales studied in [21], using sparse tensor product FEs.

Throughout the paper, by ∇ without indicating explicitly the variable, we denote the gradient with respect to x of a function of x, or the partial gradient with respect to x of a function that depends only on the time variable t and x; and by ∇_x we denote the partial gradient with respect to x of a function depending on x and/or t and the fast variables τ and y . By # we denote spaces of periodic functions. For functions depending on time t and other variables, when we only want to emphasize the time dependence, we will only indicate the time variable t.

2 Multiscale homogenization of problem (1.2)

Benssousan et al. [5] performed the multiscale asymptotic expansion

$$u^\varepsilon(t,x) = u_0\left(t, x, \frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) + \varepsilon u_1\left(t, x, \frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_2\left(t, x, \frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right) + \dots$$

where $u_i(t, x, \tau, y)$ is periodic with respect to τ and y with the period being 1 and Y respectively. They show that u_0 does not depends on τ and y, i.e. $u_0 = u_0(t, x)$. When u_0 and u_1 are smooth, Benssousan et al. [5] Theorem 2.3 page 283 show that

$$\lim_{\varepsilon\to 0} \|u^\varepsilon(\cdot,\cdot) - u_0(\cdot,\cdot) - \varepsilon u_1\left(\cdot,\cdot,\frac{\cdot}{\varepsilon^2},\frac{\cdot}{\varepsilon}\right)\|_{L^2((0,T),V)} = 0,$$

i.e., we can use u_0 and u_1 to approximate u^ε. The function u_0 satisfies the homogenized equation, and u_1 is the corrector as derived in Section 4. These functions form the solution of Eq. 2.1 below which can be derived from multiscale convergence. Thus solving (2.1) we get these functions which are necessary for approximating u^ε. Multiscale convergence was initiated by Nguetseng in [28], and developed further by Allaire [4] which is an efficient tool to find the solution of the homogenized equation and the corrector. The concept is extended to functions depending on microscopic scales with respect to both time and space in Holmbom et al. [21]. We first recall the definition of multiscale convergence in [21]. We then prove some results on time-space multiscale convergence and use them to derive the multiscale homogenized equation of (1.2).

2.1 Multiscale convergence

We first recall the definition of Holmbom et al. [21].

Definition 2.1 A sequence $\{w^\varepsilon\}_\varepsilon$ time-space multiscale (ts-ms) converges to a function $w_0 \in L^2((0,T)\times D\times(0,1)\times Y)$ if for all functions $\phi \in C((0,T)\times D\times(0,1)\times Y)$ which are $(0,1)\times Y$ periodic with respect to τ and y,

$$\lim_{\varepsilon\to 0}\int_0^T\int_D w^\varepsilon(t,x)\phi\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)dxdt = \int_0^T\int_D\int_0^1\int_Y w_0(t,x,\tau,y)\phi(t,x,\tau,y)dyd\tau dxdt.$$

We can show that (see [21]):

Proposition 2.2 *From a bounded sequence in $L^2((0,T)\times D)$, there is a time-space multiscale convergent subsequence.*

In the following propositions, we establish the time-space multiscale convergent limits of bounded sequences in $L^2((0,T),V)\bigcap H^1((0,T),V')$ that will be employed to derive the multiscale homogenization limit of the solution of (1.2). These results are first derived in [21] (see also, e.g., Woukeng [32]).

Proposition 2.3 *Let $\{w^\varepsilon\}_\varepsilon$ be a bounded sequence in $L^2((0,T),V)\bigcap H^1((0,T),V')$. Then there are functions $w_0 \in L^2((0,T),V)$ and $w_1 \in L^2((0,T)\times D\times(0,1), H^1_\#(Y))$, and a subsequence (still denoted by $\{w^\varepsilon\}$) such that*

$$\nabla w^\varepsilon \xrightarrow{ts\text{-}ms} \nabla w_0 + \nabla_y w_1.$$

Proposition 2.4 *Let $\{w^\varepsilon\}_\varepsilon$ be a bounded sequence in $L^2((0,T),H^1(D))$ such that*

$$\nabla w^\varepsilon \xrightarrow{ts\text{-}ms} \nabla w_0 + \nabla_y w_1.$$

Then for all smooth functions $\psi(t,x,\tau,y)$ which are $(0,1)\times Y$ periodic with respect to τ and y and

$$\int_Y \psi(t,x,\tau,y)dy = 0,$$

$$\lim_{\varepsilon\to 0}\int_0^T\int_D \frac{1}{\varepsilon}w^\varepsilon(t,x)\psi\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)dxdt = \int_0^T\int_D\int_0^1\int_Y u_1(t,x,\tau,y)\psi(t,x,\tau,y)dyd\tau dxdt.$$

2.2 Multiscale homogenized equation of problem (1.2)

We have the following result on the time-space multiscale limit of the solution of the multiscale problem (1.2). We denote by $V_\#$ the subspace of $H^1_\#(Y)$ which contains functions whose integrals over Y is 0, and $H_\#$ the subspace of $L^2(Y)$ which contains function whose integrals over Y is 0. As $V_\# \subset H_\# \subset V_\#'$ form a Gelfand triple, we denote by $\langle \cdot, \cdot \rangle_{H_\#}$ the inner product in $H_\#$ extended to the duality pairing between $V_\#$ and $V_\#'$.

Proposition 2.5 *There are functions $u_0 \in L^2((0,T),V) \bigcap H^1((0,T),V')$ and $u_1 \in L^2((0,T)\times D\times(0,1),V_\#) \bigcap L^2((0,T)\times D, H^1_\#((0,1),V_\#'))$ such that we can extract a subsequence from the sequence of exact solution of (1.2) $\{u^\varepsilon\}_\varepsilon$ (still denoted as $\{u^\varepsilon\}_\varepsilon$) so that*

$$\nabla u^\varepsilon \xrightarrow{ts\text{-}ms} \nabla u_0 + \nabla_y u_1.$$

The functions u_0 and u_1 satisfy the problem

$$\begin{aligned}&\left\langle \frac{\partial u_0}{\partial t}(t,\cdot),\phi_0(\cdot)\right\rangle_H + \int_D\int_0^1 \left\langle \frac{\partial u_1}{\partial \tau}(t,x,\tau,\cdot),\phi_1(x,\tau,\cdot)\right\rangle_{H_\#} d\tau dx\\ &\quad + \int_D\int_0^1\int_Y a(t,x,\tau,y)(\nabla u_0(t,x)+\nabla_y u_1(t,x,\tau,y))\\ &\quad \cdot(\nabla\phi_0(x)+\nabla_y\phi_1(x,\tau,y))dyd\tau dx\\ &= \int_D f(t,x)\phi_0(x)dx \end{aligned} \tag{2.1}$$

for all $\phi_0 \in V$ and $\phi_1 \in L^2(D\times(0,1),V_\#)$ for almost all $t\in(0,T)$, with the initial condition $u_0(0,x)=g$.

Proof Let $\psi_0 \in C_0^\infty((0,T)\times D)$ and $\psi_1 \in C_0^\infty((0,T)\times D, C_\#^\infty((0,1),C_\#^\infty(Y)))$ be such that

$$\int_Y \psi_1(t,x,\tau,y)dy = 0$$

for all $t,x,\tau \in (0,T)\times D\times(0,1)$. Let

$$\psi^\varepsilon(t,x) = \psi_0(t,x) + \varepsilon\psi_1\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right).$$

We have from (1.2) that

$$\begin{aligned}&-\int_0^T\int_D u^\varepsilon \frac{\partial\psi^\varepsilon}{\partial t}dxdt + \int_0^T\int_D a\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)\nabla u^\varepsilon(t,x)\cdot\nabla\psi^\varepsilon(t,x)dxdt\\ &\quad = \int_0^T\int_D f(t,x)\psi^\varepsilon(t,x)dxdt.\end{aligned}$$

We note that

$$\nabla\psi^\varepsilon(t,x) = \nabla\psi_0(t,x) + \varepsilon\nabla_x\psi_1\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right) + \nabla_y\psi_1\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right).$$

We then have

$$
\begin{aligned}
-\int_0^T \int_D u^\varepsilon \left(\frac{\partial \psi_0}{\partial t}(t,x) + \varepsilon \frac{\partial \psi_1}{\partial t}\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right) + \frac{1}{\varepsilon}\frac{\partial \psi_1}{\partial \tau}\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)\right) dxdt \\
+\int_0^T \int_D a\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)\nabla u^\varepsilon(t,x)\cdot\left(\nabla\psi_0(t,x)+\varepsilon\nabla_x\psi_1\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)\right. \\
\left. +\nabla_y\psi_1\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)\right) dxdt \\
= \int_0^T \int_D f(t,x)\left(\psi_0(t,x)+\varepsilon\psi_1\left(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}\right)\right) dxdt.
\end{aligned}
$$

Passing to the multiscale convergence limit, using Propositions 2.3 and 2.4, we get

$$
\begin{aligned}
-\int_0^T\int_D u_0(t,x)\frac{\partial\psi_0}{\partial t}(t,x)dxdt - \int_0^T\int_D\int_0^1\int_Y u_1(t,x,\tau,y)\frac{\partial\psi_1}{\partial\tau}(t,x,\tau,y)dyd\tau dxdt \\
+\int_0^T\int_0^1\int_D\int_Y a(t,x,\tau,y)(\nabla u_0(t,x)+\nabla_y u_1(t,x,\tau,y))\cdot(\nabla\psi_0(t,x) \\
+\nabla_y\psi_1(t,x,\tau,y))dyd\tau dxdt \\
= \int_0^T\int_D f(t,x)\psi_0(t,x)dxdt. \qquad (2.2)
\end{aligned}
$$

By a density argument, we deduce that (2.2) holds for all $\psi_0 \in L^2((0,T),V)$ and $\psi_1 \in L^2((0,T)\times D, H^1_\#((0,1),V_\#))$. From this, we deduce (2.1). To show the initial condition $u_0(0,\cdot) = g$, we first note that $\frac{\partial u^\varepsilon}{\partial t} \rightharpoonup \frac{\partial u_0}{\partial t}$ in $L^2((0,T),V')$. Let $\psi \in C^\infty((0,T)\times D)$ so that $\psi(T,\cdot)=0$. We have

$$
\begin{aligned}
\lim_{\varepsilon\to 0}\int_0^T\int_D u^\varepsilon\frac{\partial\psi}{\partial t}dxdt &= \int_0^T\int_D u_0\frac{\partial\psi}{\partial t}dxdt \\
&= -\int_0^T\int_D\frac{\partial u_0}{\partial t}\psi dxdt - \int_D u_0(0,x)\psi(0,x)dx.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\int_0^T\int_D u^\varepsilon\frac{\partial\psi}{\partial t}dxdt &= -\int_0^T\int_D\frac{\partial u^\varepsilon}{\partial t}\psi dxdt - \int_D u^\varepsilon(0,x)\psi(0,x)dx \to \\
&\quad -\int_0^T\int_D\frac{\partial u_0}{\partial t}\psi dxdt - \int_D g(x)\psi(0,x)dx
\end{aligned}
$$

when $\varepsilon\to 0$. These imply that $u_0(0,x)=g(x)$. □

Proposition 2.5 implies that Eq. 2.1 has a solution. In the next proposition, we show that a solution of (2.1) is unique.

Proposition 2.6 *Problem* (2.1) *has a unique solution. The whole sequence* $\{u^\varepsilon\}_\varepsilon$ *time-space multiscale converges to the solution* (u_0,u_1) *of* (2.1).

Proof We show that problem (2.1) has solution $u_0 = 0$ and $u_1 = 0$ when $f = 0$ and $g = 0$. From (2.1), letting $\phi_0 = u_0(t, \cdot)$ and $\phi_1 = u_1(t, \cdot, \cdot, \cdot)$, and taking the integral from 0 to T with respect to t, we have

$$\frac{1}{2}\|u_0(T)\|_H^2 + \int_0^T \int_D \int_0^1 \int_Y a(\nabla_x u_0 + \nabla_y u_1) \cdot (\nabla_x u_0 + \nabla_y u_1) dy d\tau dx dt = 0$$

where we have used the fact that $\int_D \int_0^1 \langle \frac{\partial u_1}{\partial \tau}, u_1 \rangle_{H_\#} d\tau dx = \frac{1}{2} \int_D \int_0^1 \frac{\partial}{\partial \tau} \langle u_1, u_1 \rangle_{H_\#} d\tau dx = 0$ due to the periodicity of u_1 with respect to τ. From (1.1), we deduce that $u_0 = 0$ and $u_1 = 0$. □

Remark 2.7 To derive the multiscale homogenized problem (2.1), we need to use $a\left(t, x, \frac{t}{\varepsilon^2}, \frac{x}{\varepsilon}\right)$ as a test function. Indeed, in Definition 2.1, the test function can be extended to functions in $L^2((0, T) \times D, C((0, 1) \times Y))$ which are periodic with respect to τ and y with the period being 1 and Y respectively (see [4] for a discussion in the case of time-independent functions). Therefore to derive (2.1), we only need a to belong to $L^2((0, T) \times D, C((0, 1) \times Y))$. However, to prove the regularity of u_0 and u_1 that are required for the convergence rate of the sparse tensor product FE method developed ahead, we need more regularity for a.

Remark 2.8 For the case where $k < 2$, u_1 depends on τ but τ only plays the role of a parameter in the multiscale homogenized equation. The derivative $\frac{\partial u_1}{\partial \tau}$ is not present. For the case $k > 2$, u_1 no longer depends on τ. The problem is equivalent to that of the multiscale coefficient $\int_Y a(t, x, \tau, y) d\tau$ which is independent of τ. Details can be found in [5].

3 Finite element approximations

In this section, we establish the convergence of the Crank-Nicolson scheme for solving Eq. 2.1. We first consider convergence for general finite element spaces. We then consider the case of full tensor product and sparse tensor product FE approximations for u_1.

3.1 General finite element approximation

We denote by V_1 the space $L^2(D \times (0, 1), V_\#)$. Let $V^L \subset V$ and $V_1^L \subset L^2(D \times (0, 1), V_\#) \bigcap L^2(D, H_\#^1((0, 1), V_\#'))$ be finite element spaces. Let M be an integer. Let $\Delta t = T/M$. We consider the time sequence $0 = t_0 < t_1 < \ldots < t_M = T$ where $t_m = m\Delta t$. Let $g^L \in V^L$ be an approximation of g. Let $t_{m+1/2} = t_m + \Delta t/2$. We consider the problem: Find $U_{0,m} \in V^L$, $U_{1,m} \in V_1^L$ for $m = 1, \ldots, M$ so that

$$\left\langle \frac{U_{0,m+1} - U_{0,m}}{\Delta t}, \phi_0 \right\rangle_H + \int_D \int_0^1 \left\langle \frac{\partial}{\partial \tau} \frac{U_{1,m+1} + U_{1,m}}{2}, \phi_1 \right\rangle_{H_\#} d\tau dx$$
$$+ \int_D \int_0^1 \int_Y a(t_{m+1/2}, x, \tau, y) \left(\nabla_x \frac{U_{0,m} + U_{0,m+1}}{2} + \nabla_y \frac{U_{1,m} + U_{1,m+1}}{2} \right)$$
$$\cdot (\nabla_x \phi_0 + \nabla_y \phi_1) dy d\tau dx = \int_D f(t_{m+1/2}, x) \phi_0(x) dx, \tag{3.1}$$

for all $\phi_0 \in V^L$ and $\phi_1 \in V_1^L$. Let $u_{0,m} = u_0(t_m)$ and $u_{1,m} = u_1(t_m)$. Let $z_{0,m} = u_{0,m} - U_{0,m}$, $z_{1,m} = u_{1,m} - U_{1,m}$. For a sequence $\{w_m\}$ where $m = 0, 1, \ldots, M$ such as $\{u_{0,m}\}$, $\{u_{1,m}\}$, $\{z_{0,m}\}$ and $\{z_{1,m}\}$, we denote by

$$w_{0,m+1/2} = \frac{1}{2}(w_{0,m} + w_{0,m+1}).$$

We then have the following result.

Proposition 3.1 *Problem* (3.1) *has a unique solution.*

Proof Let $\mathbf{B}$ be the Gram matrix of the basis functions of V^L in the inner product of H. Let $\mathbf{M}$ be the matrix describing the interaction of the basis functions of V_1^L with themselves in the bilinear form representing the second term on the left hand side of (3.1). Let $\mathbf{A}$ be the matrix describing the interaction of the basis functions of $V^L \times V_1^L$ in the bilinear form representing the third term on the left hand side of (3.1). Let $\mathbf{F}_{m+1/2}$ be the vector representing the interaction of $f(t_{m+1/2})$ with the basis functions of V^L in the linear form of the right hand side in (3.1). Let $\mathbf{c}_{0,m}$ be the coordinate vector of $U_{0,m}$ in the linear expansion with respect to the basis functions of V^L. Let $\mathbf{c}_{1,m}$ be the coordinate vector of $U_{1,m}$ in the linear expansion with respect to the basis functions of V_1^L. Let $\mathbf{c}_m$ be the coordinate vector of the expansion of $(U_{0,m}, U_{1,m})$ with respect to the basis functions of $V^L \times V_1^L$, i.e. $\mathbf{c}_m = (\mathbf{c}_{0,m}, \mathbf{c}_{1,m})$. Let $\mathbf{d}_0$ be the coordinate vector of ϕ_0 in the expansion with respect to the basis functions of V^L. Let $\mathbf{d}_1$ be the coordinate vector of ϕ_1 in the basis functions of V_1^L. Let $\mathbf{d}$ be the coordinate vector of (ϕ_0, ϕ_1) in the expansion with respect to the basis functions of $V^L \times V_1^L$, i.e. $\mathbf{d} = (\mathbf{d}_0, \mathbf{d}_1)$. Problem (3.1) can be written as

$$\begin{aligned} \frac{1}{\Delta t}\mathbf{B}\mathbf{c}_{0,m+1} \cdot \mathbf{d}_0 + \frac{1}{2}\mathbf{M}\mathbf{c}_{1,m+1} \cdot \mathbf{d}_1 + \frac{1}{2}\mathbf{A}\mathbf{c}_{m+1} \cdot \mathbf{d} = \mathbf{F}_{m+1/2} \cdot \mathbf{d}_0 \\ - \frac{1}{\Delta t}\mathbf{B}\mathbf{c}_{0,m} \cdot \mathbf{d}_0 - \frac{1}{2}\mathbf{M}\mathbf{c}_{1,m} \cdot \mathbf{d}_1 - \frac{1}{2}\mathbf{A}\mathbf{c}_m \cdot \mathbf{d}. \end{aligned} \tag{3.2}$$

We denote by $\mathcal{A} : \mathbb{R}^{\dim V^L + \dim V_1^L} \times \mathbb{R}^{\dim V^L + \dim V_1^L}$ be the bilinear form

$$\mathcal{A}(\mathbf{c}, \mathbf{d}) = \frac{1}{\Delta t}\mathbf{B}\mathbf{c}_0 \cdot \mathbf{d}_0 + \frac{1}{2}\mathbf{M}\mathbf{c}_1 \cdot \mathbf{d}_1 + \frac{1}{2}\mathbf{A}\mathbf{c} \cdot \mathbf{d}.$$

Let $\psi_i(x, \tau, y)$ and $\psi_j(x, \tau, y)$ be two basis functions in V_1^L. We have that

$$\mathbf{M}_{ij} = \int_D \int_0^1 \int_Y \frac{\partial \psi_i}{\partial \tau} \psi_j dy d\tau dx.$$

Due to the periodicity of ψ_i and ψ_j with respect to τ, we have

$$\mathbf{M}_{ij} + \mathbf{M}_{ji} = \int_D \int_0^1 \int_Y \frac{\partial}{\partial \tau}(\psi_i \psi_j) dy d\tau dx = 0$$

so $\mathbf{M}_{ij} = -\mathbf{M}_{ji}$. Thus for any vector $\mathbf{c}_1 \in \mathbb{R}^{\dim V_1^L}$, $\mathbf{M}\mathbf{c}_1 \cdot \mathbf{c}_1 = 0$. Hence for $\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1) \in \mathbb{R}^{\dim V^L + \dim V_1^L}$ where $\mathbf{c}_0 \in \mathbb{R}^{\dim V^L}$ and $\mathbf{c}_1 \in \mathbb{R}^{\dim V_1^L}$

$$\mathcal{A}(\mathbf{c}, \mathbf{c}) = \frac{1}{\Delta t}\mathbf{B}\mathbf{c}_0 \cdot \mathbf{c}_0 + \frac{1}{2}\mathbf{A}\mathbf{c} \cdot \mathbf{c}.$$

As $\mathbf{B}$ is the Gram matrix, and $\mathbf{A}$ is a positive definite matrix due to (1.1), $\mathcal{A}(\mathbf{c}, \mathbf{c}) \geq c|\mathbf{c}|^2$ where the constant c is independent of $\mathbf{c}$; here $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^{\mathrm{dim V^L + dim V_1^L}}$. Therefore, the bilinear form $\mathcal{A}$ is coercive. It is also bounded. Thus, Eq. 3.2 has a unique solution. □

Theorem 3.2 *Assume that* $u_0 \in C^3([0, T], H) \cap C^2([0, T], V)$*,* $u_1 \in C^2([0, T], L^2(D \times (0, 1), V_\#))$ *and* $\frac{\partial u_1}{\partial \tau} \in C^2([0, T], L^2(D \times (0, 1), V_\#'))$*. Then*

$$\begin{aligned}
&\|z_{0,M}\|_H^2 + \Delta t \sum_{m=0}^{M-1} (\|z_{0,m+1/2}\|_V^2 + \|z_{1,m+1/2}\|_{V_1}^2) \\
&\quad \leq c\Delta t \left(\sum_{m=0}^{M-1} \|(u_0 - \tilde{u}_0)_{m+1/2}\|_V^2 + \|(u_1 - \tilde{u}_1)_{m+1/2}\|_{V_1}^2 + \left\| \frac{\partial}{\partial \tau}(u_1 - \tilde{u}_1)_{m+1/2} \right\|_{V_1'}^2 \right. \\
&\qquad \left. + \sum_{m=1}^{M-1} \left\| \frac{(u_0 - \tilde{u}_0)_{m+1/2} - (u_0 - \tilde{u}_0)_{m-1/2}}{\Delta t} \right\|_H^2 \right) \\
&\qquad + \max_{m=1,\ldots,M} \|(u_0 - \tilde{u}_0)_{m-1/2}\|_H^2 + \|g - g^L\|_H^2 + c(\Delta t)^4
\end{aligned} \tag{3.3}$$

for all $\{\tilde{u}_{0m},\ m = 0, \ldots, M\} \subset V^L$ *and* $\{\tilde{u}_{1m},\ m = 1, \ldots, M\} \subset V_1^L$.

We prove this theorem in Appendix.

3.2 Full tensor finite element

To approximate $u_1(t) \in L^2(D, L^2((0, 1), V_\#)) \bigcap L^2(D, H_\#^1((0, 1), V_\#')) \cong L^2(D) \otimes L^2(0, 1) \otimes V_\# \bigcap L^2(D) \otimes H_\#^1(0, 1) \otimes V_\#'$, we use tensor product finite elements. Let $h_l = 2^{-l}$. Assuming that D is a polygonal domain in $\mathbb{R}^d$, we divide the domain D into a hierarchy of sets of triangular simplices $\{\mathcal{T}^l\}_{l\geq 0}$. Each simplex in the set $\mathcal{T}^l$ of mesh size $O(h_l)$ is obtained by dividing each simplex in $\mathcal{T}^{l-1}$ into 4 congruent triangles when $d = 2$ and 8 tedrahedra when $d = 3$. For each simplex $T \in \mathcal{T}^l$, we denote by $P^1(T)$ the set of linear polynomials in T. Similarly, we divide Y into a hierarchy $\{\mathcal{T}_\#^l\}_{l\geq 0}$ of sets of simplices with mesh size $O(h_l)$ which are distributed periodically. For each $l = 1, 2, \ldots$, the interval $(0, 1)$ for the variable τ is divided into sets $\mathcal{T}_{\tau\#}^l$ of 2^l intervals of length 2^{-l}. We define the following FE spaces:

$$V^l = \{\phi \in H_0^1(D),\ \phi \in P^1(T)\ \forall\ T \in \mathcal{T}^l\};$$
$$V_\#^l = \{\phi \in H_\#^1(Y),\ \phi \in P^1(T)\ \forall\ T \in \mathcal{T}_\#^l\};$$
$$V_{\tau\#}^l = \{\phi \in H_\#^1(0, 1),\ \phi \in P^1(T)\ \forall\ T \in \mathcal{T}_{\tau\#}^l\}.$$

We then have the following approximation properties

$$\inf_{w^l\in V^l}\|w-w^l\|_{H^1_0(D)}\le ch_l\|w\|_{H^2(D)},\quad \forall w\in H^1_0(D)\cap H^2(D);$$
$$\inf_{w^l\in V^l}\|w-w^l\|_{L^2(D)}\le ch_l\|w\|_{H^1(D)},\quad \forall w\in H^1(D);$$
$$\inf_{w^l\in V^l_\#}\|w-w^l\|_{H^1_\#(Y)}\le ch_l\|w\|_{H^2_\#(Y)},\quad \forall w\in H^2_\#(Y);$$
$$\inf_{w^l\in V^l_\#}\|w-w^l\|_{L^2(Y)}\le ch_l\|w\|_{H^1_\#(Y)},\quad \forall w\in H^1_\#(Y);$$
$$\inf_{w^l\in V^l_{\tau\#}}\|w-w^l\|_{H^1_\#((0,1))}\le ch_l\|w\|_{H^2_\#((0,1))},\quad \forall w\in H^2_\#((0,1));$$
$$\inf_{w^l\in V^l_{\tau\#}}\|w-w^l\|_{L^2((0,1))}\le ch_l\|w\|_{H^1_\#((0,1))},\quad \forall w\in H^1_\#((0,1)).$$

For approximating $u_1(t)$ we define the full tensor product FE space as

$$\bar{V}_1^L=V^L\otimes V^L_{\tau\#}\otimes V^L_\#.$$

Let $\mathcal{H}$ be the regularity space $H^1(D,H^1_\#((0,1),H^1_\#(Y)))\cap L^2(D,H^2_\#((0,1),H^1_\#(Y)))\cap L^2(D,H^1_\#((0,1),H^2_\#(Y)))$ with the norm

$$\|w\|_{\mathcal{H}}=\|w\|_{H^1(D,H^1_\#((0,1),H^1_\#(Y)))}+\|w\|_{L^2(D,H^2_\#((0,1),H^1_\#(Y)))}+\|w\|_{L^2(D,H^1_\#((0,1),H^2_\#(Y)))}.$$

We then have the following approximation properties.

Proposition 3.3 *For all $w\in\mathcal{H}$*

$$\inf_{w^L\in \bar{V}_1^L}\|w-w_1^L\|_{L^2(D,H^1((0,1),V_\#))}\le ch_L\|w\|_{\mathcal{H}}.$$

The proof of this proposition is quite standard. It is similar to the proof for similar results in Bungartz and Griebel [7] and Hoang and Schwab [20]. We refer to these references for details. We denote the solution of the Crank-Nicolson scheme (3.1) when $V_1^L=\bar{V}_1^L$ as $\bar{U}_{0,m}$ and $\bar{U}_{1,m}$ respectively, and $z_{0,M}$, $z_{0,m+1/2}$ and $z_{1,m+1/2}$ as $\bar{z}_{0,M}$, $\bar{z}_{0,m+1/2}$ and $\bar{z}_{1,m+1/2}$. We therefore have the following result.

Theorem 3.4 *Assume that $u_0\in C^3([0,T],H)\cap C^2([0,T],V)\cap C^1([0,T],H^2(D))$, $u_1\in C^2([0,T],V_1)\cap C([0,T],\mathcal{H})$. If we choose the initial condition g^L such that $\|g-g^L\|_V\le ch_L$, then*

$$\|\bar{z}_{0,M}\|_V^2+\Delta t\sum_{m=0}^{M-1}(\|\bar{z}_{0,m+1/2}\|_V^2+\|\bar{z}_{1,m+1/2}\|_{V_1}^2)\le c((\Delta t)^4+h_L^2). \tag{3.4}$$

Proof We estimate the right hand side of (3.3). As $u_1\in C([0,T],\mathcal{H})$, we can choose $\tilde{u}_{1,m}\in V_1^L$ for $m=1,\dots,M$ such that

$$\|(u_1-\tilde{u}_1)_{m+1/2}\|_{L^2(D,H^1((0,1),V_\#))}\le ch_L(\|u_1(t_m)\|_{\mathcal{H}}+\|u_1(t_{m+1})\|_{\mathcal{H}})\le ch_L,$$

where c does not depend on t. Therefore

$$\|(u_1 - \tilde{u}_1)_{m+1/2}\|_{V_1} + \left\| \frac{\partial}{\partial \tau}(u_1 - \tilde{u}_1)_{m+1/2} \right\|_{V_1'} \le ch_L.$$

As $u_0 \in C([0,T], H^2(D)) \subset C([0,T], C(\bar{D}))$ (we consider $d = 1, 2, 3$), we can define the interpolation $I^L(u_0)(t) \in V^L$ whose value at each node equals the value of $u_0(t)$. We note that

$$\|u_0(t) - I^L(u_0)(t)\|_V \le ch_L \|u_0(t)\|_{H^2(D)} \le ch_L.$$

Choosing $\tilde{u}_0(t) = I^L(u_0)(t)$, we have

$$\|(u_0 - \tilde{u}_0)_{m+1/2}\|_V \le ch_L$$

where c does not depend on m. For the other terms in the right hand side of (3.3), we have

$$\sum_{m=1}^{M-1} \left\| \frac{(u_0 - \tilde{u}_0)_{m+1/2} - (u_0 - \tilde{u}_0)_{m-1/2}}{\Delta t} \right\|_H^2$$
$$\le \sum_{m=1}^{M-1} \frac{1}{2} \left(\left\| \frac{(u_0 - \tilde{u}_0)_{m+1} - (u_0 - \tilde{u}_0)_m}{\Delta t} \right\|_H^2 + \left\| \frac{(u_0 - \tilde{u}_0)_m - (u_0 - \tilde{u}_0)_{m-1}}{\Delta t} \right\|_H^2 \right).$$

We have

$$\left\| \frac{\partial u_0}{\partial t} - \frac{\partial \tilde{u}_0}{\partial t} \right\|_H \le ch_L \left\| \frac{\partial u_0}{\partial t} \right\|_{H^1(D)}.$$

We estimate this using the procedure of [12]

$$\left\| \frac{(u_0 - \tilde{u}_0)_{m+1} - (u_0 - \tilde{u}_0)_m}{\Delta t} \right\|_H^2 = \left\| \int_{m\Delta t}^{(m+1)\Delta t} \frac{\partial (u_0 - \tilde{u}_0)}{\partial t}(t) dt \right\|_H^2 (\Delta t)^{-2}$$
$$\le \left(\int_{m\Delta t}^{(m+1)\Delta t} \left\| \frac{\partial (u_0 - \tilde{u}_0)}{\partial t}(t) \right\|_H dt \right)^2 (\Delta t)^{-2}$$
$$\le ch_L^2 \left(\int_{m\Delta t}^{(m+1)\Delta t} \left\| \frac{\partial u_0}{\partial t}(t) \right\|_{H^1(D)} dt \right)^2 (\Delta t)^{-2}$$
$$\le ch_L^2 \left(\int_{m\Delta t}^{(m+1)\Delta t} \left\| \frac{\partial u_0}{\partial t}(t) \right\|_{H^1(D)}^2 dt \right) (\Delta t)^{-1}.$$

From this, we deduce that

$$\Delta t \sum_{m=1}^{M-1} \left\| \frac{(u_0 - \tilde{u}_0)_{m+1/2} - (u_0 - \tilde{u}_0)_{m-1/2}}{\Delta t} \right\|_H^2 \le ch_L^2.$$

We then get the conclusion. □

3.3 Sparse tensor finite elements

To define the sparse tensor FE spaces, we consider the following orthogonal projections. Let

$$P^l : L^2(D) \to V^l, \;\; P^l_{\tau\#} : H^1_\#((0,1)) \to V^l_{\tau\#}, \;\; P^l_\# : H^1_\#(Y) \to V^l_\#$$

be the orthogonal projections with respect to the inner products of $L^2(D)$, $H^1_\#((0,1))$ and $H^1_\#(Y)$ respectively. We then define the following detail spaces

$$W^l = (P^l - P^{l-1})V^l, \;\; W^l_{\tau\#} = (P^l_{\tau\#} - P^{l-1}_{\tau\#})V^l_{\tau\#}, \;\; W^l_\# = (P^l_\# - P^{l-1}_\#)V^l_\#,$$

with the convention that $P^{-1} = 0$, $P^{-1}_{\tau\#} = 0$, and $P^{-1}_\# = 0$. We note that

$$V^L = \bigoplus_{0\le l\le L} W^l, \;\; V^L_{\tau\#} = \bigoplus_{0\le l\le L} W^l_{\tau\#}, \;\; V^L_\# = \bigoplus_{0\le l\le L} W^l_\#,$$

with respect to the norms of $L^2(D)$, $H^1_\#((0,1))$ and $H^1_\#(Y)$ respectively. The full tensor finite element spaces are

$$\bar{V}^L_1 = \bigoplus_{0\le l_0,l_1,l_2\le L} W^{l_0} \bigotimes W^{l_1}_{\tau\#} \bigotimes W^{l_2}_\#.$$

We define the sparse tensor product FE spaces as

$$\hat{V}^L_1 = \bigoplus_{0\le l_0+l_1+l_2\le L} W^{l_0} \bigotimes W^{l_1}_{\tau\#} \bigotimes W^{l_2}_\#.$$

To quantify the approximation of u_1 using the spaces $\hat{V}^L_1$, we define the regularity spaces $\hat{\mathcal{H}}$ that contains functions $w \in L^2(D, H^1_\#((0,1), H^1_\#(Y)))$ so that

$$\frac{\partial^{|\alpha_0|}}{\partial x^{\alpha_0}} \frac{\partial^{\alpha_1}}{\partial \tau_1^{\alpha}} \frac{\partial^{|\alpha_2|}}{\partial y^{\alpha_2}} w \in L^2(D \times (0,1) \times Y),$$

for all $\alpha_0 \in \mathbb{N}^d_0$ so that $|\alpha_0| \le 1$, $\alpha_1 \in \{0,1,2\}$ and $\alpha_2 \in \mathbb{N}^d_0$ so that $|\alpha_2| \le 2$. In other words, $\hat{\mathcal{H}} = H^1(D, H^2_\#((0,1), H^2_\#(Y)))$. We then have the following approximation result.

Proposition 3.5 *Assume that $w \in \hat{\mathcal{H}}$. Then*

$$\inf_{w^L \in \hat{V}^L_1} \|w - w^L\|_{L^2(D, H^1_\#((0,1), H^1_\#(Y)))} \le cLh_L \|w\|_{\hat{\mathcal{H}}}.$$

The proof of this proposition is similar to those for sparse tensor product FE approximations in [7] and [20]. We refer to these references for details. We denote the solution of the Crank-Nicolson scheme in (3.1) using the sparse tensor product FE space $\hat{V}^L_1$ as $\hat{U}_{0,m}$ and $\hat{U}_{1,m}$; and denote by $z_{0,M}$, $z_{0,m+1/2}$ and $z_{1,m+1/2}$ as $\hat{z}_{0,M}$, $\hat{z}_{0,m+1/2}$ and $\hat{z}_{1,m+1/2}$ respectively. We then have

Theorem 3.6 *Assume that* $u_0 \in C^3([0,T],H)\cap C^2([0,T],V)\cap C^1([0,T],H^2(D))$, $u_1 \in C^2([0,T],V_1)\cap C([0,T],\hat{\mathcal{H}})$. *If we choose the initial condition* g^L *such that* $\|g-g^L\|_V \le cLh_L$. *Then*

$$\|\hat{z}_{0,M}\|_V^2 + \Delta t\sum_{m=0}^{M-1}(\|\hat{z}_{0,m+1/2}\|_V^2 + \|\hat{z}_{1,m+1/2}\|_{V_1}^2) \le c((\Delta t)^4 + L^2h_L^2). \quad (3.5)$$

The proof is similar to that for Theorem 3.4

Remark 3.7 The dimension of the full tensor product FE space $\bar{V}_1^L$ is $O(2^{(2d+1)L})$. The dimension of the sparse tensor product FE space $\hat{V}_1^L$ is $O(L^2 2^{dL})$ which is much less than the dimension of the full tensor product FE spaces $\bar{V}_1^L$.

Remark 3.8 Another way to construct the sparse tensor product FE spaces is to use the equivalent norms of wavelet basis functions. We assume that:

(i) For each $j \in \mathbb{N}_0^d$, there exists a set $I^j \subset \mathbb{N}_0^d$ and a set of basis functions ϕ^{jk}, $k \in I^j$, such that $V^l = \text{span}\left\{\phi^{jk} : |j|_\infty \le l\right\}$. There are constants $c_2 > c_1 > 0$ which are independent of l such that if $\phi = \sum_{|j|_\infty \le l, k\in I^j}\phi^{jk}c_{jk} \in V^l$, then :

$$c_1\sum_{\substack{|j|_\infty\le l\\ k\in I^j}} c_{jk}^2 \le \|\phi\|_{L^2(D)}^2 \le c_2\sum_{\substack{|j|_\infty\le l\\ k\in I^j}} c_{jk}^2. \quad (3.6)$$

(ii) For each $j \in \mathbb{N}_0$, there exists a set $I_0^j \subset \mathbb{N}_0$ and a set of basis functions ϕ_0^{jk}, $k \in I_0^j$, such that $V_{\tau\#}^l = \text{span}\{\phi_0^{jk} : |j|_\infty \le l\}$. There are constants $c_4 > c_3 > 0$ which are independent of l such that if $\phi = \sum_{|j|_\infty \le l, k\in I_0^j}\phi_0^{jk}c_{jk} \in V_{\tau\#}^l$, then

$$c_3\sum_{\substack{|j|_\infty\le l\\ k\in I_0^j}} c_{jk}^2 \le \|\phi\|_{H_\#^1((0,1))}^2 \le c_4\sum_{\substack{|j|_\infty\le l\\ k\in I_0^j}} c_{jk}^2. \quad (3.7)$$

(iii) For each $j \in \mathbb{N}_0^d$, there exists a set $I_1^j \subset \mathbb{N}_0^d$ and a set of basis functions $\phi_1^{jk} \in H_\#^1(Y)$, $k \in I_1^j$, such that $V_\#^l = \text{span}\{\phi_1^{jk} : |j| \le l\}$. There are constants $c_6 > c_5 > 0$ which are independent of l such that if $\phi = \sum_{|j|_\infty \le l, k\in I_1^j}\phi_1^{jk}c_{jk}$, then

$$c_5\sum_{\substack{|j|_\infty\le l\\ k\in I_1^j}} c_{jk}^2 \le \|\phi\|_{H_\#^1(Y)}^2 \le c_6\sum_{\substack{|j|_\infty\le l\\ k\in I_1^j}} c_{jk}^2. \quad (3.8)$$

Using the norm equivalences, we define

$$\begin{aligned} W^l &= \text{span}\{\phi^{jk} : |j|_\infty = l\},\quad W_{\tau\#}^l = \text{span}\{\phi_0^{jk} : |j|_\infty = l\},\\ W_\#^l &= \text{span}\{\phi_1^{jk} : |j|_\infty = l\}. \end{aligned}$$

Example: (i) We construct a basis for $L^2(0,1)$ that satisfies (3.6) as follows. We first take three continuous piecewise linear functions for level $l = 0$: ψ_1^0 obtains values $(1,0)$ at $(0,1/2)$ and is 0 in $(1/2,1)$, ψ_2^0 obtains values $(0,1,0)$ at $(0,1/2,1)$, and ψ_3^0 obtains values $(0,1)$ at $(1/2,1)$ and is 0 in $(0,1/2)$. The basis functions for other levels are constructed from the function ψ that takes values $(0,-1,2,-1,0)$ at $(0,1/2,1,3/2,2)$, the left boundary function ψ^{left} taking values $(-2,2,-1,0)$ at $(0,1/2,1,3/2)$, and the right boundary function ψ^{right} taking values $(0,-1,2,-2)$ at $(1/2,1,3/2,2)$. For levels $l\geq 1$, $I_l=\{1,2,\ldots,2^l\}$, the wavelet basis functions are defined as $\psi_1^l(x)=2^{-l/2}\psi^{left}(2^lx)$, $\psi_k^l(x)=2^{-l/2}\psi(2^lx-k+3/2)$ for $k=2,\cdots,2^l-1$ and $\psi_{2^l}^l=2^{-l/2}\psi^{right}(2^lx-2^l+2)$.

(ii) For $Y=(0,1)$, we construct a hierarchical basis for $H^1_\#(Y)/\mathbb{R}$ that satisfies (3.7) from those in (i). For level 0, we exclude ψ_1^0, ψ_3^0. At other levels, the functions ψ^{left} and ψ^{right} are replaced by the continuous piecewise linear functions that take values $(0,2,-1,0)$ at $(0,1/2,1,3/2)$ and values $(0,-1,2,0)$ at $(1/2,1,3/2,2)$ respectively.

For the d dimensional cube $(0,1)^d$, the basis functions can be constructed by taking the tensor products of the basis functions in $(0,1)$. They satisfy the norm equivalence after appropriate scaling, see [18]. Examples on wavelet basis functions on regular triangular mesh can be found in, e.g., Bieri et al. [6].

Remark 3.9 We can construct the sparse tensor product by using the hierarchies $\{\mathcal{T}^l\}$, $\{\mathcal{T}^l_{\tau\#}\}$ and $\{\mathcal{T}^l_\#\}$. We denote by $\mathcal{S}^l$ the new nodes belonging to the set of simplices $\mathcal{T}^l$ but not the set of simplices $\mathcal{T}^{l-1}$. We let W_l be the set of basis functions in V^l which equals 1 at one of the nodes of $\mathcal{S}^l$ and equals 0 at other nodes. We construct the spaces $W^l_{\tau\#}$ and $W^l_\#$ similarly. The estimate for the sparse tensor product FE spaces still holds.

4 Numerical correctors

We employ the FE solutions for the multiscale homogenized problems (2.1) to construct numerical correctors in this section. We first establish the homogenized equation from (2.1). Letting $\phi_0=0$, we have

$$\int_D\int_0^1\left\langle\frac{\partial u_1}{\partial\tau}(t,x,\tau,\cdot),\phi_1(x,\tau,\cdot)\right\rangle_{H_\#}d\tau dx$$
$$+\int_D\int_0^1\int_Y a(t,x,\tau,y)(\nabla u_0(t,x)+\nabla_y u_1(t,x,\tau,y))\cdot\nabla_y\phi_1(x,\tau,y)dyd\tau dx=0$$

$\forall\,\phi_1\in L^2(D\times(0,1),V_\#)$. We therefore deduce that the solution u_1 can be written as

$$u_1(t,x,\tau,y)=\frac{\partial u_0}{\partial x_i}(t,x)N^i(t,x,\tau,y) \tag{4.1}$$

where $N^i(t,x,\tau,y)\in L^2((0,T)\times D\times(0,1),V_\#)\bigcap L^2((0,T)\times D,H^1_\#((0,1),V'_\#))$ is the unique solution of the problem

$$\frac{\partial N^i}{\partial\tau}-\nabla_y\cdot(a(e^i+\nabla_y N^i))=0. \tag{4.2}$$

Here e^i is the ith vector of the standard basis in $\mathbb{R}^d$. The existence of a unique solution of (4.2) is proved in Lions and Magenes [24] Chapter 3 Section 6.2. Due to the periodicity with respect to y, we have $\int_Y \frac{\partial N^i}{\partial \tau} dy = 0$ so fixing x and t, for all τ, $\int_Y N^i dy = c(t,x)$. Choosing the solution as $N^i - c(t,x)$, we have $\int_Y N^i(t,x,\tau,y)dy = 0$. Then letting $\phi_1 = 0$ in (2.1) we have

$$\left\langle \frac{\partial u_0}{\partial t}(t,\cdot), \phi_0(\cdot) \right\rangle_H + \int_D \int_0^1 \int_Y a(t,x,\tau,y)(\nabla u_0(t,x) + \nabla_y u_1(t,x,\tau,y)) \cdot \nabla \phi_0(x) dy d\tau dx = \int_D f(t,x)\phi_0(x)dx$$

$\forall \phi_0 \in V$. Using (4.1), we have

$$\begin{aligned}
&\int_0^1 \int_Y a_{ik}(t,x,\tau,y)(\frac{\partial u_0}{\partial x_k}(t,x) + \frac{\partial u_1}{\partial y_k}(t,x,\tau,y))dy d\tau \\
&= \int_0^1 \int_Y a_{ik}(t,x,\tau,y)(\frac{\partial u_0}{\partial x_k}(t,x) + \frac{\partial u_0}{\partial x_j}(t,x)\frac{\partial N^j}{\partial y_k}(t,x,\tau,y))dy d\tau \\
&= \int_0^1 \int_Y a_{ik}(t,x,\tau,y)(\delta_{jk} + \frac{\partial N^j}{\partial y_k}(t,x,\tau,y))\frac{\partial u_0}{\partial x_j}(t,x)dy d\tau.
\end{aligned}$$

We therefore have

$$\begin{aligned}
&\int_D \int_0^1 \int_Y a(t,x,\tau,y)(\nabla u_0(t,x) + \nabla_y u_1(t,x,\tau,y)) \cdot \nabla \phi_0(x) dy d\tau dx \\
&= \int_D \left(\int_0^1 \int_Y a_{ik}(t,x,\tau,y)(\delta_{jk} + \frac{\partial N^j}{\partial y_k}(t,x,\tau,y))dy d\tau \right) \frac{\partial u_0}{\partial x_j}(t,x)\frac{\partial \phi}{\partial x_i}(x)dx.
\end{aligned}$$

Thus the function u_0 satisfies the homogenized equation

$$\frac{\partial u_0}{\partial t} - \nabla \cdot (a^0 \nabla u_0) = f, \tag{4.3}$$

$$u_0(0,\cdot) = g \tag{4.4}$$

where the homogenized coefficient a^0 is defined as

$$a^0_{ij}(t,x) = \int_0^1 \int_Y a_{ik}(t,x,\tau,y) \left(\delta_{jk} + \frac{\partial N^j}{\partial y_k}(t,x,\tau,y) \right) dy d\tau. \tag{4.5}$$

As shown in [5], $u^\varepsilon \rightharpoonup u_0$ in $L^2((0,T),V)$. The homogenized Eq. 4.3 represents (1.2) macroscopically. It is well-posed as the coefficient a^0 in (4.5) is positively definite. Indeed, from (4.2), we have

$$\int_0^1 \int_Y \frac{\partial N^j}{\partial \tau}(t,x,\tau,y)N^i(t,x,\tau,y)dy d\tau + \int_0^1 \int_Y a_{lk}(t,x,\tau,y) \left(\delta_{jk} + \frac{\partial N^j}{\partial y_k} \right) \frac{\partial N^i}{\partial y_l} dy d\tau = 0.$$

Thus

$$a^0_{ij}(t,x) = \int_0^1 \int_Y a_{lk} \left(\delta_{jk} + \frac{\partial N^j}{\partial y_k} \right) \left(\delta_{il} + \frac{\partial N^i}{\partial y_l} \right) - \int_0^1 \int_Y \frac{\partial N^j}{\partial \tau} N^i dy d\tau.$$

Let ξ be a vector in $\mathbb{R}^d$. We have

$$\begin{aligned} a^0_{ij}(t,x)\xi_i\xi_j \;=& \int_0^1 \int_Y a_{lk} \left(\xi_k + \frac{\partial N^j}{\partial y_k}\xi_j \right) \left(\xi_l + \frac{\partial N^i}{\partial y_l}\xi_i \right) dy d\tau \\ & - \int_0^1 \int_Y \frac{\partial N^j}{\partial \tau} N^i dy d\tau \xi_i \xi_j. \end{aligned}$$

From the uniform ellipticity of a, we have

$$\begin{aligned} \int_0^1\int_Y a_{lk}\left(\xi_k + \frac{\partial N^j}{y_k}\xi_j\right)\left(\xi_l + \frac{\partial N^i}{\partial y_l}\xi_i\right) dy d\tau &\geq c\sum_{k=1}^d \int_0^1\int_Y \left(\xi_k + \frac{\partial N^j}{\partial y_k}\xi_j\right)^2 dy d\tau \\ &= c\sum_{k=1}^d \xi_k^2 + c\sum_{k=1}^d \int_0^1\int_Y \left(\frac{\partial N^j}{\partial y_k}\xi_j\right)^2 dy d\tau \\ &\geq c\sum_{k=1}^d \xi_k^2 \end{aligned}$$

due to the periodicity with respect to y of N_j. We further have

$$\begin{aligned} \int_0^1 \int_Y \frac{\partial N^j}{\partial \tau} N^i dy d\tau \xi_i \xi_j \;=& \sum_{i=1}^d \int_0^1 \int_Y \frac{\partial N^i}{\partial \tau} N^i dy d\tau \xi_i^2 \\ & + \sum_{\substack{i\neq j \\ i,j=1,\dots,d}} \int_0^1 \int_Y \left(\frac{\partial N^j}{\partial \tau} N^i + \frac{\partial N^i}{\partial \tau} N^j \right) dy d\tau \xi_i \xi_j. \end{aligned}$$

Due to the periodicity of N^i, N^j with respect to τ, we have

$$\int_0^1 \int_Y \left(\frac{\partial N^j}{\partial \tau} N^i + \frac{\partial N^i}{\partial \tau} N^j \right) dy d\tau = \int_0^1 \int_Y \frac{\partial (N^i N^j)}{\partial \tau} dy d\tau = 0.$$

Therefore

$$\int_0^1 \int_Y \frac{\partial N^j}{\partial \tau} N^i dy d\tau \xi_i \xi_j = 0.$$

We thus get the uniform coercivity of a^0 with respect to t and x.

We note that u^ε only converges weakly to u_0 in $L^2((0,T),V)$ when $\varepsilon \to 0$. We now derive a corrector result i.e. a computable function that approximates u^ε in the norm of $L^2((0,T),V)$. We first consider the operator $\mathcal{T}^\varepsilon : L^1((0,T)\times D) \to L^1((0,T)\times D \times (0,1) \times Y)$ as

$$\mathcal{T}^\varepsilon(\Phi) = \Phi\left(\varepsilon^2 \left[\frac{t}{\varepsilon^2} \right]_1 + \varepsilon^2\tau, \varepsilon\left[\frac{x}{\varepsilon}\right] + \varepsilon y \right) \tag{4.6}$$

where $[\cdot]_1$ denotes the integer part of a real number, and $[\cdot]_Y$ denotes the "interger part" with respect to the unit cube Y of a vector in $\mathbb{R}^d$. Let D^ε be the 2ε neighbourhood of D. We have

$$\int_0^T \int_D \Phi(t,x)dxdt = \int_{-2\varepsilon}^{T+2\varepsilon} \int_{D^\varepsilon} \int_0^1 \int_Y \mathcal{T}^\varepsilon(\Phi)(t,x,\tau,y)dyd\tau dxdt; \quad (4.7)$$

Φ is extended by 0 outside $(0,T)\times D$ (see, e.g., [11]). We now show that if $\{w^\varepsilon\}_\varepsilon$ time-space multiscale converges to w_0 in $L^2((0,T)\times D\times(0,1)\times Y)$ then

$$\mathcal{T}^\varepsilon(w^\varepsilon) \rightharpoonup w_0 \text{ in } L^2((0,T)\times D\times(0,1)\times Y).$$

Let $\{w^\varepsilon\}_\varepsilon$ be a bounded sequence in $L^2((0,T)\times D)$ that time-space multiscale converges to w^0. We note that $(\mathcal{T}^\varepsilon(w^\varepsilon))^2 = \mathcal{T}((w^\varepsilon)^2)$. We have from (4.7)

$$\begin{aligned}&\int_{-2\varepsilon}^{T+2\varepsilon} \int_{D^\varepsilon} \int_0^1 \int_Y \mathcal{T}^\varepsilon(w^\varepsilon)(t,x,\tau,y)^2 dyd\tau dxdt\\ &= \int_{-2\varepsilon}^{T+2\varepsilon} \int_{D^\varepsilon} \int_0^1 \int_Y \mathcal{T}^\varepsilon((w^\varepsilon)^2)(t,x,\tau,y)dyd\tau dxdt = \int_0^T \int_D w^\varepsilon(t,x)^2 dxdt.\end{aligned}$$

Thus $\mathcal{T}^\varepsilon(w^\varepsilon)$ is bounded in $L^2((0,T)\times D\times(0,1)\times Y)$ so we can extract a subsequence that weakly converges. Let $\psi(t,x,\tau,y)$ be a smooth function that is $(0,1)\times Y$ periodic with respect to τ and y. We have

$$\begin{aligned}&\lim_{\varepsilon\to0} \int_0^T \int_D w^\varepsilon(t,x)\psi(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})dxdt\\ &= \int_0^T \int_D \int_0^1 \int_Y w_0(t,x,\tau,y)\psi(t,x,\tau,y)dyd\tau dxdt.\end{aligned}$$

On the other hand, from (4.7), we have

$$\begin{aligned}&\int_0^T \int_D w^\varepsilon(t,x)\psi(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})dxdt\\ &= \int_{-2\varepsilon}^{T+2\varepsilon} \int_{D^\varepsilon} \int_0^1 \int_Y \mathcal{T}^\varepsilon(w^\varepsilon)\psi(\varepsilon^2\left[\frac{t}{\varepsilon^2}\right]_1 + \varepsilon^2\tau, \varepsilon\left[\frac{x}{\varepsilon}\right] + \varepsilon y, \tau, y)dyd\tau dxdt,\end{aligned}$$

where we have used the periodicity of ψ. Since ψ is smooth in $(0,T)\times D\times(0,1)\times Y$, we have

$$\left|\psi(\varepsilon^2\left[\frac{t}{\varepsilon^2}\right]_1 + \varepsilon^2\tau, \varepsilon\left[\frac{x}{\varepsilon}\right] + \varepsilon y, \tau, y) - \psi(t,x,\tau,y)\right| \le c\varepsilon$$

where c is independent of t, x, τ and y. Therefore

$$\begin{aligned}&\lim_{\varepsilon\to0} \int_0^T \int_D w^\varepsilon(t,x)\psi(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})dxdt\\ &= \lim_{\varepsilon\to0} \int_0^T \int_D \int_0^1 \int_Y \mathcal{T}^\varepsilon(w^\varepsilon)\psi(t,x,\tau,y)dyd\tau dxdt,\end{aligned}$$

i.e.

$$\lim_{\varepsilon\to 0}\int_0^T\int_D\int_0^1\int_Y \mathcal{T}^\varepsilon(w^\varepsilon)\psi(t,x,\tau,y)dyd\tau dxdt$$
$$=\int_0^T\int_D\int_0^1\int_Y w_0(t,x,\tau,y)\psi(t,x,\tau,y)dyd\tau dxdt.$$

As the space of smooth functions which are $(0,1)\times Y$ periodic with respect to τ and y is dence in $L^2((0,T)\times(0,1)\times D\times Y)$, we deduce that the weak limit of $\mathcal{T}^\varepsilon(w^\varepsilon)$ in $L^2((0,T)\times(0,1)\times D\times Y)$ is w_0.

To establish the numerical correctors, we define the operator $\mathcal{U}^\varepsilon : L^1((0,T)\times D\times(0,1)\times Y)\to L^1((0,T)\times D)$ as

$$\mathcal{U}^\varepsilon(\Phi)=\int_0^1\int_Y \Phi\left(\varepsilon^2\left[\frac{t}{\varepsilon^2}\right]_1+\varepsilon^2\theta,\varepsilon\left[\frac{x}{\varepsilon}\right]+\varepsilon z,\left\{\frac{t}{\varepsilon^2}\right\}_1,\left\{\frac{x}{\varepsilon}\right\}\right)dzd\theta \tag{4.8}$$

where $[\cdot]_1$ and $\{\cdot\}_1$ denote the integer and the fractional parts of a real number, and $[\cdot]$ and $\{\cdot\}$ denote the "integer" and the "fractional" part with respect to the unit cube Y of a vector in $\mathbb{R}^d$. Let $D^{2\varepsilon}$ be the 2ε neighbourhood of D. We have that

$$\int_{-2\varepsilon^2}^{T+2\varepsilon^2}\int_{D^{2\varepsilon}}\mathcal{U}^\varepsilon(\Phi)(t,x)=\int_0^T\int_D\int_0^1\int_Y\Phi(t,x,\tau,y)dyd\tau dxdt \tag{4.9}$$

for all $\Phi\in L^1((0,T)\times D\times(0,1)\times Y)$. The proof of these facts can be found in [11]. We first establish the following result.

Proposition 4.1 *For the solution of* Eq. 2.1

$$\lim_{\varepsilon\to 0}\|\nabla u^\varepsilon-\nabla u_0-\mathcal{U}^\varepsilon(\nabla_y u_1)\|_{L^2((0,T)\times D)}=0.$$

Proof First we note that as $\nabla u^\varepsilon\overset{\text{ts-ms}}{\longrightarrow}\nabla u_0+\nabla_y u_1$, $\mathcal{T}^\varepsilon(\nabla u^\varepsilon)\rightharpoonup\nabla u_0+\nabla_y u_1$ in $L^2((0,T)\times D\times(0,1)\times Y)$. Let

$$\begin{aligned}I\ =&\int_0^T\left\langle\frac{\partial u^\varepsilon}{\partial t}-\frac{\partial u_0}{\partial t},u^\varepsilon-u_0\right\rangle_H\\&+\int_0^T\int_D\int_0^1\int_Y\mathcal{T}^\varepsilon(a(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}))(\mathcal{T}^\varepsilon(\nabla u^\varepsilon)-(\nabla u_0+\nabla_y u_1))\\&\cdot(\mathcal{T}^\varepsilon(\nabla u^\varepsilon)-(\nabla u_0+\nabla_y u_1))dyd\tau dxdt.\end{aligned}$$

From (1.2), (2.1), and the fact that $\mathcal{T}^\varepsilon(a(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon}))\to a(t,x,\tau,y)$ pointwise, we have

$$\begin{aligned}\lim_{\varepsilon\to 0}I&=\lim_{\varepsilon\to 0}\int_0^T\left\langle\frac{\partial u^\varepsilon}{\partial t},u^\varepsilon\right\rangle_H-\left\langle\frac{\partial u_0}{\partial t},u_0\right\rangle_H dt+\int_0^T\int_D a(t,x,\frac{t}{\varepsilon^2},\frac{x}{\varepsilon})\nabla u^\varepsilon\cdot\nabla u^\varepsilon dxdt\\&\quad-\int_0^T\int_D\int_0^1\int_Y a(t,x,\tau,y)(\nabla u_0+\nabla_y u_1)\cdot(\nabla u_0+\nabla_y u_1)dyd\tau dxdt\\&=\lim_{\varepsilon\to 0}\int_0^T\int_D fu^\varepsilon dx-\int_0^T\int_D fu_0dx\\&=0\end{aligned}$$

due to $u^\varepsilon \rightharpoonup u_0$ in $L^2((0,T),V)$ when $\varepsilon \to 0$. As $u^\varepsilon(0) = u_0(0) = g$, using the coercivity of a, we have

$$\lim_{\varepsilon\to 0} \|u^\varepsilon(T) - u_0(T)\|_H + \|\mathcal{T}^\varepsilon(\nabla u^\varepsilon) - (\nabla u_0 + \nabla_y u_1)\|_{L^2((0,T)\times D\times(0,1)\times Y)} = 0.$$

From (4.8), we have $(\mathcal{U}^\varepsilon(\Phi)(t,x))^2 \le \mathcal{U}^\varepsilon(\Phi^2)(t,x)$. From (4.9), we have

$$\|\mathcal{U}^\varepsilon(\Phi)\|^2_{L^2((0,T)\times D)} \le \|\mathcal{U}^\varepsilon(\Phi^2)\|_{L^1((0,T)\times D))} \le \|\Phi\|^2_{L^2((0,T)\times D\times(0,1)\times Y)}.$$

We therefore have

$$\begin{aligned}&\|\mathcal{U}^\varepsilon(\mathcal{T}^\varepsilon(\nabla u^\varepsilon) - (\nabla u_0 + \nabla_y u_1))\|_{L^2((0,T)\times D)}\\ &\le \|\mathcal{T}^\varepsilon(\nabla u^\varepsilon) - (\nabla u_0 + \nabla_y u_1)\|_{L^2((0,T)\times D\times(0,1)\times Y)} \to 0\end{aligned}$$

when $\varepsilon \to 0$. Using $\mathcal{U}^\varepsilon(\mathcal{T}^\varepsilon(\nabla u^\varepsilon)) = \nabla u^\varepsilon$, we get the conclusion. □

For the full tensor product FE approximation, we define the functions $\bar{U}_0 : (0,T) \to V$ and $\bar{U}_1 : (0,T) \to L^2(D\times(0,1), V_\#)$ as

$$\bar{U}_0(t) = \frac{1}{2}(\bar{U}_{0,m} + \bar{U}_{0,m+1}), \quad \bar{U}_1(t) = \frac{1}{2}(\bar{U}_{1,m} + \bar{U}_{1,m+1}) \text{ for } t \in [t_m, t_{m+1}).$$

We then have the following approximation

Theorem 4.2 *Assume that the hypothesis of Theorem 3.4 hold. Then for the solution of the numerical scheme (3.1) using the full tensor product FEs, we have*

$$\lim_{\substack{L\to\infty\\ \varepsilon\to 0}} \left\|\nabla u^\varepsilon - \nabla\bar{U}_0 - \mathcal{U}^\varepsilon(\nabla_y(\bar{U}_1))\right\|_{L^2((0,T)\times D)} = 0,$$

i.e. for all $\delta > 0$ we can find $L_0 > 0$ and $\varepsilon_0 > 0$ such that if $L > L_0$ and $\varepsilon < \varepsilon_0$

$$\left\|\nabla u^\varepsilon - \nabla\bar{U}_0 - \mathcal{U}^\varepsilon(\nabla_y(\bar{U}_1))\right\|_{L^2((0,T)\times D)} < \delta.$$

Proof We note that

$$\begin{aligned}\int_0^T \|\nabla u_0(t) - \nabla\bar{U}_0(t)\|^2_H dt &= \sum_{m=0}^{M-1}\int_{m\Delta t}^{(m+1)\Delta t} \|\nabla u_0(t) - \nabla\bar{U}_0(t)\|^2_H dt\\ &\le \sum_{m=0}^{M-1}\left(\Delta t\|\nabla u_0(t_{m+1/2}) - \nabla\bar{U}_0(t_{m+1/2}))\|^2_H + c(\Delta t)^3\right),\end{aligned}$$

where we have used the midpoint approximation for the integral. As $u_0 \in C^2([0,T],V)$, the constant c is independent of m. We note that

$$\left\|\frac{1}{2}(\nabla u_0(t_m) + \nabla u_0(t_{m+1})) - \nabla u_0(t_{m+1/2})\right\|_H \le c(\Delta t)^2$$

Therefore

$$\int_0^T \|\nabla u_0(t)-\nabla\bar{U}_0(t)\|_H^2 dt \le \sum_{m=0}^{M-1}\Big(\Delta t\|\nabla u_{0,m+1/2}-\nabla\bar{U}_0(t_{m+1/2}))\|_H^2+c(\Delta t)^3\Big)$$
$$=\Delta t\sum_{m=0}^{M-1}\|\nabla\bar{z}_{0,m+1/2}\|_H^2+O((\Delta t)^2)$$
$$\le c((\Delta t)^2+h_L^2).$$

Similarly we have

$$\int_0^T \|\nabla_y u_1(t)-\nabla_y\bar{U}_1(t)\|^2_{L^2(D\times(0,1)\times Y)}dt \le \Delta t\sum_{m=0}^{M-1}\|\nabla_y\bar{z}_{0,m+1/2}\|^2_{L^2(D\times(0,1)\times Y)}$$
$$+O((\Delta t)^2)\le c((\Delta t)^2+h_L^2).$$

From (4.8), we have that $(\mathcal{U}^\varepsilon(\Phi)(t,x))^2\le\mathcal{U}^\varepsilon(\Phi^2)(t,x)$. Therefore, from (4.9), we have

$$\|\mathcal{U}^\varepsilon(\Phi)\|_{L^2((0,T)\times D)}\le\|\mathcal{U}^\varepsilon(\Phi^2)\|_{L^1((0,T)\times D)}\le\|\Phi\|_{L^2((0,T)\times D\times(0,1)\times Y)}.$$

From this, we have

$$\|\mathcal{U}^\varepsilon(\nabla_y u_1-\nabla_y\bar{U}_1)\|_{L^2((0,T)\times D)}\le\|\nabla_y u_1-\nabla_y\bar{U}_1\|_{L^2((0,T)\times D\times(0,1)\times Y)}\le c(\Delta t+h_L).$$

From this, we have

$$\|\nabla u^\varepsilon-\nabla\bar{U}_0-\mathcal{U}^\varepsilon(\nabla_y\bar{U}_1)\|_{L^2((0,T)\times D)}\le\|\nabla u^\varepsilon-\nabla u_0-\mathcal{U}^\varepsilon(\nabla_y u_1)\|_{L^2((0,T)\times D)}$$
$$+\|\nabla u_0-\nabla\bar{U}_0\|_{L^2((0,T)\times D)}+\|\mathcal{U}^\varepsilon(\nabla_y u_1)-\mathcal{U}^\varepsilon(\nabla_y\bar{U}_1)\|_{L^2((0,T)\times D)}\to 0$$

when $L\to\infty$ and $\varepsilon\to 0$. □

For the sparse tensor product approximations, we define

$$\hat{U}_0(t)=\frac{1}{2}(\hat{U}_{0,m}+\hat{U}_{0,m+1}),\quad \hat{U}_1(t)=\frac{1}{2}(\hat{U}_{1,m}+\hat{U}_{1,m+1})\ \text{ for } t\in[t_m,t_{m+1}).$$

We have:

Theorem 4.3 *Assume that the hypothesis of Theorem 3.6 hold. Then for the solution of the numerical scheme* (3.1) *using the sparse tensor product FEs, we have*

$$\lim_{\substack{L\to\infty\\ \varepsilon\to 0}}\left\|\nabla u^\varepsilon-\nabla\hat{U}_0-\mathcal{U}^\varepsilon(\nabla_y(\hat{U}_1))\right\|_{L^2((0,T)\times D)}=0.$$

The proof of this theorem is similar to that for Theorem 4.2.

Remark 4.4 Geng and Shen [17] deduce a corrector with a convergence rate in terms of the microscopic scale ε in the $H^1(D)$ norm for the solution u^ε. This corrector involves functions other than u_0 and u_1, which cannot be found from problem (2.1). The corrector of [17] involves a parabolic smoothness operator for the solution u_0. We are not aware of a simple corrector with an explicit homogenization rate of convergence in terms of ε similar to that for elliptic problems, i.e., a rate of convergence for the limit in Proposition 4.1 in terms of ε. However, from Theorem 1.1 of [17], if $u_0 \in C([0,T], H^1(D))$, then

$$\|u^\varepsilon - u_0\|_{L^2((0,T)\times D)} \le c\varepsilon.$$

Using this, we will have

$$\|u^\varepsilon - \bar{U}_0\|_{L^2((0,T)\times D)} \le c(\varepsilon + h_L)$$

for the solution of the full tensor product FE approximation, and

$$\|u^\varepsilon - \hat{U}_0\|_{L^2((0,T)\times D)} \le c(\varepsilon + L h_L)$$

for the solution of the sparse tensor product FE approximation.

5 Regularity

We show that the regularity required on the solution u_0 and u_1 of the multiscale homogenized problem (2.1) for obtaining the full and sparse tensor product FE errors and for obtaining the corrector hold under regularity conditions for the coefficients and the initial condition. We have the following results.

Proposition 5.1 *Assume that* $a \in C^3([0,T], C^3(\bar{D}, C([0,T]\times Y)))$, $f \in H^3((0,T), V')$, $f(0) \in H_0^3(D)$, $\frac{\partial f}{\partial t}(0) \in H_0^2(D)$, $\frac{\partial^2 f}{\partial t^2}(0) \in H$, *and* $g \in H_0^4(D)$, *then* $u_0 \in C^3([0,T], H) \cap C^2([0,T], V)$. *Further, if* $f \in H^2((0,T), H)$ *and if the domain* D *is convex, then* $u_0 \in C^1([0,T], H^2(D))$.

Proof As $a \in C^3([0,T], C^3(\bar{D}, C([0,T]\times Y)))$, from (4.5) we deduce that $a^0 \in C^3([0,T], C^3(\bar{D}))$. From the condition, we have

$$\frac{\partial}{\partial t}\frac{\partial u_0}{\partial t} - \nabla\cdot\left(a^0\nabla\frac{\partial u_0}{\partial t}\right) = \frac{\partial f}{\partial t} + \nabla\cdot\left(\frac{\partial a^0}{\partial t}\nabla u_0\right) \in L^2((0,T), V'),$$

$$\frac{\partial u_0}{\partial t}(0) = f(0) + \nabla\cdot(a^0\nabla g) \in H,$$

so $\frac{\partial u_0}{\partial t} \in L^2((0,T), V) \cap C([0,T], H)$. Therefore

$$-\nabla\cdot\left(a^0\nabla u_0\right) = f - \frac{\partial u_0}{\partial t} \in C([0,T], H)$$

so $u_0 \in C([0,T],H^2(D))$. We then have

$$\frac{\partial}{\partial t}\frac{\partial^2 u_0}{\partial t^2} - \nabla\cdot\left(a^0\nabla\frac{\partial^2 u_0}{\partial t^2}\right) = \frac{\partial^2 f}{\partial t^2} + \nabla\cdot\left(\frac{\partial^2 a^0}{\partial t^2}\nabla u_0\right) + 2\nabla\cdot\left(\frac{\partial a^0}{\partial t}\nabla\frac{\partial u_0}{\partial t}\right)$$
$$\in L^2((0,T),V'),$$
$$\frac{\partial^2 u_0}{\partial t^2}(0) = \frac{\partial f}{\partial t}(0) + \nabla\cdot\left(\frac{\partial a^0}{\partial t}(0)\nabla g\right)$$
$$+\nabla\cdot\left(a^0\nabla(f(0)+\nabla\cdot(a^0\nabla g))\right) \in H.$$

Arguing as for $\frac{\partial u_0}{\partial t}$, we deduce that $\frac{\partial^2 u_0}{\partial t^2} \in L^2((0,T),V)\cap C([0,T],H)$ and $\frac{\partial u_0}{\partial t} \in C([0,T],H^2(D))$. Continuing this process, we have

$$\frac{\partial}{\partial t}\frac{\partial^3 u_0}{\partial t^3} - \nabla\cdot(a^0\nabla\frac{\partial^3 u_0}{\partial t^3}) = \frac{\partial^3 f}{\partial t^3} + \nabla\cdot\left(\frac{\partial^3 a^0}{\partial t^3}\nabla u_0\right) + 3\nabla\cdot\left(\frac{\partial^2 a^0}{\partial t^2}\nabla\frac{\partial u_0}{\partial t}\right)$$
$$+3\nabla\cdot\left(\frac{\partial a^0}{\partial t}\nabla\frac{\partial^2 u_0}{\partial t^2}\right) \in L^2((0,T),V'),$$
$$\frac{\partial^3 u_0}{\partial t^3}(0) = \frac{\partial^2 f}{\partial t^2}(0) + \nabla\cdot\left(a^0\nabla\cdot\left(\frac{\partial f}{\partial t}(0) + \nabla\cdot\left(\frac{\partial a^0}{\partial t}(0)\nabla g\right)\right.\right.$$
$$\left.\left. + \nabla\cdot(a^0\nabla(f(0)+\nabla\cdot(a^0\nabla g)))\right)\right) + \nabla\cdot\left(\frac{\partial^2 a^0}{\partial t}(0)\nabla g\right)$$
$$+2\nabla\cdot\left(\frac{\partial a^0}{\partial t}\nabla(f(0)+\nabla\cdot(a^0\nabla g))\right) \in H.$$

Therefore $\frac{\partial^3 u_0}{\partial t^3} \in C([0,T],H)\cap L^2([0,T],V)$. As $u_0 \in H^3((0,T),V)$, we deduce that $u_0 \in C^2([0,T],V)$. We note that

$$-\nabla\cdot(a^0(t)\nabla\frac{\partial^2 u_0}{\partial t^2}(t)) = \frac{\partial^2 f}{\partial t^2} - \frac{\partial^3 u_0}{\partial t^3} + \nabla\cdot\left(\frac{\partial^2 a^0}{\partial t^2}\nabla u_0\right)$$
$$+2\nabla\cdot\left(\frac{\partial a^0}{\partial t}\nabla\frac{\partial u_0}{\partial t}\right) \in C([0,T],H).$$

Therefore for all $t \in [0,T]$, $\frac{\partial^2 u_0}{\partial t^2}(t) \in H^2(D)$ with

$$\left\|\frac{\partial^2 u_0}{\partial t^2}(t)\right\|_{H^2(D)} \le c\left(\left\|\frac{\partial^2 f}{\partial t^2}(t)\right\|_H + \left\|\frac{\partial^3 u_0}{\partial t^3}(t)\right\|_H + \left\|\frac{\partial^2 u_0}{\partial t^2}(t)\right\|_H + \|u_0(t)\|_H\right)$$

where the constant c depends only on the domain D and the Lipschitz norm of $a^0(t)$ which is uniform for all t (see [19] Theorems 3.1.3.1 and 3.2.1.2). Therefore $u_0 \in H^2((0,T),H^2(D)) \subset C^1([0,T],H^2(D))$. □

For the regularity of N^i, we have the following result.

Proposition 5.2 *Assume that* $a \in C^2([0,1]\times\bar{D}, C^3([0,1],C(\bar{Y})))$, *then* $N^i \in C^1([0,1]\times\bar{D}, H^2((0,1),H^2_\#(Y)))$.

Proof We extend N^i for all $\tau \in (0, \infty)$ periodically. It then belongs to $L^2_{\mathrm{loc}}((0,\infty), V_\#)$. Fixing t and x, we consider problem

$$\frac{\partial N^i}{\partial \tau} - \nabla_y \cdot (a \nabla_y N^i) = \nabla_y \cdot (a e^i), \tag{5.1}$$

with the initial condition at τ_1. From Theorem 5 of Section 7.1 of Evans [16] we deduce that if $N^i(\tau_1) \in V_\#$, then $N^i \in H^1((\tau_1, \Theta), H_\#)$. Indeed the theorem of [16] is for the Dirichlet boundary condition, but the it remains valid for the periodic boundary condition. As $N^i \in L^2_{\mathrm{loc}}((0,\infty), V_\#)$ we can choose a value τ, without loss of generality we let it be τ_1, so that $N^i(\tau_1) \in V_\#$ which implies $N^i \in H^1((\tau_1, \Theta), H_\#)$. Therefore

$$-\nabla_y \cdot (a \nabla_y N^i) = \nabla_y \cdot (a e^i) - \frac{\partial N^i}{\partial \tau} \in L^2((\tau_1, \Theta), H_\#). \tag{5.2}$$

Thus from elliptic regularity, we deduce that $N^i \in L^2((\tau_1, \Theta), H^2_\#(Y))$. We now consider the equation

$$\frac{\partial}{\partial \tau}\frac{\partial N^i}{\partial \tau} - \nabla_y \cdot \left(a \nabla_y \frac{\partial N^i}{\partial \tau} \right) = \nabla_y \cdot \left(\frac{\partial a}{\partial \tau} e^i \right) + \nabla_y \cdot \left(\frac{\partial a}{\partial \tau} \nabla_y N^i \right) \in L^2((\tau_1, \Theta), V'_\#). \tag{5.3}$$

As $N^i \in C_{\mathrm{loc}}([0,\infty), H)$ is uniquely determined, without loss of generality, we assume that $N^i(\tau_1) \in H^2_\#(Y)$, so Eq. 5.3 with the compatibility initial condition at τ_1 implies that $\frac{\partial N^i}{\partial \tau} \in L^2((\tau_1, \Theta), V_\#) \cap C([\tau_1, \Theta], H_\#)$. Without loss of generality, we assume that $\frac{\partial N^i}{\partial \tau}(\tau_1) \in V_\#$. By the same argument as above, we deduce that $\frac{\partial N^i}{\partial \tau} \in H^1((\tau_1, \Theta), H_\#)$, so $\frac{\partial^2 N^i}{\partial \tau^2} \in L^2((\tau_1, \Theta), H_\#)$. We then deduce from (5.3) that $\frac{\partial N^i}{\partial \tau} \in L^2((\tau_1, \Theta), H^2_\#(Y))$. We note that $\frac{\partial N^i}{\partial \tau} \in C([0,T], H_\#)$ is uniquely determined. Without loss of generality, we assume that $\frac{\partial N^i}{\partial \tau}(\tau_1) \in H^2_\#(Y)$. We consider equation

$$\begin{aligned}\frac{\partial}{\partial \tau}\frac{\partial^2 N^i}{\partial \tau^2} - \nabla_y \cdot \left(a \nabla_y \frac{\partial^2 N^i}{\partial \tau^2} \right) &= \nabla_y \cdot \left(\frac{\partial^2 a}{\partial \tau^2} e^i \right) + \nabla_y \cdot \left(\frac{\partial^2 a}{\partial \tau^2} \nabla_y N^i \right) \\ &\quad + 2\nabla_y \cdot \left(\frac{\partial a}{\partial \tau} \nabla_y \frac{\partial N^i}{\partial \tau} \right) \in L^2((\tau_1, \Theta), V'_\#),\end{aligned} \tag{5.4}$$

with the compatible initial condition at τ_1 derived from (5.3). Arguing similarly, we have $\frac{\partial^2 N^i}{\partial \tau^2} \in H^1((\tau_1, \Theta), H_\#)$. Thus from (5.4), we deduce that $\frac{\partial^2 N^i}{\partial \tau^2} \in L^2((\tau_1, \Theta), H^2_\#(Y))$ i.e $N^i \in H^2((\tau_1, \Theta), H^2_\#(Y))$. As a is twice continuously differentiable with respect to x and t, we have that $N^i \in C^2([0,T] \times \bar{D}, H^2((0,1), H^2_\#(Y))$. □

Propositions 5.1 and 5.2 and (4.1) imply that $u_1 \in C^2([0,1], V_1) \cap C([0,T], \hat{\mathcal{H}})$.

6 Numerical results

We show some numerical examples in this section to illustrate the theoretical results on the convergence of the scheme (3.1).

For a one dimensional example, we consider the domain $D = (0, 1)$. We consider the coefficient $a(t, x, \tau, y) = 3 + \cos(2\pi y) + \cos^2(2\pi\tau)$. The initial condition $u^\varepsilon(0) = 0$. Equation 4.2 cannot be solved exactly. We solve it numerically using fine mesh to compute the homogenized coefficient a^0 in (4.5). The reference solution u_1 is computed numerically. The exact solution is chosen as $u_0 = t^2(x - x^2)$. With the homogenized coefficient $a^0(t, x)$ approximated numerically as $a^0 = 3.352429824667637$, the function $f = 2t(x - x^2) + 2a^0t^2$. For the sparse tensor product FE approximation $\hat{U}_{0,m}$ and $\hat{U}_{1,m}$, we plot the errors $\|u_0 - \hat{U}_{0,m}\|_{H_0^1((0,1))}$ and $\|u_1 - \hat{U}_{1,m}\|_{L^2(D\times(0,1),H^1_\#(Y))}$ in Figs. 1 and 2 respectively where $\Delta t = \frac{1}{2}\lceil h_L^{1/2}\rceil$ at $t = 1$. The numerical results show that the errors are $O((\Delta t)^2) + O(h_L)$. When these errors hold for all t_m, we get the errors estimate (3.5). This result supports the theoretical finding. The factor L is not visible in these figures.

For a two dimensional example, we consider the case where the domain $D = (0, 1)\times(0, 1)$. We choose $a(t, x, \tau, y) = (3+\sin(2\pi y')+\sin^2(2\pi\tau))(3+\sin(2\pi y'')+\sin^2(2\pi\tau))$ for $y = (y', y'') \in Y = (0, 1)^2$. The initial condition $u^\varepsilon(0) = 0$. Cell problem (4.2) is solved numerically with fine mesh from which the homogenized coefficient $a^0 = 11.863904995808440$ is computed. We choose $u_0(t, x) = t^2x'x''(1 - x')(1 - x'')$ for $x = (x', x'')$. The function $f = 2t(x' - x'^2)(x'' - x''^2) +$

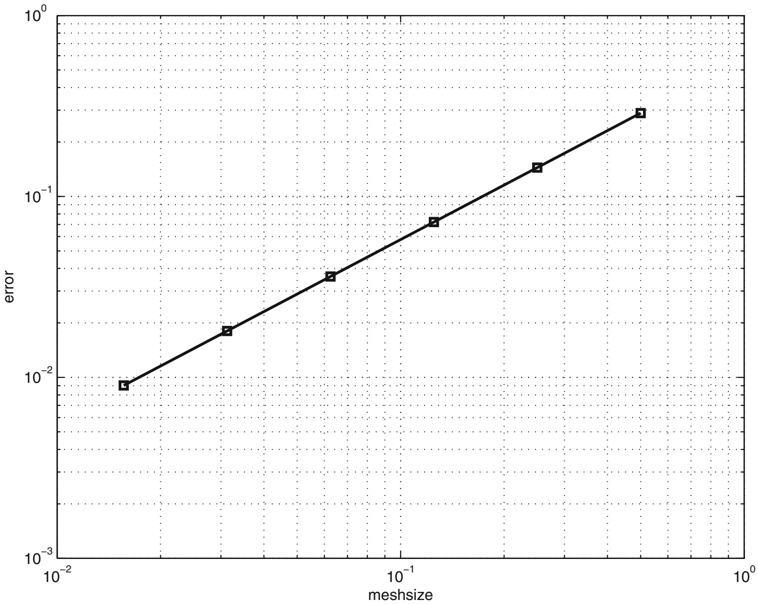

Fig. 1 The error $\|u_0 - \hat{U}_{0,m}\|_{H_0^1(D)}$ for 1 dimensional problem at $t = 1$

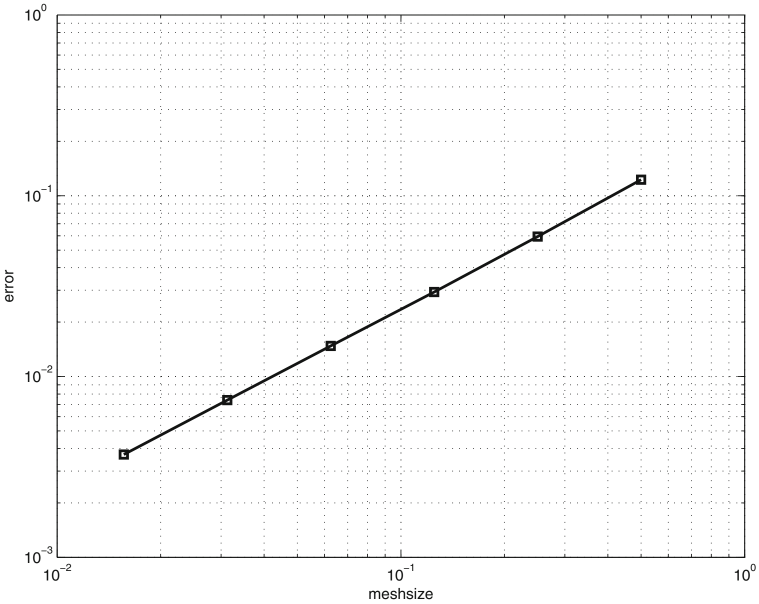

Fig. 2 The error $\|u_1 - \hat{U}_{1,m}\|_{L^2(D\times(0,1),H^1_\#(Y))}$ for 1 dimensional problem at $t = 1$

$2a^0t^2(x' - x'^2 + x'' - x''^2)$. The reference solution u_1 is computed from the numerical solution for N^i and the solution u_0. For $t = 1$, we plot the error $\|u_0 - \hat{U}_{0,m}\|_{H^1_0(D)}$ and $\|u_1 - \hat{U}_{1,m}\|_{L^2(D\times(0,1),H^1_\#(Y)/\mathbb{R})}$ for the sparse tensor product FE solutions in Figs. 3 and 4 respectively. The numerical results agree with the error estimate (3.5).

Although we only develop the theory for the case of one microscopic spatial scale, our method is capable of treating the case of multiple spatial scales. For illustration, we solve some limiting time-space multiscale homogenized equation established in [21]. Holmbom et al. [21] consider the case of two microscopic spatial scales with the coefficient

$$a^\varepsilon = a\left(t, x, \frac{t}{\varepsilon^k}, \frac{x}{\varepsilon}, \frac{x}{\varepsilon^2}\right)$$

where $a = a(t, x, \tau, y_1, y_2)$ is Y-periodic with respect to y_1 and y_2, and $(0, 1)$ periodic with respect to τ. The multiscale convergence limit of u^ε are

$$\nabla u^\varepsilon \overset{\text{ts-ms}}{\longrightarrow} \nabla_x u_0 + \nabla_{y_1} u_1 + \nabla_{y_2} u_2,$$

where $u_1 \in L^2((0, T) \times D, H^1_\#(Y)/\mathbb{R})$ and $u_2 \in L^2((0, T) \times D \times Y, H^1_\#(Y)/\mathbb{R})$. Holmbom et al. [21] establish the multiscale homogenized equation for $k > 0$ but the most interesting critical cases where both u_1 and u_2 depend on τ, and the derivatives

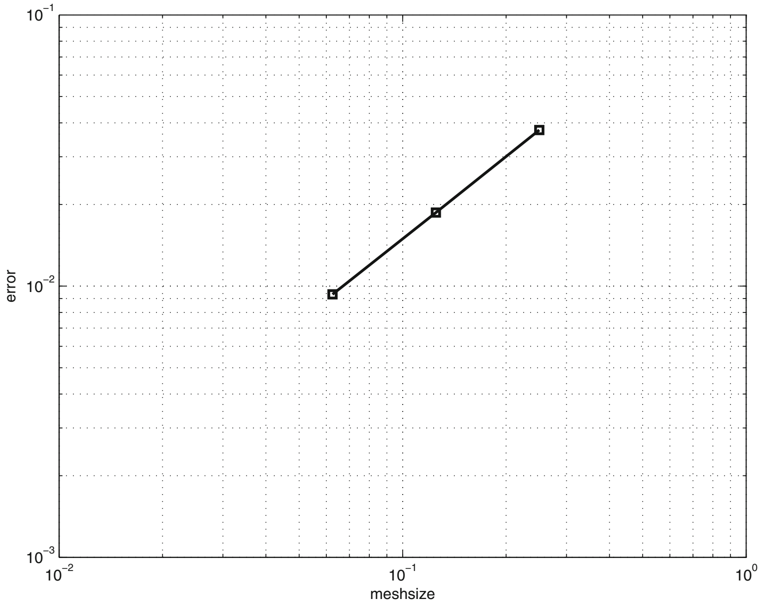

Fig. 3 The error $\|u_0 - \hat{U}_{0,m}\|_{H_0^1(D)}$ for 2 dimensional problem at $t = 1$

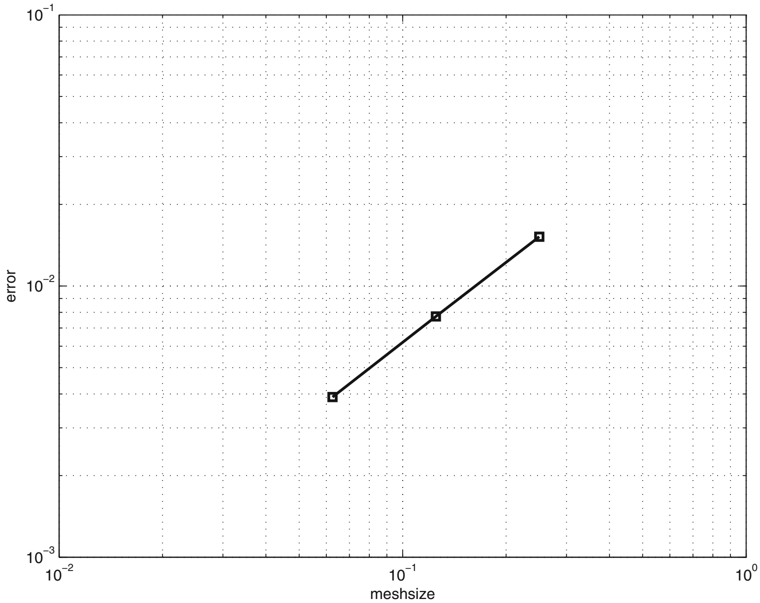

Fig. 4 The error $\|u_1 - \hat{U}_{1,m}\|_{L^2(D\times(0,1),H^1_\#(Y))}$ for 2 dimensional problem at $t = 1$

of these with respect to τ appear in the equation occur when $k = 2$ and $k = 3$. When $k = 2$, we have

$$\left\langle \frac{\partial u_0}{\partial t}(t), \phi_0 \right\rangle_H + \int_D \int_0^1 \left\langle \frac{\partial u_1}{\partial \tau}(t,x,\tau,\cdot), \phi_1 \right\rangle_{H_\#} d\tau dx$$
$$+ \int_D \int_0^1 \int_{Y_1} \int_{Y_2} a(t,x,\tau,y_1,y_2)(\nabla_x u_0(t,x)$$
$$+\nabla_{y_1} u_1(t,x,\tau,y_1) + \nabla_{y_2} u_2(t,x,\tau,y_1,y_2))$$
$$\cdot (\nabla_x \phi_0(x) + \nabla_{y_1}\phi_1(x,\tau,y_1) + \nabla_{y_2}\phi_2(x,\tau,y_1,y_2)) dy_2 dy_1 d\tau dx$$
$$= \int_D f(t,x)\phi_0(x)dx$$

$\forall \phi_0 \in H_0^1(D)$, $\phi_1 \in L^2(D \times (0,1), H_\#^1(Y)/\mathbb{R})$, $\phi_2 \in L^2(D \times (0,1) \times Y, H_\#^1(Y)/\mathbb{R})$, with the initial condition $u_0(0) = g$.

We choose the coefficient

$$a(t,x,\tau,y_1,y_2) = (3 + \sin(2\pi y_1) + \sin(2\pi\tau))(3 + \sin(2\pi y_2) + \sin(2\pi\tau)). \quad (6.1)$$

We need to solve two separate cell problems with respect to y_1 and y_2. In this case, the cell problem for y_1 is identical to (4.2) and is solved numerically, where the cell problem with respect to y_2 is the elliptic problem, i.e., without the derivative with respect to τ, and can be solved exactly. The homogenized coefficient is computed numerically as $a^0 = 8.500245683736688$, We choose $u_0(t,x) = t^2 x(1-x)$ so that $f(t,x) = t(x - x^2) + 2a^0 t^2$. In Figs. 5, 6 and 7, we plot the errors $\|u_0 - \hat{U}_{0,m}\|_{H_0^1(D)}$,

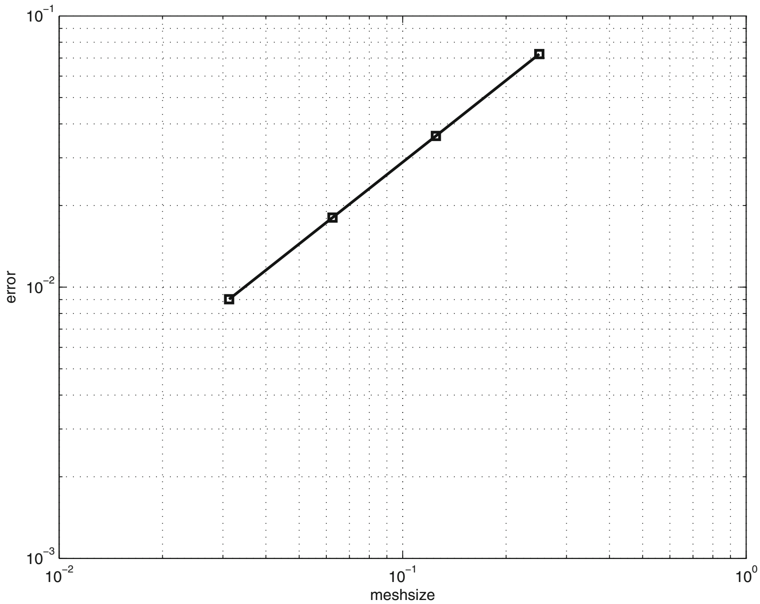

Fig. 5 The error $\|u_0 - \hat{U}_{0,m}\|_{H_0^1(D)}$ for 1 dimensional 3 spatial scales problem at $t = 1$ for $k = 2$

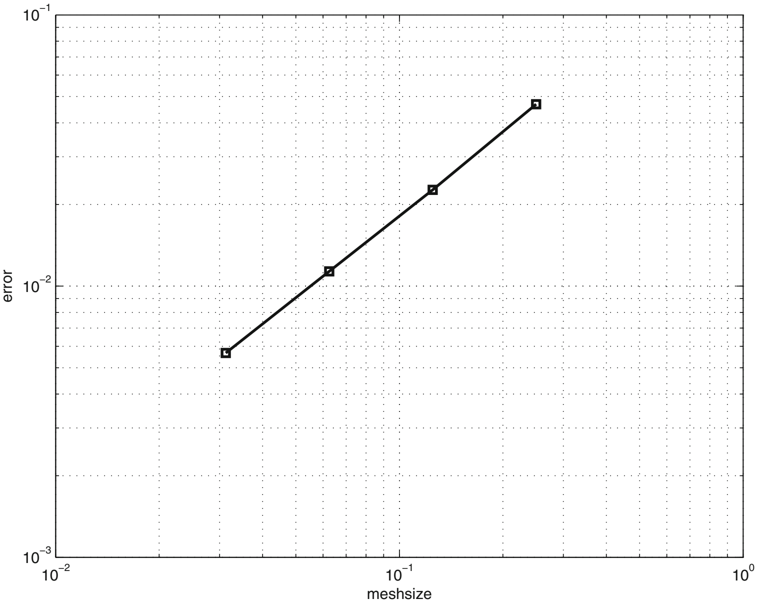

Fig. 6 The error $\|u_1 - \hat{U}_{1,m}\|_{L^2(D\times(0,1),H^1_\#(Y)/\mathbb{R})}$ for 1 dimensional 3 spatial scales problem at $t = 1$ for $k = 2$

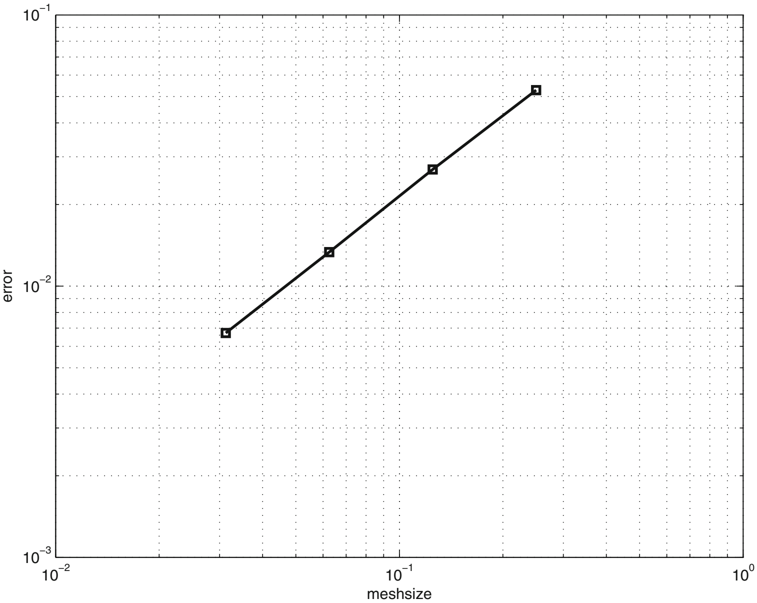

Fig. 7 The error $\|u_2 - \hat{U}_{2,m}\|_{L^2(D\times(0,1)\times Y,H^1_\#(Y)/\mathbb{R})}$ for 1 dimensional 3 spatial scales problem at $t = 1$ for $k = 2$

$\|u_1-\hat{U}_{1,m}\|_{L^2(D\times(0,1),H^1_\#(Y)/\mathbb{R})}$, and $\|u_2-\hat{U}_{2,m}\|_{L^2(D\times(0,1)\times Y,H^1_\#(Y)/\mathbb{R})}$ for the sparse tensor FE approximation respectively. The error agrees with the estimate $O((\Delta t)^2)+O(h_L)$ that we establish in this paper.

For $k=3$, the time-space multiscale homogenized equation becomes:

$$\begin{aligned}
&\left\langle \frac{\partial u_0}{\partial t}(t),\phi_0\right\rangle_H+\int_D\int_0^1\left\langle\frac{\partial u_1}{\partial \tau}(t,x,\tau,\cdot),\phi_1\right\rangle_{H_\#}d\tau dx\\
&\quad+\int_D\int_0^1\int_Y\left\langle\frac{\partial u_2}{\partial \tau}(t,x,\tau,y_1,\cdot),\phi_2\right\rangle_{H_\#}dy_1d\tau dx\\
&\quad+\int_D\int_0^1\int_{Y_1}\int_{Y_2}a(t,x,\tau,y_1,y_2)(\nabla_x u_0(t,x)\\
&\quad+\nabla_{y_1}u_1(t,x,\tau,y_1)+\nabla_{y_2}u_2(t,x,\tau,y_1,y_2))\\
&\quad\cdot(\nabla_x\phi_0(x)+\nabla_{y_1}\phi_1(x,\tau,y_1)+\nabla_{y_2}\phi_2(x,\tau,y_1,y_2))dy_2dy_1d\tau dx\\
&\quad=\int_D f(t,x)\phi_0(x)dx
\end{aligned}$$

$\forall\,\phi_0\in H^1_0(D),\phi_1\in L^2(D\times(0,1),H^1_\#(Y)/\mathbb{R}),\phi_2\in L^2(D\times(0,1)\times Y,H^1_\#(Y)/\mathbb{R})$, with the initial condition $u_0(0)=g$.

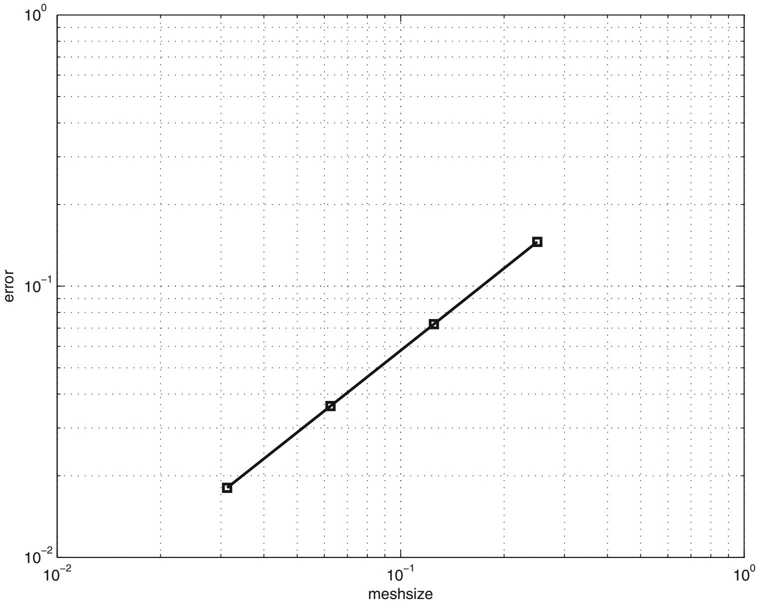

Fig. 8 The error $\|u_0-\hat{U}_{0,m}\|_{H^1_0(D)}$ for 1 dimensional 3 spatial scales problem at $t=1$ for $k=3$

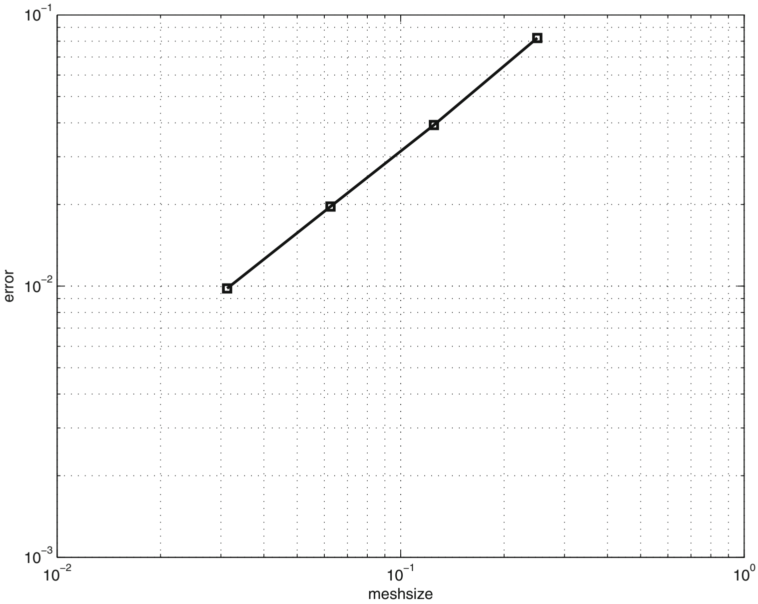

Fig. 9 The error $\|u_1 - \hat{U}_{1,m}\|_{L^2(D\times(0,1),H^1_\#(Y)/\mathbb{R})}$ for 1 dimensional 3 spatial scales problem at $t = 1$ for $k = 3$

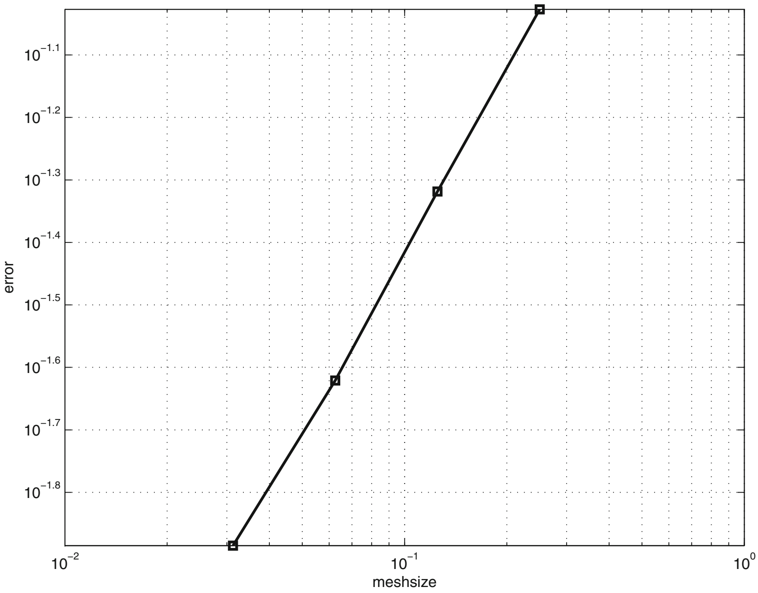

Fig. 10 The error $\|u_2 - \hat{U}_{2,m}\|_{L^2(D\times(0,1)\times Y,H^1_\#(Y)/\mathbb{R})}$ for 1 dimensional 3 spatial scales problem at $t = 1$ for $k = 3$

We choose the coefficient a as in (6.1). We need to solve two cell problems in the form (4.2) with respect to y_1 and y_2 respectively. They are solved numerically. The numerical value of the homogenized coefficient is $a^0 = 7.929947333234398$. We then choose $u_0 = t^2x(1-x)$ so that $f(t,x) = 2t(x-x^2) + 2a^0t^2$. In Figs. 8, 9 and 10, we plot the errors $\|u_0 - \hat{U}_{0,m}\|_{H_0^1(D)}$, $\|u_1 - \hat{U}_{1,m}\|_{L^2(D\times(0,1),H_\#^1(Y)/\mathbb{R})}$, and $\|u_2 - \hat{U}_{2,m}\|_{L^2(D\times(0,1)\times Y,H_\#^1(Y)/\mathbb{R})}$ respectively. The errors again agree with the estimate $O((\Delta t)^2) + O(h_L)$ that we establish in this paper.

Acknowledgments The research topic originates from a discussion with Professor Christoph Schwab, ETH, Zurich. The authors gratefully acknowledge a postgraduate scholarship of A*Star, Singapore, the AcRF Tier 1 grant 2016-T1-001-202 RG30/16, the Singapore A*Star SERC grant 122-PSF-0007, and the AcRF Tier 2 grant MOE 2013-T2-1-095 ARC 44/13.

Appendix

We prove Theorem 3.2 in this appendix. Let $\rho_{0,m} = \frac{1}{\Delta t}(u_0(t_{m+1}) - u_0(t_m)) - \frac{\partial u_0}{\partial t}(t_{m+1/2})$, $\zeta_{0,m} = \frac{1}{2}(u_0(t_{m+1}) + u_0(t_m)) - u_0(t_{m+1/2})$, $\zeta_{1,m} = \frac{1}{2}(u_1(t_{m+1}) + u_1(t_m)) - u_1(t_{m+1/2})$ and $\xi_{1,m} = \frac{1}{2}\left(\frac{\partial u_1}{\partial\tau}(t_{m+1}) + \frac{\partial u_1}{\partial\tau}(t_m)\right) - \frac{\partial u_1}{\partial\tau}(t_{m+1/2})$. Since $u_0 \in C^3([0,T],H) \cap C^2([0,T],V)$, $u_1 \in C^2([0,T],L^2(D\times(0,1),V_\#))$ and $\frac{\partial u_1}{\partial\tau} \in C^2([0,T],L^2(D\times(0,1),V_\#'))$, we deduce that

$$\|\rho_{0,m}\|_{L^2(D)} \le c(\Delta t)^2,\ \|\zeta_{0,m}\|_V \le c(\Delta t)^2,\ \|\zeta_{1,m}\|_{V_1} \le c(\Delta t)^2,\ \text{and}$$
$$\|\xi_{1,m}\|_{L^2(D\times(0,1),V_\#')} \le c(\Delta t)^2$$

where the constant c does not depend on m. From (2.1) and (3.1) considered at $t = t_{m+1/2}$ we deduce that

$$\begin{aligned}
&\left\langle \frac{z_{0,m+1}-z_{0,m}}{\Delta t}, \phi_0 \right\rangle_H + \langle \rho_{0,m}, \phi_0\rangle_H \\
&+ \int_D\int_0^1 \left\langle \frac{1}{2}\left(\frac{\partial z_{1,m+1}}{\partial\tau} + \frac{\partial z_{1,m}}{\partial\tau}\right), \phi_1\right\rangle_{H_\#} d\tau dx + \int_D\int_0^1 \langle \xi_1, \phi_1\rangle_{H_\#} d\tau dx \\
&+ \int_D\int_0^1\int_Y a(t_{m+1/2})\left(\nabla_x \frac{z_{0,m+1}+z_{0,m}}{2} + \nabla_y \frac{z_{1,m+1}+z_{1,m}}{2}\right) \\
&\cdot(\nabla_x\phi_0 + \nabla_y\phi_1) dy d\tau dx \\
&+ \int_D\int_0^1\int_Y a(t_{m+1/2})(\nabla_x\zeta_{0,m} + \nabla_y\zeta_{1,m})\cdot(\nabla_x\phi_0 + \nabla_y\phi_1) dy d\tau dx = 0.
\end{aligned} \tag{A.1}$$

Consider

$$\begin{aligned} I &= \left\langle \frac{z_{0,m+1}-z_{0,m}}{\Delta t}, \frac{z_{0,m+1}+z_{0,m}}{2} \right\rangle_H \\ &\quad + \int_D \int_0^1 \left\langle \frac{1}{2}\left(\frac{\partial z_{1,m+1}}{\partial \tau} + \frac{\partial z_{1,m}}{\partial \tau} \right), \frac{z_{1,m+1}+z_{1,m}}{2} \right\rangle_{H_\#} d\tau dx \\ &\quad + \int_D \int_0^1 \int_Y a(t_{m+1/2}) \left(\nabla_x \frac{z_{0,m+1}+z_{0,m}}{2} + \nabla_y \frac{z_{1,m+1}+z_{1,m}}{2} \right) \\ &\quad \cdot \left(\nabla_x \frac{z_{0,m+1}+z_{0,m}}{2} + \nabla_y \frac{z_{1,m+1}+z_{1,m}}{2} \right) dy d\tau dx \\ &\geq \frac{1}{2\Delta t}(\|z_{0,m+1}\|_H^2 - \|z_{0,m}\|_H^2) + \gamma(\|z_{0,m+1/2}\|_V^2 + \|z_{1,m+1/2}\|_{V_1}^2). \end{aligned} \tag{A.2}$$

For $\{\tilde u_{0,m},\ m=0,\ldots,M\} \subset V^L$ and $\{\tilde u_{1,m},\ m=1,\ldots,M\} \subset V_1^L$, we have

$$\begin{aligned} I &= \left\langle \frac{z_{0,m+1}-z_{0,m}}{\Delta t}, (u_0-\tilde u_0)_{m+1/2} \right\rangle_H + \left\langle \frac{z_{0,m+1}-z_{0,m}}{\Delta t}, (\tilde u_0-U_0)_{m+1/2} \right\rangle_H \\ &+ \int_D \int_0^1 \left\langle \frac{1}{2}\left(\frac{\partial z_{1,m+1}}{\partial \tau} + \frac{\partial z_{1,m}}{\partial \tau} \right), (u_1-\tilde u_1)_{m+1/2} \right\rangle_{H_\#} d\tau dx \\ &+ \int_D \int_0^1 \left\langle \frac{1}{2}\left(\frac{\partial z_{1,m+1}}{\partial \tau} + \frac{\partial z_{1,m}}{\partial \tau} \right), (\tilde u_1-U_1)_{m+1/2} \right\rangle_{H_\#} d\tau dx \\ &+ \int_D \int_0^1 \int_Y a(t_{m+1/2}) \left(\nabla_x \frac{z_{0,m+1}+z_{0,m}}{2} + \nabla_y \frac{z_{1,m+1}+z_{1,m}}{2} \right) \\ &\cdot \left(\nabla_x (u_0-\tilde u_0)_{m+1/2} + \nabla_y (u_1-\tilde u_1)_{m+1/2} \right) dy d\tau dx \\ &+ \int_D \int_0^1 \int_Y a(t_{m+1/2}) \left(\nabla_x \frac{z_{0,m+1}+z_{0,m}}{2} + \nabla_y \frac{z_{1,m+1}+z_{1,m}}{2} \right) \\ &\cdot \left(\nabla_x (\tilde u_0-U_0)_{m+1/2} + \nabla_y (\tilde u_1-U_1)_{m+1/2} \right) dy d\tau dx. \end{aligned}$$

From (3.1) we have

$$\begin{aligned} I &= \left\langle \frac{z_{0,m+1}-z_{0,m}}{\Delta t}, (u_0-\tilde u_0)_{m+1/2} \right\rangle_H \\ &\quad - \int_D \int_0^1 \left\langle \frac{z_{1,m+1}+z_{1,m}}{2}, \frac{\partial}{\partial \tau}(u_1-\tilde u_1)_{m+1/2} \right\rangle_{H_\#} d\tau dx \\ &\quad + \int_D \int_0^1 \int_Y a(t_{m+1/2}) \left(\nabla_x \frac{z_{0,m+1}+z_{0,m}}{2} + \nabla_y \frac{z_{1,m+1}+z_{1,m}}{2} \right) \\ &\quad \cdot \left(\nabla_x (u_0-\tilde u_0)_{m+1/2} + \nabla_y (u_1-\tilde u_1)_{m+1/2} \right) dy d\tau dx \\ &\quad - \langle \rho_{0,m}, (\tilde u_0-U_0)_{m+1/2} \rangle_H - \int_D \int_0^1 \langle \xi_{1,m}, (\tilde u_1-U_1)_{m+1/2} \rangle_{H_\#} d\tau dx \\ &\quad + \int_D \int_0^1 \int_Y a(t_{m+1/2}) \left(\nabla_x \zeta_{0,m} + \nabla_y \zeta_{1,m} \right) \\ &\quad \cdot \left(\nabla_x (\tilde u_0-U_0)_{m+1/2} + \nabla_y (\tilde u_1-U_1)_{m+1/2} \right) dy d\tau dx. \end{aligned}$$

We note that $(\tilde{u}_0 - U_0)_{m+1/2} = (\tilde{u}_0 - u_0)_{m+1/2} + z_{0,m+1/2}$ and $(\tilde{u}_1 - U_1)_{m+1/2} = (\tilde{u}_1 - u_1)_{m+1/2} + z_{1,m+1/2}$. For a positive constant $\delta > 0$, using the Cauchy-Schwartz inequality, we have

$$\begin{aligned}
I \leq & \left\langle \frac{z_{0,m+1} - z_{0,m}}{\Delta t}, (u_0 - \tilde{u}_0)_{m+1/2} \right\rangle_H \\
& + \delta \|z_{1,m+1/2}\|_{V_1}^2 + c \left\| \frac{\partial}{\partial \tau} (u_1 - \tilde{u}_1)_{m+1/2} \right\|_{V_1'}^2 \\
& + \delta \|z_{0,m+1/2}\|_V^2 + \delta \|z_{1,m+1/2}\|_{V_1}^2 + c\|(u_0 - \tilde{u}_0)_{m+1/2}\|_V^2 + c\|(u_1 - \tilde{u}_1)_{m+1/2}\|_{V_1}^2 \\
& + c\|\rho_{0,m}\|_H^2 + c\|(\tilde{u}_0 - u_0)_{m+1/2}\|_H^2 + \delta \|z_{0,m+1/2}\|_H^2 \\
& + c\|\xi_{1,m}\|_{V_1'}^2 + c\|(\tilde{u}_1 - u_1)_{m+1/2}\|_{V_1}^2 + \delta \|z_{1,m+1/2}\|_{V_1}^2 \\
& + c\|\zeta_{0,m}\|_V^2 + c\|\zeta_{1,m}\|_{V_1}^2 + c\|(\tilde{u}_0 - u_0)_{m+1/2}\|_V^2 + \delta \|z_{0,m+1/2}\|_V^2 \\
& + c\|(\tilde{u}_1 - u_1)_{m+1/2}\|_{V_1}^2 + \delta \|z_{1,m+1/2}\|_V^2.
\end{aligned}$$

From this and (A.2), choosing δ sufficiently small, we have

$$\begin{aligned}
& \frac{1}{2\Delta t}(\|z_{0,m+1}\|_H^2 - \|z_{0,m}\|_H^2) + c(\|z_{0,m+1/2}\|_V^2 + \|z_{1,m+1/2}\|_{V_1}^2) \\
& \leq \left\langle \frac{z_{0,m+1} - z_{0,m}}{\Delta t}, (u_0 - \tilde{u}_0)_{m+1/2} \right\rangle_H \\
& \quad + c \left\| \frac{\partial}{\partial \tau} (u_1 - \tilde{u}_1)_{m+1/2} \right\|_{V_1'}^2 + c\|(u_0 - \tilde{u}_0)_{m+1/2}\|_V^2 + c\|(u_1 - \tilde{u}_1)_{m+1/2}\|_{V_1}^2 + c(\Delta t)^4.
\end{aligned}$$

Fixing an integer $P \leq M$, taking the sum for $m = 0, \dots, P-1$, we have

$$\begin{aligned}
& \|z_{0,P}\|_H^2 - \|z_{0,0}\|_H^2 + c\Delta t \sum_{m=0}^{P-1} (\|z_{0,m+1/2}\|_V^2 + \|z_{1,m+1/2}\|_{V_1}^2) \\
& \leq c\Delta t \sum_{m=0}^{P-1} \left(\left\| \frac{\partial}{\partial \tau} (u_1 - \tilde{u}_1)_{m+1/2} \right\|_{V_1'}^2 + \|(u_0 - \tilde{u}_0)_{m+1/2}\|_V^2 + \|(u_1 - \tilde{u}_1)_{m+1/2}\|_{V_1}^2 \right) \\
& cP(\Delta t)^5 + 2\Delta t \sum_{m=0}^{P-1} \left\langle \frac{z_{0,m+1} - z_{0,m}}{\Delta t}, (u_0 - \tilde{u}_0)_{m+1/2} \right\rangle_H . \qquad \text{(A.3)}
\end{aligned}$$

We note that

$$
\begin{aligned}
\Delta t\sum_{m=0}^{P-1}&\left\langle\frac{z_{0,m+1}-z_{0,m}}{\Delta t},(u_0-\tilde u_0)_{m+1/2}\right\rangle_H\\
&=\left\langle z_{0,P},(u_0-\tilde u_0)_{P-1/2}\right\rangle_H-\left\langle z_{0,0},(u_0-\tilde u_0)_{1/2}\right\rangle_H\\
&\quad+\Delta t\sum_{m=1}^{P-1}\left\langle\frac{z_{0,m}}{\Delta t},(u_0-\tilde u_0)_{m-1/2}-(u_0-\tilde u_0)_{m+1/2}\right\rangle_H\\
&\le\delta\|z_{0,P}\|_H^2+c\|(u_0-\tilde u_0)_{P-1/2}\|_H^2+\|z_{0,0}\|_H^2+\|(u_0-\tilde u_0)_{1/2}\|_H^2\\
&\quad+\delta\Delta t\sum_{m=1}^{P-1}\|z_{0,m}\|_H^2+c\Delta t\sum_{m=1}^{P-1}\left\|\frac{(u_0-\tilde u_0)_{m+1/2}-(u_0-\tilde u_0)_{m-1/2}}{\Delta t}\right\|_H^2
\end{aligned}
$$

which is a consequence of the Cauchy-Schwartz inequality; δ is an arbitrary positive constant. We note that $\Delta t\sum_{m=1}^{P-1}\|z_{0,m}\|_H^2\le\Delta t P\max_{m=0,\dots,M}\|z_{0,m}\|_H^2\le T\max_{m=0,\dots,M}\|z_{0,m}\|_H^2$. From this and (A.3), choosing δ sufficiently small, we have

$$
\begin{aligned}
\|z_{0,P}\|_H^2\le c\Delta t\sum_{m=0}^{P-1}&\left(\left\|\frac{\partial}{\partial\tau}(u_1-\tilde u_1)_{m+1/2}\right\|_{V_1'}^2+\|(u_0-\tilde u_0)_{m+1/2}\|_V^2\right.\\
&\left.+\|(u_1-\tilde u_1)_{m+1/2}\|_{V_1}^2\right)\\
&+c(\Delta t)^4+c\|(u_0-\tilde u_0)_{P-1/2}\|_H^2+2\|z_{0,0}\|_H^2+\|(u_0-\tilde u_0)_{1/2}\|_H^2\\
&+c\Delta t\sum_{m=1}^{P-1}\left\|\frac{(u_0-\tilde u_0)_{m+1/2}-(u_0-\tilde u_0)_{m-1/2}}{\Delta t}\right\|_H^2+\delta T\max_{m=0,\dots,M}\|z_{0,M}\|_H^2.
\end{aligned}
$$

Choosing δ sufficiently small, we have

$$
\begin{aligned}
\max_{m=0,\dots,M}\|z_{0,m}\|_H^2\le c\Delta t\sum_{m=0}^{M-1}&\left(\left\|\frac{\partial}{\partial\tau}(u_1-\tilde u_1)_{m+1/2}\right\|_{V_1'}^2+\|(u_0-\tilde u_0)_{m+1/2}\|_V^2\right.\\
&\left.+\|(u_1-\tilde u_1)_{m+1/2}\|_{V_1}^2\right)\\
&+c(\Delta t)^4+c\max_{m=1,\dots,M}\|(u_0-\tilde u_0)_{m-1/2}\|_H^2+c\|z_{0,0}\|_H^2\\
&+\|(u_0-\tilde u_0)_{1/2}\|_H^2\\
&+c\Delta t\sum_{m=1}^{M-1}\left\|\frac{(u_0-\tilde u_0)_{m+1/2}-(u_0-\tilde u_0)_{m-1/2}}{\Delta t}\right\|_H^2.
\end{aligned}
$$

From this, we get the conclusion.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Abdulle, A., E, W., Engquist, B., Vanden-Eijnden, E.: The heterogeneous multiscale method. Acta Numer. **21**, 1–87 (2012)
2. Abdulle, A., Huber, M.E.: Finite element heterogeneous multiscale method for nonlinear monotone parabolic homogenization problems. ESAIM Math. Model. Numer Anal. **50**(6), 1659–1697 (2016)
3. Abdulle, A., Vilmart, G.: Coupling heterogeneous multiscale FEM with Runge-Kutta methods for parabolic homogenization problems: a fully discrete spacetime analysis. Math. Models Methods Appl. Sci. **22**(6), 1250002 (2012)
4. Allaire, G.: Homogenization and two-scale convergence. SIAM J. Math. Anal. **23**(6), 1482–1518 (1992)
5. Bensoussan, A., Lions, J.-L., Papanicolaou, G.: Asymptotic analysis for periodic structures, volume 5 of studies in mathematics and its applications. North-Holland Publishing Co., Amsterdam (1978)
6. Bieri, M., Andreev, R., Schwab, C.: Sparse tensor discretization of elliptic SPDEs. SIAM J. Sci. Comput. **31**(6), 4281–4304 (2009/10)
7. Bungartz, H.-J., Griebel, M.: Sparse grids. Acta Numer. **13**, 147–269 (2004)
8. Chen, S., E, W., Shu, C.-W.: The heterogeneous multiscale method based on the discontinuous Galerkin method for hyperbolic and parabolic problems. Multiscale Model Simul. **3**(4), 871–894 (2005)
9. Chu, V.T., Hoang, V.H.: High dimensional finite elements for multiscale Maxwell-type equations. IMA J. Numer. Anal. **38**(1), 227–270 (2018)
10. Chung, E.T., Efendiev, Y., Leung, W.T., Ye, S.: Generalized multiscale finite element methods for space-time heterogeneous parabolic equations. Comput. Math Appl. **76**(2), 419–437 (2018)
11. Cioranescu, D., Damlamian, A., Griso, G.: The periodic unfolding method in homogenization. SIAM J. Math. Anal. **40**(4), 1585–1620 (2008)
12. Douglas, J. Jr.., Dupont, T.: Galerkin methods for parabolic equations. SIAM J. Numer. Anal. **7**, 575–626 (1970)
13. Efendiev, Y., Hou, T.Y.: Multiscale finite element methods: theory and applications. surveys and tutorials in the applied mathematical sciences. Springer (2009)
14. Efendiev, Y., Pankov, A.: Numerical homogenization of nonlinear random parabolic operators. Multiscale Model Simul. **2**(2), 237–268 (2004)
15. Efendiev, Y., Galvis, J., Hou, T.Y.: Generalized multiscale finite element methods (GMsFEM). J. Comput. Phys. **251**, 116–135 (2013)
16. Evans, L.C.: Partial differential equations, volume 19 of Graduate Studies in Mathematics. American Mathematical Society, Providence (1998)
17. Geng, J., Shen, Z.: Convergence rates in parabolic homogenization with time-dependent periodic coefficients. J. Funct Anal. **272**(5), 2092–2113 (2017)
18. Griebel, M., Oswald, P.: Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. Adv. Comput. Math. **4**(1), 171–206 (1995)
19. Grisvard, P.: Elliptic problems in nonsmooth domains, volume 69 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2011)
20. Hoang, V.H., Schwab, C.: High-dimensional finite elements for elliptic problems with multiple scales. Multiscale Model. Simul. **3**(1), 168–194 (2004/05)
21. Holmbom, A., Svanstedt, N., Wellander, N.: Multiscale convergence and reiterated homogenization of parabolic problems. Appl. Math. **50**(2), 131–151 (2005)
22. Hou, T.Y., Wu, X.H.: A multiscale finite element method for elliptic problems in composite materials and porous media. J. Comput. Phys. **134**(1), 169–189 (1997)
23. Kazeev, V., Oseledets, I., Rakhuba, M., Schwab, C.: QTT-finite-element approximation for multiscale problems I: model problems in one dimension. Adv. Comput. Math. **43**(2), 411–442 (2017)
24. Lions, J.L., Magenes, E.: Non-homogeneous boundary value problems and applications, vol. I. Springer, Berlin (1972)
25. Målqvist, A., Persson, A.: Multiscale techniques for parabolic equations. Numer. Math. **138**(1), 191–217 (2018)
26. Målqvist, A., Peterseim, D.: Localization of elliptic multiscale problems. Math. Comp. **83**(290), 2583–2603 (2014)
27. Ming, P., Zhang, P.: Analysis of the heterogeneous multiscale method for parabolic homogenization problems. Math Comp. **76**(257), 153–177 (2007)

28. Nguetseng, G.: A general convergence result for a functional related to the theory of homogenization. SIAM J. Math. Anal. **20**(3), 608–623 (1989)
29. Owhadi, H., Zhang, L.: Homogenization of parabolic equations with a continuum of space and time scales. SIAM J. Numer. Anal. **46**(1), 1–36 (2008)
30. Tan, W.C., Hoang, V.H.: High dimensional finite element method for multiscale nonlinear monotone parabolic equations. J. Comput. Appl. Math. **345**, 471–500 (2019)
31. Wloka, J.: Partial differential equations. Cambridge University Press, Cambridge (1987)
32. Woukeng, J.L.: Periodic homogenization of nonlinear non-monotone parabolic operators with three time scales. Ann. Mat. Pura Appl. (4) **189**(3), 357–379 (2010)
33. Xia, B., Hoang, V.H.: High dimensional finite elements for multiscale wave equations. Multiscale Model. Simul. **12**(4), 1622–1666 (2014)
34. Xia, B., Hoang, V.H.: High-dimensional finite element method for multiscale linear elasticity. IMA J. Numer. Anal. **35**(3), 1277–1314 (2015)
35. Xia, B., Hoang, V.H.: Sparse tensor finite elements for elastic wave equation with multiple scales. J. Comput. Appl. Math. **282**, 179–214 (2015)