

# Multiscale methods for wave problems in heterogenous media

Assyr Abdulle<sup>1</sup> and Patrick Henning<sup>2</sup>

## Abstract

In this paper we give a survey on various multiscale methods for the numerical solution of second order hyperbolic equations in highly heterogenous media. We concentrate on the wave equation and distinguish between two classes of applications. First we discuss numerical methods for the wave equation in heterogeneous media without scale separation. Such a setting is for instance encountered in the geosciences, where natural structures often exhibit a continuum of different scales, that all need to be resolved numerically to get meaningful approximations. Approaches tailored for these settings typically involve the construction of generalized finite element spaces, where the basis functions incorporate information about the data variations. In the second part of the paper, we discuss numerical methods for the case of structured media with scale separation. This setting is for instance encountered in engineering sciences, where materials are often artificially designed. If this is the case, the structure and the scale separation can be explicitly exploited to compute appropriate homogenized/upscaled wave models that only exhibit a single coarse scale and that can be hence solved at significantly reduced computational costs.

**Keywords** finite element, multiscale method, numerical homogenization, second order hyperbolic problems, wave equation, long-time wave propagation

**AMS subject classifications** 35L05, 65M60, 65N30, 74Q10, 74Q15

## 1 Introduction

In this article we discuss recent developments of numerical methods for the multiscale wave equation

$$\begin{aligned} \partial_{tt}u_\varepsilon - \nabla \cdot (a^\varepsilon \nabla u_\varepsilon) &= F && \text{in } \Omega \times ]0, T[, \\ u_\varepsilon &= 0 && \text{on } \partial\Omega \times ]0, T[, \\ u_\varepsilon(x, 0) = g_1(x), \quad \partial_t u_\varepsilon(x, 0) &= g_2(x) && \text{in } \Omega, \end{aligned} \tag{1}$$

where  $\Omega$  is bounded domain. We make the following minimal regularity assumptions (H0)

- $\Omega \subset \mathbb{R}^d$  is a bounded Lipschitz domain with a piecewise polygonal boundary ( $d = 1, 2, 3$ );
- the data satisfy  $F \in L^2(0, T; L^2(\Omega))$ ,  $g_1 \in H_0^1(\Omega)$  and  $g_2 \in L^2(\Omega)$ ;
- the matrix-valued function  $a^\varepsilon$  is in  $\mathcal{M}(\alpha, \beta, \Omega)$  where

$$\begin{aligned} \mathcal{M}(\alpha, \beta, \Omega) := & \\ \{a \in [L^\infty(\Omega)]_{\text{sym}}^{d \times d} \mid \alpha |\xi|^2 \leq a(x) \xi \cdot \xi \leq \beta |\xi|^2 \text{ for all } \xi \in \mathbb{R}^d \text{ and almost all } x \in \Omega\}. & \end{aligned} \tag{2}$$

---

<sup>1</sup>ANMC, Section de Mathématiques, École polytechnique fédérale de Lausanne, 1015 Lausanne, Switzerland, Assyr.Abdulle@epfl.ch

<sup>2</sup>Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden, pathe@kth.se

Under (H0) there exist a unique  $u_\varepsilon \in C^0(0, T; H^1(\Omega))$ ,  $\partial_t u_\varepsilon \in C^0(0, T; L^2(\Omega))$  and  $\partial_{tt} u_\varepsilon \in L^2(0, T; H^{-1}(\Omega))$  such that for all  $v \in H_0^1(\Omega)$  and a.e.  $t > 0$  [40, Chapter 3]

$$\langle \partial_{tt} u_\varepsilon(\cdot, t), v \rangle + (a^\varepsilon \nabla u_\varepsilon(\cdot, t), \nabla v)_{L^2(\Omega)} = (F(\cdot, t), v)_{L^2(\Omega)}. \quad (3)$$

The wave speed  $a^\varepsilon$  is assumed to be a *multiscale coefficient*. By that we mean that  $a^\varepsilon$  varies on a scale of order  $\mathcal{O}(\varepsilon)$ , where  $0 < \varepsilon \ll 1$ . In general, we do not assign a particular value to  $\varepsilon$ , but we only assume that it is very small parameter. However, whenever we encounter a sufficiently regular coefficient, we assume that  $\|a^\varepsilon\|_{W^{1,\infty}(\Omega)} = \mathcal{O}(\varepsilon^{-1})$  to illustrate our arguments. For a classical numerical approximation of the problem (3) we pick a  $P1$  finite element space  $V_h$  with mesh size  $h$  (for simplicity we consider a quasi-uniform family of triangulation of the domain  $\Omega$  in simplicial elements, cf. [15]) and regard the problem: find  $u_h : [0, T] \rightarrow V_h$  such that for all  $v_h \in V_h$  and a.e.  $t > 0$

$$\langle \partial_{tt} u_h(\cdot, t), v_h \rangle + (a^\varepsilon \nabla u_h(\cdot, t), \nabla v_h)_{L^2(\Omega)} = (F(\cdot, t), v_h)_{L^2(\Omega)} \quad (4)$$

and with appropriate discrete initial values. Let us from now on use the shorthand notation  $L^p(H^s) := L^p(0, T; H^s(\Omega))$ . Following the arguments by Baker [12] leads to the standard approximation result:

$$\|u_\varepsilon - u_h\|_{L^\infty(L^2)} \leq C(\|u_\varepsilon - \Pi_h(u_\varepsilon)\|_{L^\infty(L^2)} + \|\partial_t u_\varepsilon - \partial_t \Pi_h(u_\varepsilon)\|_{L^1(L^2)}), \quad (5)$$

where  $\Pi_h : H_0^1(\Omega) \rightarrow V_h$  is the Ritz-projection on  $V_h$ , i.e., the  $(a_\varepsilon \nabla \cdot, \nabla \cdot)$ -orthogonal projection. The projection error can be further estimated by exploiting its quasi best-approximation property in  $H^1(\Omega)$ , an Aubin-Nitsche duality argument for the elliptic projection and the  $H^1$ -stability of  $\Pi_h$  to obtain

$$\|u_\varepsilon - u_h\|_{L^\infty(L^2)} \leq C h \|a^\varepsilon\|_{W^{1,\infty}(\Omega)} (\|u_\varepsilon\|_{L^\infty(H^1)} + \|\partial_t u_\varepsilon\|_{L^1(H^1)}) \leq C \frac{h}{\varepsilon}, \quad (6)$$

where  $C = C(T)$  is independent of  $\varepsilon$ . Here we assumed  $\|a^\varepsilon\|_{W^{1,\infty}(\Omega)} = \mathcal{O}(\varepsilon^{-1})$  and  $\|\partial_t u_\varepsilon\|_{L^1(H^1)} = \mathcal{O}(1)$ . At first sight, this seems not to be an optimal  $L^2$  error estimate. Indeed for the Aubin-Nitsche duality argument one usually already uses an optimal first order convergence rate in  $h$  (in the  $H^1$  norm) for the solution of the elliptic problem. In our situation this would lead to an estimate of the type  $h\|\partial_t u_\varepsilon\|_{L^1(H^2)}$ . The problem is however to bound  $\|\partial_t u_\varepsilon\|_{L^1(H^2)}$  independently of  $\varepsilon$ . For general initial condition classical a priori error bounds [25] would lead to a bound of the type  $C\varepsilon^{-2}$  [6]. This is why we avoid the  $H^2$  norm in the above estimate. This issue makes also the use of higher order spatial approximations useless when using classical FE or FD methods for the approximation of (1). Notice that the standard a priori bound for  $\|\partial_t u_\varepsilon\|_{L^1(H^1)}$  also scales as  $C/\varepsilon$  but this term can be controlled using  $G$ -convergence and perturbation arguments [6] so that the assumption  $\|\partial_t u_\varepsilon\|_{L^1(H^1)} = \mathcal{O}(1)$  can be in fact avoided. Hence, convergence usually requires  $h < \varepsilon$  and a high computational complexity if  $\varepsilon$  is small. In this contribution we consider two distinct situations that require different numerical strategies. For the first situation, we assume no scale separation in the highly heterogeneous tensor  $a^\varepsilon$ . In this case, we review numerical methods based on suitably chosen multiscale spaces. For the second situation, we assume scale separation in the tensor  $a^\varepsilon$  (e.g., periodic, locally periodic, random stationary). Here we can use classical finite element spaces for the numerical approximation but the numerical methods rely on homogenized (effective) models that must be computed ‘‘on the fly’’. A peculiar feature of highly oscillatory hyperbolic problems such as (1) is that different homogenized models must be derived according to the time span of the desired approximation. We will consider time intervals that scale as  $[0, T\varepsilon^{-2}]$ .<sup>1</sup> Notice that in this case the estimate (6) for standard finite elements reads [7]

$$\|u_\varepsilon - u_h\|_{L^\infty(0, T\varepsilon^{-2}; L^2)} \leq C \frac{h}{\varepsilon^3}, \quad (7)$$

<sup>1</sup>In the periodic homogenization setting this is the first interval of interest on which the homogenized model is not valid, e.g., for interval of the type  $[0, T\varepsilon^{-1}]$  the homogenized solution still gives adequate approximation in the  $L^\infty(L^2)$  norm.

where again we assume a bound for  $\|\partial_t u_\varepsilon\|_{L^1(H^1)}$  independent of  $\varepsilon$ . Notice that for periodic problems, well-prepared initial data can be used to obtain such a bound [6].

We define the computational complexity as the size of the linear system  $N$  required to be solved at each time-step  $\Delta t$  of the time integrator for the wave equation. For classical FEM, in view of (6), we have a computational complexity of  $N = \varepsilon^{-d}$  per time step  $\Delta t$  (short time interval) and a computational complexity of  $N = \varepsilon^{-3d}$  per time step  $\Delta t$  in view of (7) (long-time interval  $[0, T\varepsilon^{-2}]$ ). If using an explicit method such as the popular leap-frog method for the time integration, the stability constraint  $\Delta t \simeq h$  reads  $\Delta t \simeq \varepsilon$  for short-time integration and  $\Delta t \simeq \varepsilon^3$  for long-time integration. This are indeed very severe time-step constraints due the oscillatory behavior of (1). Of course implicit methods could be used, but then the additional cost due to the linear system to solve (that again scales badly with  $\varepsilon$ ) constitutes a non-trivial additional cost per time-step.

We close this introduction by recalling a fundamental homogenization result (valid in both situations described above) that will be used in both classes of numerical methods described in what follows.

The basic question of classical homogenization is the following: if  $(a^\varepsilon)_{\varepsilon>0}$  represents a sequence of tensors and  $(u_\varepsilon)_{\varepsilon>0}$  the corresponding sequence of solutions to (3), does  $u_\varepsilon$  converge in some sense to a limit function  $u_0$ ? Is there a (homogenized) equation for this limit function? The hope is that due to the average limit process for  $\varepsilon \rightarrow 0$ , the homogenized equation is cheap to solve. The abstract tool of  $G$ -convergence gives a general answer to that question.

**Definition 1.1** ( $G$ -convergence). A sequence  $(a^\varepsilon)_{\varepsilon>0} \subset \mathcal{M}(\alpha, \beta, \Omega)$  (i.e. with uniform spectral bounds in  $\varepsilon$ ) is said to be  $G$ -convergent to  $a^0 \in \mathcal{M}(\alpha, \beta, \Omega)$  if for all  $F \in H^{-1}(\Omega)$  the sequence of solutions  $v^\varepsilon \in H_0^1(\Omega)$  to

$$\int_{\Omega} a^\varepsilon \nabla v^\varepsilon \cdot \nabla v = F(v) \quad \text{for all } v \in H_0^1(\Omega)$$

satisfies  $v^\varepsilon \rightharpoonup v^0$  weakly in  $H_0^1(\Omega)$ , where  $v^0 \in H_0^1(\Omega)$  solves

$$\int_{\Omega} a^0 \nabla v^0 \cdot \nabla v = F(v) \quad \text{for all } v \in H_0^1(\Omega).$$

One of the main properties of  $G$ -convergence is the following compactness result [49, 50]: let  $(a^\varepsilon)_{\varepsilon>0}$  be a sequence of matrices in  $\mathcal{M}(\alpha, \beta, \Omega)$ , then there exists a subsequence  $(a^{\varepsilon'})_{\varepsilon'>0}$  and a matrix  $a^0 \in \mathcal{M}(\alpha, \beta, \Omega)$  such that  $(a^{\varepsilon'})_{\varepsilon'>0}$   $G$ -converges to  $a^0$ . For the wave equation, we have the following result obtained in [14, Theorem 3.2]:

**Theorem 1.2** (Homogenization of the wave equation). *Let assumptions (H0) be fulfilled and let the sequence of symmetric matrices  $(a^\varepsilon)_{\varepsilon>0} \subset \mathcal{M}(\alpha, \beta, \Omega)$  be  $G$ -convergent to some  $a^0 \in \mathcal{M}(\alpha, \beta, \Omega)$ . Let  $u_\varepsilon \in L^\infty(0, T; H_0^1(\Omega))$  denote the solution to the wave equation (3). Then it holds*

$$\begin{aligned} u_\varepsilon &\rightharpoonup u_0 \quad \text{weak-}^* \text{ in } L^\infty(0, T, H_0^1(\Omega)), \\ \partial_t u_\varepsilon &\rightharpoonup \partial_t u_0 \quad \text{weak-}^* \text{ in } L^\infty(0, T, L^2(\Omega)) \end{aligned}$$

and where  $u_0 \in L^2(0, T; H_0^1(\Omega))$  with  $\partial_{tt} u_0 \in L^2(0, T; H^{-1}(\Omega))$  is the unique weak solution to the homogenized problem

$$\begin{aligned} \langle \partial_{tt} u_0(\cdot, t), v \rangle + (a^0 \nabla u_0(\cdot, t), \nabla v)_{L^2(\Omega)} &= (F(\cdot, t), v)_{L^2(\Omega)} && \text{for all } v \in H_0^1(\Omega) \text{ and } t > 0, \\ (u_0(\cdot, 0), v)_{L^2(\Omega)} &= (g_1, v)_{L^2(\Omega)} && \text{for all } v \in H_0^1(\Omega), \\ (\partial_t u_0(\cdot, 0), v)_{L^2(\Omega)} &= (g_2, v)_{L^2(\Omega)} && \text{for all } v \in H_0^1(\Omega). \end{aligned} \tag{8}$$

This theorem and the compactness result stated above show that for *any* problem (3) based on a sequence of matrices with  $a^\varepsilon \in \mathcal{M}(\alpha, \beta, \Omega)$ , we can extract a subsequence such that the corresponding solution of the wave problem converges to a homogenized solution. Except for special situations, e.g., locally periodic coefficients  $a^\varepsilon$ , i.e. tensor  $a^\varepsilon(x) = a(x, \frac{x}{\varepsilon})$  that are  $\varepsilon$ -periodic on a fine scale or for random stationary tensors, it is not possible to construct  $a^0$  explicitly.

## 2 Numerical methods for the wave equation in heterogeneous media without scale separation

In this section we consider the setting that is encountered if the wave speed  $a^\varepsilon$  reflects the properties of a heterogeneous medium without scale separation. Such a setting is typically encountered in earth sciences such as geophysics or seismology. Here, the waves propagate through a medium that often lacks any kind of structure. Instead the medium consists of a variety of heterogeneously distributed materials, as for instance different rock and soil types, possibly interrupted by natural reservoirs of groundwater, petroleum or gas (cf. the data of the Society of Petroleum Engineering, openly accessible on <http://www.spe.org/web/csp>). In such natural structures it is typically impossible to distinguish different scales of resolution and we speak about a *lack of scale separation*. Opposite to this, problems arising from engineering applications are often artificially designed and hence exhibit a perfect scale separation. We will discuss this case in the next section.

From now on,  $a^\varepsilon$  denotes an unstructured and highly heterogeneous coefficient that lacks scale separation. In our discussion we shall also focus on the minimal possible regularity assumptions. This is an important aspect since the propagation field  $a^\varepsilon$  is discontinuous in many realistic applications. Subsequently, we let  $\varepsilon$  denote a parameter that characterizes the size of the smallest length scale on which variations of  $a^\varepsilon$  can be observed, e.g. we could define  $\varepsilon := \|a^\varepsilon\|_{W^{1,\infty}(\Omega)}^{-1}$  provided that  $a^\varepsilon$  is sufficiently regular. For the size of the computational domain we assume here that  $\text{diam}(\Omega) = \mathcal{O}(1)$ . This assumption will be relaxed in some situations for longtime wave propagation as described in Section 3.2. Let  $X_N \subset H_0^1(\Omega)$  denote a discrete space and let  $u_N : [0, T] \rightarrow X_N$  denote the corresponding semi-discrete approximation given as the solution to

$$\langle \partial_{tt} u_N(\cdot, t), v \rangle + (a^\varepsilon \nabla u_N(\cdot, t), \nabla v)_{L^2(\Omega)} = (F(\cdot, t), v)_{L^2(\Omega)}, \quad \text{for all } v \in X_N \text{ and a.e. } t > 0, \quad (9)$$

and with suitably chosen discrete initial values. The question that we want to discuss is the following. Is there for any  $N \in \mathbb{N}$  a space  $X_N = X_N(a^\varepsilon)$  that only depends on  $a^\varepsilon$ ; but not on  $t$ , the source term  $F$  or the initial values  $g_1$  and  $g_2$ ; such that two properties are fulfilled:

- (C1)  $\dim(X_N) = N$  and
- (C2)  $\|u_\varepsilon - u_N\|_{L^\infty(H^1)} \leq CH$ ; where  $H = \mathcal{O}(N^{-d})$  is a generalized mesh size and where  $C = C(T)$  is independent of the variations of  $a^\varepsilon$ , i.e. independent of  $\varepsilon$ .

The above properties would guarantee a convergence of the numerical scheme without a resolution constraint imposed by the speed of the data variations  $\varepsilon^{-1}$ . Comparing this with the setting of the classical  $P1$  FE space  $V_h$  with corresponding solution  $u_h$  given by (4) we can identify the relation  $\dim(V_h) = N$  with  $N = \mathcal{O}(h^{-d})$ , but the final error estimate reads  $\|u_\varepsilon - u_h\|_{L^\infty(L^2)} \leq C \frac{h}{\varepsilon}$  and does hence not fulfill the desired property.

Keeping these findings in mind, we next want to discuss the following problem: is there a discrete space  $X_N$  that fulfills the properties (C1) and (C2)? In fact, the question can be answered mainly positively and we can identify four different approaches to this problem in the literature. We present them in the following in chronological order. For the sake of simplicity, we shall also assume that the initial values are zero; i.e.  $g_1 = g_2 = 0$ ; and that the source term is time-independent; i.e.  $F(x, t) = F(x)$ . The case of general data is shortly discussed in Section 2.5. Also note that we

restrict the presentation of the methods to the semi-discrete setting, i.e. we do not discuss a time-discretization of (9) as this does typically not impose new problems. We just mention that equation (9) can be discretized in time in various ways (e.g. using the framework of Newmark schemes) and that a fully discrete analysis of Approach 1 [43] and Approach 4 [6] can be found in the corresponding papers.

Another approach based on operator upscaling was proposed by Minkoff et al. [52, 37, 51]. However, we will not discuss this approach since the wave equation is considered in a different form (namely of the structure  $\partial_{tt}u_\varepsilon - a^\varepsilon \Delta u_\varepsilon = F$ ), which does not involve the typical multiscale issues.

In the following, we only assume minimal regularity for  $a^\varepsilon$ , i.e.  $a^\varepsilon \in [L^\infty(\Omega)]_{\text{sym}}^{d \times d}$ .

## 2.1 Approach 1 - Harmonic coordinate transformations

Let us assume that  $\Omega \subset \mathbb{R}^2$  is a bounded and convex domain of the class  $C^2$ . The first approach devoted to the multiscale wave equation without scale separation was proposed by Owhadi and Zhang [43] (as a generalization of the elliptic case considered in [42]). The authors suggest to overcome the issue of the missing space regularity by applying a smoothing coordinate transformation  $G^\varepsilon = (G_1^\varepsilon, G_2^\varepsilon) : \Omega \rightarrow \Omega$ , whose components  $G_k^\varepsilon$  with  $k \in \{1, 2\}$  are defined as the weak solutions to the elliptic boundary value problem

$$\begin{aligned} \nabla \cdot (a^\varepsilon \nabla G_k^\varepsilon) &= 0 & \text{in } \Omega; \\ G_k^\varepsilon(x) &= x_i & \text{on } \partial\Omega. \end{aligned} \quad (10)$$

The transformation  $G^\varepsilon$  as defined above can be shown to be an automorphism over  $\Omega$  (cf. [9]) and hence maps  $\Omega$  indeed into itself.

The basic idea of the approach is to try to approximate the exact solution of (1) by  $v_0 \circ G^\varepsilon$ , where  $v_0$  denotes a smooth function that is  $\varepsilon$ -independent and that only exhibits slow variations. In contrast, the multiscale character and the low regularity part of  $u_\varepsilon$  are embedded in the components of the harmonic transformation  $G_k^\varepsilon \in H^1(\Omega)$ . Provided that this ansatz is valid, the original multiscale problem can be reinterpreted as to find the slow function  $v_0(\cdot, t) \in H_0^1(\Omega) \cap H^2(\Omega)$  such that  $v_0(\cdot, t) \circ G^\varepsilon \in H_0^1(\Omega)$  solves (3). Since  $v_0$  is a smooth single-scale function, it can be easily approximated in conventional (coarse) finite element spaces  $V_H$ . With that, we can define the coordinate transformed (and still low-dimensional) solution space as

$$V_H^\varepsilon := \{v_H \circ G^\varepsilon \mid v_H \in V_H\} \quad (11)$$

and we seek  $u_H^\varepsilon \in (0, T; V_H^\varepsilon)$  with  $u_H^\varepsilon(\cdot, 0) = \partial_t u_H^\varepsilon(\cdot, 0) = 0$  and

$$\langle \partial_{tt} u_H^\varepsilon(\cdot, t), v \rangle + (a^\varepsilon \nabla u_H^\varepsilon(\cdot, t), \nabla v)_{L^2(\Omega)} = (F, v)_{L^2(\Omega)}, \quad \text{for all } v \in V_H^\varepsilon \text{ and } t > 0. \quad (12)$$

This approach can be rigorously justified under the following geometric assumption.

**Definition 2.1** (Cordes type condition). Let  $\sigma^\varepsilon(x) := (\nabla G^\varepsilon(x))^\top a^\varepsilon(x) \nabla G^\varepsilon(x)$ , where the columns of the matrix  $\nabla G^\varepsilon$  are defined by  $\nabla G^\varepsilon := (\nabla G_1^\varepsilon, \nabla G_2^\varepsilon)$ . Furthermore, the Cordes parameter is defined as

$$\mu_{\sigma^\varepsilon} := \text{ess sup}_{x \in \Omega} \left( \frac{\lambda_{\max}(\sigma^\varepsilon(x))}{\lambda_{\min}(\sigma^\varepsilon(x))} \right),$$

where  $\lambda_{\max}(\sigma^\varepsilon(x))$  and  $\lambda_{\min}(\sigma^\varepsilon(x))$  denote the upper and lower spectral bounds of  $\sigma^\varepsilon(x)$ . We say that  $G^\varepsilon$  fulfills the Cordes type condition, if  $\mu_{\sigma^\varepsilon} < \infty$  and  $(\text{Trace}[\sigma]^{-1}) \in L^\infty(\Omega)$ .

With this, Owhadi and Zhang [43] proved the following theorem, that guarantees **(C1)** and **(C2)** for  $V_H^\varepsilon$ .

**Theorem 2.2.** *Let the Cordes type condition from Definition 2.1 be fulfilled and assume that  $g_1 = g_2 = \partial_t F = 0$ . Then we have that the coordinate-transformed solution is a regular coarse-scale function, that is  $v_0 := u_\varepsilon \circ (G^\varepsilon)^{-1} \in L^\infty(0, T; H^2(\Omega))$  and  $\|v_0\|_{L^\infty(H^2)} \leq C\|F\|_{L^2(\Omega)}$ , i.e. the  $L^\infty(H^2)$ -norm of  $v_0$  can be bounded independently of  $\varepsilon$ . Furthermore, if  $V_H$  denotes a conventional  $P1$  finite element space as in the introduction, then the numerical approximation  $u_H^\varepsilon$  to (12) (obtained in the low-dimensional multiscale space  $V_H^\varepsilon$ ) fulfills the a priori error estimate*

$$\|u_H^\varepsilon - u_\varepsilon\|_{L^\infty(H^1)} + \|\partial_t u_H^\varepsilon - \partial_t u_\varepsilon\|_{L^\infty(L^2)} \leq CH\|F\|_{L^2(\Omega)},$$

where  $C$  only depends on  $\Omega$  and the fulfillment of the Cordes type condition.

**Remark 2.3.** The error estimate in Theorem 2.2 holds with minimal regularity assumptions for  $a^\varepsilon$ , i.e.  $a^\varepsilon \in [L^\infty(\Omega)]_{\text{sym}}^{d \times d}$ . Even though the result is only proved for  $d = 2$  and under the assumption that a Cordes type condition is fulfilled, both assumptions do not seem to be necessary in practice (cf. [43] for a discussion and numerical experiments) and the method still performs well if the condition is not fulfilled. This is an important observation, since the validity of the Cordes-type condition can be hard to verify in practice.

From the numerical perspective we can identify four steps involved in this approach.

1. For  $k = 1, \dots, d$ , solve for a numerical approximation to the components  $G_k^\varepsilon \in H^1(\Omega)$  of the harmonic coordinate given by (10). Since the  $G_k^\varepsilon$  are multiscale functions, this step involves to solve  $d$  global multiscale-scale problems on a fine mesh with mesh size “ $h < \varepsilon$ ”.
2. When a sufficiently accurate approximation  $G^{\varepsilon, h}$  to  $G^\varepsilon$  is computed, we can define the basis set  $\Phi_i^\varepsilon := \Phi_i \circ G^{\varepsilon, h}$ , where  $\Phi_i$  denotes a nodal basis function of the coarse space  $V_H$ . With that it is necessary to compute the entries of the stiffness matrix  $S$  and the mass matrix  $M$  with  $S_{ij} = (a^\varepsilon \nabla \Phi_j^\varepsilon, \nabla \Phi_i^\varepsilon)_{L^2(\Omega)}$  and  $M_{ij} = (\Phi_j^\varepsilon, \Phi_i^\varepsilon)_{L^2(\Omega)}$ . Note that the basis functions  $\Phi_i^\varepsilon$  are typically non-local, i.e. they have support in the whole domain  $\Omega$ .
3. By using the precomputed coarse quantities it is now possible to march in time with a favorite time-discretization, where every step only involves operations with low-dimensional matrices and vectors.
4. For new source terms  $F$ , the results of step 1 and 2 can be reused and it is possible to directly start with step 3.

As we can see, Step 1 and 2 can be very costly since it involves global fine scale computations (step 1) and global fine scale quadrature rules (step 2). This can be considered as a one time overhead that pays off for a sufficiently high number of time steps or source terms. It should also be noted that the mass and stiffness matrices computed in step 2 are dense and typically not sparse. Solving a linear system that involves these matrices hence has a cubic computational complexity. This needs to be considered when deciding for a coarse mesh. Alternativ to using conventional  $P1$  finite elements for constructing the multiscale space  $V_H^\varepsilon$  given by (11), one could also chose  $V_H$  as the space of weighted extended B-splines (WEB, [33]). The numerical experiments in [43] indicate that using B-splines can improve the performance of the method considerably. Finally, let us mention that Step 1 introduces a numerical approximation  $G_k^{\varepsilon, h}$  to  $G_k^\varepsilon$ . So far, it has not yet been investigated analytically how such an approximation influences the validity of Theorem 2.2, which assumes that  $G_k^\varepsilon$  is available analytically.

## 2.2 Approach 2 - MsFEM using limited global information

A second approach that can be found in the literature is the *Multiscale Finite Element Method using Limited Global Information* proposed by Jiang et al. [36, 35]. This approach can be considered as

a generalization of the previously discussed harmonic coordinate transformation [43], however, from the slightly shifted point of view. Jiang et al. start from the assumption that there exist  $m$  known global fields  $G_1^\varepsilon, \dots, G_m^\varepsilon$  that are available to the user. These fields might either be precomputed (e.g. coordinate transformations as in Section 2.1) or they were inferred from measured data (for instance in the context of porous media flow). It is further assumed that there exists a smooth (unknown) function  $v_0$  that is independent of the data variations, but that allows to express the exact solution  $u_\varepsilon$  in terms of the global fields, i.e.

$$u_\varepsilon(x, t) \approx v_0(G_1^\varepsilon(x), \dots, G_m^\varepsilon(x), t).$$

Provided that this a priori knowledge is available, two approaches are proposed in [36] that we describe now. Let us consider a conventional coarse  $P1$  finite element space denoted by  $V_H$ .

**Version 1.** We define the multiscale space  $V_H^\varepsilon$  by the functions that can be expressed as products of a global field and a finite element function, i.e. we set  $G_0^\varepsilon := 1$  and let

$$V_H^\varepsilon := \{G_k^\varepsilon v_H \mid 0 \leq k \leq m; v_H \in V_H\}.$$

If  $\{\Phi_i \mid 1 \leq i \leq n\}$  denotes the nodal basis of  $V_H$ , then  $\{G_k^\varepsilon \Phi_i \mid 1 \leq i \leq n; 0 \leq k \leq m\}$  denotes a basis of  $V_H^\varepsilon$  (with dimension  $n(m+1)$ ). We see that the functions  $G_k^\varepsilon \Phi_i$  inherit their support from  $\Phi_i$ , so that we obtain as set of locally supported basis functions of  $V_H^\varepsilon$ . As before, a numerical approximation is obtained according to (9). Some analytical considerations for **Version 1** are presented in [35].

**Version 2.** The second version is restricted to the specific case that  $d = 2$  and that  $m = 1$ , i.e. there is only one global field  $G^\varepsilon = G_1^\varepsilon$  available and we consequently assume  $u_\varepsilon(x, t) \approx v_0(G_1^\varepsilon(x), t)$  for some smooth  $v_0$  and for *continuous*  $G_1^\varepsilon$ . In this case, Jiang et al. propose a different approach that is very similar to the classical Multiscale Finite Element Method (MsFEM, proposed by Hou and Wu [34]). More precisely, for every node  $z_i$  of the coarse space  $V_H$ , we define a corresponding (multiscale) nodal basis function  $\Phi_i^\varepsilon$  element-wise as follows. Let  $\omega_i := \cup\{K \in \mathcal{T}_H \mid z_i \in K\}$  denote the nodal patch that belongs to  $z_i$ . Then for every element  $K \in \mathcal{T}_H$  of the triangulation with  $K \subset \omega_i$ ,  $\Phi_i^\varepsilon$  is the solution to the elliptic problem

$$\begin{aligned} -\nabla \cdot (a^\varepsilon \nabla \Phi_i^\varepsilon) &= 0 & \text{in } K, \\ \Phi_i^\varepsilon &= g_{i,K}^\varepsilon & \text{on } \partial K. \end{aligned}$$

Here,  $g_{i,K}^\varepsilon$  is an oscillatory boundary condition induced from the global field  $G^\varepsilon$ , that we define in equation (13) below. The classical MsFEM basis function are defined in the same way, but with the difference that the boundary condition is not oscillatory, but simply the affine condition inherited from the nodal basis, i.e.  $\Phi_i^\varepsilon = \Phi_i$  on  $\partial K$  (cf. [21]). It is known that oscillatory boundary conditions typically improve the performance of the method, since the arising approximations do not suffer from so-called resonance errors (see also [31]). In the case of a single global field, meaningful oscillatory boundary values  $g_{i,K}^\varepsilon$  can be constructed in the following way. Let  $K \subset \omega_i$  be a coarse element (triangle) and let  $E_{K,0}$ ,  $E_{K,1}$  and  $E_{K,2}$  denote the three corresponding edges (with  $\partial K = E_{K,0} \cup E_{K,1} \cup E_{K,2}$ ): Furthermore, we let  $z_{K,0}$ ,  $z_{K,1}$  and  $z_{K,2}$  denote the three corresponding corners (nodes), where we assume (without loss of generality) that the numeration is such that  $z_{K,0} = z_i$ ,  $\partial E_{K,0} = \{z_{K,0}, z_{K,1}\}$ ,  $\partial E_{K,1} = \{z_{K,1}, z_{K,2}\}$  and  $\partial E_{K,2} = \{z_{K,2}, z_{K,0}\}$ . With that, we define  $g_{i,K}^\varepsilon$  on  $\partial K$  by

$$g_{i,K}^\varepsilon(x) := \begin{cases} \frac{G^\varepsilon(x) - G^\varepsilon(z_{K,1})}{G^\varepsilon(z_i) - G^\varepsilon(z_{K,1})} & \text{if } x \in E_{K,0}, \\ 0 & \text{if } x \in E_{K,1}, \\ \frac{G^\varepsilon(x) - G^\varepsilon(z_{K,2})}{G^\varepsilon(z_i) - G^\varepsilon(z_{K,2})} & \text{if } x \in E_{K,2}. \end{cases} \quad (13)$$

Assuming that the following error estimates hold

$$\|u_\varepsilon - v_0 \circ G^\varepsilon\|_{L^\infty(H^1)} + \|\partial_t(u_\varepsilon - v_0 \circ G^\varepsilon)\|_{L^\infty(L^2)} + \|\partial_{tt}(u_\varepsilon - v_0 \circ G^\varepsilon)\|_{L^2(L^2)} \leq \delta,$$

and provided that all involved functions are sufficiently regular, it is possible to derive  $\delta$ - and  $H$ -explicit and  $\varepsilon$ -independent a priori error estimates for **Version 2** of the method (see [36, Theorem 3.1]).

Observe that both versions do not follow the strategy suggested in [43]. Comparing the *MsFEM using limited global information* with the approach based on a *harmonic coordinate transformation*, we can state two crucial differences.

1. Owhadi and Zhang define  $V_H^\varepsilon$  as the coordinate transformation of  $V_H$  (hence the basis functions are concatenations  $v_H \circ G^\varepsilon$ ), whereas Jiang et al. define  $V_H^\varepsilon$  in a multiplicative way, which guarantees locally supported basis functions. This makes the assembling of stiffness and mass matrices significantly cheaper (in terms of quadrature costs) and leads to sparse matrix structures. Consequently, the method can be considered as cheaper, once the global fields are available. However, finding the global fields in the first place is required for both methods.
2. The harmonic coordinate transformation is supported by a rigorous numerical analysis and hence guarantees convergence of the method. On the other hand, the analysis for both versions of Approach 2 (see [35] for Version 1 and [36] for Version 2) rely on regularity and approximability assumptions that cannot be derived in general from the transformation  $G^\varepsilon$  given by (10). For that reason it is not clear if there exist global fields  $G_1^\varepsilon, \dots, G_m^\varepsilon$  that meet the formal requirements that are necessary for the framework of Approach 2. Nevertheless, numerical experiments in [36, 35] show a good performance of Approach 2 if  $G^\varepsilon$  is selected as a steady state solution with  $-\nabla \cdot (a^\varepsilon \nabla G^\varepsilon) = F$  and a homogenous Dirichlet boundary condition. Since this choice involves the source term, the resulting space can however not be reused for different right-hand side of (1).

### 2.3 Approach 3 - Flux-transfer transformations

The basic drawback of Approach 1 and 2 is that a considerable one-time overhead is involved, when the harmonic coordinate transformations/the global fields are computed by a global fine scale computation. An alternative that overcomes this issue was again proposed by Owhadi and Zhang in [44], who use a localizable *transfer property* as an alternative to non-local harmonic coordinate transformations discussed in Section 2.1. The transfer property says that if two fluxes  $a_1 \nabla v_1$  and  $a_2 \nabla v_2$  have the same divergence in  $L^2$ , then their potential parts (i.e. the  $\nabla H_0^1(\Omega)$ -parts in their Helmholtz-decomposition) can be approximated in discrete spaces with identical accuracy, if the discrete spaces  $X_1$  and  $X_2$  are linked through the relation  $\nabla \cdot (a_1 \nabla X_1) = \nabla \cdot (a_2 \nabla X_2)$ .

In order to make this statement precise, let  $\Omega \subset \mathbb{R}^d$  be in the following a domain with a  $C^2$ -boundary and let  $P_{\text{pot}} : [L^2(\Omega)]^d \rightarrow \nabla H_0^1(\Omega) := \{\nabla v \mid v \in H_0^1(\Omega)\}$  denote the  $L^2$ -projection onto the potential part of the Helmholtz decomposition, i.e.  $P_{\text{pot}}(\mathbf{u}) \in \nabla H_0^1(\Omega)$  fulfills

$$(P_{\text{pot}}(\mathbf{u}), \mathbf{w})_{L^2(\Omega)} = (\mathbf{u}, \mathbf{w})_{L^2(\Omega)} \quad \text{for all } \mathbf{w} \in \nabla H_0^1(\Omega).$$

With this, for any  $a \in \mathcal{M}(\alpha, \beta, \Omega)$  we define the *a-flux-norm* of a function  $w \in H_0^1(\Omega)$  by

$$\|w\|_{\text{-flux}} := \|P_{\text{pot}}(a \nabla w)\|_{L^2(\Omega)}.$$

The flux-norm can easily be seen to be equivalent to the energy-norm (with  $\alpha \|\nabla w\|_{L^2(\Omega)} \leq \|w\|_{a\text{-flux}} \leq \beta \|\nabla w\|_{L^2(\Omega)}$ ) and it can be shown that it has the following remarkable transfer property (cf. [44]).



**Lemma 2.4.** *Let  $V_H \subset H_0^1(\Omega)$  be a finite dimensional subspace and let*

$$V_{H,a^\varepsilon} := \{\nabla \cdot (a^\varepsilon \nabla v_H) \mid v_H \in V_H\}.$$

*Furthermore, for  $F \in L^2(\Omega)$  we let  $z_\varepsilon \in H_0^1(\Omega)$  and  $z \in H_0^1(\Omega)$  denote the weak solutions to the following problems*

$$\int_{\Omega} a^\varepsilon \nabla z_\varepsilon \cdot \nabla v = \int_{\Omega} F v \quad \text{and} \quad \int_{\Omega} \nabla z \cdot \nabla v = \int_{\Omega} F v \quad \text{for all } v \in H_0^1(\Omega).$$

*Then it holds the following transfer property in the flux norm*

$$\inf_{v_H \in V_H} \|z - v_H\|_{1\text{-flux}} = \inf_{v_{H,a^\varepsilon} \in V_{H,a^\varepsilon}} \|z_\varepsilon - v_{H,a^\varepsilon}\|_{a^\varepsilon\text{-flux}}, \quad (14)$$

where  $\|\cdot\|_{1\text{-flux}}$  denotes the flux-norm for  $a = 1$ , i.e.  $\nabla \cdot (a \nabla \cdot) = \Delta$ .

To emphasize the role of the transfer property, let us denote for the rest of this subsection by  $V_H$  the space of weighted extended B-splines (WEB) [33] on a uniform grid with grid width  $H$  (consequently the basis functions have a support on a domain with diameter  $\mathcal{O}(H)$ ). We stress that the basis functions are smooth (hence  $\Delta \Phi_i \in L^2(\Omega)$ ) and locally supported and that we intrinsically assume that  $V_H$  is a coarse space (i.e. the variations of  $a^\varepsilon$  are not resolved). Then from the transfer property (14) and the norm equivalence we infer the  $\varepsilon$ -independent estimate

$$\inf_{v_{H,a^\varepsilon} \in V_{H,a^\varepsilon}} \|z_\varepsilon - v_{H,a^\varepsilon}\|_{H^1(\Omega)} \leq C \inf_{v_H \in V_H} \|z - v_H\|_{H^1(\Omega)} \leq CH \|F\|_{L^2(\Omega)},$$

where  $C$  is a constant only depending on  $\Omega$ ,  $\alpha$  and  $\beta$ . The main advantage of Approach 3, compared to the harmonic coordinates transformation, is that the construction of  $V_{H,a^\varepsilon}$  can be localized in a natural way: if  $\Phi_i$  denotes a basis function of  $V_H$ , then the transferred basis function  $\Phi_i^\varepsilon$  in  $V_{H,a^\varepsilon}$  is given as the solution to  $\nabla \cdot (a^\varepsilon \nabla \Phi_i^\varepsilon) = \Delta \Phi_i$  with  $\Phi_i^\varepsilon = 0$  on  $\partial\Omega$ . Since the ‘‘source term’’  $\Delta \Phi_i$  is only locally supported, we expect  $\Phi_i^\varepsilon$  to decay to zero outside of the support of  $\Phi_i^\varepsilon$ . This justifies that computing  $\Phi_i^\varepsilon$  can be localized to smaller subdomains. We shall detail this in the following.

Let  $z_i \in \Omega$  denote the node associated with the basis function  $\Phi_i \in V_H$  and for a sufficiently large constant  $C > 0$  let  $U_i := \{x \in \mathbb{R}^d \mid |z_i - x| \leq CH^{1/2} |\log(H)|\}$  denote an environment of  $z_i$  with a diameter of order  $\mathcal{O}(H^{1/2} |\log(H)|)$ . Then we define the localized transferred basis function  $\Psi_i^\varepsilon$  as the solution to the elliptic problem

$$\begin{aligned} \frac{1}{H} \Psi_i^\varepsilon - \nabla \cdot (a^\varepsilon \nabla \Psi_i^\varepsilon) &= \Delta \Phi_i & \text{in } U_i \cap \Omega, \\ \Psi_i^\varepsilon &= 0 & \text{on } \partial(U_i \cap \Omega). \end{aligned} \quad (15)$$

At first glance, this problem formulation might be surprising since it involves the artificial zero-order term  $\frac{1}{H} \Psi_i^\varepsilon$ . In fact, this term was added to speedup the decay of the corresponding Green’s function. The decay of the Green’s function associated with the operator  $\nabla \cdot (a^\varepsilon \nabla \cdot)$  is well-known to be only algebraic, whereas adding the zero-order term  $\frac{1}{H} \Psi_i^\varepsilon$  regularizes the operator and leads to an exponential decay of the corresponding Green’s function (see also [29]). This exponential decay allows to restrict the computation of the transferred basis functions to computational domains  $U_i$  with a size of  $\text{diam}(U_i) = \mathcal{O}(H^{1/2} |\log(H)|)$ . The distortion of the transfer property caused by adding the zero-order term is basically balanced with the localization error (caused by restricting the problem to  $U_i$ ) if the constant  $C$  in the definition of  $U_i$  is chosen appropriately (cf. the numerical experiments in [44]). Note that practically (15) needs to be discretized on an additional fine mesh with mesh size ‘‘ $h < \varepsilon$ ’’. With that, we can define the multiscale space by

$$V_H^\varepsilon := \text{span}\{\Psi_i^\varepsilon \mid 1 \leq i \leq N\}, \quad (16)$$

where  $N = \dim(V_H) = \mathcal{O}(H^{-d})$ . As desired, the space is low dimensional and it has locally supported basis functions with a support of diameter  $\mathcal{O}(H^{1/2}|\log(H)|)$ . In particular the fact that only local problems (15) have to be solved is a great advantage compared to Approach 1 and 2 which both require to first compute global fields  $G^\varepsilon$  by fine scale computations on the whole domain  $\Omega$ . The problems (15) are independent and each problem individually is cheap to solve, which allows for an efficient parallel implementation of the method. Finally, the space also fulfills the properties **(C1)** and **(C2)** as shown by the following theorem [44, Theorem 5.1].

**Theorem 2.5.** *Recall that  $g_1 = g_2 = \partial_t F = 0$ , the smoothness of  $\Omega$  and that  $V_H$  denotes the space of weighted extended B-splines on a uniform grid. If  $V_H^\varepsilon$  is defined according to (16) and if  $u_H^\varepsilon \in (0, T; V_H^\varepsilon)$  with  $u_H^\varepsilon(\cdot, 0) = \partial_t u_H^\varepsilon(\cdot, 0) = 0$  solves*

$$\langle \partial_{tt} u_H^\varepsilon(\cdot, t), v \rangle + (a^\varepsilon \nabla u_H^\varepsilon(\cdot, t), \nabla v)_{L^2(\Omega)} = (F, v)_{L^2(\Omega)}, \quad \text{for all } v \in V_H^\varepsilon \text{ and } t > 0,$$

then it holds for some  $\varepsilon$ -independent constant  $C > 0$  that

$$\|u_H^\varepsilon - u_\varepsilon\|_{L^2(H^1)} + \|\partial_t u_H^\varepsilon - \partial_t u_\varepsilon\|_{L^\infty(L^2)} \leq CH \|F\|_{L^2(\Omega)}.$$

As a last remark we note that Theorem 2.5 does not rely on the usage of B-splines for  $V_H$  and that the result is also valid for alternative choices as long as the basis function of  $V_H$  are locally supported and sufficiently smooth (see [44] for details).

## 2.4 Approach 4 - Localized Orthogonal decomposition

The last approach that we shall discuss was proposed in [6] in the framework of the Localized Orthogonal Decomposition (LOD, cf. [41, 32]). Let  $V_H$  denote a (coarse)  $P1$  finite element space as in the introduction and let  $P_H : H_0^1(\Omega) \rightarrow V_H$  denote the corresponding  $L^2$ -projection, i.e.  $(P_H(v), \Phi_H)_{L^2(\Omega)} = (v, \Phi_H)_{L^2(\Omega)}$  for all  $\Phi_H \in V_H$ . The multiscale method in [6] is derived based on the following observation (see also [45]):

Let  $F \in L^2(\Omega)$  and let  $z_\varepsilon \in H_0^1(\Omega)$  denote the solution to the elliptic multiscale problem

$$\int_{\Omega} a^\varepsilon \nabla z_\varepsilon \cdot \nabla v = \int_{\Omega} F v \quad \text{for all } v \in H_0^1(\Omega).$$

Then the  $L^2$ -projection of  $z_\varepsilon$  (i.e. the  $L^2$ -best approximation of  $z_\varepsilon$  in  $V_H$ ) can be characterized as the unique solution to the Petrov-Galerkin problem

$$\int_{\Omega} a^\varepsilon \nabla P_H(z_\varepsilon) \cdot \nabla (\text{Id} + Q)(v_H) = \int_{\Omega} F (\text{Id} + Q)(v_H) \quad \text{for all } v_H \in V_H \quad (17)$$

and where  $-Q : V_H \rightarrow \ker P_H := W$  denotes the  $(a^\varepsilon \nabla \cdot, \nabla \cdot)_{L^2(\Omega)}$ -orthogonal projection into the kernel of the  $L^2$ -projection. Indeed, defining

$$V_H^\varepsilon := \{v_H + Q(v_H) \mid v_H \in V_H\},$$

we observe that  $H_0^1(\Omega) = V_H \oplus W$  with  $V_H \perp W$  with respect to the  $(\cdot, \cdot)_{L^2(\Omega)}$  scalar product while  $H_0^1(\Omega) = V_H^\varepsilon \oplus W$  with  $V_H \perp W$  with respect to the  $(a^\varepsilon \nabla \cdot, \nabla \cdot)_{L^2(\Omega)}$  scalar product.

As problem (17) is a finite dimensional problem (involving only the degrees of freedom from the coarse space  $V_H$ ), it can be solved cheaply, provided that an approximation to the operator  $Q$  is available. In particular, we have the regularity- and  $\varepsilon$ -independent error estimate

$$\|z_\varepsilon - P_H(z_\varepsilon)\|_{L^2(\Omega)} \leq CH \|z_\varepsilon\|_{H^1(\Omega)} \leq CH \|F\|_{L^2(\Omega)}.$$

Motivated by the above considerations, a low-dimensional discrete multiscale space is constructed as follows. Let  $V_h$  denote a conventional  $P1$  finite element space with fine mesh size “ $h < \varepsilon$ ” (so that all variations of  $a^\varepsilon$  are resolved) and such that  $V_H \subset V_h$ . We denote the kernel of  $L^2$ -projection  $P_H$  by  $W_h := \{w_h \in V_h \mid P_H(w_h) = 0\}$ .

1. In order to approximate the Ritz-projection  $Q_h : V_h \rightarrow W_h$  given by

$$(a^\varepsilon \nabla Q_h(v), \nabla w_h)_{L^2(\Omega)} = -(a^\varepsilon \nabla v, \nabla w_h)_{L^2(\Omega)}$$

for all  $w_h \in W_h$  in an efficient way, we make an affine decomposition of the right hand side into  $(a^\varepsilon \nabla v, \nabla w_h)_{L^2(\Omega)} = \sum_{K \in \mathcal{T}_H} (a^\varepsilon \nabla v, \nabla w_h)_{L^2(K)}$  and solve for each of the localized sources  $(a^\varepsilon \nabla v, \nabla w_h)_{L^2(K)}$  individually. Since the corresponding solutions can be shown to exhibit an exponential decay outside of  $K$ , we can replace the computational domain  $\Omega$  by a small environment  $U_k(K)$  of  $K$ . Here,  $U_k(K)$  is defined iteratively by

$$\begin{aligned} U_0(K) &:= K, \\ U_k(K) &:= \cup \{T \in \mathcal{T}_H \mid T \cap U_{k-1}(K) \neq \emptyset\} \quad k = 1, 2, \dots, \end{aligned} \tag{18}$$

i.e. the patch  $U_k(K)$  consists of the coarse element  $K$  and  $k$ -layers of coarse elements around it. With that, we define  $W_h(U_k(K)) := \{v_h \in W_h \mid v_h = 0 \text{ in } \Omega \setminus U_k(K)\}$  and solve for  $Q_h^K(v_H) \in W_h(U_k(K))$  with

$$\int_{U_k(K)} a^\varepsilon \nabla Q_h^K(v_H) \cdot \nabla w_h = - \int_K a^\varepsilon \nabla v_H \cdot \nabla w_h \quad \text{for all } w \in W_h(U_k(K)) \tag{19}$$

and we set the global approximation  $Q_{h,k}$  of  $Q_h$  to

$$Q_{h,k}(v_H) := \sum_{K \in \mathcal{T}_H} Q_h^K(v_H).$$

The stationary local problems (19) are fully independent from each other and can be hence solved in parallel. Furthermore, for small  $k$ , they are of small size and hence only involve a low computational complexity. Note that problem (19) only has to be solved for coarse basis functions that have a support in  $K$ . Details on how they can be solved practically are provided in [6] and [24].

2. With the precomputed operator  $Q_{h,k}$  we define the multiscale space by

$$V_H^{\text{ms}} := \{v_H + Q_{h,k}(v_H) \mid v_H \in V_H\}. \tag{20}$$

The space is low-dimensional and if  $k$  is large enough so that  $Q_{h,k} = Q_h$ , we obtain the  $(a^\varepsilon \nabla \cdot, \nabla \cdot)_{L^2(\Omega)}$ -orthogonal decomposition  $V_h = V_H^{\text{ms}} \oplus W_h$ . As  $k$  is typically only a small number, we speak about a *Localized Orthogonal Decomposition* (LOD).

3. March in time with a favorite time-discretization, where every step only involves operations in  $V_H^{\text{ms}}$  with  $\dim V_H^{\text{ms}} = \dim V_H$ .
4. For new source terms  $F$ , the results can be reused.

The Localized Orthogonal Decomposition approach can be shown to fulfill the following error estimates [6].

**Theorem 2.6.** *Let the localization parameter  $k \in \mathbb{N}$  be chosen such that  $k \simeq |\log(H)|$  and let  $V_H^{\text{ms}}$  denote the corresponding LOD multiscale space. Then, if  $u_H^{\text{ms}} \in (0, T; V_H^{\text{ms}})$  with  $u_H^{\text{ms}}(\cdot, 0) = \partial_t u_H^{\text{ms}}(\cdot, 0) = 0$  solves*

$$(\partial_{tt} u_H^{\text{ms}}(\cdot, t), v) + (a^\varepsilon \nabla u_H^{\text{ms}}(\cdot, t), \nabla v)_{L^2(\Omega)} = (F, v)_{L^2(\Omega)}, \quad \text{for all } v \in V_H^{\text{ms}} \text{ and } t > 0,$$

*there exists a constant  $C$  that only depends on the time  $T$ , on  $\Omega$  and on the mesh regularity, such that the following error estimates hold true*

$$\|u_\varepsilon - u_H^{\text{ms}}\|_{L^\infty(L^2)} \leq CH^2 \|F\|_{L^2(\Omega)} + e_h^{(1)}$$

and

$$\|\partial_t u_\varepsilon - \partial_t u_H^{\text{ms}}\|_{L^\infty(L^2)} + \|u_\varepsilon - u_H^{\text{ms}}\|_{L^\infty(H^1)} \leq CH + e_h^{(2)},$$

where  $e_h^{(1)} := \|u_\varepsilon - \Pi_h(u_\varepsilon)\|_{L^\infty(L^2)} + \|\partial_t u_\varepsilon - \Pi_h(\partial_t u_\varepsilon)\|_{L^1(L^2)}$  and  $e_{\text{disc}}^{(2)}(h) := \|\partial_t u_\varepsilon - \Pi_h(\partial_t u_\varepsilon)\|_{L^\infty(L^2)} + \|\partial_{tt} u_\varepsilon - \Pi_h(\partial_{tt} u_\varepsilon)\|_{L^1(L^2)} + \|u_\varepsilon - \Pi_h(u_\varepsilon)\|_{L^\infty(H^1)}$  denote the fine scale discretization errors, with  $\Pi_h$  being the Ritz-projection on  $V_h$ .

From Theorem 2.6 we see that the LOD allows for optimal error estimates in  $L^\infty(L^2)$ ,  $L^1(L^2)$  and  $L^\infty(H^1)$  without additional regularity assumptions. Furthermore, we see that the precomputations necessary to construct  $V_H^{\text{ms}}$  involves small elliptic problems in computational domains of size  $\mathcal{O}(H|\log(H)|)$ . Since the problems can be solved in parallel, the method can be implemented efficiently. As for Approach 3, the fact that the computations for Approach 4 can be localized is a considerable advantage compared to Approaches 1 and 2. In terms of the size of the local domains, we see that Approach 4 involves subdomains that are by the factor  $\sqrt{H}$  smaller than the ones necessary for Approach 3. On the down side, the local problems (19) in Approach 4 are saddle point problems which involves the computation and inversion of the Schur complement matrix associated with the constraint “ $P_H(w) = 0$ ” (cf. [24] for details). Approach 3 on the contrary only involves unconstrained local problems. As a final difference, the discrete spaces required for Approach 3 need to be of higher order and smooth (e.g. weighted extended B-splines), whereas Approach 4 can be implemented using conventional  $P1$  Lagrange finite element spaces.

Recently it has been shown that Approach 4 also intrinsically relaxes the CFL condition on adaptive meshes [47], which is very significant for corresponding time-discretizations. Generalizations of the approach to the Helmholtz equation in the context of high frequency wave propagation are given in [16, 28, 46].

## 2.5 The case of general initial values: G-convergence and perturbation arguments

In this section we shall discuss the case of general initial values, i.e.  $g_1, g_2, \partial_t F \neq 0$ . First, we easily observe that all the Approaches 1-4 can be straightforwardly modified to fit this case. Typically the discrete initial values are picked as the  $L^2$ - or Ritz-projections onto the multiscale space. Why is it therefore necessary to discuss this case independently? To understand the issue, note that if  $g_1 = g_2 = \partial_t F = 0$ , then all higher order time derivatives of  $u_\varepsilon$  in  $t = 0$  will vanish as well, i.e. we have (as far as it exists)  $\partial_t^j u_\varepsilon(\cdot, 0) = 0$  for all  $j \geq 0$ . However, if we consider the general case, we figure out that for  $j \geq 2$  the time-derivatives in  $t = 0$  are linked through  $\partial_t^j u_\varepsilon(\cdot, 0) = \partial_t^{j-2} F(\cdot, 0) + \nabla \cdot (a^\varepsilon \nabla \partial_t^{j-2} u_\varepsilon(\cdot, 0))$ . From that expression we can see that the “higher order initial values” can be rapidly oscillating with a high amplitude of order  $\mathcal{O}(\varepsilon^{1-j})$  for  $j \geq 2$ . Unfortunately, these terms will just pop up on the right hand side of the a priori error estimates for Approach 1-4, i.e. in the Theorems 2.2, 2.5 and 2.6 (see also [6, 35, 36, 43, 44]) and the convergences rates can again depend on  $H/\varepsilon$ , what we just wanted to avoid.

In fact, there is currently no positive result on the question if **(C1)** and **(C2)** can be still fulfilled simultaneously for general initial values. However, a positive answer has been recently given when we replace the  $L^\infty(H^1)$  in **(C2)** by the  $L^\infty(L^2)$  norm.

As for elliptic multiscale problems, it is easy to verify that the  $L^p$ -norm of any space derivative of  $u_\varepsilon$  that is higher than 1, will explode with decreasing  $\varepsilon$ , i.e. we can say that for any  $s > 1$  we have in general that  $\|u_\varepsilon\|_{L^1(H^s)} \xrightarrow{\varepsilon \rightarrow 0} \infty$ . It is tempting to assume that the  $L^1(H^s)$ -norms (for  $s > 1$ ) are the only norms that should be avoided. However, unfortunately this is not the case and the statement can be “often” generalized to  $\|u_\varepsilon\|_{W^{m,2}(H^s)} \rightarrow \infty$  for  $\varepsilon \rightarrow 0$ , whenever  $m + s > 1$  (cf. [18]). The vague quantification “often” is detailed by the following theorem (cf. [6]), which directly links the problem to the choice of initial values.

**Theorem 2.7** (Time-regularity and regularity estimates). *Let  $u_\varepsilon$  denote the solution to the wave equation (3) and assume that  $F \in W^{m,2}(0, T; L^2(\Omega))$  for some  $m \in \mathbb{N}$ . Recursively we define the generalized initial values  $w_j^\varepsilon \simeq \partial_t^j u_\varepsilon(\cdot, 0)$  by*

$$w_0^\varepsilon := g_1, \quad w_1^\varepsilon := g_2, \quad w_j^\varepsilon := \partial_t^{j-2} F(\cdot, 0) + \nabla \cdot (a^\varepsilon \nabla w_{j-2}^\varepsilon) \quad \text{for } j = 2, 3, \dots, m+1. \quad (21)$$

If  $w_j^\varepsilon \in H_0^1(\Omega)$  for  $0 \leq j \leq m$  and  $w_{m+1}^\varepsilon \in L^2(\Omega)$ , we have

$$\partial_t^m u_\varepsilon \in L^\infty(0, T; H_0^1(\Omega)); \quad \partial_t^{m+1} u_\varepsilon \in L^\infty(0, T; L^2(\Omega)) \quad \text{and} \quad \partial_t^{m+2} u_\varepsilon \in L^2(0, T; H^{-1}(\Omega))$$

and it holds the (optimal) regularity estimate

$$\begin{aligned} & \|\partial_t^m u_\varepsilon\|_{L^\infty(0, T; H^1(\Omega))} + \|\partial_t^{m+1} u_\varepsilon\|_{L^\infty(0, T; L^2(\Omega))} \\ & \leq C \left( \|F\|_{W^{m,2}(0, T; L^2(\Omega))} + \|w_m^\varepsilon\|_{H^1(\Omega)} + \|w_{m+1}^\varepsilon\|_{L^2(\Omega)} \right), \end{aligned} \quad (22)$$

where  $C > 0$  is a constant that only depends on  $T$  and  $\Omega$ .

From the theorem we see that we can only hope for  $\varepsilon$ -independent bounds for e.g.  $\|\partial_t u_\varepsilon\|_{L^\infty(H^1)}$  and  $\|\partial_{tt} u_\varepsilon\|_{L^\infty(L^2)}$ , if the initial values  $g_1$  and  $g_2$  are picked such that they cancel out the variations of  $a^\varepsilon$ . As considered in the previous sections, a trivial case is the case with  $g_1 = g_2 = 0$  and  $F \in L^2(\Omega)$  being constant in time, for which we conclude  $\partial_t^m u_\varepsilon \in L^\infty(0, T; H_0^1(\Omega))$  for all  $m \in \mathbb{N}$  and  $\|\partial_t^m u_\varepsilon\|_{L^\infty(0, T; H^1(\Omega))} \leq C \|F\|_{L^2(\Omega)}$ . This guarantees that  $\varepsilon$ -oscillations cannot enter through time derivatives.

In general, we observe that the right hand side of the regularity estimate (22) can be only bounded independently of  $\varepsilon$  (i.e. independent of the speed of the variations of  $a^\varepsilon$ ) if  $g_1$ ,  $g_2$  and  $\partial_t F$  are trivial or if  $g_1 = g_1^\varepsilon$ ,  $g_2 = g_2^\varepsilon$  and  $F(\cdot, 0) = F^\varepsilon(\cdot, 0)$  are  $\varepsilon$ -dependent and such that they interact adequately with the variations of  $a^\varepsilon$ . This assumption can be hard to verify in practice. Also observe that if not  $g_1 = g_2 = \partial_t F(\cdot, 0) = 0$ , the assumptions of Theorem 2.7 for  $m \geq 1$  can typically only be fulfilled if  $a^\varepsilon$  admits a higher order regularity such as  $a^\varepsilon \in W^{1,\infty}(\Omega)$ . This can be problematic in realistic applications, where the propagation field  $a^\varepsilon$  is often discontinuous. The missing smoothness of  $a^\varepsilon$  can hence be a further issue in addition to the multiscale character.

An argument to overcome the problems that arise from general initial values is presented in [6]. Here it is proposed to slightly perturb the initial value  $g_1$  in such a way, that the effect is almost “invisible” in  $L^2$ ; i.e.  $\|g_1 - g_1^\varepsilon\|_{L^2} \leq \delta \ll 1$ ; but such that the variations of the perturbed initial value  $g_1^\varepsilon$  interact adequately with the variations of  $a^\varepsilon$ ; i.e.  $\|\nabla \cdot (a^\varepsilon \nabla g_1^\varepsilon)\|_{L^2(\Omega)} = \mathcal{O}(1)$ . By stability arguments, the solution  $\hat{u}_\varepsilon$  to the perturbed initial value is close to the original solution  $u_\varepsilon$ , in the sense that  $\|\hat{u}_\varepsilon - u_\varepsilon\|_{L^\infty(L^2)} \leq \|g_1 - g_1^\varepsilon\|_{L^2} \leq \delta$ . Since the second order initial value  $\partial_{tt} u_\varepsilon(\cdot, 0) = F(\cdot, 0) + \nabla \cdot (a^\varepsilon \nabla g_1^\varepsilon)$  is harmless (it does no longer blow up with decreasing  $\varepsilon$ ), the arguments for the trivial case apply again.

The construction of  $g_1^\varepsilon$  is obtained via  $G$ -convergence (see Definition 1.1). For instance, for the LOD-approach (Section 2.4) the following result has been proved in [6].

**Theorem 2.8.** *With the assumptions and the notations of Theorem 2.6, we assume that  $a^\varepsilon$  is  $G$ -convergent with  $G$ -limit  $a^0$  and we let  $g_1^\varepsilon \in H_0^1(\Omega)$  denote the solution to*

$$\int_{\Omega} a^\varepsilon \nabla g_1^\varepsilon \cdot \nabla v = \int_{\Omega} a^0 \nabla g_1 \cdot \nabla v \quad \text{for all } v \in H_0^1(\Omega), \quad (23)$$

i.e.  $g_1^\varepsilon$  is constructed such that its homogenized limit coincides with  $g_1$ . If  $g_2 \in H_0^1(\Omega)$ ,  $\partial_t F \in L^2(0, T; L^2(\Omega))$  and  $\nabla \cdot (a^0 \nabla g_1) + F(\cdot, 0) \in L^2(\Omega)$  then it holds

$$\lim_{h \rightarrow 0} \|u_\varepsilon - u_H^{\text{ms}}\|_{L^\infty(L^2)} \leq CH \left( \|F\|_{W^{1,2}(L^2)} + \|g_1\|_{H^1} + \|g_2\|_{H^1} + \|\nabla \cdot (a^0 \nabla g_1)\|_{L^2} \right) + \|g_1^\varepsilon - g_1\|_{L^2},$$

where  $C = C(T)$ .

Since  $\|g_1^\varepsilon - g_1\|_{L^2} \rightarrow 0$  for  $\varepsilon \rightarrow 0$  by the definition of  $G$ -convergence, we can assume that  $\|g_1^\varepsilon - g_1\|_{L^2} \leq CH$  and hence the total observable convergence rate will be linear in  $H$ , i.e.  $\lim_{h \rightarrow 0} \|u_\varepsilon - u_H^{\text{ms}}\|_{L^\infty(L^2)} \lesssim H$ . Note that  $g_1^\varepsilon$  in (23) is never computed, but is just a tool useful for the analysis. We close this section by mentioning that the above arguments do no longer apply for  $L^\infty(H^1)$ -errors.

### 3 Numerical methods for the wave equation in heterogeneous media with scale separation

When the oscillatory tensor  $a^\varepsilon$  in the wave equation (1) exhibit scale separation, numerical homogenization methods with much lower computational cost than the methods described in Section 2 can be constructed. A typical situation of scale separation for the oscillatory coefficient  $a^\varepsilon$  is for example a locally periodic structure, i.e., when  $a^\varepsilon$  is of the form  $a^\varepsilon(x) = a(x, x/\varepsilon) = a(x, y)$ , where  $a(x, y)$  is  $Y$ -periodic in  $y$ . Here  $Y$  is a unit cell, e.g.,  $Y = (0, 1)^d$ . Another situation is the case of a tensor  $a^\varepsilon$  modeled by a random field  $a^\varepsilon = a(x/\varepsilon, \omega)$ , where the analog of the periodicity in this case is the stationarity of the field. In this review we will focus on locally periodic tensor and refer to [27] for a recent account on the theory and numerics for wave in random media.

The theoretical framework for the development of numerical methods in locally periodic media is that of  $G$ -convergence. Indeed in locally periodic media, it is known that for each  $x \in \Omega$  the whole sequence of symmetric oscillatory tensors  $(a^\varepsilon)_{\varepsilon > 0}$   $G$ -converges to a unique effective tensor  $a^0(x)$ . Hence the whole family of solutions  $u_\varepsilon$  of the wave equation (1) converges to the solution of an effective wave equation (8). It can be shown that  $a^0$  is again in  $\mathcal{M}(\alpha, \beta, \Omega)$  and

$$a_{ij}^0(x) = \frac{1}{|Y|} \int_Y e_i^T a(x, y) (\nabla \chi_j(x, y) + e_j) dy, \quad (24)$$

where  $\chi^j(x, y)$ ,  $j = 1, \dots, d$  are the solutions in  $H_{per}^1(Y)$  of so-called cell problems

$$\int_Y a(x, y) (\nabla \chi_j(x, y) + e_j) \cdot \nabla w(y) dy = 0 \quad \text{for all } w \in H_{per}^1(Y), \quad (25)$$

where  $e_j$  are the vectors of the canonical basis of  $\mathbb{R}^d$ . Observe that the solution of problem (25) is unique up to a constant that needs to be fixed. Except for the case when  $a^\varepsilon$  is periodic, the full computation of the map  $x \rightarrow a^0(x)$  is not possible as it relies on infinitely many PDE solutions. Hence a numerical homogenization scheme must rely on a finite set of well chosen homogenized tensors  $a^0(x_j)_{j=1}^M$ .

#### 3.1 Effective model and numerical homogenization method for short-time wave propagation

A number of numerical methods based on homogenization theory and the asymptotic expansions [13] have been recently proposed. We mention the method given in [20] that applies to problems with uniformly periodic oscillatory tensors that are assumed to be symmetric with respect to the center of the periodic cell. Furthermore the macroscopic computational domain is assumed to be the union of an integer number of periodic cells. Numerical homogenization methods for the wave equation have also appeared in the geoscience community. In [38] a numerical homogenization method based on asymptotic expansion [13] is derived for elastic waves. Unlike the FE-HMM algorithm presented below that couples macro and micro scales in a global numerical scheme, the procedure in [38] is sequential and consists in first finding an effective tensor (analytical expression are used for 1d periodic problems, or spectral element method are used for 2d problems) and then solving the effective wave equation. It should be noted that a filtering technique to compute effective parameters for highly oscillatory non-periodic wave equations is discussed. While no rigorous theoretical foundation of such procedure

seems to be available, this technique is shown to give good results for some problems. Such filtering techniques could also be used in a pre-processing step for the numerical homogenization algorithm described below. Recently, a finite difference method (FDM) [22] and a finite element method [3] for wave problems in highly oscillatory media have been proposed in the framework of heterogeneous multiscale methods (HMM). These HMM methods are general algorithms to approximate numerically the homogenized solution in case of scale separation in the oscillatory tensor  $a^\varepsilon$  that we will describe in more detail. We discuss here finite element or finite difference algorithms but note that HMM methods have also been coupled with spectral element methods [2]. Such a coupling could readily be implemented for the methods described below.

### 3.1.1 Approach 1 - Finite-element numerical homogenization method

In this section we describe the finite element heterogeneous multiscale method (FE-HMM) proposed in [3] that can be seen as a general numerical homogenization method. We pick a standard macroscopic finite element space  $V_H$  made of piecewise polynomial on each macro element  $K$  of a partition of the computational domain  $\Omega = \cup_{K \in \mathcal{T}_H} K$ . We next define a sampling domain  $K_\delta$  (of size  $\delta$  comparable to  $\varepsilon$ ) within each macro element  $K$ . A micro finite element space with a triangulation that resolves the fine scale  $\varepsilon$  is defined in each sampling domain. We consider then the following problem: find  $u_H : [0, T] \rightarrow V_H$  such that

$$(\partial_{tt} u_H(\cdot, t), v_H) + B_H(u_H(\cdot, t), v_H) = (F(\cdot, t), v_H) \quad \text{for all } v_H \in V_H, \quad (26)$$

with appropriate projection of the true initial conditions, where

$$B_H(u_H, v_H) := \sum_{K \in \mathcal{T}_H} \frac{|K|}{|K_\delta|} \int_{K_\delta} a^\varepsilon(x) \nabla u_K^h \cdot \nabla v_K^h dx, \quad (27)$$

and  $u_K^h$  (respectively  $v_K^h$ ) are solutions of the following micro problems: for  $K \in \mathcal{T}_H$  find  $(u_K^h - u_H) \in V_h(K_\delta)$  such that

$$\int_{K_\delta} a^\varepsilon(x) \nabla u_K^h \cdot \nabla z^h dx = 0 \quad \text{for all } z^h \in V_h(K_\delta). \quad (28)$$

**Remark 3.1.** The above method is a particular case of a more general algorithm proposed in [3]. Indeed,  $V_H$  can be replaced by  $V_H^\ell$  in which  $v_H|_K$  is either a polynomial of total degree  $\ell$  if  $K$  is a simplex or a polynomial of degree at most  $\ell$  in each variable if  $K$  is a parallelogram. We then need  $j = 1, \dots, J$  integration points and sampling domains  $K_{\delta_j} = x_{K_j} + \delta I$ , where  $I = (-1/2, 1/2)^d$  and  $\delta_j$  the size of the sampling domain is such that  $\varepsilon \leq \delta_j \ll H$ . The bilinear form (27) becomes

$$B_H(u_H, v_H) = \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{K_j}}{|K_{\delta_j}|} \int_{K_{\delta_j}} a^\varepsilon(x) \nabla u_{K_j}^h \cdot \nabla v_{K_j}^h dx, \quad (29)$$

where  $u_{K_j}^h$  are solutions of (28) in  $K_{\delta_j}$  such that  $(u_{K_j}^h - u_{H,\text{lin}}) \in V_h^q(K_{\delta_j})$ , where  $u_{H,\text{lin}}$  is a piecewise linear approximation of  $u_H$  in  $K$  around  $x_{K_j}$  and  $V_h^q(K_{\delta_j})$  is a  $q$ -th order finite element space.

A fully discrete analysis of the method (26) has been obtained in [3]. Let the assumptions (H0) be fulfilled and consider a sequence  $(a^\varepsilon)_{\varepsilon>0} \subset \mathcal{M}(\alpha, \beta, \Omega)$  that  $G$ -converges to  $a^0 \in \mathcal{M}(\alpha, \beta, \Omega)$ . Then for the approximation of the solution  $u_0$  of the corresponding wave equation (8) by the FE-HMM solution (26) with the general modified bilinear form (29) it holds

**Theorem 3.2.** *Under suitable regularity of the solutions  $u_{K_j}^h$  to the micro-cell problems and of the homogenized tensor  $a^0$  the error  $e_H = u_0 - u_H$  satisfies*

$$\|\partial_t e_H\|_{L^\infty(0,T;L^2(\Omega))} + \|e_H\|_{L^\infty(0,T;H^1(\Omega))} \leq C \left( H^l + \left(\frac{h}{\varepsilon}\right)^{2q} + \text{err}_{\text{mod}} \right), \quad (30)$$

$$\|e_H\|_{L^\infty(0,T;L^2(\Omega))} \leq C \left( H^{l+1} + \left(\frac{h}{\varepsilon}\right)^{2q} + \text{err}_{\text{mod}} \right), \quad (31)$$

where  $C$  is independent of  $H, h, \varepsilon$  but depends on  $T$ .

The term  $err_{mod}$  stands for the modeling error and quantifies how well the micro averaging procedure captures the homogenized tensor at a given quadrature point  $x_{K_j}$ . This error depends on the size of the sampling domain  $K_{\delta_j}$ , the boundary condition used to solve (28), the nature of the scale separation in  $a^\varepsilon$  (e.g., local periodicity, random stationary, etc.). For example, for locally periodic coefficients, if the measure of  $K_{\delta_j}$  is an integer number of cubes of size  $\varepsilon$  and  $V_h^q(K_{\delta_j})$  is a subspace of a Sobolev space of periodic functions, we have  $err_{mod} = 0$ . If  $V_h^q(K_{\delta_j})$  is a subspace of  $H_0^1(K_{\delta_j})$  and  $\delta > \varepsilon$  arbitrary, then  $err_{mod} = C(\delta + \varepsilon/\delta)$  [3]. Further results and recent improvement of this bounds also for random stationary problems are reported in [30].

We conclude this subsection by discussing the computational complexity per time-step  $\Delta t$  for the FE-HMM. For simplicity of the discussion, let  $\ell = q = 1$ . Assuming that  $N_{macro} = 1/H$ ,  $H$  is the diameter of the macroscopic FE, the size of the system of ODEs that need to be solved at each time step  $\Delta t$  is  $N = H^{-d}$  for piecewise linear macro FEs. We denote this computational complexity per time step as  $\text{cost}(\Delta t, H^{-d})$ . Notice that the stability constraint if using an explicit time-integrator reads  $\Delta t \simeq H \gg \varepsilon$ . Hence the macroscopic time-step can be chosen independently of  $\varepsilon$  in sharp contrast with the resolved FEM for the original highly oscillatory wave equation (1). Next we discuss the offline cost for the FE-HMM: this is the cost involved in solving the micro problems in each sampling domain  $K_{\delta_j}$ . We note that this computation has only to be performed once at the beginning of the macroscopic time integration, unless the highly oscillatory tensor  $a^\varepsilon$  is time dependent. Assume that we use  $N_{micro}$  elements in each space dimension for the discretization of the sampling domains  $K_{\delta_j}$ , then  $h = \delta/N_{micro}$  and  $h/\varepsilon = \delta/(\varepsilon N_{micro}) = \mathcal{O}(1/N_{micro})$  as  $\delta \simeq \varepsilon$ . Hence the microscopic degrees of freedom read  $M_{micro} = (N_{micro})^d$  and the total offline cost is given by the solution of  $\mathcal{O}(H^{-d})$  linear systems of size  $(N_{micro})^d$ . As micro and macro error must be balanced, in view of the estimates (30), (31) we can choose  $N_{micro} = \sqrt{N_{macro}}$  to guarantee a macroscopic linear convergence rate in the energy norm and  $N_{micro} = N_{macro}$  for a macroscopic quadratic convergence rate in the  $L^2(L^\infty)$  norm.

### 3.1.2 Approach 2 - Finite-difference numerical homogenization method

In this section we describe a finite difference heterogeneous multiscale method (FD-HMM) proposed in [22] that can be seen as a general finite difference numerical homogenization method. The FD-HMM follows the methodology first given in [1] and is also related to the HMM for first order hyperbolic problems given in [17]. Consider the homogenized equation (8) in strong form (where we set  $F \equiv 0$  for simplicity)

$$\begin{aligned} \partial_{tt}u_0 &= \nabla \cdot (a^0 \nabla u_0) && \text{in } \Omega \times ]0, T[, \\ u_0 &= 0 && \text{on } \partial\Omega \times ]0, T[, \\ u_0(x, 0) &= g_1(x), \quad \partial_t u_0(x, 0) = g_2(x) && \text{in } \Omega. \end{aligned} \quad (32)$$

Consider a spatial grid  $x_{i_1, \dots, i_d}, i_1 \dots, i_d = 1, \dots, N_{macro}$  with mesh size  $H$  for  $\Omega$ . The first step is the macroscopic discretization of the flux formulation for the effective equation. For simplicity of exposition we set  $d = 2$  and  $(i_1, i_2) = (i, j)$ . We seek the evolution of a function  $U(t) = \{U_{i,j}(t)\}$  such that

$$\frac{d^2}{dt^2} U_{i,j}(t) = \frac{1}{H} (J_{i+\frac{1}{2},j}^H(t) - J_{i-\frac{1}{2},j}^H(t)) + \frac{1}{H} (J_{i,j+\frac{1}{2}}^H(t) - J_{i,j-\frac{1}{2}}^H(t)), \quad (33)$$

where  $J_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^H(t) \simeq a^0(x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}) \nabla u^0(x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}, t)$ . The approximation of  $\nabla u_0(\cdot, t)$  at the points  $x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}$  is denoted by  $P_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^H(t)$ . For example

$$P_{i,j+\frac{1}{2}}^H = \left( \frac{(U_{i+1,j+1} - U_{i+1,j}) - (U_{i-1,j+1} - U_{i-1,j})}{4H} \right), \quad (34)$$



and similarly for the other gradients  $P_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^H$ .

The evaluation of  $J_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^H(t)$  is obtained by solving a micro problem that relies, as for the FE-HMM, on the original multiscale data. For each domain  $K_{mic} = K_\delta(x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}) := x_{i\pm\frac{1}{2},j\pm\frac{1}{2}} + \delta I$  we solve for the oscillatory wave equation (1)

$$\begin{aligned} \partial_{ss} u_{mic}(x, s) &= \nabla \cdot (a^\varepsilon(x) \nabla u_{mic}(x, s)) \quad \text{in } K_\delta(x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}) \times [0, \tau], \\ u_{mic}(x, 0) &= P_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^H(t) \cdot x, \quad \partial_s u_{mic}(x, 0) = 0, \quad \text{on } K_\delta(x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}) \times \{s = 0\}, \\ u_{mic}(x, s) - u_{mic}(x, 0) &\text{ is } K_\delta(x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}) \text{-periodic.} \end{aligned} \quad (35)$$

The computation of the macroscopic flux  $J_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^H$  is then obtained by computing the space time average of  $a^\varepsilon \nabla u_{mic}$  over the domains  $K_\delta(x_{i\pm\frac{1}{2},j\pm\frac{1}{2}}) \times [0, \tau]$ . The computation of this macroscopic flux can be further improved by computing suitable weighted averages of the microscopic flux. This can reduce the error coming from the artificial boundary conditions (the so called resonance or modeling error) and the error arising from the time averaging of the oscillatory functions (we refer to [22] for details).

**Remark 3.3.** Of course in practice (35) needs to be discretized numerically. In [22], the same numerical method as for the macro problem (33) is used. Precisely, we consider a spatial grid  $x_{k,l}$ ,  $k, l = 1, \dots, N_{micro}$  with meshsize  $h$  for each domain  $K_\delta$  and seek the evolution of a micro function  $u = \{u_{k,l}\}$  such that

$$\frac{d^2}{ds^2} u_{k,l}(s) = \frac{1}{h} (J_{k+\frac{1}{2},l}^h(s) - J_{k-\frac{1}{2},l}^h(s)) + \frac{1}{h} (J_{k,l+\frac{1}{2}}^h(s) - J_{k,l-\frac{1}{2}}^h(s)), \quad (36)$$

together with the initial and boundary condition of (35). Here  $J_{k\pm\frac{1}{2},l\pm\frac{1}{2}}^h = a^\varepsilon(x_{k\pm\frac{1}{2},l\pm\frac{1}{2}}) p_{k\pm\frac{1}{2},l\pm\frac{1}{2}}$ , where  $p_{k\pm\frac{1}{2},l\pm\frac{1}{2}}$  is the micro gradient that can be computed with a similar formula as in (34) with  $u_{k,l}$  instead of  $U_{i,j}$  and  $h$  instead of  $H$ . The weighted average of the discrete micro gradients is then used to compute the macro fluxes  $J_{i\pm\frac{1}{2},j\pm\frac{1}{2}}^H$ . We note that  $h < \varepsilon$  is required to compute these micro problems.

In [22], the method presented above has been analyzed for periodic problems in one dimension assuming  $H/\varepsilon \in \mathbb{N}$  and the following error estimate has been derived

$$\sup_{0 \leq t_n \leq T} |U_i^n - u^0(x_i, t_n)| \leq C(H^2 + err_{mic} + err_{mod}) \quad \text{for all } x_i \in \Omega, \quad (37)$$

where  $C$  depends on  $T$ . The analysis presented in [22] assumes an exact micro flux  $a^\varepsilon \nabla u_{mic}$ . A generalized analysis was recently provided in [11]. Taking into account the discretization error introduced by (36) would allow further to quantify  $err_{mic}$ . Notice that the modeling error can be quantified for periodic problems and depends on the choice of the weighted average used to compute the micro flux averages.

Similarly as for the FE-HMM, assuming a macroscopic mesh size  $H$  the size of the system of ODEs that need to be solved at each time step  $\Delta t$  is  $N = H^{-d}$ . However at first sight, the evaluation of the fluxes need to be performed at each macro time step, representing an additional cost of  $\tau/(\delta t)$  steps of a micro time integrator solving an ODE of size  $M_{micro} = (N_{micro})^d$  corresponding to (35) ( $N_{micro}$  is independent of  $\varepsilon$  and corresponds to the number of grid points in each spatial dimension of the mesh used for the sampling domain  $K_\delta$ ). But as the computation of the macroscopic flux  $J^H$  depends linearly on the macroscopic gradient  $P^H$  through (35), we observe that  $J^H$  can be expanded in a linear combination of precomputed macroscopic fluxes  $J^H$  solution of (35) with initial conditions  $u_{mic}(x, 0) = e_i \cdot x$ , where  $e_i$ ,  $i = 1, \dots, d$  is the canonical basis of  $\mathbb{R}^d$ . This is of course only valid if the oscillatory tensor  $a^\varepsilon$  is time-independent, in which case the solutions of the micro problems in the FD-HMM can be considered as an offline cost.

### 3.2 Effective model and numerical homogenization method for long-time wave propagation

It has been observed for some time [48, 26] that while classical homogenization describes well wave propagation in heterogeneous medium for short time, i.e.,

$$\|u_\varepsilon - u_0\|_{L^\infty(0,T;L^2(\Omega))} \leq C\varepsilon, \quad (38)$$

dispersive effects accumulate for longer time in the oscillatory wave  $u_\varepsilon$  that are not captured by the homogenized wave  $u_0$ . Precisely, the estimate (38) is no longer valid for time interval  $[0, T\varepsilon^{-2}]$ . For periodic problems, using formal Bloch waves techniques, a higher order effective model has been derived in [48] capable of capturing the dispersive effects of the true wave over longer time intervals. Based on this effective model, a generalization of the FD-HMM method (3.1.2) has been derived in [23]. However, as the effective model derived in [48, 26] is ill-posed, the related numerical method needs to be regularization. The whole numerical scheme is not trivial to analyze, and while in [10] the flux error of the generalized FD-HMM scheme has been analyzed for one-dimensional periodic problems, a full analysis of the numerical scheme is yet to be done. Recently another dispersive limit has been derived in [39], again for periodic problems, in the form of a Boussinesq type equation and the first rigorous error analysis for time intervals  $[0, T\varepsilon^{-2}]$  for the error between  $u_\varepsilon$  and the solution of the Boussinesq equation has been derived. Multi-dimensional problems have been analyzed in [19] using Bloch waves techniques. The approach in [39] has been generalized in [7], where a whole family of effective equations has been derived and rigorous error estimates of the corresponding solutions towards the highly oscillatory wave  $u^\varepsilon$  have been established. Generalization for multi-dimensional problems have appeared in [8]. A generalization of the FE-HMM, called FE-HMM-L valid over long-time has been proposed in [4, 5]. In [7] a rigorous analysis of the FE-HMM-L has been given. In the following we briefly describe various effective models valid over long-time intervals and discuss corresponding numerical schemes.

#### 3.2.1 A family of effective equations for the wave equation over long-time

Consider (1) with  $a^\varepsilon = a(x/\varepsilon) = a(y)$  is  $Y$ -periodic ( $Y$  is a unit cube, e.g.,  $(0, 1)^d$ ). In this situation the homogenized model is given according to (8) by

$$\partial_{tt}u_0 + a_{ij}^0 \partial_{ij}^2 u_0 = F, \quad (39)$$

that we write here in a slightly different form using Einstein summation rule. The constant homogenized tensor  $a^0$  is given by (24) and the initial conditions for (39) are those given in (8). For time intervals of length  $[0, T\varepsilon^{-1}]$  it is still possible to show that [7]

$$\|u_\varepsilon - u_0\|_{L^\infty(0,T\varepsilon^{-1};L^2(\Omega))} \leq C\varepsilon, \quad (40)$$

where  $C$  depends on  $T$  but not on  $\varepsilon$ . Hence the numerical methods described in Sections 3.1.1, 3.1.2 are still a good approximation of  $u^\varepsilon$  in the  $L^\infty(L^2)$ . For example for the FE-HMM we obtain

$$\|u_\varepsilon - u_H\|_{L^\infty(0,T\varepsilon^{-1};L^2(\Omega))} \leq C \left( H^{l+1} + \left(\frac{h}{\varepsilon}\right)^{2q} \right), \quad (41)$$

where for periodic problem, the modeling error error in (31) vanishes if one uses periodic FE spaces for the micro solver with an integer number of oscillatory periods in each direction of the sampling domains. The model (39) however does not longer capture effective behavior of the true oscillatory wave  $u_\varepsilon$  for time intervals  $[0, T\varepsilon^{-2}]$  [48, 26]. In turn, an estimate such as (40) is no longer valid. Building on [39, 19] and following [8] we consider for a domain  $\Omega$  that is a union of cells of volume  $\varepsilon^d|Y|$  the following effective equation

$$\partial_{tt}\tilde{u} - a_{ij}^0 \partial_{ij}^2 \tilde{u} + \varepsilon^2 (a_{ijkl}^2 \partial_{ijkl}^4 \tilde{u} - b_{ij}^0 \partial_{ij}^2 \partial_{tt}\tilde{u}) = F, \quad (42)$$

such that  $x \mapsto \tilde{u}(t, x)$  is  $\Omega$ -periodic with initial conditions given in (8). Notice that this is an effective model for the equation (1) with periodic boundary conditions on  $\Omega$ . We also note that most of the results below are valid on arbitrary large domain  $\Omega$  with appropriate changes [7, 19]. Indeed two related model for wave propagation over long-time can be considered: in the first model the wave propagates on long-time interval through a large spatial domain without hitting its boundary, in the second model, the spatial domain is  $\mathcal{O}(1)$  and the wave enters and leaves many times the domain thanks to the periodic boundary conditions. One crucial difference in the analysis is that in the former case, one needs to control the Poincaré inequality arising from the large diameter of  $\Omega$ . This can be done by choosing an appropriate weak norm in space. We refer to [7, 19] for details. In the above approximation, the coefficients  $a_{ij}^0$  are the homogenized coefficients that appear in (39). The coefficients  $b_{ij}^0$  are the entries of a matrix  $b^0 \in \text{Ten}^2(\mathbb{R}^d)$ , while the coefficients  $a_{ijkl}^2$  are the entries of a tensor of order four  $a^2 \in \text{Ten}^4(\mathbb{R}^d)$ , where  $\text{Ten}^n(\mathbb{R}^d)$  is the space of tensors of order  $n$  in  $\mathbb{R}^d$ . We also denote by  $\text{Sym}^n(\mathbb{R}^d)$  the subspace of  $\text{Ten}^n(\mathbb{R}^d)$  of symmetric tensors. If we assume that

$$b^0 \in \text{Sym}^2(\mathbb{R}^d), \quad b^0 \eta \cdot \eta \geq 0 \quad \forall \eta \in \mathbb{R}^d; \quad a_{ijkl}^2 = a_{ikji}^2, \quad a^2(\eta\eta^T) : (\eta\eta^T) \geq 0 \quad \forall \eta \in \mathbb{R}^d, \quad (43)$$

then (recall that  $a^0$  is elliptic and bounded) the problem (42) is well posed (see [8]). The approximation property of a solution to (42) is summarized in the following theorem proved in [8]. Recall that  $\chi_i$  is a solution to Problem (25). To analyze the solution to (42) over long-time intervals, further families of cell-problems need to be considered. First define a so-called adaptation operator of the form

$$\mathcal{B}^\varepsilon \tilde{u}(t, x) = \tilde{u}(t, x) + \varepsilon \chi_i(y) \partial_i \tilde{u}(t, x) + \varepsilon^2 \theta_{ij}(y) \partial_{ij}^2 \tilde{u}(t, x) + \dots, \quad (44)$$

where we note that an expansion of up to order four in  $\varepsilon$  is needed, that we skip here for simplicity. Plugging this approximation into (42) and separating the different powers of  $\varepsilon$  allows to define appropriate equations for  $\chi_i, \theta_{ij}$ . First we see that  $\chi_i$ , for  $i = 1, \dots, d$ , is a solution to Problem (25) (order  $\varepsilon^{-1}$ ), second at order  $\varepsilon^0$  we obtain that  $\theta_{ij}$ , for  $i, j = 1, \dots, d$ , are solutions of

$$\int_Y a(x, y) \nabla \theta_j(y) \cdot \nabla w(y) dy = G_{ij}(w) \quad \forall w \in H_{per}^1(Y), \quad (45)$$

where  $G_{ij}(w) = S_{ij}^2 \{ - (a(x, \cdot) e_i \chi_j, \nabla w)_{L^2(Y)} + (a(x, \cdot) (\nabla \chi_j + e_j) - a^0(x) e_j, e_i w)_{L^2(Y)} \}$  and  $S_{ij}^n(\cdot)$  denotes the symmetrization operator of a given tensor. Continuing this procedure up to order  $\varepsilon^2$  leads to the following conditions on the tensors  $b^0$  and  $a^2$  that guarantee an accurate approximation of the solution of (1) by the solution of (42) over long-time. We summarize results obtained in [7, 8].

**Theorem 3.4.** *Under appropriate regularity assumptions on the data  $a^\varepsilon, g_1, g_2$  and if  $b^0$  and  $a^2$  satisfy the relation*

$$S_{ijkl}^4 \{ a_{ijkl}^2 - a_{ij}^0 b_{kl}^0 \} = S_{ijkl}^4 \{ \langle a_{jk} \chi_l \chi_i \rangle_Y - \langle a \nabla \theta_{ji} \cdot \nabla \theta_{kl} \rangle_Y - a_{jk}^0 \langle \chi_l \chi_i \rangle_Y \}, \quad (46)$$

where  $\chi_l$  is any solution of (25) and  $\theta_{kl}$  is any corresponding solution of (45). Then, the following error estimate holds

$$\|u_\varepsilon - \tilde{u}\|_{L^\infty(0, T\varepsilon^{-2}; L^2(\Omega))} \leq C\varepsilon \quad (47)$$

where  $C = C(T)$  is independent of  $\varepsilon$ .

For one-dimensional problems, among the class of effective equations, there is a very simple representative that reads [39, 7]

$$\partial_{tt} \tilde{u} - a^0 \partial_{xx} \tilde{u} - \varepsilon^2 b^0 \partial_{ttxx} \tilde{u} = F, \quad (48)$$

where  $b^0 = \frac{1}{|Y|} \int_Y \chi^2 dy$  and  $\chi$  is the solution of the cell problem (25).

### 3.2.2 Numerical homogenization methods for the wave equations over long-time

We start by describing the generalization of the FE-HMM introduced in [4, 5]. Recall that the FE-HMM is given by (26). Next we replace the  $L^2$  scalar product  $(\partial_{tt}u^H, v^H)$  with  $(\partial_{tt}u^H, v^H)_Q$  where

$$(u_H, v_H)_Q := (u_H, w_H) + \sum_{K \in \mathcal{T}_H} \frac{|K|}{|K_\delta|} \int_{K_\delta} (u_K^h - v_H)(v_K^h - v_H) dx,$$

where  $u_K^h$  (respectively  $v_K^h$ ) are the micro functions already used in (27). The FE-HMM-L method is then defined by: find  $u_H : [0, T\varepsilon^{-2}] \rightarrow V_H$  such that

$$(\partial_{tt}u_H, v_H)_Q + B_H(u_H, v_H) = (F, v_H) \quad \forall v_H \in V_H. \quad (49)$$

We observe that the cost of this method is similar to the original FE-HMM as the modification of the  $L^2$  scalar product relies on micro functions that need anyway to be computed to assemble  $B_H(\cdot, \cdot)$ . As for the FE-HMM we can replace  $B_H(\cdot, \cdot)$  by the more general bilinear form (29) and likewise  $(u_H, v_H)_Q$  can be replaced by

$$(u_H, v_H)_Q = (u_H, w_H) + \sum_{K \in \mathcal{T}_H} \sum_{j=1}^J \frac{\omega_{K_j}}{|K_{\delta_j}|} \int_{K_{\delta_j}} (u_{K_j}^h - u_{H,\text{lin}})(v_{K_j}^h - v_{H,\text{lin}}) dx.$$

For one-dimensional problems, an error analysis over long-time of the FE-HMM-L has been given in [7]. Under suitable regularity assumptions it has been shown that

$$\|u^\varepsilon - u_H\|_{L^\infty(0, \varepsilon^{-2}T; L^2(\Omega))} \leq C \left( \varepsilon + (h/\varepsilon^2)^2 + H^{\ell+1}/\varepsilon^2 + H^\ell/\varepsilon \right) \quad (50)$$

where  $C$  is independent of  $\varepsilon$ . In fact it can be seen that for one-dimensional problems, the FE-HMM-L approximates the effective equation (48). We emphasize that in sharp contrast with the full fine-scale approximation with error estimate (7), we can take higher order FEs for the FE-HMM-L and do not require that  $H < \varepsilon$ , while higher order FEs in (7) would result in additional negative power of  $\varepsilon$  as discussed in the introduction. Furthermore even in the regime  $H < \varepsilon$  the computational cost of the FE-HMM-L is much smaller than the cost of applying a classical FEM applied to (1) as described below. For multi-dimensional problems, however, the fourth order term in the effective equation (42) does not vanish and the FE-HMM-L cannot approximate the full effective equation. In [8], an efficient numerical algorithm to approximate the effective coefficients in (42) based on the solution of  $d + \binom{d+1}{2}$  cell problems is given. The resolution of the effective equation relies on a fast Fourier transform algorithm.

A generalization of the FD-HMM proposed in [22] has also been derived in order to capture long-time dispersive effects [23, 10]. This method has only been investigated for one-dimensional problems and we briefly describe the method for this case. Similarly to (33) we assume  $F = 0$  and consider a FD macroscopic flux formulation for the evaluation of  $U = \{U_i\}$  that reads

$$\frac{d^2}{dt^2} U_i = \frac{1}{24H} (J_{i-3/2}^H - 27J_{i-1/2}^H + 27J_{i+1/2}^H - J_{i+3/2}^H). \quad (51)$$

For a given smooth function, the finite difference approximation based on the above scheme ensure the fourth order approximation of its derivative. The evaluation of  $J_{i\pm 3/2}^H, J_{i\pm 1/2}^H$  relies on micro problems similar to (35) but with higher order initial value, i.e.,  $u_{mic}(x, 0) = q(x)$ , where  $q(x)$  is a cubic polynomial obtained by interpolating the current macroscopic solution at four points around the micro sampling domains. Some care is required to define the actual  $q(x)$  used for computation as it is based on a subtle post-processing of the actual cubic interpolation polynomial [23, 10]. Finally, the actual computation of  $J_{i\pm 3/2}^H, J_{i\pm 1/2}^H$  is again based on a weighted average of  $a^\varepsilon \nabla u_{mic}$  over the

domains  $K_\delta(x_{i\pm\frac{1}{2}}) \times [0, \tau]$ . A more accurate procedure as for short-time approximation is also required for this averaging procedure. At a given point  $x$ , the computed flux  $J_x^H$  flux is shown to approximate the flux

$$\hat{F} = a^0 \partial_x \hat{u} + \varepsilon^2 b^0 \partial_{xxx} \hat{u}$$

that is the flux of the ill-posed effective equation derived in [48] that reads for  $F = 0$  (compare with (48))

$$\partial_{tt} \hat{u} - a^0 \partial_{xx} \hat{u} - \varepsilon^2 b^0 \partial_{xxxx} \hat{u} = 0. \quad (52)$$

Finally, we discuss the computational complexity of the FE-HMM-L for one-dimensional problems, as this is the only method for which an a priori error analysis is available that in turn gives an estimate on the size of the spatial mesh size needed for long-time integration. In view of the estimate (41) for  $\ell = 1$  the size of the linear system to be solved per time-step  $\Delta t$  is  $N = \varepsilon^{-1}$ . The CFL constraint reads here  $\Delta t \simeq \varepsilon$ . This is significantly less expensive than the cost per time step over long-time intervals by a classical FE solver, for which  $N = \varepsilon^{-3}$  and the CFL constrain reads  $\Delta t \simeq \varepsilon^3$ . Furthermore, for the FE-HMM-L higher order macro solvers can be used and hence  $H$  can be chosen larger than  $\varepsilon$  (see the error estimate (50)). Here again, as for the FE-HMM, the solution of the micro problems is a one shot offline cost that is negligible in view of the macroscopic cost for each time step  $\Delta t$  over a long-time time interval  $[0, T\varepsilon^{-2}]$ . In contrast, higher order FEM for a direct fine scale solver applied to (1) does not improve the time cost as already discussed in Section 1.

**Acknowledgements.** This work is partially supported by the Swiss National Foundation.

## References

- [1] A. Abdulle and W. E. Finite difference heterogeneous multi-scale method for homogenization problems. *J. Comput. Phys.*, 191(1):18–39, 2003.
- [2] A. Abdulle and B. Engquist. Finite element heterogeneous multiscale methods with near optimal computational complexity. *Multiscale Model. Simul.*, 6(4):1059–1084, 2007.
- [3] A. Abdulle and M. J. Grote. Finite element heterogeneous multiscale method for the wave equation. *Multiscale Model. Simul.*, 9(2):766–792, 2011.
- [4] A. Abdulle, M. J. Grote, and C. Stohrer. FE heterogeneous multiscale method for long-time wave propagation. *C. R. Math. Acad. Sci. Paris*, 351(11-12):495–499, 2013.
- [5] A. Abdulle, M. J. Grote, and C. Stohrer. Finite element heterogeneous multiscale method for the wave equation: long-time effects. *Multiscale Model. Simul.*, 12(3):1230–1257, 2014.
- [6] A. Abdulle and P. Henning. Localized orthogonal decomposition method for the wave equation with a continuum of scales. *ArXiv e-print 1406.6325*, to appear in *Math. Comp.*, 2014.
- [7] A. Abdulle and T. Pouchon. A priori error analysis of the finite element heterogeneous multiscale method for the wave equation over long time. submitted, 2015.
- [8] A. Abdulle and T. Pouchon. Effective models for the multidimensional wave equation in heterogeneous media over long time. preprint, 2016.
- [9] G. Alessandrini and V. Nesi. Univalent  $\sigma$ -harmonic mappings: connections with quasiconformal mappings. *J. Anal. Math.*, 90:197–215, 2003.

- [10] D. Arjmand and O. Runborg. Analysis of heterogeneous multiscale methods for long time wave propagation problems. *Multiscale Model. Simul.*, 12(3):1135–1166, 2014.
- [11] D. Arjmand and O. Runborg. Estimates for the upscaling error in heterogeneous multiscale methods for wave propagation problems in locally-periodic media. ArXiv e-print 1605.02386, 2016.
- [12] G. A. Baker. Error estimates for finite element methods for second order hyperbolic equations. *SIAM J. Numer. Anal.*, 13(4):564–576, 1976.
- [13] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*. North-Holland Publishing Co., Amsterdam, 1978.
- [14] S. Brahim-Otsmane, G. A. Francfort, and F. Murat. Correctors for the homogenization of the wave and heat equations. *J. Math. Pures Appl.*, 71(3):197–231, 1992.
- [15] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.
- [16] D. Brown, D. Gallistl, and D. Peterseim. Multiscale Petrov-Galerkin method for high-frequency heterogeneous Helmholtz equations. *ArXiv e-prints*, 2015.
- [17] S. Chen, W. E, and C.-W. Shu. The heterogeneous multiscale method based on the discontinuous Galerkin method for hyperbolic and parabolic problems. *Multiscale Model. Simul.*, 3(4):871–894, 2005.
- [18] D. Cioranescu and P. Donato. *An introduction to homogenization*, volume 17 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, New York, 1999.
- [19] T. Dohnal, A. Lamacz, and B. Schweizer. Dispersive homogenized models and coefficient formulas for waves in general periodic media. *Asymptot. Anal.*, 93(1-2):21–49, 2015.
- [20] Q.-L. Dong and L.-Q. Cao. Multiscale asymptotic expansions and numerical algorithms for the wave equations of second order with rapidly oscillating coefficients. *Appl. Numer. Math.*, 59(12):3008–3032, 2009.
- [21] Y. Efendiev and T. Y. Hou. *Multiscale finite element methods. Theory and applications*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, New York, 2009.
- [22] B. Engquist, H. Holst, and O. Runborg. Multi-scale methods for wave propagation in heterogeneous media. *Commun. Math. Sci.*, 9(1), 2011.
- [23] B. Engquist, H. Holst, and O. Runborg. Multiscale methods for wave propagation in heterogeneous media over long time. In *Numerical analysis of multiscale computations*, pages 167–186. Springer, 2012.
- [24] C. Engwer, P. Henning, A. Målqvist, and D. Peterseim. Efficient implementation of the localized orthogonal decomposition method. ArXiv e-print 1602.01658 (submitted), 2016.
- [25] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
- [26] J. Fish, W. Chen, and G. Nagai. Non-local dispersive model for wave propagation in heterogeneous media: multi-dimensional case. *Internat. J. Numer. Methods Engrg.*, 54(3):347–363, 2002.

- [27] J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sølna. *Wave propagation and time reversal in randomly layered media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.
- [28] D. Gallistl and D. Peterseim. Stable multiscale Petrov-Galerkin finite element method for high frequency acoustic scattering. *Comp. Meth. Appl. Mech. Eng.*, 295:1–17, 2015.
- [29] A. Gloria. Reduction of the resonance error. Part 1: Approximation of homogenized coefficients. *Math. Models Methods Appl. Sci.*, 21(8):1601–1630, 2011.
- [30] A. Gloria and Z. Habibi. Reduction of the resonance error - part 2: Approximation of correctors, extrapolation, and spectral theory. Preprint, hal-00933234, 2014.
- [31] I. G. Graham and R. Scheichl. Robust domain decomposition algorithms for multiscale PDEs. *Numer. Methods Partial Differential Equations*, 23(4):859–878, 2007.
- [32] P. Henning and D. Peterseim. Oversampling for the Multiscale Finite Element Method. *SIAM Multiscale Model. Simul.*, 11(4):1149–1175, 2013.
- [33] K. Höllig. *Finite element methods with B-splines*, volume 26 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2003.
- [34] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997.
- [35] L. Jiang and Y. Efendiev. A priori estimates for two multiscale finite element methods using multiple global fields to wave equations. *Numer. Methods Partial Differential Equations*, 28(6):1869–1892, 2012.
- [36] L. Jiang, Y. Efendiev, and V. Ginting. Analysis of global multiscale finite element methods for wave equations with continuum spatial scales. *Appl. Numer. Math.*, 60(8):862–876, 2010.
- [37] O. Korostyshevskaya and S. E. Minkoff. A matrix analysis of operator-based upscaling for the wave equation. *SIAM J. Numer. Anal.*, 44(2):586–612 (electronic), 2006.
- [38] J. M. L. Guillot, Y. Capdeville. 2-d non periodic homogenization for the sh wave equation. *Geophys. J. Int.*, 182:1438–1454, 2010.
- [39] A. Lamacz. Dispersive effective models for waves in heterogeneous media. *Math. Models Methods Appl. Sci.*, 21(9):1871–1899, 2011.
- [40] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I*. Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.
- [41] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014.
- [42] H. Owhadi and L. Zhang. Metric-based upscaling. *Comm. Pure Appl. Math.*, 60(5):675–723, 2007.
- [43] H. Owhadi and L. Zhang. Numerical homogenization of the acoustic wave equations with a continuum of scales. *Comput. Methods Appl. Mech. Engrg.*, 198(3):397–406, 2008.
- [44] H. Owhadi and L. Zhang. Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast. *Multiscale Model. Simul.*, 9(4):1373–1398, 2011.

- [45] D. Peterseim. Variational Multiscale Stabilization and the Exponential Decay of Fine-scale Correctors. ArXiv e-prints 2015. to appear.
- [46] D. Peterseim. Eliminating the pollution effect in Helmholtz problems by local subscale correction. *to appear in Math. Comp.*, 2016+.
- [47] D. Peterseim and M. Schedensack. Relaxing the CFL condition for the wave equation on adaptive meshes. ArXiv e-prints 1601.04812 2016.
- [48] F. Santosa and W. Symes. A dispersive effective medium for wave propagation in periodic composites. *SIAM J. Appl. Math.*, 51(4):984–1005, 1991.
- [49] S. Spagnolo. Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 22(4):571–597, 1968.
- [50] L. Tartar. Cours Peccot. Collège de France, 1977.
- [51] T. Vdovina and S. E. Minkoff. An a priori error analysis of operator upscaling for the acoustic wave equation. *Int. J. Numer. Anal. Model.*, 5(4):543–569, 2008.
- [52] T. Vdovina, S. E. Minkoff, and O. Korostyshevskaya. Operator upscaling for the acoustic wave equation. *Multiscale Model. Simul.*, 4(4):1305–1338, 2005.