Comptes Rendus

Mécanique

Rutger A. Biezemans, Claude Le Bris, Frédéric Legoll and Alexei Lozinski

Non-intrusive implementation of a wide variety of Multiscale Finite Element Methods

The scientific legacy of Roland Glowinski / *L'héritage scientifique de Roland Glowinski*

Non-intrusive implementation of a wide variety of Multiscale Finite Element Methods

Rutger A. Biezemans [*, a, b], **Claude Le Bris** [c, d], **Frédéric Legoll** [c, d] and **Alexei Lozinski** [®, d, e]

[a] École Nationale des Ponts et Chaussées, 6 et 8 avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2, France

[b] MATHERIALS project-team, Inria Paris, 2 rue Simone Iff, CS 42112, 75589 Paris Cedex 12, France

[c] Université de Franche-Comté, CNRS, LmB, F-25000 Besançon, France

E-mails: rutger.biezemans@enpc.fr (R. A. Biezemans), claude.le-bris@enpc.fr (C. Le Bris), frederic.legoll@enpc.fr (F. Legoll), alexei.lozinski@univ-fcomte.fr (A. Lozinski)

This work is dedicated to the memory of R. Glowinski, a wonderful individual, a leading scientist in computational methods, and a pioneer in the transfer between research and industry.

Abstract. Multiscale Finite Element Methods (MsFEMs) are now well-established finite element type approaches dedicated to multiscale problems. They first compute local, oscillatory, problem-dependent basis functions that generate a suitable discretization space, and next perform a Galerkin approximation of the problem on that space. We investigate here how these approaches can be implemented in a non-intrusive way, in order to facilitate their dissemination within industrial codes or non-academic environments. We develop an abstract framework that covers a wide variety of MsFEMs for linear second-order partial differential equations. Non-intrusive MsFEM approaches are developed within the full generality of this framework, which may moreover be beneficial to steering software development and improving the theoretical understanding and analysis of MsFEMs.

Keywords. Partial differential equations, Finite element methods, Multiscale problems, Non-intrusive implementation.

* Corresponding author.

1. Introduction

In this article, we consider highly oscillatory partial differential equations (PDEs) of the form $\mathscr{L}^\varepsilon u^\varepsilon = f$, posed in a bounded domain $\Omega \subset \mathbb{R}^d$, where $\mathscr{L}^\varepsilon$ is a second-order linear differential operator with (possibly rough) coefficients that oscillate on a microscopic length scale of size ε much smaller than the diameter of Ω. (See Sections 2 and 4 for a complete description of the problems that we study.) We seek a numerical approximation of u^ε by applying a Galerkin approach.

It is well-known that standard (say $\mathbb{P}_1$) finite element methods (FEMs) yield a poor approximation as a consequence of the highly oscillatory nature of the problem, unless a prohibitively expensive fine mesh is employed. An explicit example of this phenomenon is given, e.g., in [1, Example 1.1]. Dedicated multiscale approaches have thus been introduced, which provide a reasonably accurate approximation of u^ε at a limited computational cost. Among the many multiscale approaches that have been proposed in the literature, we mention the *Heterogeneous Multiscale Method* (abbreviated HMM) [2], the *Localized Orthogonal Decomposition* (LOD) [3] method, and the *Multiscale Finite Element Method* (MsFEM) [4], on which we focus here. We refer the reader to [1, 5–7] for more comprehensive expositions of these multiscale methods.

The MsFEM is a finite element type approach that relies on the adaptation of the finite element space to the highly oscillatory differential operator, an idea that was first introduced in [8]. The MsFEM was introduced in [4]. It consists of two steps:

(1) An "offline" stage, where highly oscillatory, problem-dependent basis functions are computed numerically as solutions to local problems (that mimic the reference problem on a subdomain). The local problems serve as a preprocessing step of the microstructure.
(2) An "online" stage, where a Galerkin approximation of the reference problem, performed in the finite-dimensional space generated by the basis functions that are computed in the offline stage, is solved. This constitutes the global coupling between the local computations performed in the offline stage.

The MsFEM approach is particularly interesting for multi-query contexts, where the PDE of interest is to be solved repeatedly for multiple right-hand sides f (think e.g. of optimization problems or time-dependent problems where the time discretization results in a PDE in space to advance from one time step to the next). In this case, the basis functions, which depend on $\mathscr{L}^\varepsilon$, remain unchanged, so the offline stage is performed only once. The online stage solves a problem on a space of much lower dimension than a high fidelity space that fully resolves the microscale, resulting in a significant computational gain.

The specific choice of problem-dependent basis functions has led to various MsFEM variants in the literature (see, for instance, [4, 6, 9, 10]). Although their implementation is rightfully considered relatively easy [6, 11], these methods dictate an intrusive workflow, i.e., substantial modifications have to be made to the code of some existing finite element software to implement the MsFEM. For instance, all automatic computations involving integrals of standard (e.g. polynomial) basis functions have to be replaced by the corresponding computations for specialized basis functions for each specific problem. We explore this further in Section 2.3.

The intrusive character of the MsFEM hinders the use of the method in many industrial contexts, because the time and tools required to adapt a legacy code that is currently in use may not be available. In this work, we propose a "non-intrusive" MsFEM strategy that uses an existing legacy code for single-scale problems (based on standard finite elements), without any modifications, to obtain an accurate resolution of oscillatory PDEs.

The intrusiveness of the MsFEM is due to the following fact: the microstructure is preprocessed through the computation of specialized basis functions, thereby coupling the microstruc-

ture explicitly to the global numerical model. We circumvent this coupling with a novel formulation of the MsFEM basis functions. The resulting MsFEM strategy can be summarized as follows:

(1) In the offline stage, so-called numerical correctors are computed as the solution to problems mimicking the reference problem on a subdomain. The numerical correctors are then used to average the microstructure in the form of effective, piecewise constant coefficients, leading to an effective PDE.

(2) In the online stage, the effective PDE is solved by a standard FEM *with the legacy code.*

(3) A post-processing stage is introduced to restore microscopic features in the macroscopic FEM result (obtained in the online stage) with the help of the numerical correctors computed in the offline stage.

Our non-intrusive MsFEM approach was introduced on a prototypical example in [12]. In this article, we extend the findings of [12] to a large class of MsFEM variants that can be "related" to a $\mathbb{P}_1$ FEM, in a sense that will be made precise in Remark 7. This requires the formulation of a general MsFEM framework, covering a generic definition of the MsFEM for the approximation of an abstract variational formulation of second-order linear PDEs. To the best of our knowledge, the question of how to make MsFEM approaches less intrusive has not been studied in the literature, except for the preliminary study in [12], and this work is a first step in that direction. We comment on the intrusiveness of other multiscale methods in Section 5.4.

An overview of the contents of this article is as follows. In Section 2, we recall the basic principles of the FEM and the MsFEM and we explain the intrusive character of the MsFEM. We also review the non-intrusive MsFEM approach that was proposed in [12] for the simplest MsFEM variant on the example of a diffusion problem. We also highlight a link between the non-intrusive MsFEM approach and classical homogenization here. Then we summarize in Section 3 which properties of the MsFEM are essential for the non-intrusive workflow, and motivate the development of a general framework covering a wide variety of MsFEMs that we present in Section 4. We extend the non-intrusive MsFEM approach of [12] to our general MsFEM framework in Section 5. The non-intrusive MsFEM approach of [12] was found to be equivalent to a Petrov–Galerkin MsFEM (with $\mathbb{P}_1$ test functions). This is no longer true for all MsFEMs covered by our general framework, and we obtain two non-intrusive MsFEMs: the Petrov–Galerkin MsFEM, which is completely equivalent to its non-intrusive implementation, and an approximate version of the Galerkin MsFEM that can be implemented in a non-intrusive way. The three essential formulas for the formulation of the non-intrusive MsFEM are highlighted in special boxes, both for the diffusion problem in Section 2 and for the general framework in Sections 4 and 5. We then study the general MsFEM framework applied to diffusion problems in Section 6, where we obtain a number of convergence results for the difference between the intrusive and non-intrusive MsFEM approaches. We conclude the article in Section 7 by a numerical comparison of the intrusive and non-intrusive MsFEM approaches for diffusion problems, in order to assess the efficiency of our approaches for cases that are not covered by the convergence results of Section 6. Our results show that the Petrov–Galerkin MsFEM as well as the non-intrusive approximation of the Galerkin MsFEM are close to the original Galerkin MsFEM. Any possible additional error introduced by making the MsFEM non-intrusive is thus negligible.

Notation

In this article, we shall adopt standard notation for Sobolev spaces. In particular, the dual space of $H_0^1(\Omega)$ is denoted $H^{-1}(\Omega)$. Further, for a given simplicial mesh $\mathcal{T}_H$ of Ω, we use the notation $H^1(\mathcal{T}_H)$ to denote the broken Sobolev space

$$H^1(\mathcal{T}_H) = \left\{ u \in L^2(\Omega) \mid u|_K \in H^1(K) \text{ for all mesh elements } K \in \mathcal{T}_H \right\}.$$

The standard norm for the space $H^1(\Omega)$ is

$$\|u\|_{H^1(\Omega)} = \sqrt{\|u\|_{L^2(\Omega)}^2 + \|\nabla u\|_{L^2(\Omega)}^2}$$

and the corresponding broken norm is

$$\|u\|_{H^1(\mathcal{T}_H)} = \sqrt{\sum_{K \in \mathcal{T}_H} \|u\|_{H^1(K)}^2}.$$

The space of functions whose restriction to each element of $\mathcal{T}_H$ is a polynomial of degree k is denoted $\mathbb{P}_k(\mathcal{T}_H)$.

2. The (intrusive) multiscale finite element method for diffusion problems

2.1. *Discrete variational formulation*

Let $d \geq 1$ denote the space dimension of interest and let $\Omega \subset \mathbb{R}^d$ be a bounded polytope (e.g. a polygon in dimension $d = 2$, a polyhedron in dimension $d = 3$). Convexity of Ω can be assumed for elliptic regularity results to hold, for which we refer to [13]. This technical assumption is not necessary for the algorithmic aspects of the MsFEM that are the main focus of this article.

By way of example, we consider first the diffusion equation with homogeneous Dirichlet boundary conditions. In a second step, from Section 4 onwards, we will also consider more general problems, and we will mention other types of boundary conditions in Section 5.3. More precisely, we focus here on the boundary value problem

$$\begin{cases} -\operatorname{div}(A^\varepsilon \nabla u^\varepsilon) = f \text{ in } \Omega, \\ u^\varepsilon = 0 \text{ on } \partial\Omega, \end{cases} \tag{1}$$

where the diffusion tensor $A^\varepsilon \in L^\infty(\Omega, \mathbb{R}^{d \times d})$ satisfies the uniform bounds

$$\forall \, \xi \in \mathbb{R}^d, \quad m|\xi|^2 \leq \xi \cdot A^\varepsilon(x)\xi \quad \text{a.e. in } \Omega,$$

$$\text{and} \qquad \forall \, \xi, \eta \in \mathbb{R}^d, \quad |\eta \cdot A^\varepsilon(x)\xi| \leq M|\xi||\eta| \quad \text{a.e. in } \Omega, \tag{2}$$

for some $M \geq m > 0$ independent of ε. The right-hand side f does not vary on the microscopic scale ε. We denote the diffusion tensor with a superscript ε to keep in mind that A^ε might be highly oscillatory on a typical length scale of size ε much smaller than the diameter of Ω (assumed to be of order 1). No further structural assumptions on A^ε are made. In particular, A^ε need not be the rescaling of a fixed periodic matrix of the form $A^\varepsilon(x) = A(x/\varepsilon)$. We will specialize to this periodic setting in Section 6.3 only to obtain convergence results, but this assumption is of no relevance for the practical implementation of the MsFEM. Let us also mention that none of the considerations in this article require symmetry of the diffusion tensor. Our development of non-intrusive MsFEMs also generalizes to linear *systems* of PDEs. The analysis we provide is also expected to extend to e.g. the system of linear elasticity up to some technicalities that we do not consider here.

For simplicity of exposition, we assume that $f \in L^2(\Omega)$ (rather than $f \in H^{-1}(\Omega)$, for which the problem (1) is in fact well-posed). We do so to avoid unnecessary technicalities. Our proposed non-intrusive MsFEM carries over to the more general case. For some convergence results, the condition $f \in L^2(\Omega)$ cannot be relaxed. In this case, this is also explicitly stated.

Problem (1) admits a unique solution in the space $H_0^1(\Omega)$. This solution is also characterized by the variational formulation

$$\begin{cases} \text{Find } u^\varepsilon \in H_0^1(\Omega) \qquad \text{such that} \\ a^{\varepsilon,\text{diff}}(u^\varepsilon, v) = F(v) \quad \text{for all } v \in H_0^1(\Omega), \end{cases} \tag{3}$$

where the bilinear form $a^{\varepsilon,\text{diff}}$ and the linear form F are defined, for any $u, v \in H_0^1(\Omega)$, by

$$a^{\varepsilon,\text{diff}}(u, v) = \int_\Omega \nabla v \cdot A^\varepsilon \nabla u, \quad F(v) = \int_\Omega f v. \tag{4}$$

The coercivity hypothesis in (2) ensures that the bilinear form $a^{\varepsilon,\text{diff}}$ is coercive on the space $H_0^1(\Omega)$. Then the Lax–Milgram Theorem [14, Theorem 5.8] shows that (3) is indeed well-posed.

The numerical approximation of (3) with a finite element method starts by the introduction of a mesh $\mathcal{T}_H$ for Ω. The subscript H denotes the typical size of the mesh elements. We assume $\mathcal{T}_H$ to be a simplicial, conformal mesh. For some convergence results, we shall assume quasi-uniformity. These assumptions are standard in finite element analysis. We refer, e.g., to [15–17] for a general exposition and various examples. Again, these regularity properties of the mesh do not have any impact on the implementation of the MsFEM on a given mesh. The regularity plays a role only to obtain convergence results.

A finite element method for (1) is obtained by restricting the equivalent formulation (3) to a finite-dimensional subspace of $H_0^1(\Omega)$, typically consisting of functions that are piecewise polynomial on the mesh $\mathcal{T}_H$. We suppose that we are in the regime where H is larger than or comparable to the microscale ε. In this case, it is well known that a Galerkin approximation of (3) on, say, the standard (conforming) Lagrange $\mathbb{P}_1$ space on $\mathcal{T}_H$ provides only a poor, not to say an incorrect approximation of u^ε. See [1, Example 1.1], for instance, for an explicit example where the $\mathbb{P}_1$ approximation on a coarse mesh fails. At the same time, the use of a finite element method on a fine mesh of size $H \ll \varepsilon$ might be unfeasible from a computational point of view because of its prohibitive computational cost. To remedy this issue, we shall next introduce the *multiscale finite element method* (MsFEM) [4, 6].

2.2. *A simple multiscale finite element method*

The MsFEM is a Galerkin approximation of (3) for which the approximation space is adapted in order to achieve satisfactory accuracy even on a coarse mesh. The correct choice of approximation space yields a numerical approximation that is much closer to u^ε than a standard $\mathbb{P}_1$-approximation when ε is smaller than H, and especially when ε becomes asymptotically small. To begin with, we introduce here the simplest variant of the MsFEM, which originally appeared in [4], before moving on to other MsFEM variants in Section 4.

Let $x_1, \ldots, x_{N_0}$ be an enumeration of the interior vertices of the mesh $\mathcal{T}_H$, i.e., the vertices that do not lie on $\partial\Omega$. We denote by $\phi_i^{\mathbb{P}_1}$ the unique piecewise $\mathbb{P}_1$ function such that $\phi_i^{\mathbb{P}_1}(x_j) = \delta_{i,j}$ for all $1 \le j \le N_0$. (These are the basis functions for the standard $\mathbb{P}_1$ Lagrange finite element.) We define the *multiscale basis functions* ϕ_i^ε (for $1 \le i \le N_0$) by

$$\forall\ K \in \mathcal{T}_H, \quad \begin{cases} -\operatorname{div}\left(A^\varepsilon \nabla \phi_i^\varepsilon\right) = 0 & \text{in } K, \\ \phi_i^\varepsilon = \phi_i^{\mathbb{P}_1} & \text{on } \partial K. \end{cases} \tag{5}$$

All these problems, on each mesh element K, are again well-posed by coercivity of A^ε and the Lax–Milgram Theorem. The functions ϕ_i^ε so defined belong to the global space $H_0^1(\Omega)$ because the local boundary conditions on ∂K imply continuity across all mesh elements K. It is also immediately seen that ϕ_i^ε is supported by exactly the same mesh elements as $\phi_i^{\mathbb{P}_1}$.

Remark 1. On each mesh element K, problem (5) defines at most $d+1$ non-trivial basis functions. Let $i_1, \ldots, i_{d+1}$ be the indices of the vertices of K. It is easily inferred from (5) that

$$\phi_{i_{d+1}}^\varepsilon\Big|_K = 1 - \sum_{j=1}^d \phi_{i_j}^\varepsilon\Big|_K.$$

Thus, one only has to compute d basis functions by the resolution of the PDE (5) on K.

The multiscale approximation space is defined as $V_{H,0}^\varepsilon = \text{span}\{\phi_i^\varepsilon \mid 1 \le i \le N_0\}$. This is a finite-dimensional space of the same dimension as the one used for a $\mathbb{P}_1$ Lagrange finite element approximation on the mesh $\mathcal{T}_H$. The MsFEM consists in computing the approximation $u_H^\varepsilon \in V_{H,0}^\varepsilon$ defined by the problem

$$\forall \ v_H^\varepsilon \in V_{H,0}^\varepsilon, \qquad a^{\varepsilon,\text{diff}}\left(u_H^\varepsilon, v_H^\varepsilon\right) = F\left(v_H^\varepsilon\right). \tag{6}$$

Since $V_{H,0}^\varepsilon$ is a subspace of $H_0^1(\Omega)$, the bilinear form $a^{\varepsilon,\text{diff}}$ is coercive on $V_{H,0}^\varepsilon$ and the discrete problem (6) is again well-posed by virtue of the Lax–Milgram Theorem.

The computation of the multiscale basis functions ϕ_i^ε is called the *offline stage* of the MsFEM and only has to be carried out once if (1) has to be solved multiple times for different right-hand sides. Also note that all problems (5) are independent of each other, and can thus be solved in parallel. Once all basis functions are known, one can compute the stiffness matrix of the MsFEM (see Section 2.3 for more details), which is also part of the offline stage. In practice, the ϕ_i^ε are approximated numerically on a fine mesh of $K \in \mathcal{T}_H$ of mesh size $h \le \varepsilon$ that resolves the oscillations of A^ε. We omit these details here because they have no importance for the non-intrusive strategy that we shall propose later in this article.

The resolution of the global problem (6), each time the right-hand side F changes, is called the *online stage*. The computational cost for this problem is the same as for a standard $\mathbb{P}_1$ approximation on the same mesh. A further discussion of the practical implementation of the MsFEM is provided in Section 2.3. This discussion partially reproduces some elements of [12]. We include it here to clarify and motivate the developments in the sequel.

2.3. *Intrusive workflow*

The practical resolution of the global problem (6) consists in the construction and resolution of the following linear system:

$$\mathbb{A}^\varepsilon U^\varepsilon = \mathbb{F}^\varepsilon, \tag{7}$$

with the stiffness matrix $\mathbb{A}^\varepsilon$ and the right-hand side $\mathbb{F}^\varepsilon$ of the linear system given by

$$\forall \ 1 \le i, j \le N_0, \quad \mathbb{A}_{j,i}^\varepsilon = a^{\varepsilon,\text{diff}}\left(\phi_i^\varepsilon, \phi_j^\varepsilon\right), \quad \mathbb{F}_j^\varepsilon = F\left(\phi_j^\varepsilon\right), \tag{8}$$

where we recall that N_0 denotes the number of interior vertices of $\mathcal{T}_H$. The MsFEM approximation u_H^ε is given by $u_H^\varepsilon = \sum_{i=1}^{N_0} U_i^\varepsilon \phi_i^\varepsilon$. The MsFEM can then be written (as it is traditionally presented) as in Algorithm 1. We use the notation $a_K^{\varepsilon,\text{diff}}(u,v) = \int_K \nabla v \cdot A^\varepsilon \nabla u$ for all $u, v \in H^1(K)$ and we write $F_K(v) = \int_K f v$ for any $v \in L^2(K)$.

Lines (1)-(12) of Algorithm 1 (resp. (13)-(17)) constitute the offline (resp. online) stage of the MsFEM. Note that the computation of the stiffness matrix $\mathbb{A}^\varepsilon$ in line (11) only depends on the multiscale basis functions (and not on the right-hand side f) and can therefore be carried out once and for all in the offline stage. Also note that, for an efficient computation of the ϕ_i^ε in line (6), one should apply Remark 1. Only the online stage is to be repeated when problem (1) is to be solved multiple times for various right-hand sides f.

Implementing Algorithm 1 in an industrial code is challenging. Indeed, the practical implementation of any finite element method relies on (i) the construction of a mesh, (ii) the construction of the linear system associated to the discrete variational formulation and (iii) the resolution of the linear system. An efficient implementation of the second step heavily relies on the choice of the discretization space.

Regarding the construction of the linear system (performed in line (11) of Algorithm 1), it is by no means obvious to adapt existing finite element codes based on generic approximation spaces (for instance spaces of piecewise polynomial functions, such as the piecewise affine functions that we will introduce in Definition 10 below) to a different, problem-dependent choice of space

---

Algorithm 1 MsFEM approach for problem (1) (see comments in the text)

1: Construct a mesh $\mathcal{T}_H$ of Ω, denote N_0 the number of internal vertices and $\mathcal{N}(n, K)$ the global index of the vertex of $K \in \mathcal{T}_H$ that has local index $1 \le n \le d+1$ in K

2: Set $\mathbb{A}^\varepsilon := 0$ and $\mathbb{F}^\varepsilon := 0$

3: **for all** $K \in \mathcal{T}_H$ **do**

4:     **for** $1 \le n \le d+1$ **do**

5:         Set $i := \mathcal{N}(n, K)$

6:         Solve for $\phi_i^\varepsilon\big|_K$ in (5)

7:     **end for**

8:     **for** $1 \le l \le d+1$ **do**

9:         Set $j := \mathcal{N}(l, K)$

10:         **for** $1 \le n \le d+1$ **do**

11:             Set $i := \mathcal{N}(n, K)$ and $\mathbb{A}_{j,i}^\varepsilon \mathrel{+}= a_K^{\varepsilon, \mathrm{diff}}\left(\phi_i^\varepsilon, \phi_j^\varepsilon\right)$

12:         **end for**

13:         Set $\mathbb{F}_j^\varepsilon \mathrel{+}= F_K\left(\phi_j^\varepsilon\right)$

14:     **end for**

15: **end for**

16: Solve the linear system $\mathbb{A}^\varepsilon U^\varepsilon = \mathbb{F}^\varepsilon$

17: Obtain the MsFEM approximation $u_H^\varepsilon = \sum\limits_{i=1}^{N_0} U_i^\varepsilon \phi_i^\varepsilon$

---

such as V_H^ε. No analytic expressions for the basis functions ϕ_i^ε are available (and thus a fine mesh should be used to approximate them), the computation of $a_K^\varepsilon(\phi_i^\varepsilon, \phi_j^\varepsilon)$ should be performed by quadrature rules on the fine mesh because the integrands are highly oscillatory, one should have at hand the correspondence between element and vertex indices of the coarse mesh ($\mathcal{N}(n, K)$ in Algorithm 1), the assembly of the global stiffness matrix $\{\mathbb{A}_{j,i}^\varepsilon\}_{1 \le i, j \le N_0}$ should be executed by a dedicated new piece of software, etc. To alleviate these obstacles, we introduce below a way of implementing the MsFEM that capitalizes on an existing code for solving (1) by a $\mathbb{P}_1$ approximation on $\mathcal{T}_H$ in the case of slowly varying diffusion coefficients. The three central identities for our approach that we aim to generalize to other MsFEMs in this article are framed in distinctive boxes.

2.4. *Effective problem on the macroscopic scale*

Let us consider the construction of the stiffness matrix of the MsFEM in more detail. The stiffness matrix defined in (8) requires the computation of the quantities

$$\mathbb{A}_{j,i}^\varepsilon = a^{\varepsilon, \mathrm{diff}}\left(\phi_i^\varepsilon, \phi_j^\varepsilon\right) = \sum_{K \in \mathcal{T}_H} \int_K \nabla \phi_j^\varepsilon \cdot A^\varepsilon \nabla \phi_i^\varepsilon, \tag{9}$$

for all $1 \le i, j \le N_0$.

Following [12], we rewrite the multiscale basis functions as

$$\boxed{\quad \forall\, K \in \mathcal{T}_H, \qquad \phi_i^\varepsilon = \phi_i^{\mathbb{P}_1} + \sum_{\alpha=1}^d \left(\partial_\alpha \phi_i^{\mathbb{P}_1}\right)\Big|_K \chi_K^{\varepsilon, \alpha} \quad \text{in } K, \quad} \tag{10}$$

for all $1 \le i \le N_0$, where, for each mesh element K, we define the numerical corrector $\chi_K^{\varepsilon,\alpha} \in H_0^1(\Omega)$ $(1 \le \alpha \le d)$ as the function supported by K that is the unique solution to the local problem

$$
\begin{cases}
-\operatorname{div}\left(A^\varepsilon \nabla \chi_K^{\varepsilon,\alpha}\right) = \operatorname{div}\left(A^\varepsilon e_\alpha\right) \text{ in } K, \\
\qquad\qquad \chi_K^{\varepsilon,\alpha} = 0 \text{ on } \partial K.
\end{cases}
\tag{11}
$$

Here, e_α denotes the α-th canonical unit vector of $\mathbb{R}^d$. The expansion (10) is obtained upon rewriting (5) as a PDE for $\phi_i^\varepsilon - \phi_i^{\mathbb{P}_1}$, and then using linearity of the PDE and the fact that $\nabla \phi_i^{\mathbb{P}_1}$ is constant in K to show that

$$
\sum_{\alpha=1}^d \left(\partial_\alpha \phi_i^{\mathbb{P}_1}\right)\Big|_K \chi_K^{\varepsilon,\alpha}
$$

is indeed the unique solution to this PDE.

Inserting (10) for the trial and test functions in (9) and again exploiting the fact that all $\phi_i^{\mathbb{P}_1}$ have piecewise constant gradients, we obtain

$$
\begin{aligned}
\mathbb{A}_{j,i}^\varepsilon &= \sum_{K \in \mathcal{T}_H} \sum_{\alpha,\beta=1}^d \left(\partial_\beta \phi_j^{\mathbb{P}_1}\right)\Big|_K \left(\int_K \left(e_\beta + \nabla \chi_K^{\varepsilon,\beta}\right) \cdot A^\varepsilon \left(e_\alpha + \nabla \chi_K^{\varepsilon,\alpha}\right) \right) \left(\partial_\alpha \phi_i^{\mathbb{P}_1}\right)\Big|_K \\
&= \sum_{K \in \mathcal{T}_H} \sum_{\alpha,\beta=1}^d \left(\partial_\beta \phi_j^{\mathbb{P}_1}\right)\Big|_K a_K^{\varepsilon,\mathsf{diff}}\left(x^\alpha + \chi_K^{\varepsilon,\alpha}, x^\beta + \chi_K^{\varepsilon,\beta}\right) \left(\partial_\alpha \phi_i^{\mathbb{P}_1}\right)\Big|_K.
\end{aligned}
$$

Next we define the piecewise constant effective diffusion tensor $\overline{A} \in \mathbb{P}_0(\mathcal{T}_H, \mathbb{R}^{d \times d})$ by

$$
\overline{A}_{\beta,\alpha}\Big|_K = \frac{1}{|K|} a_K^{\varepsilon,\mathsf{diff}}\left(x^\alpha + \chi_K^{\varepsilon,\alpha}, x^\beta + \chi_K^{\varepsilon,\beta}\right) \quad \text{for each } K \in \mathcal{T}_H \text{ and each } 1 \le \alpha, \beta \le d,
\tag{12}
$$

where $|K|$ denotes the measure of the mesh element K. Then (9) can be written as

$$
\mathbb{A}_{j,i}^\varepsilon = \int_\Omega \nabla \phi_j^{\mathbb{P}_1} \cdot \overline{A} \nabla \phi_i^{\mathbb{P}_1}.
\tag{13}
$$

Motivated by (13), we introduce the coarse-scale problem

$$
\begin{cases}
-\operatorname{div}\left(\overline{A} \nabla u\right) = f \text{ in } \Omega, \\
\qquad\qquad u = 0 \text{ on } \partial\Omega,
\end{cases}
\tag{14}
$$

and its Galerkin discretization with $\mathbb{P}_1$ Lagrange elements: with $V_{H,0} = \operatorname{span}\{\phi_i^{\mathbb{P}_1} \mid 1 \le i \le N_0\}$ (note that the definition of $V_{H,0}$ will be generalized in Definition 11), find $u_{H,0} \in V_{H,0}$ such that

$$
\forall \; v_H \in V_{H,0}, \quad \overline{a}^{\mathsf{diff}}(u_H, v_H) = F(v_H),
\tag{15}
$$

where the linear form F is defined in (3) and the bilinear form $\overline{a}^{\mathsf{diff}}$ is defined as

$$
\forall \; u, v \in H_0^1(\Omega), \quad \overline{a}^{\mathsf{diff}}(u, v) = \int_\Omega \nabla v \cdot \overline{A} \nabla u.
\tag{16}
$$

Problem (15) equivalently writes

$$
\mathbb{A}^{\mathbb{P}_1} U^{\mathbb{P}_1} = \mathbb{F}^{\mathbb{P}_1},
\tag{17}
$$

with

$$
\forall \; 1 \le i, j \le N_0, \quad \mathbb{A}_{j,i}^{\mathbb{P}_1} = \overline{a}^{\mathsf{diff}}\left(\phi_i^{\mathbb{P}_1}, \phi_j^{\mathbb{P}_1}\right), \quad \mathbb{F}_j^{\mathbb{P}_1} = F\left(\phi_j^{\mathbb{P}_1}\right).
\tag{18}
$$

Comparing the expressions (13) and (18), we deduce that $\mathbb{A}^\varepsilon = \mathbb{A}^{\mathbb{P}_1}$. In other words:

Lemma 2. *The stiffness matrix of the MsFEM problem* (6) *is identical to the stiffness matrix of the* $\mathbb{P}_1$ *problem* (15).

This lemma immediately implies that the $\mathbb{P}_1$ problem (15) is well-posed, since the MsFEM problem (6) itself is well-posed.

Let us point out that problems (14) and (15) are defined entirely in terms of quantities that vary only on the macroscopic scale H. The finite element problem (15) can thus be solved using a legacy code that is designed for standard FEMs. Lemma 2 then suggests including the $\mathbb{P}_1$ approximation (15) of the effective, coarse-scale problem (14) as an integral part of the MsFEM approach. We do so in Algorithm 2 below.

The right-hand side vector $\mathbb{F}^\varepsilon$ in (8) is, in general, different from $\mathbb{F}^{\mathbb{P}_1}$ in (18). Indeed, we integrate the product of f with highly oscillatory basis functions in the former problem and with $\mathbb{P}_1$ basis functions in the latter. The solutions U^ε and $U^{\mathbb{P}_1}$ to (7) and (17), respectively, are thus different a priori.

2.5. *Non-intrusive workflow*

We propose the following non-intrusive MsFEM variant:

$$\text{Set } u_H^\varepsilon = u_H + \sum_{K \in \mathcal{T}_H} (\partial_\alpha u_H)|_K \, \chi_K^{\varepsilon,\alpha} \in V_{H,0}^\varepsilon \text{ where } u_H \in V_{H,0} \text{ is the unique solution to (15).} \quad (19)$$

The MsFEM approximation u_H^ε is well-defined, since we have seen above that problem (15) is well-posed.

Note that the symbol u_H^ε shall be used here and in the sequel for the solution to various MsFEMs variants to alleviate the notation. The exact MsFEM will be specified by the context. We will use distinct notation for different MsFEM variants when required for clarity.

The most efficient way to compute u_H^ε from u_H is not as stated here, however. The *evaluation* of $u_H(x)$ may require the determination of the degrees of freedom associated to the simplex K to which x belongs. This demands the use of the internal mechanisms of the legacy code that is used to compute u_H. The use of the legacy code can be avoided by expanding u_H as follows. For any affine function φ on K, we have

$$\varphi(x) = \varphi\left(x_{c,K}\right) + \sum_{\alpha=1}^{d} \partial_\alpha \varphi \left(x^\alpha - x_{c,K}^\alpha\right) \quad \text{on } K, \quad (20)$$

where x^α denotes the function that to a point $x \in \Omega$ associates its αth coordinate, and $x_{c,K} = (x_{c,K}^1, \ldots, x_{c,K}^d)$ is the centroid of K. If one uses the legacy code to store the values of $u_H(x_{c,K})$ and $\partial_\alpha u_H$ element by element at the end of the online stage, then u_H^ε defined in (19) can be computed element by element according to

$$\forall \, K \in \mathcal{T}_H, \quad u_H^\varepsilon(x) = u_H(x_{c,K}) + \sum_{\alpha=1}^{d} (\partial_\alpha u_H)|_K \left(x^\alpha - x_{c,K}^\alpha + \chi_K^{\varepsilon,\alpha}(x)\right) \quad \text{on } K, \quad (21)$$

without using the legacy code.

The above observations culminate in the computational approach presented in Algorithm 2. We can distinguish

   (1) the offline stage consisting of lines (1)-(7),
   (2) the online stage being executed entirely in line (8),
   (3) a post-processing step in line (9).

The superiority of Algorithm 2 over the classical MsFEM Algorithm 1 is that the global problem of the online stage can completely be constructed and solved by the use of a pre-existing $\mathbb{P}_1$ PDE solver. The only requirements for the legacy code are the functionality to provide piecewise constant diffusion coefficients to the solver and the existence of a procedure to store the value of the $\mathbb{P}_1$ solution and its gradient at the centroids of the mesh. An additional advantage in the online stage is that the construction of the right-hand side $\mathbb{F}^{\mathbb{P}_1}$ (see (18)) for the global problem

---

Algorithm 2 Non-intrusive MsFEM approach for problem (1)

1: Let $\mathcal{T}_H$ be the mesh used by the legacy code

2: **for all** $K \in \mathcal{T}_H$ **do**
3:     **for** $1 \le \alpha \le d$ **do**
4:         Solve for $\chi_K^{\varepsilon,\alpha}$ defined by (11)
5:     **end for**
6:     Compute $\overline{A}|_K$ defined by (12)
7: **end for**

8: Use the legacy code to solve for u_H defined by (15) and to save $\{u_H(x_{c,K})\}_{K \in \mathcal{T}_H}$ and $\{(\partial_\alpha u_H)|_K\}_{K \in \mathcal{T}_H, 1 \le \alpha \le d}$

9: Obtain the MsFEM approximation u_H^ε by (21)

---

only requires a numerical quadrature on the coarse mesh and is therefore cheaper than the construction of $\mathbb{F}^\varepsilon$ (see (8)), involving the multiscale basis functions and requiring numerical quadratures at the microscale.

The part of the offline stage that manipulates fine meshes (lines (2)-(7)) and the post-processing step can be developed independently of the legacy code used in line (8). The requirement for these fine-scale solvers is that they have access to the coarse mesh $\mathcal{T}_H$ used by the global solver. Note also that the local problem (11) is only indexed by the coarse mesh element K, in contrast to the local problem (5) that is indexed both by the coarse mesh element K and the vertex index i. For the latter problems, one has to know, for each element K, the global index that corresponds to the vertices of K, a piece of information that may be difficult to access in a legacy code. For the problems (11), this correspondence is not needed to compute $\overline{A}$, nor for the computation of the fine-scale solution u_H^ε in (21), both of which are entirely defined element-wise.

Remark 3 (Quantities of interest).   In the post-processing step of Algorithm 2, it is easy to compute pointwise values of the approximation u_H^ε by (21) and to use these for further computational steps, such as the evaluation of the energy or other quantities of interest. This task can be carried out element wise, hence Eq. (21) can easily be used. (See also Remark 51.)

Remark 4 (Visualization).   We focus our attention here on the visualization of the MsFEM approximation u_H^ε, which can be an important tool in engineering practices. Visualization requires the combination of information on neighbouring mesh elements, and this can in general not be carried out by the legacy code since it does not have access to the fine meshes used to compute the numerical correctors. Even if this were the case, the fine meshes may not yield a globally conformal mesh when combined. The question of a global visualization then becomes a complex one that requires innovations beyond the contributions of this article. Instead, we propose the following two-step visualization approach:

- One can visualize the coarse part u_H of the MsFEM approximation for a global view of the solution with the tools provided by the legacy code that is used to compute u_H;
- The fine scale details of u_H^ε in regions of interest can be studied through zooms inside mesh elements, using the code that is used for computations at the microscale.

2.6. *Interpretation of the non-intrusive MsFEM*

We emphasized above that the right-hand sides of the linear system for the MsFEM in (8) and the linear system solved for the non-intrusive MsFEM in (18) are different in general. This motivates

the comparison of the non-intrusive MsFEM approach (19) to the following Petrov–Galerkin MsFEM:

$$\text{Find } u_H^\varepsilon \in V_{H,0}^\varepsilon \text{ such that } a^{\varepsilon,\text{diff}}\left(u_H^\varepsilon, \phi_j^{\mathbb{P}_1}\right) = F\left(\phi_j^{\mathbb{P}_1}\right) \quad \text{for all } 1 \le j \le N_0, \tag{22}$$

based on the trial space $V_{H,0}^\varepsilon$ and the test space $V_{H,0}$ for both the bilinear and the linear form. The following result was shown in [12].

Lemma 5. *The non-intrusive MsFEM variant* (19) *coincides with the Petrov-Galerkin Ms-FEM* (22).

The non-intrusive MsFEM approach is generalized in Section 5 after the development of a general framework to define a wide variety of MsFEMs in Section 4. Lemma 5 does not generalize to the full framework. We will see the conditions under which the non-intrusive approach leads to a Petrov–Galerkin MsFEM in Lemma 38.

2.7. *Relation to homogenization theory*

We highlight in this section the fact that many ingredients of our non-intrusive MsFEM approach are reminiscent of standard quantities of homogenization theory, or the theory of H-convergence, which studies the limit of a sequence of solutions u^ε to a PDE as ε tends to 0. This relation to H-convergence provides an interesting interpretation of the effective tensor $\overline{A}$ introduced in (12).

Let us suppose in this section (and in this section only, except for Section 6.3) that $A^\varepsilon(x) = A^{\text{per}}(x/\varepsilon)$ for some bounded, $\mathbb{Z}^d$-periodic matrix A^{per} satisfying the coercivity property in (2). In this case, the sequence of matrices A^ε has a homogenized limit that is explicitly known. (An explicit characterization of the limit is not available for H-convergence in general.) We summarize the main results below. See, for instance, [18, 19] or [20, Chapter 1] for details on periodic homogenization.

Due to H-convergence of A^ε, the functions u^ε, solution to (1), converge to a limit function $u^\star$ (weakly in $H^1(\Omega)$, strongly in $L^2(\Omega)$) as $\varepsilon \to 0$. The homogenized limit $u^\star$ is the solution to the homogenized equation (25) below.

Let Q denote the unit cube of $\mathbb{R}^d$. We introduce the corrector functions $w_1, \dots, w_d \in H^1_{per}(Q)$ solution to

$$\begin{cases} -\operatorname{div}(A^{\text{per}}\nabla w_\alpha) = \operatorname{div}(A^{\text{per}} e_\alpha) \text{ in } \mathbb{R}^d, \\ \qquad\qquad w_\alpha \text{ is } Q\text{-periodic}, \end{cases} \tag{23}$$

which uniquely defines w_α up to an irrelevant additive constant. The entries of the (constant) homogenized diffusion tensor $A^\star \in \mathbb{R}^{d \times d}$ are given by

$$A_{\beta,\alpha}^\star = \int_Q \left(e_\beta + \nabla w_\beta\right) \cdot A^{\text{per}}\left(e_\alpha + \nabla w_\alpha\right), \qquad 1 \le \alpha, \beta \le d. \tag{24}$$

The homogenized limit $u^\star$ of u^ε is the unique solution in $H_0^1(\Omega)$ to the boundary value problem

$$\begin{cases} -\operatorname{div}(A^\star \nabla u^\star) = f \text{ in } \Omega, \\ \qquad\qquad u^\star = 0 \text{ on } \partial\Omega. \end{cases} \tag{25}$$

The truncated reconstruction of $u^\star$ that is called the first-order two-scale expansion takes the form

$$u^{\varepsilon,1}(x) = u^\star(x) + \varepsilon \sum_{\alpha=1}^d \partial_\alpha u^\star(x)\, w_\alpha\left(\frac{x}{\varepsilon}\right). \tag{26}$$

Under suitable regularity assumptions, the difference $u^\varepsilon - u^{\varepsilon,1}$ converges to 0 strongly in $H^1(\Omega)$ as $\varepsilon \to 0$. This property will be used for the convergence results in Section 6.3.

In the periodic setting, the expansion (26) can be used to construct a numerical approximation of u^ε, without the need of any computations at the fine scale. This approximation is presumably valid only in the regime of very small parameters ε and deteriorates if ε grows. Moreover, in more general settings, the corrector functions are not local nor explicit, for their definition involves a PDE posed on the whole domain Ω and that depends on an effective tensor that is itself defined in terms of the corrector functions. Details can be found, e.g., in [19–21]. This prevents the H-convergence theory from being directly applicable for the numerical approximation of u^ε.

Numerical homogenization techniques, that draw their inspiration from the various elements above, offer an alternative for the approximation of u^ε that can be applied in much more general contexts. We can see the similarities between the corrector functions w_α in (23) and the numerical correctors $\chi_K^{\varepsilon,\alpha}$ in (11). Note that the $\chi_K^{\varepsilon,\alpha}$ solve problems similar to (23), but that they need to be solved at the microscale and on each mesh element K. Similarly, we note the resemblance between the reconstruction (26) and the definition of u_H^ε in (19), and between the homogenized coefficient $A^\star$ defined in (24) and the effective macroscopic coefficient $\overline{A}$ from (12). However, contrary to $A^\star$, the MsFEM quantity $\overline{A}$ has to be computed on an element-by-element basis, and it is not necessarily constant throughout Ω. Finally, the MsFEM analogue of the homogenized problem (25) is the resolution of the effective macroscale problem (14).

Example 6. A very particular setting, although academic in nature and only useful for pedagogical purposes, actually leads to an MsFEM approximation that is exactly equivalent to a discretization of the periodic homogenization setting. Consider (1) in 2D posed on the unit square. Let us consider a mesh consisting of squares that are perfectly aligned with the periodicity of A^ε. We solve the corrector problems (11) on all square mesh elements with periodic boundary conditions and subsequently compute the effective diffusion tensor $\overline{A}$ according to (12).

In this case, the problems for the numerical correctors all reduce to (23) and $\overline{A}$ is constant and equal to the homogenized coefficient $A^\star$ as defined by (24). A $\mathbb{Q}_1$ discretization of the effective problem (14) thus constitutes a non-intrusive MsFEM that is equivalent to the $\mathbb{Q}_1$ approximation of the homogenized equation (25).

3. Why develop a general framework?

In the sequel we develop a general framework for a wide variety of MsFEMs in an abstract setting. We motivate here why this general framework for MsFEMs is useful.

3.1. *Local boundary conditions*

First, let us explain why various MsFEMs have been proposed in the literature. One reason is that different equations than (1) (e.g. advection-diffusion equations) give rise to different choices of the local problem (5), depending on which terms of the global PDE are included (see, for instance, [10, 22].)

The other reason is that, even for the pure diffusion problem (1), the choice of the basis functions defined in (5) has an important drawback. The definition of the multiscale basis functions requires a choice of *arbitrary* boundary conditions on the mesh element boundary ∂K, since the exact boundary condition satisfied by u^ε is unknown. In (5), affine boundary conditions are imposed. In view of this choice, we shall refer to the MsFEM defined above as the "MsFEM-lin".

The MsFEM-lin cannot yield an accurate representation of u^ε near ∂K if A^ε is highly oscillatory and the mesh $\mathcal{T}_H$ is coarse. Variations on the definition of the functions ϕ_i^ε have been proposed to improve the MsFEM. Here we summarize the ideas of *oversampling* and of MsFEM *à la Crouzeix–Raviart*, which together inspire the formulation of a general MsFEM framework in Section 4.

The oversampling variant of the MsFEM was introduced along with the variant based on (5) at the time of its first appearance in [4]. For this method, an oversampling domain S_K is associated to each mesh element K (details are provided in Section 4.3.1). The problems (5) are solved on the larger domain S_K rather than K, so the inadequate boundary conditions are pushed away from the actual mesh elements. To construct the multiscale basis functions, the resulting functions on S_K are restricted to the actual mesh elements K and suitably combined around each vertex x_i. The new multiscale basis functions oscillate on ∂K if the oversampling patch is taken large enough. We note that, in general, this strategy leads to discontinuous basis functions. Hence, the finite element space obtained is no longer conforming.

The MsFEM with Crouzeix–Raviart type boundary conditions for the local problems (which we shall abbreviate as "MsFEM-CR") was introduced in [9]. It uses basis functions associated to the edges of the mesh (in contrast to the MsFEM-lin presented above, and its oversampling variant, where basis functions are associated to the vertices of the mesh). A typical basis function satisfies the following on ∂K: the flux through each face of K is constant, and the constants are determined by the condition that the average of the basis function be 1 over one particular face and 0 over all other faces. Again, this is a way to avoid imposing any conditions on the trace of the basis function directly. The multiscale functions can thus be oscillatory on the faces of the mesh. As is the case for oversampling methods, the resulting finite element space is nonconforming.

All of these variations, applied to any MsFEM for linear second-order PDEs, are covered by the general MsFEM framework that we develop in Section 4.

3.2. *The non-intrusive approach*

The intrusiveness of the specific MsFEM-lin variant introduced in Section 2.2 is exemplary for all MsFEMs described in Section 3.1. It turns out that the non-intrusive MsFEM approach introduced in [12] and recalled in Section 2.5 can also be generalized to all these MsFEM variants. We summarize the key ingredients that allow for the formulation of the non-intrusive MsFEM approach of Algorithm 2 (corresponding to the identities in boxes in Section 2.4).

The non-intrusive MsFEM follows from the expansion (10), namely the expression of the multiscale basis function as a $\mathbb{P}_1$ basis function and a linear combination of numerical correctors that are fully localized. We note that

- the full localization of the numerical correctors defined in (11) allows the preprocessing of the microstructure independently of the global approximation indices related to the finite element method;
- the expansion (10) follows from the fact that $\nabla \phi_i^{\mathbb{P}_1}$ is piecewise constant combined with linearity of the local problems (5);
- the stiffness matrix can be formulated in terms of a piecewise constant effective diffusion tensor in (13) thanks to full localization of the corrector functions, the piecewise constant gradient of $\phi_i^{\mathbb{P}_1}$ in the expansion (10) and bilinearity of the global problem (3).

These observations provide the main structure of the general framework. First, we choose an underlying, low-dimensional space of *piecewise affine* functions to which the MsFEM is associated (Definition 10). This will be the standard conforming Lagrange space of order 1 (for the MsFEM-lin), or the Crouzeix–Raviart space of order 1 (for the MsFEM-CR). Second we need to formulate the local problems for the numerical correctors (Definition 17 and Definition 22). This involves the definition of oversampling patches (for MsFEMs with oversampling, Definition 13), and an extension of the notion of degrees of freedom to define the boundary conditions for the numerical correctors (Definitions 11, 14 and 15) on oversampling patches. It is then possible to define the multiscale basis functions as a generalization of (10) (see Definition 26) and finally to define the MsFEM for our general framework in Definition 33.

Remark 7. We note that our development of non-intrusive MsFEM approaches relies to a great extent on the fact that (10), and its generalization (36) in the general framework developed below, provide a description of the multiscale basis functions in terms of $\mathbb{P}_1$ basis functions, without the need of higher-order functions, in a linear manner. Higher-order MsFEMs can be found in [23,24] (see also [25]). Possible analogues of (36) for such MsFEMs and the subsequent techniques to design a non-intrusive MsFEM variant are more involved and may be the topic of future work. See [22].

3.3. *Other motivations for the general framework*

Besides a unified formulation of our non-intrusive MsFEM approach, our general framework can also be beneficial to concrete code development for the MsFEM. Common features among various multiscale methods have previously been used to design flexible and efficient software for the implementation of such methods on the DUNE platform [26, 27] within the Exa-Dune project [28]. For example, the distribution of local problems over multiple processors and subsequent coupling in a global problem are handled by designated software components [29]. Our work may contribute to the efficient implementation of all MsFEMs covered by our general framework in such a project and similar endeavours yet to come.

When formulating the general framework, we also clarify a few practical matters that are often left pending in the various research articles we are aware of. In particular, we give a rigorous definition of the oversampling procedure near the boundary $\partial\Omega$ of the global domain.

As we explore the general framework, we will also propose an MsFEM variant that has not yet appeared in the literature: the MsFEM-CR combined with the oversampling technique (see Example 30). We hope that our framework may also further the development of new MsFEM variants in an attempt to improve on the shortcomings of the methods known today.

Finally, the present study may also uncover a deeper understanding of MsFEMs by paving the way to a unified convergence analysis of different variants. This work is currently in preparation.

4. Abstract definition of the MsFEM

We develop here a general framework for multiscale finite element methods. The ultimate aim is to generalize the key identities of Section 2.4. This is done in Definition 17 and 22 for the numerical correctors introduced in (11), and in Definition 26 for the expansion (10) of the multiscale basis functions. This allows the reformulation of the linear system of the MsFEM as the linear system of an effective problem in (48) (for a Petrov–Galerkin MsFEM) and (53) (for a Galerkin MsFEM) in Section 5. The other notions introduced in this section, although rather technical and abstract, are necessary tools to capture a wide variety of MsFEMs in our general framework.

4.1. *The continuous problem*

The abstract variational problem for our general MsFEM framework is as follows. Let a^ε be a continuous bilinear form on $H^1(\Omega) \times H^1(\Omega)$. We are interested in the solution to the problem

$$\text{Find } u^\varepsilon \in H_0^1(\Omega) \text{ such that } a^\varepsilon\left(u^\varepsilon, v\right) = F(v) \text{ for any } v \in H_0^1(\Omega), \tag{27}$$

where F is defined as in (4) for any $f \in L^2(\Omega)$. To ensure well-posedness of (27), we suppose that the bilinear form a^ε is coercive on $H_0^1(\Omega)$. The bilinear form a^ε may contain coefficients that oscillate on a microscopically small scale.

The oversampling and Crouzeix–Raviart variants of the MsFEM introduced in Section 3.1 show that we need to accommodate for approximation spaces with discontinuities at the interfaces. This requires some additional assumptions on the formulation of the abstract problem. We suppose that the bilinear form a^ε is in fact defined on the broken Sobolev space $H^1(\mathcal{T}_H) \times H^1(\mathcal{T}_H)$. More precisely, we assume that we can represent it as $a^\varepsilon = \sum_{K \in \mathcal{T}_H} a_K^\varepsilon$, where, for each $K \in \mathcal{T}_H$, a_K^ε is a continuous bilinear form defined on $H^1(K) \times H^1(K)$.

To ensure well-posedness of MsFEMs, which may use nonconforming approximation spaces, coercivity on $H_0^1(\Omega)$ may be insufficient. Therefore, we add the following coercivity hypothesis for the bilinear forms a_K^ε:

$$\text{for all } K \in \mathcal{T}_H, \text{ there exists } \alpha_K > 0 \text{ such that } \forall\ u \in H^1(K), \quad a_K^\varepsilon(u, u) \geq \alpha_K \|\nabla u\|_{L^2(K)}^2. \tag{28}$$

In order to perform a convergence analysis, one also has to assume that the α_K are bounded from below by some $\widetilde{\alpha} > 0$ that does not depend on H. We provide convergence results in Section 6 for the pure diffusion problem (1), in which case we have $\alpha_K = m$ from (2).

As an example, the introductory problem (1) with the associated bilinear form $a^{\varepsilon,\text{diff}}$ is covered by this framework as is made explicit in Example 8 below. Other second-order PDEs that fit in our abstract variational formulation are given in Example 9.

Example 8. The diffusion problem (1) is covered by the abstract variational formulation above. Indeed, we can set

$$a^\varepsilon = a^{\varepsilon,\text{diff}},$$

where $a^{\varepsilon,\text{diff}}$ is the bilinear form defined in (4). Further, we define

$$a_K^\varepsilon(u, v) = a_K^{\varepsilon,\text{diff}}(u, v) := \int_K \nabla v \cdot A^\varepsilon \nabla u,$$

for all $u, v \in H^1(\mathcal{T}_H)$, so that we have indeed $a^{\varepsilon,\text{diff}} = \sum_{K \in \mathcal{T}_H} a_K^{\varepsilon,\text{diff}}$. Clearly, each $a_K^{\varepsilon,\text{diff}}$ satisfies (28) with $\alpha_K = m$ the coercivity constant from (2).

Example 9. The reaction-advection-diffusion equation,

$$-\operatorname{div}\left(A^\varepsilon \nabla u^\varepsilon\right) + b \cdot \nabla u^\varepsilon + \sigma u^\varepsilon = f,$$

with a divergence-free advection field $b : \Omega \mapsto \mathbb{R}^d$ and a non-negative reaction coefficient $\sigma : \Omega \mapsto \mathbb{R}$, can be modelled (under some regularity hypotheses that we do not state here) with the bilinear forms

$$a_K^\varepsilon(u, v) = \int_K \nabla v \cdot A^\varepsilon \nabla u + v\, b \cdot \nabla u + \sigma u v.$$

However, these bilinear forms a_K^ε do not satisfy (28) even though the bilinear form a^ε is coercive on $H_0^1(\Omega)$. To this end, a skew-symmetrized formulation of the transport term can be used. The skew-symmetrized formulation uses the bilinear form

$$a_K^\varepsilon(u, v) = \int_K \nabla v \cdot A^\varepsilon \nabla u + \frac{1}{2} v\, b \cdot \nabla u - \frac{1}{2} u\, b \cdot \nabla v + \sigma u v, \tag{29}$$

which does satisfy (28). Assumption (28) is used for proving well-posedness of the MsFEM in Lemma 34, but note that both choices for a_K^ε mentioned here can be studied in practice. We refer e.g. to [22, 30, 31] for more details. Within the general MsFEM framework, b and σ are allowed to be highly oscillatory, and this may impact the specific MsFEM strategy to be preferred.

4.2. *Piecewise affine structure*

In Section 2.2, we have seen that the relation between multiscale basis functions and piecewise affine functions is essential for the development of our non-intrusive MsFEM. For the MsFEM definition in the general framework, we start by choosing such a structure in the following definition.

Definition 10. *Let a mesh $\mathcal{T}_H$ be given. The underlying $\mathbb{P}_1$ space for the MsFEM, denoted V_H, is one of the following two spaces: the Lagrange approximation space*

$$V_H^L = \left\{ v \in \mathbb{P}_1(\mathcal{T}_H) \,\middle|\, v \text{ is continuous on } \Omega \right\},$$

in which case we shall refer to the associated MsFEM as the MsFEM-lin, or the Crouzeix–Raviart approximation space

$$V_H^{CR} = \left\{ v \in \mathbb{P}_1(\mathcal{T}_H) \,\middle|\, \forall\ K \in \mathcal{T}_H,\ \forall\ e \in \mathcal{F}(K) \text{ such that } e \subset \Omega : \int_e [\![v]\!] = 0 \right\},$$

in which case the associated MsFEM shall be called the MsFEM-CR. We use the notation $\mathcal{F}(K)$ for the set of faces of K and $[\![v]\!]$ denotes the jump of v over the face e. The space V_H^L is a subspace of $H^1(\Omega)$, but V_H^{CR} is not. Note that no restrictions apply on faces lying on $\partial\Omega$.

We note that the underlying $\mathbb{P}_1$ space has the following property: if $v \in V_H$ is piecewise constant on the mesh $\mathcal{T}_H$, then v is constant in Ω. Contrary to the space V_H^L, functions in the Crouzeix–Raviart space V_H^{CR} are discontinuous in general. They are continuous, however, at the centroids of all faces of the mesh.

For standard finite elements, the notion of degrees of freedom allows to characterize any finite element function. The idea of the MsFEM is to preserve this notion of degrees of freedom (in a suitable way made precise below) in the definition of a multiscale approximation space, while adapting the piecewise affine structure to the microstructure of the PDE. We formalize this notion for the two underlying $\mathbb{P}_1$ spaces that we introduced in Definition 10. The definition involves an arbitrary simplex K, which is typically an element of the mesh $\mathcal{T}_H$, or an associated oversampling patch (for the oversampling technique of the MsFEM) that we shall define in Definition 13. The latter is not always a simplex, and we extend Definition 11 to such oversampling patches in Definitions 14 and 15.

Definition 11. *A degree of freedom operator (DOF operator) Γ associates to any simplex $K \subset \mathbb{R}^d$ and $v \in \mathbb{P}_1(K)$ a vector $\Gamma(K, v) \in \mathbb{R}^{d+1}$, whose components are called the degrees of freedom of v on K, in such a way that the application $v \mapsto \Gamma(K, v)$ is a linear bijection from $\mathbb{P}_1(K)$ to $\mathbb{R}^{d+1}$. More precisely, $\Gamma(K, \cdot)$ will denote in the sequel one of the following two operators:*

    (1) *(DOF operator for the MsFEM-lin.) Let $x_0, \ldots, x_d$ denote the vertices of K. We set*

$$\forall\ v \in \mathbb{P}_1(K), \qquad \Gamma^L(K, v) = (v(x_0), \ldots, v(x_d)).$$

    For $K \in \mathcal{T}_H$, the degree of freedom $[\Gamma^L(K, \cdot)]_j$ is said to be associated to the boundary if, for all $v \in \mathbb{P}_1(K)$, $[\Gamma^L(K, v)]_j = v(x)$ for a vertex x of the mesh that lies on $\partial\Omega$.

    (2) *(DOF operator for the MsFEM-CR.) Let $e_0, \ldots, e_d$ denote the faces of K. We set*

$$\forall\ v \in \mathbb{P}_1(K), \qquad \Gamma^{CR}(K, v) = \left(\frac{1}{|e_0|} \int_{e_0} v, \ldots, \frac{1}{|e_d|} \int_{e_d} v \right).$$

    For $K \in \mathcal{T}_H$, the degree of freedom $[\Gamma^{CR}(K, \cdot)]_j$ is said to be associated to the boundary if, for all $v \in \mathbb{P}_1(K)$, $[\Gamma^{CR}(K, v)]_j = \frac{1}{|e|} \int_e v$ for a face e of the mesh that lies on $\partial\Omega$.

The $\mathbb{P}_1$ test space is defined as

$$V_{H,0} = \left\{ v \in V_H \,\middle|\, \begin{array}{l} \forall\ K \in \mathcal{T}_H,\ \forall\ 1 \le j \le d+1, [\Gamma(K, v)]_j = 0 \text{ if the degree} \\ \text{of freedom } [\Gamma(K, \cdot)]_j \text{ is associated to the boundary} \end{array} \right\}.$$

The $\mathbb{P}_1$ test space is used in practice to approximate the subspace $H_0^1(\Omega)$ of $H^1(\Omega)$. The degrees of freedom are defined element per element and are thus local. Global properties of the underlying $\mathbb{P}_1$ space V_H are most easily made explicit through the identification of a basis for V_H.

Definition 12. *Let V_H be an underlying $\mathbb{P}_1$ space as in Definition 10, and let Γ be the associated DOF operator. We shall denote by N the dimension of V_H. The $\mathbb{P}_1$ basis functions $\phi_1^{\mathbb{P}_1}, \ldots, \phi_N^{\mathbb{P}_1}$ are defined as follows:*

- *For the MsFEM-lin, let $x_1, \ldots, x_N$ be an enumeration of the (internal and boundary) vertices of $\mathcal{T}_H$. Then $\phi_i^{\mathbb{P}_1}$ is defined by $\phi_i^{\mathbb{P}_1}(x_j) = \delta_{i,j}$ for all $1 \leq i, j \leq N$.*
- *For the MsFEM-CR, let $e_1, \ldots, e_N$ be an enumeration of the (internal and boundary) faces of $\mathcal{T}_H$. Then $\phi_i^{\mathbb{P}_1}$ is defined by $\frac{1}{|e_j|} \int_{e_j} \phi_i^{\mathbb{P}_1} = \delta_{i,j}$ for all $1 \leq i, j \leq N$.*

In both cases, these functions form a basis of the corresponding space V_H of Definition 10.

4.3. *Local problems*

4.3.1. *Oversampling patches*

To replace the (standard) underlying $\mathbb{P}_1$ space by a space of the same (low) dimension, adapted to the microstructure of a^ε, we associate to each mesh element $K \in \mathcal{T}_H$ an oversampling patch. It serves to avoid imposing artificial, non-oscillatory boundary conditions on K directly when computing numerical correctors to process the microstructure.

Definition 13. *Let $K \in \mathcal{T}_H$ be any mesh element and let S'_K be a simplex obtained from K by homothety around the centroid of K with homothety ratio $\rho \geq 1$. The oversampling patch S_K is defined as $S_K = S'_K \cap \Omega$.*

See Figure 1 for an illustration of the construction of oversampling patches in dimension 2. In this work, we allow for the trivial homothety ratio $\rho = 1$. In this case, the patch S_K coincides with K.

We will call an MsFEM *without oversampling* an MsFEM for which all oversampling patches satisfy $S_K = K$. Otherwise, the MsFEM is called an MsFEM *with oversampling*. We speak simply of an MsFEM when there are no assumptions on the oversampling patches.

(a)	**(b)**
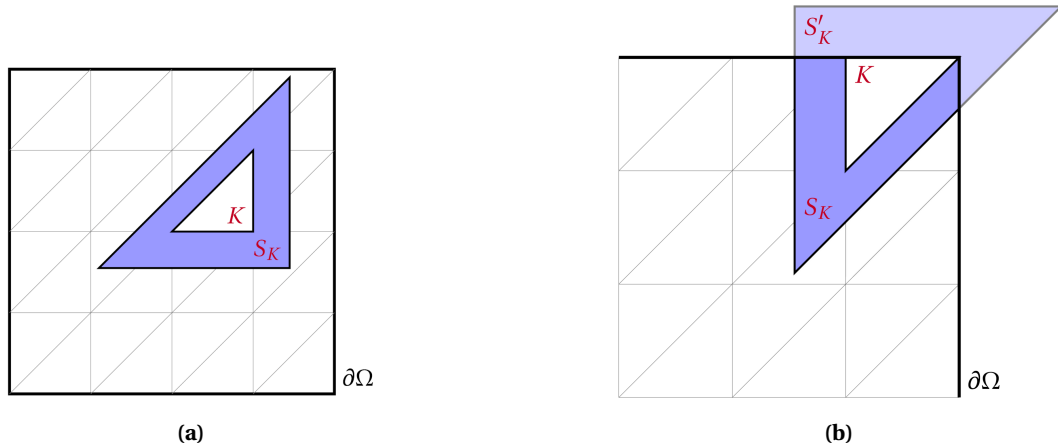
Figure 1. Oversampling patches for MsFEM in 2D. Left: the patch for the mesh element K is obtained from K by homothety. Right: The triangle S'_K partially lies outside the domain Ω and the oversampling patch S_K is not homothetic to K. It is not even a triangle.

For most mesh elements K, the patch S_K in Definition 13 is a simplex. However, for mesh elements close to the boundary $\partial\Omega$, alternative constructions should be considered. We have not found any explicit description of such a construction in the literature. This complicates the reproducibility of the method as well as a rigorous convergence analysis. The precise definitions of this section provide a first step to address these issues. A fully rigorous convergence analysis of the MsFEM with oversampling as described here is the subject of ongoing investigations [22].

4.3.2. *Degrees of freedom on oversampling patches*

Definition 11 provides the definition of *DOF* operators on any simplex. For the MsFEM, we wish to compute multiscale functions on oversampling patches S_K, in which case Definition 11 may be insufficient. We illustrated this in Figure 1b. Indeed, the number of vertices/faces of the oversampling patch may be larger than $d + 1$. In order to associate a multiscale basis function to every $\mathbb{P}_1$ basis function, we still need a notion of *DOF* operator such that $\Gamma(S_K, \cdot)$ is a linear bijection from $\mathbb{P}_1(S_K)$ to $\mathbb{R}^{d+1}$. Therefore, we extend the definition of the degrees of freedom operators Γ^L and Γ^{CR} in Definitions 14 and 15.

Definition 14. *Let $K \in \mathcal{T}_H$ and let S_K be its associated oversampling patch. Let $x_0, \ldots, x_d$ be a selection of $d + 1$ distinct vertices of S_K. We define the DOF operator Γ^L by*

$$\forall \ v \in \mathbb{P}_1(S_K), \qquad \Gamma^L(S_K, v) = (v(x_0), \ldots, v(x_d)).$$

We note that any choice of $d + 1$ nodal values unequivocally characterizes an affine function on S_K. Hence, $\Gamma^L(S_K, \cdot)$ is indeed a bijection. Now the precise choice of the vertices in Definition 14 is unimportant, because $\Gamma^L(S_K, \cdot)$ will only be used in the sequel to describe the trace of $\mathbb{P}_1$ functions on ∂S_K in boundary value problems. For any $\mathbb{P}_1$ function, this trace is uniquely defined by its values in $d + 1$ distinct vertices of S_K. Finally, when S_K is a simplex, it has only $d + 1$ vertices and Definition 14 reduces to Definition 11.

To generalize the notion of degrees of freedom for the Crouzeix–Raviart space to non-simplicial patches, we need to introduce some additional notation. On the boundary of a non-simplicial oversampling patch, we can identify some faces that collapse to a single vertex if we shrink S_K to K. We call these faces the additional faces and denote the set containing them by $\mathcal{F}_a(S_K)$. The other faces of S_K are referred to as the dilated faces, collected in the set $\mathcal{F}_d(S_K)$. When the patch S_K does not touch $\partial\Omega$, we have $\mathcal{F}_d(S_K) = \mathcal{F}(S_K)$ and $\mathcal{F}_a(S_K) = \emptyset$. In Figure 2a, for example, the additional faces are exactly those faces that lie on $\partial\Omega$. This is not always the case, as is illustrated by Figure 2b.

For the definition of $\Gamma^{CR}(S_K, \cdot)$, we shall rely on the existence of $d + 1$ dilated faces, because we need $\Gamma^{CR}(S_K, \cdot)$ to be a bijection between $\mathbb{P}_1(S_K)$ and $\mathbb{R}^{d+1}$. This imposes a constraint on the choice of the homothety ratio used to construct S_K. For example, in the case of Figure 2a, the lower right dilated face falls outside Ω if the homothety ratio is too large, and the oversampling patch S_K only has two dilated faces (edges here) and two additional faces. We do not consider this case hereafter.

Definition 15. *Let $K \in \mathcal{T}_H$ and let S_K be its associated oversampling patch. We assume that S_K has $d + 1$ dilated faces, and we denote them by $e_0, \ldots, e_d$. We define the DOF operator Γ^{CR} by*

$$\forall \ v \in \mathbb{P}_1(S_K), \qquad \Gamma^{CR}(S_K, v) = \left(\frac{1}{|e_0|} \int_{e_0} v, \ldots, \frac{1}{|e_d|} \int_{e_d} v \right).$$

When S_K is a simplex, we have $\mathcal{F}_d(S_K) = \mathcal{F}(S_K)$, and Definition 15 coincides with the respective elements of Definition 11.

4.3.3. *Numerical correctors: first oversampling strategy*

We now provide the precise assumptions under which we will consider local problems, i.e., the analogues of (5) defining the MsFEM-lin basis functions and the definition of the numerical correctors in (11). In fact, since the numerical correctors play an essential role in the construction of non-intrusive MsFEM approaches, we define the numerical correctors first and use them to define the multiscale basis functions in Definition 26.

We discuss two definitions of the numerical correctors, corresponding to two ways to define the oversampling technique for the MsFEM. The functional settings for these constructions
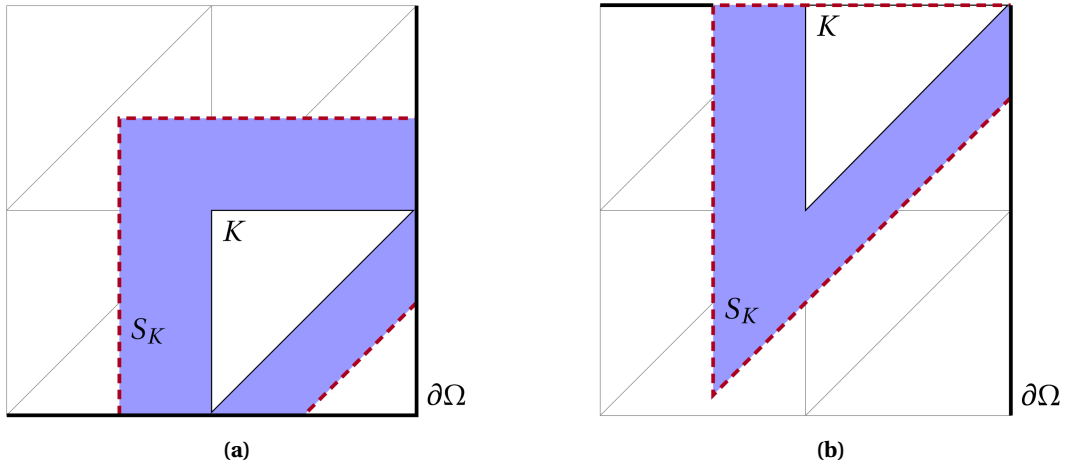
Figure 2. Non-simplicial oversampling patches in 2D. The dilated edges of the patch S_K, those that "correspond" to the edges of the original triangle K, are dashed and drawn in red.

are provided by Definitions 16 and 21. These definitions involve a "sampling space", whose name is inspired by the idea that only a limited number of local problems will be solved to encode the microstructure of the PDE in the numerical model. The choice of sampling space has to accommodate for the boundary conditions that one wishes to impose on the numerical correctors and basis functions (e.g. essential or natural; see Examples 18 and 19).

Definition 16. *Let $K \in \mathcal{T}_H$, let S_K be its associated oversampling patch and let Γ be a DOF operator from Definitions 11, 14 or 15. A subspace V_K of $H^1(S_K)$ and bilinear form $s_K^\varepsilon : V_K \times V_K \to \mathbb{R}$ are called* sampling space *and* sampling form, *respectively, if they satisfy the following:*

(1) *the space V_K contains the space of affine functions $\mathbb{P}_1(S_K)$;*

(2) *the operator $\Gamma(S_K, \cdot)$ is well-defined on V_K;*

(3) *the DOF-extended local problem: find $v \in V_K$ such that*

$$\begin{cases} s_K^\varepsilon(v, w) & = \langle g, w \rangle \text{ for all } w \in V_{K,0}, \\ \Gamma(S_K, v) & = \text{given}, \end{cases} \tag{30}$$

has a unique solution for any $g \in (H^1(S_K))'$. Here, $V_{K,0} = \{w \in V_K \mid \Gamma(S_K, w) = 0\}$ is the sampling test space.

Problem (30) is called "*DOF*-extended" because the degrees of freedom, controlling the boundary conditions associated to the local problem, are imposed on the oversampling patch S_K rather than the (generally smaller) mesh element K.

The sampling form s_K^ε shall be used to encode the oscillations of the bilinear form a^ε and thus the microstructure of the problem in the multiscale finite element functions. There is some flexibility in choosing the sampling form; one may choose to include all the same terms as those in the bilinear form a_K^ε of the original problem (27), or only some of them. When the MsFEM was first proposed in [4], it was suggested that s_K^ε should include those terms that correspond to the highest-order terms of the PDE that is to be solved. In the context of the advection-diffusion equation, one may thus choose to include in our MsFEM framework only the diffusion terms, or both the diffusion and advection terms. Both options have been studied e.g. in [10, 31].

In the functional setting of Definition 16, the generalization of (11) to define the numerical correctors for the general MsFEM framework is as follows.

Definition 17. *For all $K \in \mathcal{T}_H$, for any $0 \le \alpha \le d$, we introduce the function $\chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}} \in V_{K,0}$ as the unique solution to the* corrector problem

$$\forall\, w \in V_{K,0}, \quad s_K^\varepsilon\left(\chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}}, w\right) = \begin{cases} -s_K^\varepsilon\left(1, w\right) & \text{if } \alpha = 0, \\ -s_K^\varepsilon\left(x^\alpha - x_{c,K}^\alpha, w\right) & \text{if } 1 \le \alpha \le d. \end{cases} \tag{31}$$

The DOF-extended numerical corrector $\chi_K^{\varepsilon,\alpha,\mathsf{e}}$ is defined as the restriction of $\chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}}$ to K, extended to all of Ω by 0.

Note that the above definition introduces one more numerical corrector than introduced in (11) (namely the corrector for $\alpha = 0$). The precise definition of the numerical correctors is chosen such that the analogous expansion of (10) for the general framework (see (36)) leads to a PDE for the multiscale basis functions analogous to (5); we show this in Lemma 27 and (for a second oversampling strategy introduced below) in Lemma 28. In the following example, we see that Definition 17 is indeed a generalization of the numerical correctors defined by (11) in Section 2.4.

Example 18 (MsFEM-lin for diffusion problems). We consider $V_H = V_H^L$ and $\Gamma = \Gamma^L$ from Definition 11. For the diffusion problem (1), we have $a^\varepsilon = a^{\varepsilon,\mathsf{diff}}$ and we set $s_K^\varepsilon = a_K^{\varepsilon,\mathsf{diff}}$ (see Example 8). The sampling space for the MsFEM-lin is defined as

$$V_K = V_K^L := \left\{ v \in H^1(S_K) \,\middle|\, \exists\, w \in \mathbb{P}_1(S_K) \text{ such that } v|_{\partial S_K} = w|_{\partial S_K} \right\}.$$

Then the sampling test space $V_{K,0}^L$ is the space $H_0^1(S_K)$. In this case, it holds $a_K^{\varepsilon,\mathsf{diff}}(1, w) = 0$ for all $w \in V_{K,0}^L$. Consequently, the *DOF*-extended numerical corrector $\chi_K^{\varepsilon,0,\mathsf{e}}$ is identically equal to 0; we obtain indeed exactly d numerical correctors as in Section 2.4. For the non-trivial numerical correctors, Definition 17 corresponds to the weak formulation of the following boundary value problem:

$$-\operatorname{div}\left(A^\varepsilon \nabla \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}}\right) = \operatorname{div}\left(A^\varepsilon e_\alpha\right) \text{ in } S_K, \quad \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}} = 0 \text{ on } \partial S_K, \tag{32}$$

which is clearly well-posed.

Example 19 (MsFEM-CR for diffusion problems). Taking a^ε, a_K^ε and s_K^ε as in the previous example, we construct the MsFEM-CR with the sampling space $V_K^{CR} := H^1(S_K)$. With $V_H = V_H^{CR}$ and $\Gamma = \Gamma^{CR}$ from Definition 11, the corrector problem (31) for $\alpha = 0$ reduces to $\chi_K^{\varepsilon,0,\mathsf{e}} = 0$, as in Example 18. For $1 \le \alpha \le d$, the *DOF*-extended numerical corrector $\chi_K^{\varepsilon,\alpha,\mathsf{e}}$ is obtained from the boundary value problem:

$$\begin{cases} -\operatorname{div}\left(A^\varepsilon \nabla \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}}\right) & = \operatorname{div}\left(A^\varepsilon e_\alpha\right) \text{ in } S_K, \\ \vec{n} \cdot A^\varepsilon \nabla \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}} & = -\vec{n} \cdot A^\varepsilon e_\alpha \text{ on each } h \in \mathscr{F}_a(S_K), \\ \vec{n} \cdot A^\varepsilon \nabla \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}} & = c_h - \vec{n} \cdot A^\varepsilon e_\alpha \text{ on each } h \in \mathscr{F}_d(S_K), \\ \dfrac{1}{|h|} \displaystyle\int_h \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}} & = 0 \text{ for each } h \in \mathscr{F}_d(S_K), \end{cases} \tag{33}$$

where $\vec{n}$ denotes the outward unit vector on ∂S_K and c_h is a constant whose value is uniquely determined by the above problem. We note that the condition for the flux on the additional faces of S_K is entirely determined by the right-hand side in (31), whereas the flux on the dilated faces of S_K involves an additional constant, due to the fact that the test functions in $V_{K,0}^{CR}$ cannot take arbitrary values on the dilated faces. Indeed, their mean vanishes on these faces according to Definition 16.

When $S_K = K$ and when the faces of K do not lie on $\partial\Omega$, this corresponds to the setting of the original MsFEM-CR defined in [9]. The latter work also provides an alternative characterization of the multiscale Crouzeix–Raviart space.

When a face e of K lies on $\partial\Omega$, the basis functions that we will define below do not satisfy $\phi_e^\varepsilon = 0$ on e, but only satisfy a weak boundary condition in the average sense on e (and so does the corresponding MsFEM approximation to (27) defined below). This does not correspond to the original definition of the MsFEM-CR in [9, 32]. The MsFEM-CR with local boundary conditions as defined here was studied in [33–35].

Remark 20. In both Examples 18 and 19, the numerical corrector $\chi_K^{\varepsilon,0,\mathsf{e}}$ vanishes, because $s_K^\varepsilon(1, w) = 0$ for all w in the sampling test space. This is no longer the case e.g. for an MsFEM for advection-diffusion problems in which the sampling problem uses the skew-symmetrized bilinear form defined in (29). In this case, the numerical corrector $\chi_K^{\varepsilon,0,\mathsf{e}}$ does not vanish. In the corresponding effective numerical scheme that we will derive in (52), this leads to a term of order 0 even if such a term is not present in the advection-diffusion equation itself.

When $S_k = K$ (i.e., in the absence of oversampling), the *DOF* operator allows us to prescribe certain continuity properties on the faces of the mesh elements K. More precisely, when the MsFEM-lin with *DOF* operator Γ^L is employed, the numerical correctors $\chi_K^{\varepsilon,\alpha,\mathsf{e}}$ vanish at the vertices of the mesh, and, with the correct choice of sampling space (see Example 18), they vanish on all faces of K and are thus continuous on Ω. When the MsFEM-CR with *DOF* operator Γ^{CR} is considered, we obtain weak continuity of the numerical correctors over all faces of the mesh. The definition of the multiscale basis functions that we give below (see Definition 26, in the vein of the expansion (10)) shows that the continuity properties of the $\mathbb{P}_1$ basis functions of the underlying $\mathbb{P}_1$ space are not perturbed when building the multiscale basis functions.

In the general case, when the oversampling patch S_K is larger than K, we cannot preserve any of these continuity properties if we use *DOF*-extended local problems for our local computations, since the values on ∂K are not controlled by the degrees of freedom $\Gamma(S_K, \cdot)$ on ∂S_K. Therefore, we introduce another variant of the local problems to define *DOF*-continuous numerical correctors in the next section.

4.3.4. *Numerical correctors: second oversampling strategy*

Definition 21. *Let $K \in \mathcal{T}_H$ and let V_K and s_K^ε be a sampling space and sampling form, respectively, according to Definition 16. Additionally, suppose that the operator $\Gamma(K, \cdot)$ is well-defined on V_K. Then a DOF-continuous local problem is to find $v \in V_K$ such that*

$$\begin{cases} s_K^\varepsilon(v, w) & = \langle g, w \rangle \text{ for all } w \in V_{K,0}, \\ \Gamma(K, v) & = given, \end{cases} \tag{34}$$

for some $g \in (H^1(S_K))'$.

Definition 22. *Suppose any DOF-continuous local problem in Definition 21 is well-posed. Then we introduce, for all $K \in \mathcal{T}_H$ and all $0 \le \alpha \le d$, the functions $\chi_{S_K}^{\varepsilon,\alpha,\mathsf{c}}$ as the unique functions in V_K with $\Gamma(K, \chi_{S_K}^{\varepsilon,\alpha,\mathsf{c}}) = 0$ satisfying the corrector problem (31). We define the DOF-continuous numerical correctors $\chi_K^{\varepsilon,\alpha,\mathsf{c}}$ as the restriction of $\chi_{S_K}^{\varepsilon,\alpha,\mathsf{c}}$ to K, extended to all of Ω by 0.*

We emphasize that the local problems of Definitions 17 and 22 use test functions w in *the same space* $V_{K,0}$. This means that the test functions satisfy $\Gamma(S_K, w) = 0$ rather than $\Gamma(K, w) = 0$. The difference between *DOF*-extended and *DOF*-continuous numerical correctors is that the former satisfy $\Gamma(S_K, \chi_K^{\varepsilon,\alpha,\mathsf{e}}) = 0$, whereas the latter satisfy $\Gamma(K, \chi_K^{\varepsilon,\alpha,\mathsf{c}}) = 0$.

Remark 23. Clearly, when $S_K = K$, there is no difference between the *DOF*-extended and *DOF*-continuous problems (30) and (34). We shall in this case simply refer (30) (or (34)) as

local problems, and we write $\chi_K^{\varepsilon,\alpha,\bullet} = \chi_K^{\varepsilon,\alpha}$ for the numerical correctors of MsFEMs without oversampling.

Example 24 (MsFEM-lin for diffusion problems). Continuing Example 18, consider now the *DOF*-continuous numerical corrector $\chi_K^{\varepsilon,\alpha,\mathsf{c}}$. Equation (34) solves the following problem for $1 \le \alpha \le d$: there exists $w \in \mathbb{P}_1(S_K)$ such that

$$-\mathrm{div}\left(A^\varepsilon \nabla \chi_K^{\varepsilon,\alpha,\mathsf{c}}\right) = \mathrm{div}\left(A^\varepsilon e_\alpha\right) \ \text{ in } S_K, \quad \chi_K^{\varepsilon,\alpha,\mathsf{c}} = w \ \text{ on } \partial S_K, \quad \chi_K^{\varepsilon,\alpha,\mathsf{c}} = 0 \ \text{ at the vertices of } K.$$

The boundary condition on ∂S_K is complemented by a condition at the vertices of K. Except when A^ε is constant (and a solution is $\chi_K^{\varepsilon,\alpha,\mathsf{c}} = 0$), it is not evident whether a solution to this problem exists. For $\alpha = 0$, the numerical corrector $\chi_K^{\varepsilon,0,\mathsf{c}}$ vanishes, as in the *DOF*-extended case.

Example 25 (MsFEM-CR for diffusion problems). For the MsFEM-CR considered in Example 19, the *DOF*-continuous numerical correctors satisfy the same problem (33) (for $1 \le \alpha \le d$) as the *DOF*-extended numerical correctors, but with the average condition (the final equation in (33)) replaced by $\frac{1}{|h|} \int_h \chi_K^{\varepsilon,\alpha,\mathsf{c}} = 0$ for each $h \in \mathscr{F}(K)$. As we saw for the MsFEM-lin in Example 24, this is not a standard boundary value problem on S_K. For the case $\alpha = 0$, we have $\chi_K^{\varepsilon,0,\mathsf{c}} = 0$, which clearly satisfies the constraints $\frac{1}{|h|} \int_h \chi_K^{\varepsilon,0,\mathsf{c}} = 0$ for each $h \in \mathscr{F}(K)$.

Examples 18 and 19 show that a *DOF*-extended local problem is typically equivalent to a PDE with boundary conditions on S_K. Under reasonable assumptions, these problems have a unique solution as required by Definition 16. We have seen in Examples 24 and 25 that this is not the case for *DOF*-continuous problems, for which one finds some boundary conditions on ∂S_K (because the degrees of freedom of test functions in $V_{K,0}$ are prescribed on S_K) and another set of conditions on ∂K that are explicitly imposed through the degrees of freedom on K in (34). Well-posedness is not obvious in general, and cannot always be deduced from well-posedness of the *DOF*-extended counterpart (30). We address the well-posedness of *DOF*-continuous problems in more detail in Section 4.3.5. The advantage of *DOF*-continuous oversampling is that it imposes certain continuity properties on the multiscale basis functions, and we will see in Section 7 that it yields better numerical approximations than *DOF*-extended oversampling.

4.3.5. *Well-posedness of DOF-continuous numerical correctors*

We have seen in Examples 24 and 25 that *DOF*-continuous local problems lead to non-standard boundary conditions. This poses not only a theoretical issue, but also a computational challenge. To complete our study of the general MsFEM framework, we now present a computational strategy to obtain the *DOF*-continuous numerical correctors, and we use this strategy to discuss the well-posedness of the associated local problems.

In Definition 16 we assume the well-posedness of *DOF*-extended problems, and we have seen in Examples 18 and 19 that this is a natural assumption. It is also natural to assume that we can compute *DOF*-extended numerical correctors numerically. We compute the *DOF*-continuous numerical correctors from the *DOF*-extended numerical correctors, by subtracting a linear combination of suitable functions W^β from the *DOF*-extended numerical correctors. The W^β must all satisfy the homogeneous equation $s_K^\varepsilon(W^\beta, w) = 0$ for all $w \in V_{K,0}$, in order not to perturb the local problem (31) that is already satisfied by both types of numerical correctors. We shall use the functions $W^0 := 1 + \chi_{S_K}^{\varepsilon,0,\mathsf{e}}$ and $W^\beta := x^\beta - x_{c,K}^\beta + \chi_{S_K}^{\varepsilon,\beta,\mathsf{e}}$ for $1 \le \beta \le d$, where $\chi_{S_K}^{\varepsilon,\beta,\mathsf{e}}$ is defined in Definition 17. The precise strategy is as follows.

Fix $0 \le \alpha \le d$. We look for coefficients $c_0^\alpha, \ldots, c_d^\alpha$ such that $\chi_{S_K}^{\varepsilon,\alpha,\mathsf{c}} = \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}} - \sum_{\beta=0}^d c_\beta^\alpha W^\beta$ on S_K, where we recall that $\chi_{S_K}^{\varepsilon,\alpha,\mathsf{c}}$ is defined by Definition 22. Note that both sides of the equation clearly

satisfy (31). The desired equality thus holds if and only if $\Gamma(K, \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}} - \sum_{\beta=0}^{d} c_\beta^\alpha W^\beta) = 0$. Since the *DOF* operators are linear, this leads to the linear system

$$\underbrace{\begin{bmatrix} | & | & & | \\ \Gamma(K, W^0) & \Gamma(K, W^1) & \dots & \Gamma(K, W^d) \\ | & | & & | \end{bmatrix}}_{=:\, \mathbb{M}} \begin{bmatrix} c_0^\alpha \\ c_1^\alpha \\ \vdots \\ c_d^\alpha \end{bmatrix} = \Gamma\left(K, \chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}}\right). \tag{35}$$

Invertibility of the matrix $\mathbb{M}$ is thus a sufficient condition for the existence of all *DOF*-continuous numerical correctors, and the resolution of the linear system (35) for each α (where all *DOF*-extended numerical correctors are replaced by their numerical approximation) allows to compute the *DOF*-continuous numerical correctors numerically.

Before studying the invertibility of the matrix $\mathbb{M}$ in a few special cases, let us consider the matrix composed of the degrees of freedom on S_K, i.e., the matrix

$$\widetilde{\mathbb{M}} := \begin{bmatrix} | & | & & | \\ \Gamma\left(S_K, W^0\right) & \Gamma\left(S_K, W^1\right) & \dots & \Gamma\left(S_K, W^d\right) \\ | & | & & | \end{bmatrix}.$$

By definition of the functions $\chi_{S_K}^{\varepsilon,\beta,\mathsf{c}}$, we have $\Gamma(S_K, W^\beta) = \Gamma(S_K, x^\beta - x_{c,K}^\beta)$ for $1 \le \beta \le d$, and $\Gamma(S_K, W^0) = \Gamma(S_K, 1)$. Note that the constant function together with the coordinate functions $x^\beta - x_{c,K}^\beta$ ($1 \le \beta \le d$) span $\mathbb{P}_1(S_K)$. Since $\Gamma(S_K, \cdot)$ is a bijection, the vectors $\Gamma(S_K, W^0), \dots, \Gamma(S_K, W^d)$ are linearly independent. Hence the matrix $\widetilde{\mathbb{M}}$ is invertible. One may hope that the linear independence of the vectors $\Gamma(S_K, W^\beta)$ is preserved for the degrees of freedom on the interior boundary ∂K instead of ∂S_K, yielding invertibility of $\mathbb{M}$. We found this to hold for all numerical tests that we performed, involving both the MsFEM-lin and the MsFEM-CR.

We can prove invertibility of $\mathbb{M}$ in a few special cases. When s_K^ε is the sampling form that was used in Example 18 (corresponding to a diffusion problem; we will consider this case until the end of this section) and if A^ε is constant, all numerical correctors vanish on S_K and the foregoing argument for the matrix $\widetilde{\mathbb{M}}$ shows invertibility of $\mathbb{M}$.

In the periodic setting (see Section 2.7), even though A^ε itself is not constant, its homogenized limit $A^\star$ is. In this case, the $\chi_{S_K}^{\varepsilon,\beta,\mathsf{e}}$ converge to zero weakly in $H^1(S_K)$. (We show this in Lemma 45 in the absence of oversampling, but the argument can be generalized to *DOF*-extended oversampling.) Now consider the MsFEM-CR. The weak convergence of the $\chi_{S_K}^{\varepsilon,\beta,\mathsf{e}}$ in $H^1(S_K)$ ensures weak convergence on each face of K in the $H^{1/2}$-norm by continuity of the trace operator. Since the embedding of $H^{1/2}(\partial K)$ in $L^2(\partial K)$ is compact, the $\chi_{S_K}^{\varepsilon,\beta,\mathsf{e}}$ converge to 0 strongly in L^2 on each face of K. Consequently, the degrees of freedom $\Gamma(K, \chi_{S_K}^{\varepsilon,\beta,\mathsf{e}})$ (the averages on the faces of K) converge to zero as $\varepsilon \to 0$. Thus,

$$\Gamma\left(K, W^0\right) \to \Gamma(K, 1) \text{ and } \Gamma\left(K, W^\beta\right) \to \Gamma\left(K, x^\beta - x_{c,K}^\beta\right) \text{ as } \varepsilon \to 0 \text{ for all } 1 \le \beta \le d$$

and, by the above argument for the matrix $\widetilde{\mathbb{M}}$, the matrix $\mathbb{M}$ is invertible in this limit. By continuity of the determinant function, the matrix $\mathbb{M}$ is invertible when ε is small enough, and the *DOF*-continuous basis functions exist in this regime.

The study of the *DOF*-continuous numerical correctors for the MsFEM-lin is more delicate, since pointwise operations are involved in evaluating the degrees of freedom, which are ill-defined on $H^1(S_K)$. One can invoke the De Giorgi–Nash result, which can be found e.g. in [14, Theorem 8.22], to see that the multiscale basis functions, obtained from the numerical correctors in Definition 26 below, are in fact continuous for any bounded diffusion tensor. (See Example 29 for a definition of the multiscale basis functions for the MsFEM-lin independent of the numerical correctors.) Pointwise evaluation is then justified. It would therefore be convenient to study

the *DOF*-continuous basis functions directly, without the intermediate step of the numerical correctors. We do not further pursue this topic here.

4.4. *The multiscale basis functions*

We can now define the multiscale basis functions for the approximation of the abstract problem (27) in terms of the numerical correctors. We recall that in Section 2.2, the numerical correctors were derived from the definition of the basis functions. We give an equivalent definition of the multiscale basis functions, independent of the numerical correctors, in Lemmas 27 and 28. Recall that $\phi_1^{\mathbb{P}_1}, \dots, \phi_N^{\mathbb{P}_1}$ is a basis of the space V_H (see Definition 12). We can suppose that the first N_0 basis functions form a basis of $V_{H,0}$. The following definition is the generalization of (10) to the general MsFEM framework.

Definition 26. *For each $i = 1, \dots, N$, the* multiscale basis function ϕ_i^ε *is defined by*

$$\forall\, K \in \mathcal{T}_H, \qquad \phi_i^\varepsilon\big|_K = \phi_i^{\mathbb{P}_1}\Big|_K + \phi_i^{\mathbb{P}_1}(x_{c,K})\,\chi_K^{\varepsilon,0,\bullet} + \sum_{\alpha=1}^d \partial_\alpha\left(\phi_i^{\mathbb{P}_1}\Big|_K\right)\chi_K^{\varepsilon,\alpha,\bullet}, \tag{36}$$

where $\bullet = \mathsf{e}$ corresponds to DOF-extended multiscale basis functions and $\bullet = \mathsf{c}$ corresponds to DOF-continuous multiscale basis functions.

The *DOF*-extended multiscale basis functions satisfy a variational problem on the oversampling patches S_K as shown by the following lemma.

Lemma 27. *Let K be any mesh element and let $1 \le i \le N$. Consider an MsFEM with DOF-extended basis functions. Define an extension of ϕ_i^ε from K to S_K by*

$$\widehat{\phi_i^\varepsilon} = \widehat{\phi_i^{\mathbb{P}_1}\Big|_K} + \phi_i^{\mathbb{P}_1}(x_{c,K})\,\chi_{S_K}^{\varepsilon,0,\mathsf{e}} + \sum_{\alpha=1}^d \partial_\alpha\left(\phi_i^{\mathbb{P}_1}\Big|_K\right)\chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}}, \quad in\ S_K, \tag{37}$$

where $\widehat{\phi_i^{\mathbb{P}_1}|_K}$ denotes the affine extension of $\phi_i^{\mathbb{P}_1}|_K$ to S_K, and $\chi_K^{\varepsilon,\alpha,\mathsf{e}}$ is as in Definition 17. Then $\widehat{\phi_i^\varepsilon}$ is the unique solution in V_K to

$$\begin{cases} s_K^\varepsilon\left(\widehat{\phi_i^\varepsilon}, w\right) &= 0\ \text{for all}\ w \in V_{K,0}, \\ \Gamma\left(S_K, \widehat{\phi_i^\varepsilon}\right) &= \Gamma\left(S_K, \widehat{\phi_i^{\mathbb{P}_1}\Big|_K}\right). \end{cases} \tag{38}$$

In the case of the MsFEM-lin for the diffusion problem (1), problem (38) with $S_K = K$ coincides with the definition of the multiscale basis functions in (5); see Example 29.

Proof. Problem (38) has a unique solution in view of Definition 16. It thus suffices to show that $\widehat{\phi_i^\varepsilon}$ satisfies (38). Since the numerical correctors $\chi_K^{\varepsilon,\alpha,\mathsf{e}}$ belong to $V_{K,0}$ for all $0 \le \alpha \le d$, it is clear from (37) that

$$\Gamma\left(S_K, \widehat{\phi_i^\varepsilon}\right) = \Gamma\left(S_K, \widehat{\phi_i^{\mathbb{P}_1}\Big|_K}\right).$$

Inserting (37) into (38) and applying (31) to all $\chi_K^{\varepsilon,\alpha,\mathsf{e}}$, we find, for any test function $w \in V_{K,0}$,

$$s_K^\varepsilon\left(\widehat{\phi_i^\varepsilon}, w\right) = s_K^\varepsilon\left(\widehat{\phi_i^{\mathbb{P}_1}\Big|_K}, w\right) + \phi_i^{\mathbb{P}_1}(x_{c,K})\,s_K^\varepsilon\left(\chi_{S_K}^{\varepsilon,0,\mathsf{e}}, w\right) + \sum_{\alpha=1}^d \left(\partial_\alpha\phi_i^{\mathbb{P}_1}\right)\Big|_K\,s_K^\varepsilon\left(\chi_{S_K}^{\varepsilon,\alpha,\mathsf{e}}, w\right)$$

$$= s_K^\varepsilon\left(\widehat{\phi_i^{\mathbb{P}_1}\Big|_K}, w\right) - \phi_i^{\mathbb{P}_1}(x_{c,K})\,s_K^\varepsilon(1, w) - \sum_{\alpha=1}^d \left(\partial_\alpha\phi_i^{\mathbb{P}_1}\right)\Big|_K\,s_K^\varepsilon\left(x^\alpha - x_{c,K}^\alpha, w\right)$$

$$= s_K^\varepsilon\left(\widehat{\phi_i^{\mathbb{P}_1}\Big|_K}, w\right) - s_K^\varepsilon\left(\phi_i^{\mathbb{P}_1}(x_{c,K}) + \sum_{\alpha=1}^d \left(\partial_\alpha\phi_i^{\mathbb{P}_1}\right)\Big|_K\left(x^\alpha - x_{c,K}^\alpha\right), w\right).$$

Here we use that s_K^ε is a bilinear form on V_K, that all piecewise affine functions are contained in V_K according to Definition 16 (this ensures that $\widehat{\phi_i^{\tilde\varepsilon}}$ indeed lies in the domain of s_K^ε), and the property that $\nabla\phi_i^{\mathbb{P}1}$ is piecewise constant. Finally, we use (20) for $\varphi = \widehat{\phi_i^{\mathbb{P}1}}$ to conclude that

$$s_K^\varepsilon\left(\widehat{\phi_i^{\tilde\varepsilon}}, w\right) = s_K^\varepsilon\left(\left.\widehat{\phi_i^{\mathbb{P}1}}\right|_K, w\right) - s_K^\varepsilon\left(\left.\widehat{\phi_i^{\mathbb{P}1}}\right|_K, w\right) = 0,$$

which establishes the desired variational formulation satisfied by $\widehat{\phi_i^{\tilde\varepsilon}}$. $\qquad\square$

If the *DOF*-continuous problems (34) are well-posed, we obtain by the same arguments the following result for *DOF*-continuous multiscale basis functions.

Lemma 28. *Let K be any mesh element and let $1 \le i \le N$. Assume that any DOF-continuous local problem* (34) *is well-posed. Consider an MsFEM with DOF-continuous basis functions. Define an extension of ϕ_i^ε from K to S_K by*

$$\widehat{\phi_i^{\tilde\varepsilon}} = \left.\widehat{\phi_i^{\mathbb{P}1}}\right|_K + \phi_i^{\mathbb{P}1}\left(x_{c,K}\right)\chi_{S_K}^{\varepsilon,0,\mathsf{c}} + \sum_{\alpha=1}^d \partial_\alpha\left(\left.\phi_i^{\mathbb{P}1}\right|_K\right)\chi_{S_K}^{\varepsilon,\alpha,\mathsf{c}}, \quad in\ S_K,$$

where $\chi_K^{\varepsilon,\alpha,\mathsf{c}}$ is as in Definition 22, and $\widehat{\phi_i^{\mathbb{P}1}}|_K$ is as defined in Lemma 27. Then $\widehat{\phi_i^{\tilde\varepsilon}}$ is the unique solution in V_K to

$$\begin{cases} s_K^\varepsilon\left(\widehat{\phi_i^{\tilde\varepsilon}}, w\right) & = 0\ \text{for all}\ w \in V_{K,0}, \\ \Gamma\left(K, \widehat{\phi_i^{\tilde\varepsilon}}\right) & = \Gamma\left(K, \phi_i^{\mathbb{P}1}\right). \end{cases} \tag{39}$$

Example 29 (MsFEM-lin for diffusion problems). In the setting of Example 18, any *DOF*-extended multiscale basis function ϕ_i^ε for the MsFEM-lin constructed in (38) is obtained, in each mesh element K, as the restriction of a function $\widehat{\phi_i^{\tilde\varepsilon}}$, which is the unique solution in $H^1(S_K)$ to

$$-\operatorname{div}\left(A^\varepsilon\nabla\widehat{\phi_i^{\tilde\varepsilon}}\right) = 0\ \text{in}\ S_K, \quad \widehat{\phi_i^{\tilde\varepsilon}} = \widehat{\phi_i^{\mathbb{P}1}}\ \text{on}\ \partial S_K.$$

For a *DOF*-continuous basis function, $\widehat{\phi_i^{\tilde\varepsilon}}$ solves the same PDE in S_K, is affine on ∂S_K, and satisfies $\widehat{\phi_i^{\tilde\varepsilon}}(x_j) = \phi_i^{\mathbb{P}1}(x_j)$ at all vertices x_j of K.

Example 30 (MsFEM-CR for diffusion problems). In the continuation of Example 19, the *DOF*-extended multiscale basis function ϕ_i^ε for the MsFEM-CR is the restriction to K of $\widehat{\phi_i^{\tilde\varepsilon}}$, the unique solution in $H^1(S_K)$ to

$$\begin{cases} -\operatorname{div}\left(A^\varepsilon\nabla\widehat{\phi_i^{\tilde\varepsilon}}\right) & = 0\ \text{in}\ S_K, \\ \vec{n}\cdot A^\varepsilon\nabla\widehat{\phi_i^{\tilde\varepsilon}} & = 0\ \text{on each}\ h \in \mathscr{F}_a(S_K), \\ \vec{n}\cdot A^\varepsilon\nabla\widehat{\phi_i^{\tilde\varepsilon}} & = c_h\ \text{on each}\ h \in \mathscr{F}_d(S_K), \\ \dfrac{1}{|h|}\displaystyle\int_h \widehat{\phi_i^{\tilde\varepsilon}} & = \dfrac{1}{|h|}\displaystyle\int_h \widehat{\phi_i^{\mathbb{P}1}}\ \text{for each}\ h \in \mathscr{F}_d(S_K), \end{cases}$$

where the constants c_h are uniquely determined by the problem. We recall that the sets of faces $\mathscr{F}_a(S_K)$ and $\mathscr{F}_d(S_K)$ are defined in Section 4.3.2. For *DOF*-continuous basis functions, the last condition is applied to the faces $h \in \mathscr{F}(K)$ (and all other conditions remain unchanged).

Our general framework allows two characterizations of the multiscale basis functions, namely (36) and (38) or (39), as was the case for the MsFEM studied in Section 2.2 (where ϕ_i^ε is given by (5) or (10)). The essential advantage of (36) is that the microscale is fully encoded in the numerical correctors $\chi_K^{\varepsilon,\alpha,\bullet}$, that can be computed element per element without any global information. In particular, the *global* index i of the multiscale basis function ϕ_i^ε is irrelevant for the computation of the numerical correctors. The expression in (36) is therefore the crucial relationship that we will employ to develop non-intrusive MsFEMs within the general framework in Section 5, just as was (10) in Sections 2.4 and 2.5.

The second formulation of the multiscale basis functions, as solutions to the local problems (38) or (39), provides a more direct interpretation of the multiscale basis functions in terms of the sampling form chosen. It also gives a relation between the degrees of freedom of the $\mathbb{P}_1$ basis functions and the associated multiscale basis function. This is useful in particular for the well-posedness of the MsFEM, that we study in Lemma 34.

Remark 31. Our definition of the multiscale basis functions in (36) is reminiscent of the Variational Multiscale Method, a framework developed in [36,37] to adapt Galerkin approximations on low-dimensional spaces to the presence of multiscale features. In this context, our formulation of the MsFEM also exhibits a link with residual-free bubbles, see e.g. [37–39].

Remark 32. The first introduction of the MsFEM in [4] corresponds to the idea of oversampling with *DOF*-continuous basis functions. Although their existence cannot be established in general, they are computed numerically by taking linear combinations of *DOF*-extended basis functions (following an analogous strategy to the one we discussed in Section 4.3.5). The MsFEM with *DOF*-extended basis functions is studied in the works [40, 41] dealing with the convergence analysis of the MsFEM-lin with oversampling.

Let us also note that the combination of Crouzeix-Raviart MsFEM and oversampling has, to the best of our knowledge, not yet been proposed in the literature. This method, for which the basis functions are given explicitly in Example 30, is a natural by-product of the identification of the abstract MsFEM framework.

4.5. *The global problem*

We can now define the multiscale trial and test spaces, respectively V_H^ε and $V_{H,0}^\varepsilon$, as follows:

$$V_H^\varepsilon = \left\{ \phi_i^\varepsilon \mid 1 \le i \le N \right\}, \qquad V_{H,0}^\varepsilon = \left\{ \phi_i^\varepsilon \mid 1 \le i \le N_0 \right\}.$$

We recall that we have assumed the first N_0 basis functions of V_H to form a basis of $V_{H,0}$ in Section 4.4. Note that we only use $V_{H,0}^\varepsilon$ in the present section, because (27) is posed with homogeneous Dirichlet boundary conditions, but that the larger space V_H^ε is useful for more general boundary conditions (see Section 5.3). Applying (36), we have the equivalent characterization in terms of the $\mathbb{P}_1$ space V_H,

$$V_H^\varepsilon = \left\{ v_H^\varepsilon = v_H + \sum_{K \in \mathcal{T}_H} \left(v_H(x_{c,K}) \, \chi_K^{\varepsilon,0,\bullet} + \sum_{\alpha=1}^d \partial_\alpha \left(v_H\|_K \right) \chi_K^{\varepsilon,\alpha,\bullet} \right) \middle| \, v_H \in V_H \right\}.$$

Definition 33. *Let V_H be an underlying $\mathbb{P}_1$ space defined in Definition 10 with the associated DOF operator Γ from Definition 11 and Definitions 14-15. Define for each mesh element $K \in \mathcal{T}_H$ an oversampling patch (Definition 13), a sampling space and sampling form in accordance with Definition 16. Let the multiscale basis functions ϕ_i^ε be given as in Definition 26. Then a* Multiscale Finite Element Method *(MsFEM) for problem* (27) *is: find $u_H^\varepsilon \in V_{H,0}^\varepsilon$ such that*

$$\forall \; v_H^\varepsilon \in V_{H,0}^\varepsilon, \qquad \sum_{K \in \mathcal{T}_H} a_K^\varepsilon \left(u_H^\varepsilon, v_H^\varepsilon \right) = F \left(v_H^\varepsilon \right). \tag{40}$$

In the following lemma, we investigate the well-posedness of the MsFEM.

Lemma 34. *Consider an MsFEM without oversampling, or an MsFEM with oversampling using DOF-continuous basis functions (assuming the associated basis functions are well-defined). When a^ε satisfies* (28)*, the MsFEM* (40) *has a unique solution.*

Proof. Note that, with *DOF*-continuous oversampling, but also without oversampling, the multiscale basis functions satisfy (39). In particular, all degrees of freedom of u_H^ε related to the boundary vanish. Also note that, the dimension of $V_{H,0}^\varepsilon$ being finite, it suffices to show that $u_H^\varepsilon = 0$ is the unique solution to problem (40) with $F = 0$.

If $0 = F(u_H^\varepsilon) = a^\varepsilon(u_H^\varepsilon, u_H^\varepsilon)$, it follows from (28) that u_H^ε is piecewise constant. Let us write $u_H^\varepsilon = \sum_{i=1}^{N_0} \alpha_i \phi_i^\varepsilon$ for some coefficients $\alpha_i \in \mathbb{R}$ and introduce the function $u_H = \sum_{i=1}^{N_0} \alpha_i \phi_i^{\mathbb{P}_1} \in V_H$. Because of (39), we have $\Gamma(K, u_H^\varepsilon) = \Gamma(K, u_H)$ for all mesh elements K. Since u_H^ε is piecewise constant and $\Gamma(K, \cdot)$ is a bijection from $\mathbb{P}_1(K)$ to $\mathbb{R}^{d+1}$ (recall Definition 11), it follows that $u_H^\varepsilon = u_H$. In particular, the multiscale function u_H^ε in fact belongs to the underlying $\mathbb{P}_1$ space V_H.

We remarked immediately below Definition 10 that, for either of the two spaces $V_H = V_H^L$ or V_H^{CR}, the above implies that u_H^ε is constant throughout Ω. Since the degrees of freedom of u_H^ε associated to the boundary vanish, we readily deduce that $u_H^\varepsilon = 0$. □

We do not know of the existence of a result on the well-posedness of MsFEMs with oversampling using *DOF*-extended multiscale basis functions. In [41], the authors establish an inf-sup result for a variant of the MsFEM-lin-OS with $\mathbb{P}_1$ test functions (see also Definition 35). This result is obtained for a periodic diffusion coefficient in the limit of sufficiently small ε.

5. Non-intrusive MsFEM for the general framework

We show in this section how to develop a non-intrusive approach for the general MsFEM framework of Section 4. We have seen in Lemma 5 that, for a particular MsFEM variant, the non-intrusive Galerkin MsFEM approach coincides with a Petrov-Galerkin MsFEM. This does not hold for all MsFEMs in the general framework. We first develop a non-intrusive MsFEM approach for a Petrov–Galerkin MsFEM in the general framework. We show that the non-intrusive approach for the Petrov–Galerkin MsFEM is actually equivalent to the Petrov–Galerkin MsFEM itself. In a second step, we introduce a non-intrusive approximation of the Galerkin MsFEM. Before doing so, let us summarize the main steps of Sections 2.4 and 2.5 to obtain a non-intrusive MsFEM approach:

(1) the expansion (10) allows to recast the matrix $\mathbb{A}^\varepsilon$ of the linear system for the MsFEM as the matrix $\mathbb{A}^{\mathbb{P}_1}$ associated to the $\mathbb{P}_1$ discretization of an effective problem;

(2) we approximate the right-hand side $\mathbb{F}^\varepsilon$ of the MsFEM problem by the right-hand side $\mathbb{F}^{\mathbb{P}_1}$ of this $\mathbb{P}_1$ discretization;

(3) the post-processing step (21) applied to the $\mathbb{P}_1$ approximation of the effective problem yields the MsFEM approximation.

5.1. *The Petrov–Galerkin MsFEM*

We recall that the abstract continuous problem for which we developed the MsFEM in Section 4 is given by (27) and that it can be rewritten in terms of the bilinear forms a_K^ε satisfying (28). Petrov-Galerkin variants of the multiscale finite element method with $\mathbb{P}_1$ test functions were previously studied in [24, 41]. In our general MsFEM framework, the adaptation of Definition 33 to a Petrov-Galerkin MsFEM is the following.

Definition 35. *Let V_H be an underlying $\mathbb{P}_1$ space defined in Definition 10 with the associated DOF operator Γ from Definition 11 and Definitions 14-15. Define for each mesh element $K \in \mathcal{T}_H$ an oversampling patch (Definition 13), a sampling space and sampling form in accordance with Definition 16. Let the multiscale basis functions ϕ_i^ε be given as in Definition 26. Then a* Petrov-Galerkin Multiscale Finite Element Method *(PG-MsFEM) for problem* (27) *is: find $u_H^\varepsilon \in V_{H,0}^\varepsilon$ such that*

$$\forall \; v_H \in V_{H,0}, \qquad \sum_{K \in \mathcal{T}_H} a_K^\varepsilon\left(u_H^\varepsilon, v_H\right) = F(v_H). \tag{41}$$

When confusion may arise, we shall refer to the MsFEM defined in Definition 33 as the Galerkin MsFEM (G-MsFEM). To study well-posedness of the PG-MsFEM, it is most convenient to relate this method to the G-MsFEM. Therefore, we postpone well-posedness of (41) to Lemma 38.

We now execute step 1 of the summary of the non-intrusive MsFEM approach at the beginning of this section. The matrix $\mathbb{A}^\varepsilon$ of the linear system associated to (41) is defined by

$$\mathbb{A}^\varepsilon_{j,i} = \sum_{K \in \mathscr{T}_H} a^\varepsilon_K\left(\phi^\varepsilon_i, \phi^{\mathbb{P}_1}_j\right), \quad 1 \le i, j \le N_0. \tag{42}$$

To find an effective $\mathbb{P}_1$ formulation with the same linear system, we will use the definition (36) of the multiscale basis functions in the general framework, but first we combine it with (20) applied to $\varphi = \phi^{\mathbb{P}_1}_i$ to rewrite (36) as

$$\phi^\varepsilon_i\big|_K = \phi^{\mathbb{P}_1}_i\left(x_{c,K}\right) + \phi^{\mathbb{P}_1}_i\left(x_{c,K}\right)\chi^{\varepsilon,0,\bullet}_K + \sum_{\alpha=1}^{d} \partial_\alpha \phi^{\mathbb{P}_1}_i\Big|_K \left(x^\alpha - x^\alpha_{c,K} + \chi^{\varepsilon,\alpha,\bullet}_K\right)$$
$$= \phi^{\mathbb{P}_1}_i\left(x_{c,K}\right)\Lambda^{\varepsilon,0}_K + \sum_{\alpha=1}^{d} \partial_\alpha \phi^{\mathbb{P}_1}_i\Big|_K \Lambda^{\varepsilon,\alpha}_K, \tag{43}$$

where

$$\Lambda^{\varepsilon,0}_K := 1 + \chi^{\varepsilon,0,\bullet}_K, \quad \Lambda^{\varepsilon,\alpha}_K := x^\alpha - x^\alpha_{c,K} + \chi^{\varepsilon,\alpha,\bullet}_K, \tag{44}$$

for all $1 \le \alpha \le d$ and each $K \in \mathscr{T}_H$. We recall that $\bullet \in \{e, c\}$ indicates the choice of *DOF*-extended or *DOF*-continuous basis functions. Inserting (43) into (42) for ϕ^ε_i and (20) for $\varphi = \phi^{\mathbb{P}_1}_j$ yields

$$\mathbb{A}^\varepsilon_{j,i} = \sum_{K \in \mathscr{T}_H} \left(\phi^{\mathbb{P}_1}_i\left(x_{c,K}\right) a^\varepsilon_K\left(\Lambda^{\varepsilon,0}_K, 1\right) \phi^{\mathbb{P}_1}_j\left(x_{c,K}\right) + \sum_{\alpha=1}^{d} \left(\partial_\alpha \phi^{\mathbb{P}_1}_i\right)\Big|_K a^\varepsilon_K\left(\Lambda^{\varepsilon,\alpha}_K, 1\right) \phi^{\mathbb{P}_1}_j\left(x_{c,K}\right) \right.$$
$$+ \sum_{\beta=1}^{d} \phi^{\mathbb{P}_1}_i\left(x_{c,K}\right) a^\varepsilon_K\left(\Lambda^{\varepsilon,0}_K, x^\beta - x^\beta_{c,K}\right) \left(\partial_\beta \phi^{\mathbb{P}_1}_j\right)\Big|_K$$
$$\left. + \sum_{\alpha,\beta=1}^{d} \left(\partial_\alpha \phi^{\mathbb{P}_1}_i\right)\Big|_K a^\varepsilon_K\left(\Lambda^{\varepsilon,\alpha}_K, x^\beta - x^\beta_{c,K}\right) \left(\partial_\beta \phi^{\mathbb{P}_1}_j\right)\Big|_K \right),$$

and therefore,

$$\mathbb{A}^\varepsilon_{j,i} =$$
$$\sum_{K \in \mathscr{T}_H} |K| \left(\overline{M} \phi^{\mathbb{P}_1}_i \phi^{\mathbb{P}_1}_j\right)(x_{c,K}) + \int_K \phi^{\mathbb{P}_1}_j \overline{B}^1 \cdot \nabla \phi^{\mathbb{P}_1}_i + \phi^{\mathbb{P}_1}_i \overline{B}^2 \cdot \nabla \phi^{\mathbb{P}_1}_j + \nabla \phi^{\mathbb{P}_1}_j \cdot \overline{A} \nabla \phi^{\mathbb{P}_1}_i, \tag{45}$$

where we have defined the effective mass $\overline{M}$, (adjoint) advection vector $\overline{B}^1$ and $\overline{B}^2$, and the effective diffusion tensor $\overline{A}$, for all $1 \le \alpha, \beta \le d$ and for each $K \in \mathscr{T}_H$, as

$$\overline{M}\big|_K = \frac{1}{|K|} a^\varepsilon_K\left(\Lambda^{\varepsilon,0}_K, 1\right), \qquad \overline{B}^1_\alpha\big|_K = \frac{1}{|K|} a^\varepsilon_K(\Lambda^{\varepsilon,\alpha}_K, 1),$$
$$\overline{B}^2_\beta\big|_K = \frac{1}{|K|} a^\varepsilon_K\left(\Lambda^{\varepsilon,0}_K, x^\beta - x^\beta_{c,K}\right), \quad \overline{A}_{\beta,\alpha}\big|_K = \frac{1}{|K|} a^\varepsilon_K\left(\Lambda^{\varepsilon,\alpha}_K, x^\beta - x^\beta_{c,K}\right). \tag{46}$$

Note that $\overline{M}$, $\overline{B}^1$, $\overline{B}^2$ and $\overline{A}$ are all piecewise constant quantities. All integrals in (45) can be computed exactly by evaluating the integrand at the centroid. With this quadrature rule, we observe that the term $|K|(\overline{M}\phi^{\mathbb{P}_1}_i \phi^{\mathbb{P}_1}_j)(x_{c,K})$ also equals the numerical approximation of the integral $\int_K \overline{M}\phi^{\mathbb{P}_1}_i \phi^{\mathbb{P}_1}_j$.

The new expression (45) for the matrix of the linear system motivates us to introduce the effective bilinear forms $\overline{a}_K$ defined on $H^1(K) \times H^1(K)$ by

$$\overline{a}_K(u,v) = \int_K \nabla v \cdot \overline{A} \nabla u + v \left(\overline{B}^1 \cdot \nabla u \right) + u \left(\overline{B}^2 \cdot \nabla v \right) + \overline{M} \, u \, v, \quad \text{for all } u, v \in H^1(K), \qquad (47)$$

and the associated $\mathbb{P}_1$ Galerkin approximation on the space $V_{H,0}$:

$$\text{Find } u_H \in V_{H,0} \text{ such that } \sum_{K \in \mathcal{T}_H} \overline{a}_K(u_H, v_H) = F(v_H) \quad \text{for all } v_H \in V_{H,0}. \qquad (48)$$

This discrete problem leads to a linear system with the matrix

$$\mathbb{A}_{j,i}^{\mathbb{P}_1} = \overline{a} \left(\phi_i^{\mathbb{P}_1}, \phi_j^{\mathbb{P}_1} \right) = \sum_{K \in \mathcal{T}_H} \overline{a}_K \left(\phi_i^{\mathbb{P}_1}, \phi_j^{\mathbb{P}_1} \right), \quad 1 \le i, j \le N_0.$$

The identity (45) thus implies the following result, which generalizes Lemma 2 to the PG-MsFEM in the general framework.

Lemma 36. *The matrices $\mathbb{A}^\varepsilon$ and $\mathbb{A}^{\mathbb{P}_1}$ are identical if the integrals in* (47) *are evaluated at the centroid of each mesh element K for the computation of $\mathbb{A}^{\mathbb{P}_1}$. Then the PG-MsFEM* (41) *coincides with the resolution of the effective problem* (48) *combined with the post-processing step*

$$u_H^\varepsilon \big|_K = u_H(x_{c,K}) \Lambda_K^{\varepsilon,0} + \sum_{\alpha=1}^d \partial_\alpha u_H \big|_K \Lambda_K^{\varepsilon,\alpha}. \qquad (49)$$

Note that Step 2 of the summary at the beginning of this section is irrelevant for the PG-MsFEM. The computation of the right-hand side in (41) is clearly part of any standard FEM software. We refer to Remarks 3 and 4 for some additional comments on the post-processing step.

The computational approach described by Lemma 36 naturally fits within the non-intrusive workflow of Algorithm 2. The numerical correctors on line (4) are, of course, replaced by those of Definition 17 or Definition 22. Line (6) is replaced by the computation of all effective quantities in (46), where $\Lambda_K^{\varepsilon,\alpha}$ is related to the numerical correctors by (44). The online phase in line (8) amounts to solving the $\mathbb{P}_1$ problem (48), where all integrations to construct the matrix of the linear system are to be performed by evaluation at the centroid. (This is not the case for the construction of the right-hand side, however.) Finally, in the post-processing phase, we construct u_H^ε from u_H by virtue of (49).

Next we generalize the above expansions to design a non-intrusive approximation of the G-MsFEM.

5.2. *The non-intrusive Galerkin MsFEM*

For the G-MsFEM (introduced in Definition 33), we need to replace the $\mathbb{P}_1$ test space $V_{H,0}$ of the PG-MsFEM by the multiscale test space $V_{H,0}^\varepsilon$. The matrix of the linear system associated to (40) is given by

$$\mathbb{A}_{j,i}^{\varepsilon,\mathsf{G}} = \sum_{K \in \mathcal{T}_H} a_K^\varepsilon \left(\phi_i^\varepsilon, \phi_j^\varepsilon \right), \quad 1 \le i, j \le N_0.$$

Upon inserting (36) for the test function ϕ_j^ε, we find, for all $1 \le i, j \le N_0$,

$$\mathbb{A}_{j,i}^{\varepsilon,\mathsf{G}} = \mathbb{A}_{j,i}^\varepsilon + \sum_{K \in \mathcal{T}_H} \left(\phi_j^{\mathbb{P}_1}(x_{c,K}) \, a_K^\varepsilon \left(\phi_i^\varepsilon, \chi_K^{\varepsilon,0,\bullet} \right) + \sum_{\beta=1}^d \left(\partial_\beta \phi_j^{\mathbb{P}_1} \right) \Big|_K \, a_K^\varepsilon \left(\phi_i^\varepsilon, \chi_K^{\varepsilon,\beta,\bullet} \right) \right),$$

where $\mathbb{A}^\varepsilon$ is the matrix of the Petrov–Galerkin MsFEM, see (42) and (45).

An effective formulation can again be derived by inserting (43) for the ϕ_i^ε. We obtain

$$
\mathbb{A}_{j,i}^{\varepsilon,\mathsf{G}} =
$$
$$
\sum_{K\in\mathscr{T}_H} |K| \left(\overline{M}^{\mathsf{G}}\phi_i^{\mathbb{P}_1}\phi_j^{\mathbb{P}_1}\right)(x_{c,K}) + \int_K \phi_j^{\mathbb{P}_1}\,\overline{B}^{1,\mathsf{G}}\cdot\nabla\phi_i^{\mathbb{P}_1} + \phi_i^{\mathbb{P}_1}\,\overline{B}^{2,\mathsf{G}}\cdot\nabla\phi_j^{\mathbb{P}_1} + \nabla\phi_j^{\mathbb{P}_1}\cdot\overline{A}^{\mathsf{G}}\,\nabla\phi_i^{\mathbb{P}_1}, \quad (50)
$$

where the effective mass, (adjoint) advection vectors and diffusion tensor are given by (using those defined in (46))

$$
\overline{M}^{\mathsf{G}}\Big|_K = \overline{M}\Big|_K + \frac{1}{|K|}\,a_K^\varepsilon\left(\Lambda_K^{\varepsilon,0},\chi_K^{\varepsilon,0,\bullet}\right), \quad \overline{B}_\alpha^{1,\mathsf{G}}\Big|_K = \overline{B}_\alpha^1\Big|_K + \frac{1}{|K|}\,a_K^\varepsilon\left(\Lambda_K^{\varepsilon,\alpha},\chi_K^{\varepsilon,0,\bullet}\right),
$$
$$
\overline{B}_\beta^{2,\mathsf{G}}\Big|_K = \overline{B}_\beta^2\Big|_K + \frac{1}{|K|}\,a_K^\varepsilon\left(\Lambda_K^{\varepsilon,0},\chi_K^{\varepsilon,\beta,\bullet}\right), \quad \overline{A}_{\beta,\alpha}^{\mathsf{G}}\Big|_K = \overline{A}_{\beta,\alpha}\Big|_K + \frac{1}{|K|}\,a_K^\varepsilon\left(\Lambda_K^{\varepsilon,\alpha},\chi_K^{\varepsilon,\beta,\bullet}\right).
$$
$$(51)$$

Again, these quantities are all piecewise constant.

The above computations lead to the introduction of the effective bilinear form $\overline{a}^{\mathsf{G}} = \sum\limits_{K\in\mathscr{T}_H} \overline{a}_K^{\mathsf{G}}$ with

$$
\overline{a}_K^{\mathsf{G}}(u,v) = \int_K \nabla v\cdot\overline{A}^{\mathsf{G}}\,\nabla u + v\left(\overline{B}^{1,\mathsf{G}}\cdot\nabla u\right) + u\left(\overline{B}^{2,\mathsf{G}}\cdot\nabla v\right) + \overline{M}^{\mathsf{G}}\,u\,v. \tag{52}
$$

We formulate the following effective variational problem:

$$
\text{Find } u_H \in V_{H,0} \text{ such that } \sum_{K\in\mathscr{T}_H} \overline{a}_K^{\mathsf{G}}(u_H,v_H) = F(v_H) \quad \text{for all } v_H \in V_{H,0}. \tag{53}
$$

The associated linear system has coefficients $\mathbb{A}_{j,i}^{\mathbb{P}_1,\mathsf{G}} = \overline{a}^{\mathsf{G}}(\phi_i^{\mathbb{P}_1},\phi_j^{\mathbb{P}_1})$. We have the following analogue of Lemma 36, which generalizes Lemma 2 to the G-MsFEM in the general framework.

Lemma 37. *The matrices $\mathbb{A}^{\varepsilon,\mathsf{G}}$ and $\mathbb{A}^{\mathbb{P}_1,\mathsf{G}}$ are identical if the integrals in* (52) *are evaluated at the centroid of each mesh element K in the computation of $\mathbb{A}^{\mathbb{P}_1,\mathsf{G}}$.*

Contrary to the matrices, the right-hand sides of the effective problem (53) and the Galerkin MsFEM (40) are not equal in general. We apply step 2 formulated at the beginning of this section: the right-hand side of the G-MsFEM is approximated by the right-hand side of the effective problem to obtain an approximate, but non-intrusive, MsFEM. The non-intrusive G-MsFEM becomes:

$$
\text{Find } u_H^\varepsilon \in V_{H,0}^\varepsilon \text{ such that } \sum_{K\in\mathscr{T}_H} a_K^\varepsilon\left(u_H^\varepsilon,\phi_j^\varepsilon\right) = F\left(\phi_j^{\mathbb{P}_1}\right) \quad \text{for all } 1 \le j \le N_0. \tag{54}
$$

This problem is no longer a Galerkin approximation of (1), because different test spaces are used for the bilinear and for the linear form. In view of Lemma 37, the non-intrusive MsFEM can equivalently be formulated as

compute $u_H \in V_{H,0}$ solution to (53) and compute u_H^ε from u_H by (49),

provided all integrals in (52) are evaluated at the centroid for the construction of the matrix of the linear system in (53).

The latter formulation of the non-intrusive MsFEM immediately suggests how to effectively implement the non-intrusive MsFEM in a non-intrusive way similar to Algorithm 2. For completeness, we provide the algorithm for the non-intrusive G-MsFEM in Algorithm 3.

The discussion surrounding Algorithm 2 regarding the advantages for the implementation of this non-intrusive MsFEM approach also applies here.

Let us now comment on the well-posedness of the MsFEMs for the general framework introduced above. We recall that the hypotheses of the general framework without oversampling, or with *DOF*-continuous oversampling, provide well-posedness of the G-MsFEM (40) by Lemma 34.

Algorithm 3 Non-intrusive G-MsFEM for the general framework

1: Let $\mathcal{T}_H$ be the mesh used by the legacy code, let $\bullet \in \{e, c\}$ be the chosen oversampling variant

2: **for all** $K \in \mathcal{T}_H$ **do**

3: **for** $0 \le \alpha \le d$ **do**

4:    Solve for the applicable $\chi_K^{\varepsilon, \alpha, \bullet}$ from Definition 17 or 22

5: **end for**

6: Compute the effective tensors defined in (51)

7: **end for**

8: Use the legacy code to construct the matrix $\mathbb{A}^{\mathbb{P}_1}$ by evaluating (52) at the centroid of each mesh element and to solve for u_H defined by (53)

9: Save $\left\{u_H(x_{c,K})\right\}_{K \in \mathcal{T}_H}$ and $\{(\partial_\alpha u_H)|_K\}_{K \in \mathcal{T}_H, 1 \le \alpha \le d}$

10: Obtain the MsFEM approximation u_H^ε from (49)

In this case, the non-intrusive approximation (54) is also well-posed, because the matrices associated to both MsFEM variants are the same. Regarding the PG-MsFEM (41), we can only establish well-posedness if the associated matrix coincides with the matrix of the corresponding Galerkin MsFEM. This is stated in the following lemma, which generalizes Lemma 5 to the general framework.

Lemma 38. *Consider a G-MsFEM as defined by Definition 33 without oversampling and suppose that the sampling form s_K^ε equals the local bilinear form a_K^ε. Then the matrix associated to this G-MsFEM coincides with the matrix associated to the corresponding PG-MsFEM of Definition 35. Consequently, the non-intrusive Galerkin MsFEM (54) coincides with the Petrov–Galerkin MsFEM (41) and in particular, the Petrov–Galerkin MsFEM is well-posed.*

Proof. To prove the lemma, we show that the matrices corresponding to the linear problems defined in (54) and (41) are equal. Using that $s_K^\varepsilon = a_K^\varepsilon$, we have for all $1 \le i, j \le N_0$,

$$\mathbb{A}_{j,i}^{\varepsilon, \mathsf{G}} - \mathbb{A}_{j,i}^\varepsilon = \sum_{K \in \mathcal{T}_H} a_K^\varepsilon \left(\phi_i^\varepsilon, \phi_j^\varepsilon - \phi_j^{\mathbb{P}_1}\right) = \sum_{K \in \mathcal{T}_H} s_K^\varepsilon \left(\phi_i^\varepsilon, \phi_j^\varepsilon - \phi_j^{\mathbb{P}_1}\right) = 0. \tag{55}$$

The last equality stems from the fact that the multiscale basis functions satisfy $\Gamma(K, \phi_i^\varepsilon) = \Gamma(K, \phi_i^{\mathbb{P}_1})$ for all K, so that $\phi_j^\varepsilon - \phi_j^{\mathbb{P}_1} \in V_{K,0}$ (see (38) with $S_K = K$ and recall Definition 16 for the sampling test space $V_{K,0}^\varepsilon$), and the variational problem in (38) (with $S_K = K$) shows that the above quantity vanishes. $\square$

5.3. *Further extensions of the non-intrusive MsFEM*

We sketch some other FEM settings to which we have applied the above strategy to develop non-intrusive MsFEM approaches. For more details, we refer to [22].

Stabilized finite element formulations. In the context of advection-diffusion problems, stabilized finite element formulations add mesh-dependent terms to a discrete variational formulation (such as (40)) to remove numerical instabilities, for example caused by sharp boundary layers of the exact solution. See [10] for such a variant of the MsFEM and see [15, 42, 43] for the stabilization of single-scale problems. The expansion (43) can also be inserted in these additional terms to find a non-intrusive implementation of the associated MsFEM.

Petrov–Galerkin formulations. Other test spaces than the $\mathbb{P}_1$ space $V_{H,0}$ can be considered in Petrov-Galerkin formulations. An example would be to use multiscale test functions that locally solve the adjoint problem rather than the direct problem, introducing yet another bilinear form

than s_K^ε in (38) or (39). See e.g. [44]. An expansion of the kind (43) can still be found for such test functions, with a suitably adapted definition of the numerical correctors. This way, a non-intrusive formulation can be found using the techniques of this work.

Non-homogeneous Dirichlet conditions. Suppose that a legacy FEM code can provide a solution to an effective problem such as (48) posed on the space $V_{H,0}$ and complemented with non-homogeneous Dirichlet conditions for u_H on $\partial\Omega$. This solution can directly be used to construct a multiscale approximation $u_H^\varepsilon \in V_H^\varepsilon$ from (49). The translation of the Dirichlet condition to the MsFEM approximation is as follows: if *DOF*-continuous oversampling is applied, the function u_H^ε satisfies $[\Gamma(K, u_H^\varepsilon)]_j = [\Gamma(K, u_H)]_j$ for all degrees of freedom associated to the boundary. Here, $[\Gamma(K, u_H)]_j$ is determined by the legacy code. When *DOF*-extended oversampling is used, the degrees of freedom associated to the boundary are equal to the sum of $[\Gamma(K, u_H)]_j$ and a perturbation due to the fact that the degrees of freedom of the numerical correctors do not vanish.

Neumann conditions. To apply Neumann conditions on $\partial\Omega$, one solves a Galerkin approximation of the variational formulation in the space V_H^ε. The suitable adaptation of (40) can be approximated by a non-intrusive Galerkin MsFEM following the same methodology as above. The effective $\mathbb{P}_1$ approximation that is obtained corresponds to the resolution of an effective PDE with Neumann conditions, for which a legacy code can be used. In the case of the diffusion problem (1), the Neumann boundary condition in the effective problem is imposed on the effective flux $\vec{n} \cdot \overline{A} \nabla u_H$, where $\overline{A}$ is defined in (12).

Parabolic equations. When a parabolic equation is discretized in time, problems of the form (27) are typically obtained for each time step, but with a right-hand side that depends on the solution of the previous time step. This term belongs to the space V_H^ε, so it varies on the microscale and cannot be integrated numerically by the legacy code that operates on the coarse mesh. The non-intrusive strategy of the foregoing sections cannot be applied directly to find a non-intrusive MsFEM. In the vein of our non-intrusive approach, one could introduce an additional approximation by replacing the multiscale solution of the previous time step by its underlying $\mathbb{P}_1$ representation in the $\mathbb{P}_1$ space V_H. Studying the effect of this approximation is beyond the scope of the present work.

5.4. *Intrusiveness of other multiscale methods*

Some work on the formulation of effective $\mathbb{P}_1$ problems in multiscale methods, and the related question of non-intrusive approaches, can be found in the literature. We discuss here the case of the HMM and the LOD method in the context of numerical homogenization, and provide some additional references to other fields at the end of the section.

First, the HMM is less intrusive than the original MsFEM, because its main objective is to approximate u^ε on the coarse scale. The HMM directly proposes to solve a $\mathbb{P}_1$ problem on the coarse scale, where effective coefficients of the $\mathbb{P}_1$ problem are defined in terms of the solutions to local problems. This workflow corresponds to our non-intrusive MsFEM approach, and when the local problems of the HMM coincide with the computation of the numerical correctors introduced in this work, the HMM and the MsFEM for the pure diffusion problem are identical. For more general problems, there is an important difference between the two methods. In the MsFEM, the form of the effective equation and the definition of the effective coefficients follow directly from the choice of basis functions, and thus from the choice of local problems. For the HMM, the local problems and the effective equation are formulated independently, and the link between the two is only justified heuristically, drawing inspiration from homogenization theory.

The LOD method aims at approximating u^ε at both the coarse and the microscale by the use of multiscale basis functions, like the MsFEM. It is shown in [45] that a Petrov-Galerkin LOD method (see also [46]) can, with some additional approximations, be recast as the $\mathbb{P}_1$ discretization of an

appropriate coarse-scale problem. This opens the way to non-intrusive implementations in the spirit of the present article. The LOD method and the MsFEM notably differ in the fact that the LOD basis functions are defined on a patch around the vertices of the mesh that should generally be taken larger than the support of the associated $\mathbb{P}_1$ functions. In contrast, the MsFEM uses fully localized basis functions (even though they may have been computed using oversampling patches), each of which has the same support as the corresponding $\mathbb{P}_1$ basis functions.

The question of non-intrusive implementations of multiscale algorithms is an interesting and relevant question in many more fields of scientific computing than we can discuss here. Beyond the field of finite element methods, we mention the multiscale finite volume method [47, 48], in which a non-intrusive coupling between the local and global computations is natural, since the local computations lead to transmissibilities that can be used in a separate, global finite volume simulator. Other than numerical homogenization methods, there are mixed finite element methods for multiscale modelling [49–51] and domain decomposition techniques (such as the generalized FEM, patches of finite elements, numerical zoom; see [52–54]), for which non-intrusive approaches can e.g. be found in [55, 56]. Finally, we would like to mention the reduced basis method for the efficient resolution of parameterized PDEs. Non-intrusive adaptations of this method (both for finite element and finite volume schemes) have been proposed and analyzed e.g. in [57–59].

6. Comparison of the classical and non-intrusive MsFEM for diffusion problems

We study in this section a particular setting within the general MsFEM framework, namely that of MsFEMs for diffusion problems. We set in this section $a_K^\varepsilon = a_K^{\varepsilon,\mathrm{diff}}$ defined in Example 8, and we choose the sampling form $s_K^\varepsilon = a_K^{\varepsilon,\mathrm{diff}}$.

6.1. *The general framework for diffusion problems*

For the convenience of the reader, we first give an explicit description of the simplifications of the general framework in the diffusion setting. In Definitions 17 and 22 for the numerical correctors, Equation (31) reduces to

$$a_K^{\varepsilon,\mathrm{diff}}\left(\chi_{S_K}^{\varepsilon,\alpha,\bullet}, w\right) = -a_K^{\varepsilon,\mathrm{diff}}\left(x^\alpha, w\right), \tag{56}$$

for all $w \in V_{K,0}$ (where $V_{K,0}$ is the sampling test space for either the MsFEM-lin or the MsFEM-CR; see Examples 18 and 19) when $1 \le \alpha \le d$, whereas $\chi_{S_K}^{\varepsilon,0,\bullet} = 0$. (The notation $\chi_K^{\varepsilon,\alpha}$ will be used in the absence of oversampling, see Remark 23.) This means that $\Lambda_K^{\varepsilon,0} = 1$ in (43). Consequently, regarding the formulation of the effective $\mathbb{P}_1$ problem, only the effective diffusion coefficient does not vanish in (46) and (51). Its definition in (51) is identical to the formula in (12) for the applicable choice of the numerical correctors.

The definition of the multiscale basis functions by (36) reduces to (10) (again upon replacing the numerical correctors $\chi_K^{\varepsilon,\alpha}$ by the relevant ones for the MsFEM under consideration). Hence, we can associate a multiscale counterpart in V_H^ε to any $v_H \in V_H$, given by

$$v_H^\varepsilon = v_H + \sum_{K \in \mathscr{T}_H} \sum_{\alpha=1}^d (\partial_\alpha v_H)|_K \, \chi_K^{\varepsilon,\alpha,\bullet}. \tag{57}$$

The non-intrusive MsFEM (54) becomes

$$\text{Find } u_H^\varepsilon \in V_{H,0}^\varepsilon \text{ such that } \sum_{K \in \mathscr{T}_H} a_K^{\varepsilon,\mathrm{diff}}\left(u_H^\varepsilon, \phi_j^\varepsilon\right) = F\left(\phi_j^{\mathbb{P}_1}\right) \quad \text{for all } 1 \le j \le N_0. \tag{58}$$

Lemma 38 now amounts to the following.

Lemma 39. *Let "MsFEM" refer to the MsFEM-lin or the MsFEM-CR, both without oversampling. The non-intrusive Galerkin MsFEM* (58) *coincides with the following Petrov-Galerkin MsFEM:*

$$Find\ u_H^\varepsilon \in V_{H,0}^\varepsilon\ such\ that\ \sum_{K \in \mathcal{T}_H} a_K^{\varepsilon,\mathrm{diff}}\left(u_H^\varepsilon, \phi_j^{\mathbb{P}_1}\right) = F\left(\phi_j^{\mathbb{P}_1}\right) \quad for\ all\ 1 \le j \le N_0, \tag{59}$$

We will specify for all results in this section to which specific MsFEMs they apply among the MsFEM-lin and the MsFEM-CR, with or without oversampling. Lemmas 40, 42 and 48 are generalizations of results in [12], where the MsFEM-lin without oversampling is considered.

6.2. *Convergence results*

We estimate here the difference between the solutions to the (intrusive) Galerkin approximation (40) and the non-intrusive MsFEM (58), which coincides with the Petrov-Galerkin Ms-FEM (59). We first show coercivity of the effective diffusion tensor $\overline{A}$.

Lemma 40. *Consider the MsFEM-lin or the MsFEM-CR, without oversampling, or the MsFEM-CR with DOF-continuous oversampling. The effective tensor $\overline{A}$ defined by* (12) *with the appropriate numerical correctors satisfies*

$$\forall\ \xi \in \mathbb{R}^d, \quad m|\xi|^2 \le \xi \cdot \overline{A}\xi.$$

Here, m is the same coercivity constant as in (2).

Proof. Let $\xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$, and let K be any simplex of the mesh $\mathcal{T}_H$. We have

$$|K| \xi \cdot \overline{A}\Big|_K \xi = \sum_{\alpha,\beta=1}^d a_K^{\varepsilon,\mathrm{diff}}\left(\xi_\alpha\left(x^\alpha + \chi_K^{\varepsilon,\alpha}\right), \xi_\beta\left(x^\beta + \chi_K^{\varepsilon,\beta}\right)\right) = \int_K \left(\xi + \nabla\chi^\xi\right) \cdot A^\varepsilon\left(\xi + \nabla\chi^\xi\right),$$

denoting by χ^ξ the function $\chi^\xi = \sum_{\alpha=1}^d \xi_\alpha \chi_K^{\varepsilon,\alpha}$. Using (2), we obtain

$$|K| \xi \cdot \overline{A}\Big|_K \xi \ge m \int_K \left|\xi + \nabla\chi^\xi\right|^2 \ge m|K||\xi|^2 + 2m \int_K \xi \cdot \nabla\chi^\xi.$$

Using an integration by parts, we see that $\int_K \xi \cdot \nabla\chi^\xi = \int_{\partial K} \chi^\xi n \cdot \xi$, where n is the unit outward normal vector on ∂K. In the case of the MsFEM-lin, the function χ^ξ vanishes on ∂K. In the case of the MsFEM-CR with *DOF*-continuous oversampling, or without oversampling, the function χ^ξ has average zero on each face of K. Since the factor $n \cdot \xi$ is constant on each face, the integral again vanishes. In conclusion, we have $\int_K \xi \cdot \nabla\chi^\xi = 0$.

We thus obtain the inequality $\xi \cdot \overline{A}|_K\xi \ge m|\xi|^2$. Since $K \in \mathcal{T}_H$ is arbitrary here, this shows coercivity of $\overline{A}$ and completes the proof. $\qquad \square$

Coercivity of the effective tensor $\overline{A}$ implies coercivity of the bilinear form $\overline{a}^{\mathrm{diff}}$ on $H_0^1(\Omega)$. By an application of the Lax-Milgram Theorem, we conclude that the (continuous) effective problem (14) is well-posed for the MsFEM-lin and the MsFEM-CR without oversampling, and for the MsFEM-CR with *DOF*-continuous oversampling.

Remark 41. The proof of the above lemma does not extend to the MsFEM-lin with oversampling, because there is no global information about $\chi_K^{\varepsilon,\alpha}$ on the faces of K.

The following lemma provides a variational characterization of the bijection (57).

Lemma 42. *Consider the MsFEM-lin or the MsFEM-CR, both without oversampling. Let $v_H^\varepsilon \in V_H^\varepsilon$. The unique $v_H \in V_H$ for which* (57) *holds, is the unique solution in V_H to the problem*

$$\forall\ w_H \in V_H, \quad \overline{a}^{\mathrm{diff}}(v_H, w_H) = a^{\varepsilon,\mathrm{diff}}\left(v_H^\varepsilon, w_H\right). \tag{60}$$

In addition, we have, with the constants m and M from (2), *the estimate*

$$\|\nabla v_H\|_{L^2(\mathcal{T}_H)} \le \frac{M}{m} \left\|\nabla v_H^\varepsilon\right\|_{L^2(\mathcal{T}_H)}.$$

Proof. Let $v_H \in V_H$ be the unique element of V_H such that v_H^ε and v_H satisfy (57). Take any $w_H \in V_H$. Using that ∇v_H and ∇w_H are piecewise constant, we compute

$$a^{\varepsilon,\text{diff}}\left(v_H^\varepsilon, w_H\right) = \sum_{K \in \mathcal{T}_H} \sum_{\alpha,\beta=1}^d \left(\partial_\beta w_H\right)\big|_K a_K^{\varepsilon,\text{diff}}\left(x^\alpha + \chi_K^{\varepsilon,\alpha}, x^\beta\right)(\partial_\alpha v_H)\big|_K.$$

For the MsFEM without oversampling, the numerical correctors belong to the sampling test space $V_{K,0}$. We can thus use (56) to obtain

$$\forall\, 1 \le \alpha, \beta \le d, \quad a_K^{\varepsilon,\text{diff}}\left(x^\alpha + \chi_K^{\varepsilon,\alpha}, x^\beta\right) = a_K^{\varepsilon,\text{diff}}\left(x^\alpha + \chi_K^{\varepsilon,\alpha}, x^\beta + \chi_K^{\varepsilon,\beta}\right).$$

Using the definitions of $\overline{A}$ in (12) and of $\overline{a}^{\text{diff}}$ in (16) (we recall that these expressions hold true here upon replacing the numerical correctors by those under consideration), we conclude that

$$a^{\varepsilon,\text{diff}}\left(v_H^\varepsilon, w_H\right) = \sum_{K \in \mathcal{T}_H} \sum_{\alpha,\beta=1}^d \int_K \partial_\beta w_H\, \overline{A}_{\beta,\alpha}\, \partial_\alpha v_H = \overline{a}^{\text{diff}}(v_H, w_H).$$

It follows that v_H satisfies (60). In addition, in view of the coercivity of $\overline{A}$ established in Lemma 40 and by the Lax-Milgram Theorem, problem (60) uniquely characterizes v_H.

The estimate on v_H follows by testing the characterization (60) against $w_H = v_H$. This yields

$$m\|\nabla v_H\|_{L^2(\mathcal{T}_H)}^2 \le \overline{a}^{\text{diff}}(v_H, v_H) = a^{\varepsilon,\text{diff}}\left(v_H^\varepsilon, v_H\right) = \sum_{K \in \mathcal{T}_H} \int_K \nabla v_H \cdot A^\varepsilon \nabla v_H^\varepsilon$$
$$\le M \sum_{K \in \mathcal{T}_H} \|\nabla v_H\|_{L^2(K)} \left\|\nabla v_H^\varepsilon\right\|_{L^2(K)}.$$

The first inequality follows from coercivity of $\overline{A}$ and the second inequality from the upper bound on A^ε in (2) and the Cauchy-Schwarz inequality. With a discrete Cauchy-Schwarz inequality, we obtain

$$m\|\nabla v_H\|_{L^2(\mathcal{T}_H)}^2 \le M \sum_{K \in \mathcal{T}_H} \|\nabla v_H\|_{L^2(K)} \left\|\nabla v_H^\varepsilon\right\|_{L^2(K)} \le M \|\nabla v_H\|_{L^2(\mathcal{T}_H)} \left\|\nabla v_H^\varepsilon\right\|_{L^2(\mathcal{T}_H)}.$$

The proof of Lemma 42 is completed upon simplifying by $\|\nabla v_H\|_{L^2(\mathcal{T}_H)}$. $\qquad\square$

For the remainder of this section, we consider MsFEMs without oversampling. Let $u_H^{\varepsilon,\mathsf{G}}$ denote the solution to the MsFEM approximation (40) (we use the superscript G to stress that this is a *Galerkin* approximation) and let $u_H^{\varepsilon,\mathsf{PG}}$ denote the solution to the non-intrusive MsFEM (58) (which is equivalent to the *Petrov-Galerkin* MsFEM (59), since we do not apply the oversampling technique).

We first study the error $u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$ when $\varepsilon \to 0$. In this case, we do not need a rate of convergence in H and we shall relax the condition $f \in L^2(\Omega)$ to the condition $f \in H^{-1}(\Omega)$. Then the definition of the linear form F in (4) has to be adapted. Given $f \in H^{-1}(\Omega)$, there exist $f_0, f_1, \ldots, f_d \in L^2(\Omega)$ such that

$$F(v) = \sum_{K \in \mathcal{T}_H} \left(\int_K f_0\, v + \sum_{\beta=1}^d \int_K f_\beta\, \partial_\beta v \right),$$

which is in fact well-defined for any $v \in H^1(\mathcal{T}_H)$ and thus in particular on V_H, the underlying affine space for the MsFEM, and the multiscale space V_H^ε.

We consider in Lemma 43 a sequence of diffusion tensors A^ε that H-converges to a constant diffusion tensor. This means that u^ε converges weakly in $H^1(\Omega)$ as $\varepsilon \to 0$ towards a function $u^\star \in H_0^1(\Omega)$, solution to the homogenized problem (25), and $A^\varepsilon \nabla u^\varepsilon \rightharpoonup A^\star \nabla u^\star$ weakly in $L^2(\Omega)$.

Lemma 43. *Consider the MsFEM-lin or the MsFEM-CR, both without oversampling. Suppose that $(A^\varepsilon)_{\varepsilon>0}$ is a sequence of matrices satisfying (2) that H-converges to a constant matrix. Let $f \in H^{-1}(\Omega)$. Then*

$$\left\| u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}} \right\|_{H^1(\mathcal{T}_H)} \to 0 \text{ as } \varepsilon \to 0.$$

Remark 44. A rate of convergence can be obtained under some additional structural assumptions on A^ε; see Lemma 50.

We need a few auxiliary results to establish Lemma 43. The first result below concerns the convergence of the numerical correctors as $\varepsilon \to 0$.

Lemma 45. *Suppose that A^ε H-converges to a constant homogenized matrix $A^\star$. Consider the MsFEM-lin or the MsFEM-CR, both without oversampling. Then, for all $K \in \mathcal{T}_H$ and all $1 \le \alpha \le d$, we have $\chi_K^{\varepsilon,\alpha} \rightharpoonup 0$ weakly in $H^1(K)$ as $\varepsilon \to 0$.*

Proof. We introduce for each $\alpha = 1, \ldots, d$ the function $\tau^{\varepsilon,\alpha} = x^\alpha + \chi_K^{\varepsilon,\alpha}$. Then (56) implies the equation $-\operatorname{div}(A^\varepsilon \nabla \tau^{\varepsilon,\alpha}) = 0$ in K. For the MsFEM-lin, the boundary conditions of the local problems for $\chi_K^{\varepsilon,\alpha}$ (see (32) with $S_K = K$) lead to $\tau^{\varepsilon,\alpha} = x^\alpha$ on ∂K. The boundary conditions associated to the MsFEM-CR follow from (33) and are as follows: the flux $\vec{n} \cdot A^\varepsilon \nabla \tau^{\varepsilon,\alpha}$ is constant on each face of K (but may depend on ε) and $\int_h \tau^{\varepsilon,\alpha} = \int_h x^\alpha$ for all faces h of K.

It follows that the homogenized limit $\tau^{\star,\alpha}$ of $\tau^{\varepsilon,\alpha}$ satisfies the equation $-\operatorname{div}(A^\star \nabla \tau^{\star,\alpha}) = 0$ in K. For the MsFEM-lin, the boundary condition for the homogenized problem is $\tau^{\star,\alpha} = x^\alpha$ on ∂K. The boundary conditions associated to the MsFEM-CR are a constant flux $\vec{n} \cdot A^\star \nabla \tau^{\alpha,\star}$ on each face of K and $\int_h \tau^{\star,\alpha} = \int_h x^\alpha$ for all faces h of K.

Both for the MsFEM-lin and the MsFEM-CR, the homogenized equation has a unique solution, which is easily seen to be $\tau^{\star,\alpha} = x^\alpha$, because $A^\star$ is constant. Therefore, $\tau^{\varepsilon,\alpha} \rightharpoonup x^\alpha$ weakly in $H^1(K)$. Subtracting the function x^α, we deduce the desired convergence. $\qquad\square$

We will also use the following result, which is a straightforward generalization of the extended Poincaré inequality in [17, Lemma 3.31].

Lemma 46. *Let W be the subspace of $H^1(\mathcal{T}_H)$ defined by*

$$W = \left\{ v \in H^1(\mathcal{T}_H) \,\middle|\, \int_h [\![v]\!] = 0 \text{ for each face } h \text{ of } \mathcal{T}_h, \int_h v = 0 \text{ for each face } h \subset \partial\Omega \right\}.$$

There exists a constant $C > 0$ depending only on Ω but not on H such that

$$\forall\, v \in W, \qquad \|v\|_{L^2(\Omega)} \le C \|\nabla v\|_{L^2(\mathcal{T}_H)}.$$

Note that the multiscale space $V_{H,0}^\varepsilon$ is contained in W for both the MsFEM-lin and the MsFEM-CR, without oversampling. Finally, we provide a number of useful bounds for the difference between $u_H^{\varepsilon,\mathsf{G}}$ and $u_H^{\varepsilon,\mathsf{PG}}$.

Lemma 47. *Let $f \in H^{-1}(\Omega)$ and consider the MsFEM-lin or the MsFEM-CR, both without oversampling. Let $e_H^\varepsilon = u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$. There exists a unique $e_H^{\mathbb{P}_1} \in V_H$ and a linear combination of the numerical correctors, that we denote by e_H^{osc}, such that $e_H^\varepsilon = e_H^{\mathbb{P}_1} + e_H^{\mathsf{osc}}$, and it holds, with the constants m, M from (2) and the constant C from Lemma 46,*

$$a^{\varepsilon,\mathrm{diff}}\left(e_H^\varepsilon, e_H^\varepsilon\right) = F\left(e_H^{\mathsf{osc}}\right), \tag{61}$$

$$\left\| \nabla e_H^{\mathsf{osc}} \right\|_{L^2(K)} \le \frac{M}{m} \left\| \nabla e_H^{\mathbb{P}_1} \right\|_{L^2(K)} \quad \text{for all } K \in \mathcal{T}_H, \tag{62}$$

$$\left\| \nabla e_H^{\mathbb{P}_1} \right\|_{L^2(\mathcal{T}_H)} \le \frac{M}{m} \left\| \nabla e_H^\varepsilon \right\|_{L^2(\mathcal{T}_H)}, \tag{63}$$

$$\left\| \nabla e_H^\varepsilon \right\|_{L^2(\mathcal{T}_H)} \le \sqrt{1 + C^2} \frac{M^2}{m^3} \|F\|_{\mathscr{L}(H^1(\mathcal{T}_H))}, \tag{64}$$

where $\|\cdot\|_{\mathscr{L}(H^1(\mathcal{T}_H))}$ is the operator norm on $\mathscr{L}(H^1(\mathcal{T}_H))$.

Proof. Since the numerical approximations $u_H^{\varepsilon,\mathsf{G}}$ and $u_H^{\varepsilon,\mathsf{PG}}$ both belong to the multiscale approximation space V_H^ε, it follows that $e_H^\varepsilon \in V_H^\varepsilon$, and we are in a position to use (57): there exists a unique $e_H^{\mathbb{P}_1} \in V_H$ such that

$$e_H^\varepsilon = e_H^{\mathbb{P}_1} + e_H^{\mathsf{osc}}, \quad e_H^{\mathsf{osc}} = \sum_{K \in \mathscr{T}_H} \sum_{\alpha=1}^d \left(\partial_\alpha e_H^{\mathbb{P}_1} \right)\Big|_K \chi_K^{\varepsilon,\alpha}. \tag{65}$$

Applying Lemma 42 to $v_H^\varepsilon = e_H^\varepsilon$, we immediately obtain (63).

Now recall that the numerical correctors are defined by (56). Using the fact that $\nabla e_H^{\mathbb{P}_1}$ is piecewise constant, this implies that e_H^{osc} satisfies the following variational problem in each $K \in \mathscr{T}_H$:

$$\forall\ w \in V_{K,0}, \quad a_K^{\varepsilon,\mathsf{diff}} \left(e_H^{\mathsf{osc}}, w \right) = -a_K^{\varepsilon,\mathsf{diff}} \left(e_H^{\mathbb{P}_1}, w \right).$$

Without oversampling, it holds $\chi_K^{\varepsilon,\alpha} \in V_{K,0}$ for each $1 \le \alpha \le d$, so e_H^{osc} can be used as a test function here. With the bounds in (2), implying continuity and coercivity of $a_K^{\varepsilon,\mathsf{diff}}$, we obtain (62).

Next using (65), we can write

$$a^{\varepsilon,\mathsf{diff}} \left(e_H^\varepsilon, e_H^\varepsilon \right) = a^{\varepsilon,\mathsf{diff}} \left(u_H^{\varepsilon,\mathsf{G}}, e_H^\varepsilon \right) - a^{\varepsilon,\mathsf{diff}} \left(u_H^{\varepsilon,\mathsf{PG}}, e_H^{\mathbb{P}_1} \right) - a^{\varepsilon,\mathsf{diff}} \left(u_H^{\varepsilon,\mathsf{PG}}, e_H^{\mathsf{osc}} \right).$$

We deduce from (55) that $a^{\varepsilon,\mathsf{diff}}(u_H^{\varepsilon,\mathsf{PG}}, e_H^{\mathsf{osc}}) = 0$. Since e_H^ε can be used as a test function in the discrete problem (40) and $e_H^{\mathbb{P}_1}$ in (59), we have

$$a^{\varepsilon,\mathsf{diff}} \left(u_H^{\varepsilon,\mathsf{G}}, e_H^\varepsilon \right) - a^{\varepsilon,\mathsf{diff}} \left(u_H^{\varepsilon,\mathsf{PG}}, e_H^{\mathbb{P}_1} \right) = F \left(e_H^{\mathsf{osc}} \right),$$

which shows (61). It follows that

$$a^{\varepsilon,\mathsf{diff}} \left(e_H^\varepsilon, e_H^\varepsilon \right) \le \|F\|_{\mathscr{L}(H^1(\mathscr{T}_H))} \left\| e_H^{\mathsf{osc}} \right\|_{H^1(\mathscr{T}_H)} \le \|F\|_{\mathscr{L}(H^1(\mathscr{T}_H))} \sqrt{1+C^2} \left\| \nabla e_H^{\mathsf{osc}} \right\|_{L^2(\mathscr{T}_H)},$$

where C is the Poincaré constant from Lemma 46. Now applying (62) and (63) on the right, and using coercivity of $a^{\varepsilon,\mathsf{diff}}$ on the left, we find

$$m \left\| \nabla e_H^\varepsilon \right\|_{L^2(\mathscr{T}_H)}^2 \le \sqrt{1+C^2} \left(\frac{M}{m} \right)^2 \|F\|_{\mathscr{L}(H^1(\mathscr{T}_H))} \left\| \nabla e_H^\varepsilon \right\|_{L^2(\mathscr{T}_H)},$$

from which we deduce (64). $\qquad\square$

Proof of Lemma 43. Let $e_H^\varepsilon = u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$. We will use (65). By Lemma 47, we have (61). Combined with (2) and Lemma 46, this implies

$$C \left\| u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}} \right\|_{H^1(\mathscr{T}_H)}^2 \le a^{\varepsilon,\mathsf{diff}} \left(e_H^\varepsilon, e_H^\varepsilon \right) = F \left(e_H^{\mathsf{osc}} \right) = \sum_{K \in \mathscr{T}_H} \sum_{\alpha=1}^d \left(\partial_\alpha e_H^{\mathbb{P}_1} \right)\Big|_K F \left(\chi_K^{\varepsilon,\alpha} \right).$$

By Lemma 45, we know that $\chi_K^{\varepsilon,\alpha} \rightharpoonup 0$ as $\varepsilon \to 0$ weakly in $H^1(K)$ for each K and for each α. Therefore, $F(\chi_K^{\varepsilon,\alpha}) \to 0$ as $\varepsilon \to 0$. In view of (63) and (64), every derivative $(\partial_\alpha e_H^{\mathbb{P}_1})|_K$ is bounded independently of ε. It follows that $F(e_H^{\mathsf{osc}}) \to 0$ as $\varepsilon \to 0$. The conclusion now follows from the above inequality. $\qquad\square$

We next study the convergence of $u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$ as $H \to 0$. To this end, we return to the original hypotheses of Section 2, i.e., $f \in L^2(\Omega)$. Note that for the next result, the additional convergence hypothesis of Lemma 43 for A^ε is not needed.

Lemma 48. *Consider the MsFEM-lin or the MsFEM-CR, both without oversampling. Assume that $f \in L^2(\Omega)$. Then there exists a constant C independent of ε, H and f such that*

$$\left\| u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}} \right\|_{H^1(\mathscr{T}_H)} \le CH\|f\|_{L^2(\Omega)}.$$

To prove this lemma, we will use some Poincaré–Friedrichs inequalities, for which we refer e.g. to [9, Lemma 4.3], [17, Lemma B.66].

Proof. Let $e_H^\varepsilon = u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$ and recall the results of Lemma 47. We have $e_H^\varepsilon = e_H^{\mathbb{P}_1} + e_H^{\mathsf{osc}}$ (see (65)), and (61) provides, for $f \in L^2(\Omega)$, the equality $a^{\varepsilon,\mathsf{diff}}(e_H^\varepsilon, e_H^\varepsilon) = (f, e_H^{\mathsf{osc}})_{L^2(\Omega)}$. Hence, by the Cauchy–Schwarz inequality,

$$a^{\varepsilon,\mathsf{diff}}(e_H^\varepsilon, e_H^\varepsilon) \leq \|f\|_{L^2(\Omega)} \left\| e_H^{\mathsf{osc}} \right\|_{L^2(\Omega)}. \tag{66}$$

For the MsFEM-lin (without oversampling), it holds that $\chi_K^{\varepsilon,\alpha} = 0$ on ∂K for all mesh elements K and all $1 \leq \alpha \leq d$, and it follows that $e_H^{\mathsf{osc}} = 0$ on the boundaries of all mesh elements. In the case of the MsFEM-CR (without oversampling), it holds that $\int_h \chi_K^{\varepsilon,\alpha} = 0$ for all faces h of the mesh and all $1 \leq \alpha \leq d$. (Note that the average of $\chi_K^{\varepsilon,\alpha}$ over any face h is well-defined even if $\chi_K^{\varepsilon,\alpha}$ is in general discontinuous along faces.) Since $\partial_\alpha e_H^{\mathbb{P}_1}$ is constant on each mesh element K, we also have $\int_h e_H^{\mathsf{osc}} = 0$. Hence, both for the MsFEM-lin and for the MsFEM-CR, an appropriate variant of the Poincaré–Friedrichs inequality yields a constant C independent of K but dependent on the regularity of the mesh, such that

$$\left\| e_H^{\mathsf{osc}} \right\|_{L^2(K)} \leq CH \left\| \nabla e_H^{\mathsf{osc}} \right\|_{L^2(K)}. \tag{67}$$

Upon inserting the inequalities (67), (62) and (63) into (66), it follows that

$$a^{\varepsilon,\mathsf{diff}}(e_H^\varepsilon, e_H^\varepsilon) \leq CH \left(\frac{M}{m} \right)^2 \left\| \nabla e_H^\varepsilon \right\|_{L^2(\mathcal{T}_H)} \|f\|_{L^2(\Omega)}.$$

One more time using the lower bound in (2), we find

$$\left\| \nabla e_H^\varepsilon \right\|_{L^2(\mathcal{T}_H)} \leq CH \frac{M^2}{m^3} \|f\|_{L^2(\Omega)}.$$

The proof is concluded by application of Lemma 46 to e_H^ε. $\qquad \square$

6.3. *Convergence results in the periodic setting*

We now study the MsFEM-lin applied to the periodic setting introduced in Section 2.7 in some more detail. To the best of our knowledge, all convergence results known for the MsFEM are obtained in this periodic setting (see e.g. [6, 9, 23, 24, 31, 32, 40, 41, 60]). The analysis in these works relies on the explicit description of the microstructure that we summarized in Section 2.7. In particular, recall the existence of a homogenized diffusion coefficient given by (24) and the first-order two-scale expansion (26). We emphasize, however, that the *application* of the MsFEM does not require the periodic setting, nor does it even suppose the PDE under consideration to be embedded in a sequence of PDEs for a family of parameters ε that tend 0. We refer to Section 7 for examples of such numerical experiments.

Applying the MsFEM to a sequence of matrices $A^\varepsilon = A^{\mathsf{per}}(\cdot/\varepsilon)$, we obtain a sequence of effective tensors $\overline{A}(\varepsilon)$. Each $\overline{A}(\varepsilon)$ is defined by (12) for a fixed value of ε. We have the following convergence result.

Lemma 49. *Let $\overline{A}(\varepsilon)$ be the sequence of effective tensors obtained in (12) by applying the MsFEM-lin without oversampling to $A^\varepsilon = A^{\mathsf{per}}(\cdot/\varepsilon)$. We have $\overline{A}(\varepsilon) \to A^\star$ as $\varepsilon \to 0$.*

Proof. We fix a mesh element $K \in \mathcal{T}_H$. First observe that $\overline{A}(\varepsilon)$ and $A^\star$ satisfy

$$\overline{A}_{\beta,\alpha}(\varepsilon)\Big|_K = \frac{1}{|K|} a_K^{\varepsilon,\mathsf{diff}}\left(x^\alpha + \chi_K^{\varepsilon,\alpha}, x^\beta \right), \qquad A_{\beta,\alpha}^\star = \int_Q e_\beta \cdot A^{\mathsf{per}}(e_\alpha + \nabla w_\alpha), \tag{68}$$

for each $1 \leq \alpha, \beta \leq d$, in view of the variational formulations satisfied by $\chi_K^{\varepsilon,\alpha}$ (solution to the PDE (11)) and w_α (solution to the PDE (23)). We recall that Q is the unit cube of $\mathbb{R}^d$.

Now let $\tau^{\varepsilon,\alpha} = x^\alpha + \chi_K^{\varepsilon,\alpha}$. In view of Lemma 45, $\tau^{\varepsilon,\alpha} \rightharpoonup \tau^{\star,\alpha}$ as $\varepsilon \to 0$ weakly in $H^1(K)$, with $\tau^{\star,\alpha}(x) = x^\alpha$. Writing the two-scale expansion (26) of $\tau^{\varepsilon,\alpha}$, we thus have, when ε is small,

$$\tau^{\varepsilon,\alpha}(x) \approx \tau^{\star,\alpha}(x) + \varepsilon \sum_{\gamma=1}^d w_\gamma\left(\frac{x}{\varepsilon} \right) \partial_\gamma \tau^{\star,\alpha}(x) = x^\alpha + \varepsilon\, w_\alpha\left(\frac{x}{\varepsilon} \right),$$

and the difference tends to zero in $H^1(K)$ as $\varepsilon \to 0$. Inserting this convergence in (68), we deduce that

$$\lim_{\varepsilon \to 0} \overline{A}_{\beta,\alpha}(\varepsilon)\Big|_K = \lim_{\varepsilon \to 0} \frac{1}{|K|} \int_K e_\beta \cdot A^{\mathsf{per}}\left(\frac{x}{\varepsilon}\right)\left(e_\alpha + \nabla w_\alpha\left(\frac{x}{\varepsilon}\right)\right) \mathrm{d}x = A^\star_{\beta,\alpha}.$$

The convergence to the mean on the unit cube in the last equality follows from the Q-periodicity of the function $e_\beta \cdot A^{\mathsf{per}}(e_\alpha + \nabla w_\alpha)$. $\qquad\square$

The following lemma studies the convergence of $u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$ towards 0 as $\varepsilon \to 0$ for the MsFEM-lin without oversampling. As was stated in Remark 44, thanks to the periodic setting, we now obtain a rate for the convergence stated in Lemma 43.

Lemma 50. *Let $f \in L^2(\Omega)$. Suppose that the family of meshes $(\mathscr{T}_H)_{H>0}$ is quasi-uniform. Consider the MsFEM-lin without oversampling. For $A^\varepsilon = A^{\mathsf{per}}(\cdot/\varepsilon)$ sufficiently regular, we have*

$$\left\|u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}\right\|_{H^1(\Omega)} \le C\varepsilon \|f\|_{L^2(\Omega)},$$

where the constant C depends on the dimension d and the constants m, M in (2), but not on ε, H or f.

Proof. Let $e_H^\varepsilon = u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$. Lemma 47 applies, so we can use (61) and a Cauchy–Schwarz inequality to find

$$a^{\varepsilon,\mathsf{diff}}\left(e_H^\varepsilon, e_H^\varepsilon\right) \le \|f\|_{L^2(\Omega)} \left\|e_H^{\mathsf{osc}}\right\|_{L^2(\Omega)} \le \|f\|_{L^2(\Omega)} \left\|\sum_{K \in \mathscr{T}_H} \sum_{\alpha=1}^d \left(\partial_\alpha e_H^{\mathbb{P}_1}\right)\Big|_K \chi_K^{\varepsilon,\alpha}\right\|_{L^2(\Omega)}. \tag{69}$$

Next we seek a bound on $\chi_K^{\varepsilon,\alpha}$ in $L^2(K)$. Using (11) and (23), we have

$$\mathrm{div}\left(A^{\mathsf{per}}\left(\frac{\cdot}{\varepsilon}\right)\nabla\left[\chi_K^{\varepsilon,\alpha} - \varepsilon w_\alpha\left(\frac{\cdot}{\varepsilon}\right)\right]\right) = 0 \quad \text{in } K.$$

Since $\chi_K^{\varepsilon,\alpha}$ vanishes on ∂K (recall that we consider the MsFEM-lin without oversampling), the maximum principle [14, Theorem 8.1] yields

$$\left\|\chi_K^{\varepsilon,\alpha} - \varepsilon w_\alpha\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(K)} \le \sup_{\partial K}\left|\chi_K^{\varepsilon,\alpha} - \varepsilon w_\alpha\left(\frac{\cdot}{\varepsilon}\right)\right| \sqrt{\int_K 1} = \varepsilon |K|^{1/2} \sup_{\partial K}\left|w_\alpha\left(\frac{\cdot}{\varepsilon}\right)\right|.$$

When A^{per} is sufficiently regular, the corrector functions w_α are uniformly bounded. Then the mesh regularity provides a constant C such that for each $K \in \mathscr{T}_H$ and each $1 \le \alpha \le d$, we have

$$\left\|\chi_K^{\varepsilon,\alpha}\right\|_{L^2(K)} \le \left\|\chi_K^{\varepsilon,\alpha} - \varepsilon w_\alpha\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(K)} + \varepsilon\left\|w_\alpha\left(\frac{\cdot}{\varepsilon}\right)\right\|_{L^2(K)} \le C\varepsilon H^{d/2}.$$

Since all $\chi_K^{\varepsilon,\alpha}$ have disjoint supports, we can use the latter estimate to bound

$$\begin{aligned}
\left\|\sum_{K \in \mathscr{T}_H} \sum_{\alpha=1}^d \left(\partial_\alpha e_H^{\mathbb{P}_1}\right)\Big|_K \chi_K^{\varepsilon,\alpha}\right\|_{L^2(\Omega)}^2 &= \sum_{K \in \mathscr{T}_H} \left\|\sum_{\alpha=1}^d \left(\partial_\alpha e_H^{\mathbb{P}_1}\right)\Big|_K \chi_K^{\varepsilon,\alpha}\right\|_{L^2(K)}^2 \\
&\le C\varepsilon^2 \sum_{K \in \mathscr{T}_H} \sum_{\alpha=1}^d \left(H^{d/2}\left(\partial_\alpha e_H^{\mathbb{P}_1}\right)\Big|_K\right)^2 \\
&\le C\varepsilon^2 \sum_{K \in \mathscr{T}_H} \sum_{\alpha=1}^d \left\|\partial_\alpha e_H^{\mathbb{P}_1}\right\|_{L^2(K)}^2 \\
&= C\varepsilon^2 \left\|\nabla e_H^{\mathbb{P}_1}\right\|_{L^2(\Omega)}^2.
\end{aligned} \tag{70}$$

The last inequality relies on the quasi-uniformity of the mesh.

We insert (70) combined with (63) into (69) to find

$$a^{\varepsilon,\mathsf{diff}}\left(e_H^\varepsilon, e_H^\varepsilon\right) \le C\varepsilon \|f\|_{L^2(\Omega)} \left\|\nabla e_H^\varepsilon\right\|_{L^2(\Omega)}.$$

Applying the coercivity property in (2) on the left-hand side and a Poincaré inequality on Ω, we obtain the desired result. $\qquad\qquad\square$

The classical error estimate for the Galerkin MsFEM approach (6) is obtained in the periodic setting and under some regularity assumption on A^{per} and on the homogenized limit $u^\star$. The bound obtained in [6, Theorem 6.5] reads

$$\left\| u^\varepsilon - u_H^{\varepsilon,\mathsf{G}} \right\|_{H^1(\Omega)} \le C\left(H + \varepsilon + \sqrt{\varepsilon/H} \right), \tag{71}$$

for some C independent of ε and H. Lemma 50 shows that the same estimate holds true for $u_H^{\varepsilon,\mathsf{PG}}$, the Petrov-Galerkin MsFEM approximation, under the correct regularity assumptions. We note that the bound for $u_H^{\varepsilon,\mathsf{PG}}$ can also be inferred from Lemma 48. However, since the MsFEM is applied in the regime where $\varepsilon < H$, the result of Lemma 50 is more precise, thanks to the extra structural assumptions made on the diffusion tensor A^ε.

7. Numerical comparison

We now compare the Galerkin MsFEM (40), its non-intrusive approximation (54) and the Petrov–Galerkin MsFEM (41) on a concrete numerical example in 2D ($d = 2$). The numerical approximations obtained for these various MsFEMs shall be denoted $u_H^{\varepsilon,\mathsf{G}}$, $u_H^{\varepsilon,\mathsf{G\text{-}ni}}$ and $u_H^{\varepsilon,\mathsf{PG}}$, respectively.

7.1. *Description of the numerical experiments*

We consider the pure diffusion equation (1) on the domain $\Omega = (0,1) \times (0,1)$. Thus, the local bilinear forms are $a_K^\varepsilon = a_K^{\varepsilon,\mathsf{diff}}$ defined in Example 8, where we consider the three diffusion tensors

$$A^{\varepsilon,\mathsf{per}}(x) = \nu^\varepsilon(x)\,\mathrm{Id}, \quad \nu^\varepsilon(x) = 1 + 100\cos^2(\pi x_1/\varepsilon)\sin^2(\pi x_2/\varepsilon), \tag{72a}$$

$$A^{\varepsilon,\mathsf{lp}}(x) = \left(1 + \cos^2(2\pi x_1)\right) A^{\varepsilon,\mathsf{per}}(x), \tag{72b}$$

$$A^{\varepsilon,\mathsf{np}}(x) = \nu^{\varepsilon,\mathsf{np}}(x)\,\mathrm{Id}, \quad \nu^{\varepsilon,\mathsf{np}}(x) = 1 + \left(1 + 100\cos^2(\pi x_1/\varepsilon)\sin^2(\pi x_2/\varepsilon)\right)\cos^2\left(\frac{x_1^2 + x_2^2}{\varepsilon}\right). \tag{72c}$$

We fix $f(x) = \sin(x_1)\sin(x_2)$.

The coefficient $A^{\varepsilon,\mathsf{per}}$ is ε-periodic with period $\varepsilon = \pi/150 \approx 0.02$. The coefficient $A^{\varepsilon,\mathsf{lp}}$ is locally periodic and, although a homogenized coefficient exists (see [18]), it is not constant. Consequently, a certain number of lemmas established in Section 6 are not known to hold true. Finally, we include the coefficient $A^{\varepsilon,\mathsf{np}}$ as an example of a multiscale problem for which we are not aware of any explicit homogenization results. We will see nevertheless that the non-intrusive MsFEMs that we introduced above provide good approximations compared to their intrusive G-MsFEM counterparts for all test cases.

A reference solution u_h^ε is computed on a uniform 1024×1024 mesh $\mathcal{T}_h$ by means of a standard $\mathbb{P}_1$ finite element method using FreeFem++ [61]. The mesh $\mathcal{T}_h$ (as well as the coarse mesh introduced below) consists of squares cut in two along a diagonal that is in the same direction for all squares, i.e., such as the meshes in Figure 1. The FreeFem++ scripts to perform all different MsFEMs can be found at [62].

We compare the reference solution u_h^ε to MsFEM solutions obtained on a coarse mesh $\mathcal{T}_H$ for varying H. The mesh $\mathcal{T}_H$ is a uniform $1/H \times 1/H$ triangulation of Ω. We test the MsFEM-lin and the MsFEM-CR using the sampling operator $s_K^\varepsilon = a_K^{\varepsilon,\mathsf{diff}}$. All oversampling methods in this section use a homothety ratio of 3 for the construction of the oversampling patches in Definition 13. A precise definition of the associated basis functions can be found in Examples 29

and 30. The mesh $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$ for all values of H. Therefore, for each $K \in \mathcal{T}_H$, we use the corresponding submesh of $\mathcal{T}_h$ (consisting of all triangles included in K) for the numerical approximation of the numerical correctors in (31) by $\mathbb{P}_1$ Lagrange finite elements.

Remark 51. We provide a few remarks on the computation of the error, which takes place in the post-processing step of the MsFEM. Evidently, these computations have to be carried out by integration on the fine scale and one may try to perform these computations on the global mesh $\mathcal{T}_h$. However, the legacy code does not, in general, operate on the global fine mesh. Moreover, we stress that the approximation u_H^ε is in general discontinuous across element edges (for the MsFEM-CR, and for all MsFEMs with oversampling), and can therefore not be represented globally by e.g. a piecewise $\mathbb{P}_1$ function on the fine mesh $\mathcal{T}_h$ (even if one supposes that $\mathcal{T}_h$ is conformal). Thus, one has to compute the error element by element, using the code for the microscale, according to the sum

$$\left\| u^\varepsilon - u_H^\varepsilon \right\|_{H^1(\mathcal{T}_H)}^2 = \sum_{K \in \mathcal{T}_H} \left\| u^\varepsilon - u_H^\varepsilon \right\|_{H^1(K)}^2.$$

To do so, Equation (49) can be used on each element K to find the correct values of u_H^ε, and the global fine mesh $\mathcal{T}_h$ is never used.

7.2. *Results*

We first compare the approximations $u_H^{\varepsilon,\mathsf{G}}$ and $u_H^{\varepsilon,\mathsf{G}\text{-ni}}$ for varying H in Figure 3 for MsFEMs without oversampling and MsFEMs with *DOF*-continuous oversampling. Without oversampling (OS), the approximation $u_H^{\varepsilon,\mathsf{G}\text{-ni}}$ equals $u_H^{\varepsilon,\mathsf{PG}}$ due to Lemma 38. We also report the error committed by the G-MsFEM. We observe that, without oversampling, the difference $u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{G}\text{-ni}}$ is much smaller than this error. As a result, the errors obtained with the G-MsFEM and its non-intrusive approximation are of the same size. Indeed, the error of the non-intrusive G-MsFEM-lin deviates from the error of the G-MsFEM-lin by at most 0.05% for all tests that we report here. For the MsFEM-CR, this is at most 1.2%. In both cases, the two MsFEM variants thus have practically the same accuracy. This is in agreement with the theoretical result of Lemma 48.

The estimates obtained in Section 6 do not apply to MsFEMs with oversampling. From Figure 3, we can see that the difference $u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{G}\text{-ni}}$ is still small with respect to the error committed by the G-MsFEM when *DOF*-continuous oversampling is applied. The approximation errors for the non-intrusive G-MsFEMs with *DOF*-continuous oversampling differ by at most 1.3% from the error of the G-MsFEM. Similar conclusions hold for the MsFEM-lin with *DOF*-extended oversampling. The difference between the G-MsFEM and the non-intrusive G-MsFEM is larger for the MsFEM-CR with *DOF*-extended oversampling. We do not include these results in the comparison of Figure 3 because both methods perform particularly badly when compared to the G-MsFEM without oversampling.

Let us also point out the qualitative and quantitative similarities between the performance of the MsFEM for the periodic and the non-periodic diffusion coefficients. Although the study of the homogenized limit of u^ε becomes increasingly difficult for the various coefficients (72a) to (72c), the non-intrusive approximation does not deteriorate the accuracy of the MsFEM in these numerical tests.

Before moving on to a comparison with the Petrov–Galerkin MsFEMs with oversampling, let us discuss a phenomenon in Figures 3 and 4 known as the "resonance effect" in the literature, preventing convergence of the MsFEM if the coarse scale H is close to ε. Upon further decreasing H, convergence is found only when H is sufficiently small with respect to the microscale ε, in which case we are in the regime of classical FEMs. From a theoretical point of view, this is explained by the term $\sqrt{\varepsilon/H}$ in the error estimate (71) (or ε/H for the MsFEM-lin with oversampling; see [40]).

We note that the same error estimate was obtained in [9] for the MsFEM-CR (without oversampling). Figure 4 shows that the resonance effect is more pronounced for the MsFEM-lin with oversampling than for the MsFEM-CR with oversampling.
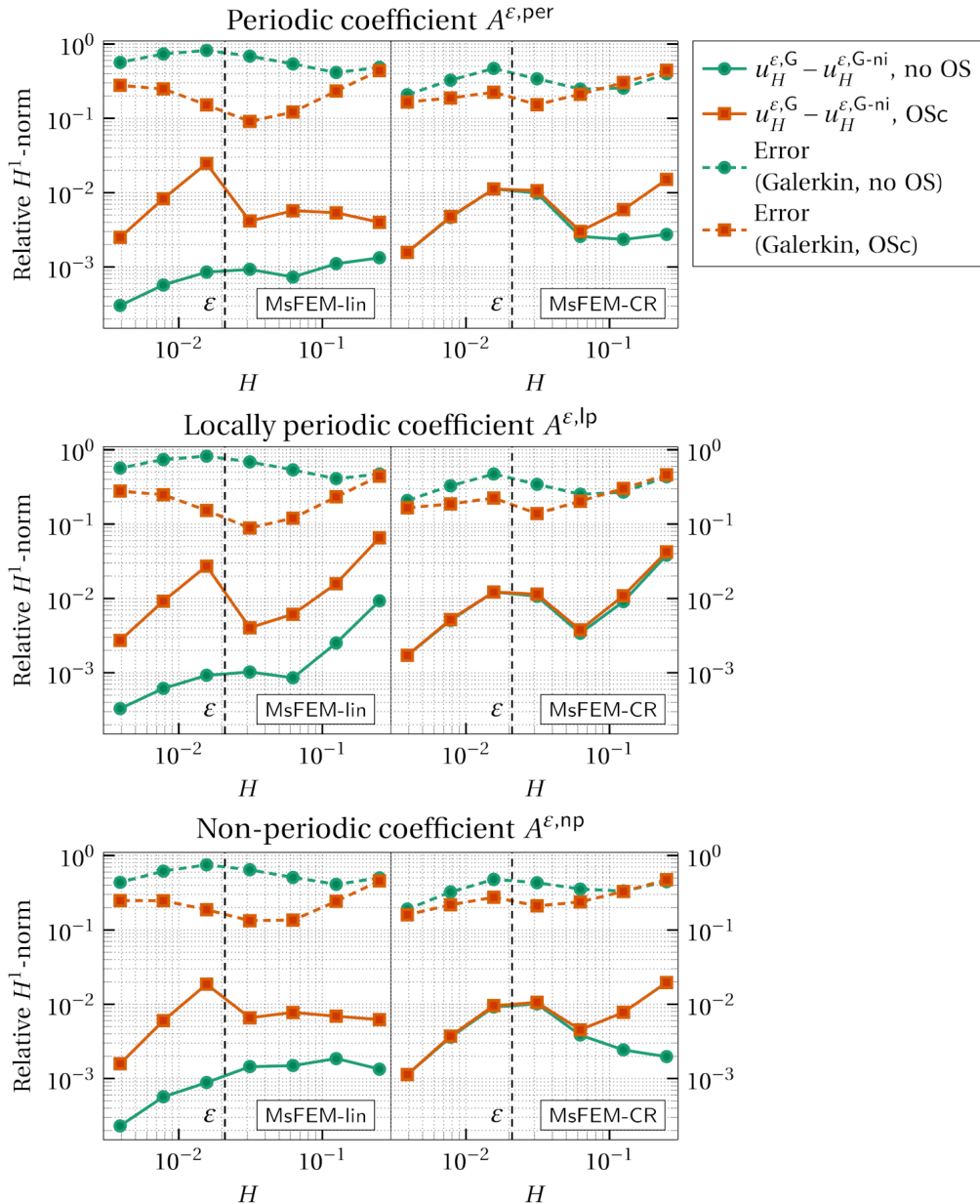
Figure 3. Solid lines: difference between the Galerkin MsFEM approximation ($u_H^{\varepsilon,\mathsf{G}}$ defined by (40)) and the non-intrusive Galerkin MsFEM approximation ($u_H^{\varepsilon,\mathsf{G\text{-}ni}}$ defined by (54)), without oversampling (no OS) and with *DOF*-continuous oversampling (OSc), for the diffusion coefficients in (72) as the mesh size H varies. Dashed lines: error of the Galerkin MsFEM with respect to the reference solution. All values are normalized with respect to the H^1 norm of the reference solution.
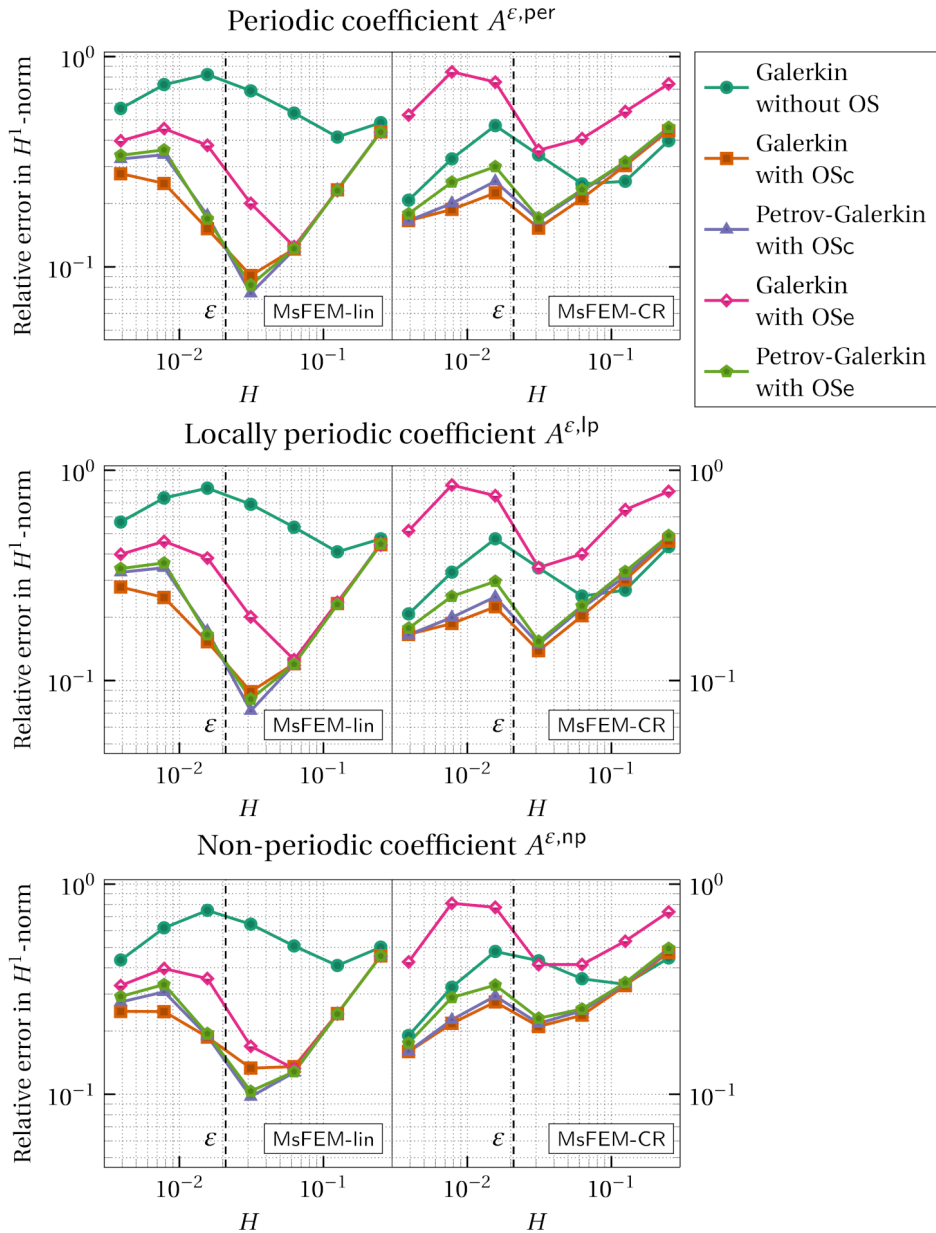
Figure 4. Comparison of the errors of the (intrusive) Galerkin MsFEM (40) and the (non-intrusive) Petrov-Galerkin MsFEM (41) for the diffusion coefficients in (72) as the mesh size H varies. Different oversampling strategies are applied: *DOF*-continuous (OSc, Definition 22) and *DOF*-extended (OSe, Definition 17). The Galerkin MsFEM without OS is included to illustrate the effect of the OS strategies.

We consider next in Figure 4 MsFEMs with the two different oversampling strategies of Section 4: *DOF*-continuous and *DOF*-extended oversampling. The PG-MsFEM variant, with or without oversampling, is completely equivalent to its non-intrusive implementation by virtue of Lemma 36. With oversampling, however, it does not coincide with the (intrusive or non-intrusive) G-MsFEM.

With oversampling, the matrices of the linear systems for the G-MsFEM and PG-MsFEM are different; Lemma 38 does not apply. The result is that the differences $u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{PG}}$ are larger than the differences $u_H^{\varepsilon,\mathsf{G}} - u_H^{\varepsilon,\mathsf{G\text{-}ni}}$. This is reflected in the numerical errors of the methods. We show the errors of the PG-MsFEM and the G-MsFEM with respect to the reference solution u_h^{ε} in Figure 4. (The non-intrusive G-MsFEM is too close to the G-MsFEM to be distinguishable on the scale of Figure 4 for all MsFEMs except the MsFEM with *DOF*-extended oversampling.) The G-MsFEM without oversampling is also shown to highlight the effect of oversampling.

Let us first consider the two different oversampling strategies. For all Galerkin MsFEMs, it is clear that the *DOF*-continuous variant performs (much) better than the *DOF*-extended variant. For the Petrov–Galerkin MsFEMs, the difference between the two oversampling strategies is smaller, but the *DOF*-continuous version of oversampling continues to perform better over all.

Although clear differences in the performance of the Galerkin and Petrov–Galerkin MsFEMs with *DOF*-continuous oversampling can be observed, these differences are small and both MsFEM approaches have a comparable accuracy. There is no systematic disadvantage in choosing the non-intrusive PG-MsFEM over the (intrusive or non-intrusive) G-MsFEM. Moreover, the non-periodic test cases again show the robustness of all MsFEM variants when going beyond the setting of periodic homogenization. In particular, this demonstrates the robustness of the non-intrusive approaches for the MsFEM developed in this article.

References

[1] R. Altmann, P. Henning, D. Peterseim, "Numerical homogenization beyond scale separation", *Acta Numer.* **30** (2021), p. 1-86.

[2] W. E, B. Engquist, "The Heterogeneous Multiscale Methods", *Commun. Math. Sci.* **1** (2003), no. 1, p. 87-132.

[3] A. Målqvist, D. Peterseim, "Localization of elliptic multiscale problems", *Math. Comput.* **83** (2014), no. 290, p. 2583-2603.

[4] T. Y. Hou, X.-H. Wu, "A Multiscale Finite Element Method for Elliptic Problems in Composite Materials and Porous Media", *J. Comput. Phys.* **134** (1997), no. 1, p. 169-189.

[5] A. Abdulle, E. Weinan, B. Engquist, E. Vanden-Eijnden, "The heterogeneous multiscale method", *Acta Numer.* **21** (2012), p. 1-87.

[6] Y. Efendiev, T. Y. Hou, *Multiscale Finite Element Methods*, Surveys and Tutorials in the Applied Mathematical Sciences, vol. 4, Springer, 2009.

[7] C. Le Bris, F. Legoll, "Examples of computational approaches for elliptic, possibly multiscale PDEs with random inputs", *J. Comput. Phys.* **328** (2017), p. 455-473.

[8] I. Babuška, J. E. Osborn, "Generalized Finite Element Methods: Their Performance and Their Relation to Mixed Methods", *SIAM J. Numer. Anal.* **20** (1983), no. 3, p. 510-536.

[9] C. Le Bris, F. Legoll, A. Lozinski, "MsFEM à la Crouzeix–Raviart for Highly Oscillatory Elliptic Problems", *Chin. Ann. Math., Ser. B* **34** (2013), no. 1, p. 113-138.

[10] C. Le Bris, F. Legoll, F. Madiot, "A numerical comparison of some Multiscale Finite Element approaches for advection-dominated problems in heterogeneous media", *ESAIM, Math. Model. Numer. Anal.* **51** (2017), no. 3, p. 851-888.

[11] L. H. Nguyen, D. Schillinger, "A residual-driven local iterative corrector scheme for the multiscale finite element method", *J. Comput. Phys.* **377** (2019), p. 60-88.

[12] R. A. Biezemans, C. Le Bris, F. Legoll, A. Lozinski, "Non-intrusive implementation of Multiscale Finite Element Methods: An illustrative example", *J. Comput. Phys.* **477** (2023), article no. 111914.

[13] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monographs and Studies in Mathematics, vol. 24, Pitman Publishing Inc., 1985.

[14] D. Gilbarg, N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 3rd ed., Classics in Mathematics, Springer, 2001.

[15] A. Quarteroni, *Numerical Models for Differential Problems*, 3rd ed., MS&A. Modeling, Simulation and Applications, vol. 16, Springer, 2018.

[16] P. G. Ciarlet, *The finite element method for elliptic problems*, Studies in Mathematics and its Applications, no. 4, North-Holland, 1978.

[17] A. Ern, J.-L. Guermond, *Theory and Practice of Finite Elements*, Applied Mathematical Sciences, vol. 159, Springer, 2004.

[18] A. Bensoussan, J.-L. Lions, G. Papanicolaou, *Asymptotic analysis for periodic structures*, Studies in Mathematics and its Applications, vol. 5, North-Holland, 1978.

[19] V. V. Zhikov, S. M. Kozlov, O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer Berlin, 1994.

[20] G. Allaire, *Shape Optimization by the Homogenization Method*, Applied Mathematical Sciences, vol. 146, Springer, 2002.

[21] F. Murat, L. Tartar, "*H*-Convergence", in *Topics in the Mathematical Modelling of Composite Materials*, Progress in Nonlinear Differential Equations and their Applications, vol. 31, Birkhäuser, 1997, p. 21-43.

[22] R. A. Biezemans, *Multiscale methods: non-intrusive implementation, advection-dominated problems and related topics*, PhD thesis, École des Ponts ParisTech, Paris, France, 2023, in preparation.

[23] G. Allaire, R. Brizzi, "A Multiscale Finite Element Method for Numerical Homogenization", *Multiscale Model. Simul.* **4** (2005), no. 3, p. 790-812.

[24] J. S. Hesthaven, S. Zhang, X. Zhu, "High-Order Multiscale Finite Element Method for Elliptic Problems", *Multiscale Model. Simul.* **12** (2014), no. 2, p. 650-666.

[25] F. Legoll, P.-L. Rothé, C. Le Bris, U. Hetmaniuk, "An MsFEM Approach Enriched Using Legendre Polynomials", *Multiscale Model. Simul.* **20** (2022), no. 2, p. 798-834.

[26] P. Bastian, M. Blatt, A. S. Dedner, C. Engwer, R. Klöfkorn, M. Ohlberger, O. Sander, "A generic grid interface for parallel and adaptive scientific computing. Part I: abstract framework", *Computing* **82** (2008), no. 2-3, p. 103-119.

[27] P. Bastian, M. Blatt, A. S. Dedner, C. Engwer, R. Klöfkorn, R. Kornhuber, M. Ohlberger, O. Sander, "A generic grid interface for parallel and adaptive scientific computing. Part II: implementation and tests in DUNE", *Computing* **82** (2008), no. 2-3, p. 121-138.

[28] P. Bastian, C. Engwer, D. Göddeke, O. Iliev, O. Ippisch, M. Ohlberger, S. Turek, J. Fahlke, S. Kaulmann, S. Müthing, D. Ribbrock, "EXA-DUNE: Flexible PDE Solvers, Numerical Methods and Applications", in *Euro-Par 2014: Parallel Processing Workshops*, Lecture Notes in Computer Science, vol. 8806, Springer, 2014, p. 530-541.

[29] P. Bastian, C. Engwer, J. Fahlke, M. Geveler, D. Göddeke, O. Iliev, O. Ippisch, R. Milk, J. Mohring, S. Müthing, M. Ohlberger, D. Ribbrock, S. Turek, "Advances Concerning Multiscale Methods and Uncertainty Quantification in EXA-DUNE", in *Software for Exascale Computing - SPPEXA 2013-2015*, Lecture Notes in Computational Science and Engineering, vol. 113, Springer, 2016, p. 25-43.

[30] V. John, J. M. Maubach, L. Tobiska, "Nonconforming streamline-diffusion-finite-element-methods for convection-diffusion problems", *Numer. Math.* **78** (1997), no. 2, p. 165-188.

[31] C. Le Bris, F. Legoll, F. Madiot, "Multiscale Finite Element Methods for Advection-Dominated Problems in Perforated Domains", *Multiscale Model. Simul.* **17** (2019), no. 2, p. 773-825.

[32] C. Le Bris, F. Legoll, A. Lozinski, "An MsFEM type approach for perforated domains", *Multiscale Model. Simul.* **12** (2014), no. 3, p. 1046-1077.

[33] P. Degond, A. Lozinski, B. P. Muljadi, J. Narski, "Crouzeix-Raviart MsFEM with Bubble Functions for Diffusion and Advection-Diffusion in Perforated Media", *Commun. Comput. Phys.* **17** (2015), no. 4, p. 887-907.

[34] B. P. Muljadi, J. Narski, A. Lozinski, P. Degond, "Nonconforming Multiscale Finite Element Method for Stokes Flows in Heterogeneous Media. Part I: Methodologies and Numerical Experiments", *Multiscale Model. Simul.* **13** (2015), no. 4, p. 1146-1172.

[35] G. Jankowiak, A. Lozinski, "Non-Conforming Multiscale Finite Element Method for Stokes Flows in Heterogeneous Media. Part II: error estimates for periodic microstructure", https://arxiv.org/abs/1802.04389, submitted, 2018.

[36] T. J. R. Hughes, "Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods", *Comput. Methods Appl. Mech. Eng.* **127** (1995), no. 1-4, p. 387-401.

[37] T. J. R. Hughes, G. R. Feijóo, L. Mazzei, J.-B. Quincy, "The variational multiscale method—a paradigm for computational mechanics", *Comput. Methods Appl. Mech. Eng.* **166** (1998), no. 1-2, p. 3-24.

[38] F. Brezzi, A. Russo, "Choosing bubbles for advection-diffusion problems", *Math. Models Methods Appl. Sci.* **4** (1994), no. 4, p. 571-587.

[39] F. Brezzi, L. P. Franca, T. J. R. Hughes, A. Russo, "$b = \int g$", *Comput. Methods Appl. Mech. Eng.* **145** (1997), p. 329-339.

[40] Y. Efendiev, T. Y. Hou, X.-H. Wu, "Convergence of a nonconforming multiscale finite element method", *SIAM J. Numer. Anal.* **37** (2000), no. 3, p. 888-910.

[41] T. Y. Hou, X.-H. Wu, Y. Zhang, "Removing the Cell Resonance Error in the Multiscale Finite Element Method via a Petrov–Galerkin Formulation", *Commun. Math. Sci.* **2** (2004), no. 2, p. 185-205.

[42] A. N. Brooks, T. J. R. Hughes, "Streamline Upwind/Petrov–Galerkin Formulations for Convection Dominated Flow with Particular Emphasis on the Incompressible Navier–Stokes Equation", *Comput. Methods Appl. Mech. Eng.* **32** (1982), p. 199-259.

[43] T. J. R. Hughes, L. E. Franca, G. M. Hulbert, "A New Finite Element Method Formulation for Computational Fluid Dynamics: VIII. The Galerkin/Least-Squares Method for Advective-Diffusive Equations", *Comput. Methods Appl. Mech. Eng.* **73** (1989), p. 173-189.

[44] L. P. Franca, A. Russo, "Recovering SUPG using Petrov–Galerkin formulations enriched with adjoint residual-free bubbles", *Comput. Methods Appl. Mech. Eng.* **182** (2000), no. 3-4, p. 333-339.

[45] D. Gallistl, D. Peterseim, "Computation of Quasi-Local Effective Diffusion Tensors and Connections to the Mathematical Theory of Homogenization", *Multiscale Model. Simul.* **15** (2017), no. 4, p. 1530-1552.

[46] D. Elfverson, V. Ginting, P. Henning, "On multiscale methods in Petrov–Galerkin formulation", *Numer. Math.* **131** (2015), no. 4, p. 643-682.

[47] P. Jenny, S. H. Lee, H. A. Tchelepi, "Multi-scale finite-volume method for elliptic problems in subsurface flow simulation", *J. Comput. Phys.* **187** (2003), no. 1, p. 47-67.

[48] H. Hajibeygi, G. Bonfigli, M. A. Hesse, P. Jenny, "Iterative multiscale finite-volume method", *J. Comput. Phys.* **227** (2008), no. 19, p. 8604-8621.

[49] Z. Chen, T. Y. Hou, "A mixed multiscale finite element method for elliptic problems with oscillating coefficients", *Math. Comput.* **72** (2002), no. 242, p. 541-576.

[50] T. Arbogast, "Implementation of a Locally Conservative Numerical Subgrid Upscaling Scheme for Two-Phase Darcy Flow", *Comput. Geosci.* **6** (2002), p. 453-481.

[51] T. Arbogast, G. Pencheva, M. F. Wheeler, I. Yotov, "A Multiscale Mortar Mixed Finite Element Method", *Multiscale Model. Simul.* **6** (2007), no. 1, p. 319-346.

[52] F. Brezzi, J.-L. Lions, O. Pironneau, "The Chimera method for a model problem", in *Numerical Mathematics and Advanced Applications* (F. Brezzi, A. Buffa, S. Corsaro, A. Murli, eds.), Springer, 2003, p. 817-825.

[53] R. Glowinski, J. He, A. Lozinski, J. Rappaz, J. Wagner, "Finite element approximation of multi-scale elliptic problems using patches of elements", *Numer. Math.* **101** (2005), no. 4, p. 663-687.

[54] J.-B. Apoung Kamga, O. Pironneau, "Numerical zoom for multiscale problems with an application to nuclear waste disposal", *J. Comput. Phys.* **224** (2007), no. 1, p. 403-413.

[55] M. Duval, J.-C. Passieux, M. Salaün, S. Guinard, "Non-intrusive Coupling: Recent Advances and Scalable Nonlinear Domain Decomposition", *Arch. Comput. Methods Eng.* **23** (2016), no. 1, p. 17-38.

[56] P. Gupta, J. P. Pereira, D.-J. Kim, C. A. Duarte, T. Eason, "Analysis of three-dimensional fracture mechanics problems: A non-intrusive approach using a generalized finite element method", *Eng. Fract. Mech.* **90** (2012), p. 41-64.

[57] R. Chakir, Y. Maday, "Une méthode combinée d'éléments finis à deux grilles/bases réduites pour l'approximation des solutions d'une E.D.P. paramétrique (A two-grid finite-element/reduced basis scheme for the approximation of the solution of parametric dependent P.D.E.)", *C. R., Math., Acad. Sci. Paris* **347** (2009), no. 7-8, p. 435-440.

[58] R. Chakir, Y. Maday, P. Parnaudeau, "A non-intrusive reduced basis approach for parametrized heat transfer problems", *J. Comput. Phys.* **376** (2019), p. 617-633.

[59] E. Grosjean, Y. Maday, "Error estimate of the non-intrusive reduced basis method with finite volume schemes", *ESAIM, Math. Model. Numer. Anal.* **55** (2021), no. 5, p. 1941-1961.

[60] C. Le Bris, F. Legoll, F. Thomines, "Multiscale Finite Element approach for "weakly" random problems and related issues", *ESAIM, Math. Model. Numer. Anal.* **48** (2014), no. 3, p. 815-858.

[61] F. Hecht, "New development in FreeFem++", *J. Numer. Math.* **20** (2012), no. 3-4, p. 251-265.

[62] R. A. Biezemans, "MsFEM in FreeFEM: Release version 1.0.0", 2023, DOI: 10.5281/zenodo.7525059.