

## CAPTURING SMALL SCALES IN ELLIPTIC PROBLEMS USING A RESIDUAL-FREE BUBBLES FINITE ELEMENT METHOD\*

GIANCARLO SANGALLI<sup>†</sup>

**Abstract.** In this work we study the residual-free bubbles (RFB) finite element method for solving second order elliptic equations with rapidly varying coefficients. The RFB technique is closely related to both the multiscale finite element method (MsFEM) introduced by Hou, Wu, and Cai [*Math. Comp.*, 68 (1999), pp. 913–943] and the *upscaling* procedures which are very common in the engineering literature for solving this kind of partial differential equation. We also introduce a variation of the RFB method, based on macrobubbles and referred to as the residual-free macrobubbles (RFMB) method, which gives more accurate numerical solutions. In the case of periodic coefficients we are able to prove a priori error estimates for the methods. Eventually, we test the numerical methods on model problems.

**Key words.** multiscale, finite element, homogenization

**AMS subject classifications.** 65F10, 65N30, 65N55

**PII.** S1540345902411402

**1. Introduction.** This paper is devoted to the numerical treatment of second order elliptic boundary value problems with highly oscillatory coefficients. Our model problem is

$$(1.1) \quad \begin{cases} \mathcal{L}u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where the differential operator is  $\mathcal{L}(\cdot) := \nabla \cdot (\mathbf{a} \cdot \nabla(\cdot))$ ,  $u$  is the solution defined on the convex polygonal domain  $\Omega \in \mathbb{R}^2$  with unitary diameter,  $\mathbf{a} = \{a_{ij}\}$  are the coefficients of the differential operator, and  $f \in L^2(\Omega)$  is the source term. Classical numerical methods such as the plain Galerkin finite element method are very effective only if the mesh is fine enough compared to the scale of  $\mathbf{a}$ . When  $\mathbf{a}$  is highly oscillatory, standard methods may require a too much refined mesh in order to achieve a satisfactory approximation of  $u$ . This may occur actually in practical applications, as the simulation of flow in a porous media, governed by Darcy's law: typically  $\mathbf{a}$ , which represents the permeability, has some small scale properties that we do not want to solve on a very fine mesh, but we have to somehow take into account those small scale details in order to deal with coarser meshes. Therefore in the last two decades many remedies have been proposed, particularly in the engineering literature, under the common name of *upscaling* procedures (see [12] for a review).

In this paper we discuss the advantages of using the residual-free bubbles (RFB) approach for solving problem (1.1), especially for rapidly varying  $\mathbf{a}$ . We consider the effect of the enrichment of the usual finite element space by means of *bubbles* (functions whose support stays within the elements) or *macro-bubbles* (whose support stays inside macroelements). We shall analyze the pros and cons of those approaches, and for the case of periodic  $\mathbf{a}$  we will be able to prove a priori error estimates.

---

\*Received by the editors July 19, 2002; accepted for publication (in revised form) December 8, 2002; published electronically DATE.

<http://www.siam.org/journals/mms/x-x/41140.html>

<sup>†</sup>Istituto di Matematica Applicata e Tecnologie Informatiche, Via Ferrata 1, 27100 Pavia, Italy (sangalli@imati.cnr.it).

The RFB method has been first proposed by Brezzi and Russo in [9] for advection-diffusion problems; it is based on previous ideas (see [1, 4]) and has been further analyzed in [5, 6, 7, 14, 18]. Its validity in a more general framework has been advocated by Brezzi and Marini in the recent work [8].

As shown in [8] and in the following section, the RFB method when applied to (1.1) has a close relation to the multiscale finite element method (MsFEM) proposed by Hou, Wu, and Cai in [15], where they also develop a theoretical analysis which is a mathematical justification of the upscaling approach (we notice for the interested reader that the MsFEM has been very effectively improved in the recent work [11] by means of the so-called *oversampling* technique, whose analogous in the RFB framework is not investigated in the present paper). Our analysis (in section 4) could be considered, as well, as a mathematical justification of the upscaling from a different point of view; further we shall see that using macrobubbles makes the method more accurate.

The outline of the paper is as follows: in section 2 we present the notation, the RFB method, and we discuss its relation to the MsFEM and to upscaling procedures; in section 3 we extend the formulation to allow macrobubbles; section 4 is devoted to the theoretical analysis of the proposed numerical methods, while in section 5 we discuss the numerical implementation of the methods and confirm, by means of numerical testing, the effectiveness of this methodology.

**2. The RFB formulation.** Given a regular subset  $\omega$  (possibly one-dimensional) of  $\Omega$ , we follow the usual notation for denoting Lebesgue spaces  $L^P(\omega)$ , endowed with the norm  $\|\cdot\|_{L^P(\omega)}$ , and Sobolev spaces  $H^s(\omega) \equiv W^{2,s}(\omega)$ , with norm  $\|\cdot\|_{H^s(\omega)}$  and seminorms  $|\cdot|_{H^s(\omega)}$ , where  $s \in \mathbb{R}$  (see [16] for details).  $H_0^1(\omega)$  is the space of  $H^1(\omega)$  functions with null trace on the boundary  $\partial\omega$ , equipped with the norm  $\|\cdot\|_{H_0^1(\omega)} := |\cdot|_{H^1(\omega)}$ . The standard notation  $\langle \cdot, \cdot \rangle$  is used for the pairing between spaces in duality.

In what follows,  $C$  and  $C_i$  denote generic constants whose value, possibly different at any occurrence, can depend only on the explicitly indicated quantities. We also adopt the notational convention

$$\alpha \simeq \beta \iff \alpha \leq C_1\beta \text{ and } \beta \leq C_2\alpha.$$

We introduce the bilinear form  $a : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$ , defined as

$$a(w, v) := \int_{\Omega} \nabla w(\mathbf{x})^t \cdot \mathbf{a}(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x}.$$

We assume in the paper that  $\mathbf{a}$  is symmetric, uniformly bounded, and positive definite:

$$(2.1) \quad \|\mathbf{a}\|_{L^\infty} \leq C_1,$$

$$(2.2) \quad \|\xi\| = 1 \Rightarrow \xi^t \cdot \mathbf{a}(\mathbf{x}) \cdot \xi \geq C_2 > 0;$$

thanks to the Poincaré inequality  $a(\cdot, \cdot)$  is coercive on  $H_0^1(\Omega)$ :

$$(2.3) \quad a(v, v) \geq C\|v\|_{H_0^1(\Omega)}^2 \quad \forall v \in H_0^1(\Omega),$$

where  $C > 0$ . Therefore the variational formulation of (1.1) is

$$(2.4) \quad \begin{cases} \text{Find } u \in H_0^1(\Omega) \text{ such that} \\ a(u, v) = \langle f, v \rangle \quad \forall v \in H_0^1(\Omega). \end{cases}$$

As usual in the finite element framework, we assume to have a partition  $\mathcal{T}_h$  of the domain  $\Omega$  into elements  $T$ , either triangles or quadrilaterals. Let  $\mathcal{T}_h$  be *admissible* (i.e., nonoverlapping elements, their union reproduces the domain, etc.), *shape regular* (i.e., the elements verify a minimum angle condition), and *quasi-uniform*, where the maximum diameter of elements is  $h \leq 1$ .

Let  $V_P$  denote the usual space of continuous piecewise linear (for triangular elements) or bilinear (for quadrilateral elements) finite element functions. The case of higher order elements could be considered as well. The plain Galerkin finite element method for (1.1) is based on (2.4), taking  $V_P$  instead of  $H_0^1(\Omega)$ ; its accuracy is very poor when the mesh-size  $h$  is larger than the smallest scale of  $\mathbf{a}$ , which is actually the case we want to deal with.

We now introduce the *bubbles* space

$$(2.5) \quad V_B := \{v \in H_0^1(\Omega) \text{ such that } v_{\partial T} = 0 \ \forall T \in \mathcal{T}_h\}$$

and the augmented space

$$(2.6) \quad V_h := V_P \oplus V_B.$$

At the abstract level, the RFB method reads

$$(2.7) \quad \begin{cases} \text{Find } u_h \in V_h \text{ such that} \\ a(u_h, v) = \langle f, v \rangle \quad \forall v \in V_h. \end{cases}$$

The method was first proposed by Brezzi and Russo (in [9]) for advection-diffusion problems. In that context the degrees of freedom of  $V_B$  were *statically condensed* by hand, because of the particular structure of that differential operator. This is not the case here: actually we need to compute numerically the degrees of freedom of  $V_B$  by introducing sufficiently fine *subgrids* within each element  $T \in \mathcal{T}_h$ ; we postpone the details of the choice of such subgrids until section 4.

In the next sections we shall show the advantages and disadvantages of using the RFB methodology for solving (1.1). We discuss now the structure of the algorithm one can derive from (2.7). The crucial point is that we can separate the *bubbles* degrees of freedom (inside the elements) from the degrees of freedom on the element interfaces  $\partial T$ . Doing that, we decouple (2.7) into *local* problems (inside the elements, i.e., on the subgrids mentioned before) and a single global problem on the coarse grid  $\mathcal{T}_h$ . We present, following [8], two different decoupling of (2.7): in the first case the global problem corresponds to a usual FEM for a modified differential operator (see section 2.1), while in the second case we end up with a FEM with special shape functions.

In the rest of the present section we restrict ourselves to piecewise constants source terms  $f$ , and we denote by  $F = \{f\}$  the set of all the admissible (piecewise constant) source terms. One could also consider (with minor modification) any finite dimensional set  $F$ .

**2.1. RFB method as an upscaling procedure.** *Upscaling* procedures modify the differential operator, usually by changing  $\mathbf{a}$  into  $\mathbf{a}^*$ , in such a way that the numerical solution by a plain Galerkin scheme (or other kinds of standard numerical methods) on the coarse grid (i.e.,  $\mathcal{T}_h$ ) is close to a *filtered* solution. The upscaled  $\mathbf{a}^*$  takes into accounts the effect of the *small scales* which are not contained into the finite element space  $V_P$ . The construction of  $\mathbf{a}^*$  is performed by solving *local* problems at the scale of the element size on a fine grid.

We are going to show here that a direct solution for the polynomial degrees of freedom of (2.7) leads to an upscaling procedure; in other words, the RFB method could be seen as a justification from a different point of view of the *upscaling* approach.

We consider here triangular elements. By (2.6), we can decompose  $u_h = u_P + u_B$  in a unique way, and analogously we can split the system of equations (2.7) into

$$(2.8) \quad a(u_P, v) + a(u_B, v) = \langle f, v \rangle \quad \forall v \in V_P,$$

$$(2.9) \quad a(u_P, v) + a(u_B, v) = \langle f, v \rangle \quad \forall v \in V_B.$$

We recognize that the first equation (2.8) gives the polynomial (coarse scale) approximation  $u_P$ , once the effect of  $u_B$ , which is  $a(u_B, v)$ , is known. Indeed we can follow the general procedure of [8, equations (3.20)–(3.23)]: after solving (2.9) for  $u_B$ , with  $u_P$  and  $f$  as data, and substituting back in (2.8), we end up with an upscaled formulation for  $u_P$ , as stated in the next proposition.

PROPOSITION 2.1. *The  $u_P$ , given by (2.8)–(2.9), verifies*

$$(2.10) \quad a^*(u_P, v) = \langle f^*, v \rangle \quad \forall v \in V_P,$$

where

$$(2.11) \quad a^*(w, v) = \int_{\Omega} \nabla w(\mathbf{x})^t \cdot \mathbf{a}^*(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x},$$

the coefficients  $\mathbf{a}^* \equiv \{a_{ij}^*\}$  are given by

$$(2.12) \quad a_{ij|T}^* = \frac{1}{|T|} \int_T \left( a_{ij}(\mathbf{x}) + \sum_k a_{ik}(\mathbf{x}) \frac{\partial \chi_j(\mathbf{x})}{\partial x_k} \right) d\mathbf{x} \quad \forall T \in \mathcal{T}_h,$$

the source term is given by

$$(2.13) \quad \langle f^*, v \rangle := - \sum_{T \in \mathcal{T}_h} \int_{\partial T} f(\mathbf{x}) (\mathbf{n}_T(\mathbf{x})^t \cdot \mathbf{a}(\mathbf{x}) \cdot \nabla \chi_0(\mathbf{x})) v(\mathbf{x}) \, d\mathbf{x} \quad \forall v \in V_P,$$

and, eventually,  $\chi_i \in V_B$  are in each element  $T \in \mathcal{T}_h$  the solutions of the local problems

$$(2.14) \quad -\nabla \cdot (\mathbf{a} \cdot \nabla \chi_0(\mathbf{x})) = 1,$$

$$(2.15) \quad -\nabla \cdot (\mathbf{a} \cdot \nabla \chi_j(\mathbf{x})) = \sum_i \frac{\partial a_{ij}(\mathbf{x})}{\partial x_i}.$$

*Proof.* We focus attention on a typical element  $T \in \mathcal{T}_h$  and, in what follows, each function, operator, and form is restricted to  $T$ ; with an abuse of notation, the restriction to  $T$  will not be indicated. Recall that both  $f$  and  $\frac{\partial u_P}{\partial x_j}$  are assumed to be constant (on  $T$ ). From (2.9) we get

$$\begin{aligned} \mathcal{L}u_B(\mathbf{x}) &= f - \mathcal{L}u_P(\mathbf{x}) \\ &= f + \sum_{ij} \frac{\partial a_{ij}(\mathbf{x})}{\partial x_i} \frac{\partial u_P}{\partial x_j}; \end{aligned}$$

using this and (2.14)–(2.15), we obtain

$$u_B(\mathbf{x}) = f \chi_0(\mathbf{x}) + \sum_j \frac{\partial u_P}{\partial x_j} \chi_j(\mathbf{x}),$$

whence

$$a(u_B, v) = f a(\chi_0, v) + \sum_j \frac{\partial u_P}{\partial x_j} a(\chi_j, v) = I + II.$$

We recognize that  $-I$  gives the contribution of the *bubble elimination* (on  $T$ ) on the right-hand side of (2.8), while  $II$  gives the contribution on the left-hand side of (2.8). Therefore (using integration by parts and (2.14))

$$\begin{aligned} \langle f^*, v \rangle &= \langle f, v \rangle - f a(\chi_0, v) \\ &= f \left( \int_T v(\mathbf{x}) d\mathbf{x} - \int_T \nabla v(\mathbf{x})^t \cdot \mathbf{a}(\mathbf{x}) \cdot \nabla \chi_0(\mathbf{x}) d\mathbf{x} \right) \\ &= -f \int_{\partial T} (\mathbf{n}_T(\mathbf{x})^t \cdot \mathbf{a}(\mathbf{x}) \cdot \nabla \chi_0(\mathbf{x})) v(\mathbf{x}) d\mathbf{x} \end{aligned}$$

and analogously (still using integration by parts and (2.15))

$$\begin{aligned} a^*(u_P, v) &= a(u_P, v) + II \\ &= \sum_{ij} \frac{\partial u_P}{\partial x_j} \frac{\partial v}{\partial x_i} \int_T a_{ij}(\mathbf{x}) d\mathbf{x} \\ &\quad + \sum_j \frac{\partial u_P}{\partial x_j} \int_T \sum_{ik} a_{ik}(\mathbf{x}) \frac{\partial \chi_j(\mathbf{x})}{\partial x_k} \frac{\partial v}{\partial x_i} d\mathbf{x} \\ &= \sum_{ij} \frac{\partial u_P}{\partial x_j} \frac{\partial v}{\partial x_i} \int_T \left( a_{ij}(\mathbf{x}) + \sum_k a_{ik}(\mathbf{x}) \frac{\partial \chi_j(\mathbf{x})}{\partial x_k} \right) d\mathbf{x}; \end{aligned}$$

this proves (2.10)–(2.13).  $\square$

**2.2. RFB method as a multiscale method.** In the present subsection we still restrict ourselves to piecewise constant source terms  $f \in F$ . Then we introduce the subspace of bubbles  $V_F \subset V_B$  (where  $\dim(V_F) = \dim(F)$ ) containing local solutions for source terms in  $F$ :

$$(2.16) \quad V_F := \{w \in V_B : a(w, v) = \langle \psi, v \rangle, \psi \in F \forall v \in V_B\}.$$

In our case, the construction of a basis  $\{\phi_i\}$  for  $V_F$  is just a matter of solving  $\mathcal{L}\phi_i = 1$  in each element  $T_i \in \mathcal{T}_h$ . We also introduce the space  $V_{\mathcal{L}} \subset V_h$  of locally  $\mathcal{L}$ -homogeneous functions, which means

$$(2.17) \quad V_{\mathcal{L}} := \{w \in V_h : a(w, v) = 0 \forall v \in V_B\}.$$

We have  $\dim(V_{\mathcal{L}}) = \dim(V_P)$ , and again the construction of a basis for  $V_{\mathcal{L}}$  is based on solving local problems: one needs to modify a basis for  $V_P$  in the interior of the elements; moreover,  $V_{\mathcal{L}}$  and  $V_F$  are orthogonal w.r.t.  $a(\cdot, \cdot)$ . We can merge the two bases in a basis for  $V_{\mathcal{L}} \oplus V_F$ , which is a finite dimensional coarse space ( $\dim(V_{\mathcal{L}} \oplus V_F) = \dim(V_{\mathcal{L}}) + \dim(V_F)$ ). Note that the RFB solution  $u_h$  belongs to  $V_{\mathcal{L}} \oplus V_F$ , and we can restate (2.7) as

$$(2.18) \quad \begin{cases} \text{Find } u_h \in V_{\mathcal{L}} \oplus V_F \text{ such that} \\ a(u_h, v) = \langle f, v \rangle \quad \forall v \in V_{\mathcal{L}} \oplus V_F. \end{cases}$$

Therefore we can look at the RFB method as a Galerkin formulation on a coarse grid, with special shape functions, the shape being determined by solving local problems.

As discussed in [8], (2.18) is similar to the MsFEM. If compared to (2.18), the MsFEM formulation is just the restriction of that formulation to the sole space  $V_{\mathcal{L}}$ :

in other words, the numerical approximation given by the MsFEM, here denoted by  $u_{MsFEM}$ , belongs to  $V_{\mathcal{L}}$  and verifies

$$(2.19) \quad a(u_{MsFEM}, v) = \langle f, v \rangle \quad \forall v \in V_{\mathcal{L}}.$$

The two methods, i.e., RFB and MsFEM, are actually very closely related. Indeed, due to the orthogonality between  $V_{\mathcal{L}}$  and  $V_B$ ,  $u_{MsFEM}$  coincides with  $u_h$  on the boundaries  $\partial T$  of the elements  $T \in \mathcal{T}_h$ ; see [8, Theorem 1] for more details.

**3. The formulation with macrobubbles.** As we shall see in following sections, the RFB method is not completely satisfactory, because it is affected by the so called *resonance error*, which means that it has a poor accuracy when the characteristic scale of  $\mathbf{a}$  is comparable to the mesh-size  $h$ . This behavior is typical of upscaling procedures too (see [12]).

For that reason we propose an improvement of the RFB method, based on macrobubbles instead of bubbles; we shall refer to this improvement as the residual-free macrobubbles (RFMB) method. This is an extension of the *macrobubbles* approach proposed by Russo and Franca for the Stokes problem in [13]. The macrodomains  $K$  are either triangles or quadrilaterals of a macropartition  $\mathcal{T}_H$  of  $\Omega$ ; for the sake of simplicity, we shall consider  $K$  to be either triangles or quadrilaterals. We assume  $\mathcal{T}_H$  to be shape regular and quasi-uniform, and the maximum diameter of  $K \in \mathcal{T}_H$  is denoted by  $H$ , where  $h \leq H \leq 1$ . We also assume  $\mathcal{T}_h$  to be a refinement of  $\mathcal{T}_H$ , which means that any element  $T \in \mathcal{T}_h$  is included in a single macroelement  $K \in \mathcal{T}_H$ .

The space of macrobubbles  $V_{MB}$  is

$$(3.1) \quad V_{MB} := \{v \in H_0^1(\Omega) \text{ such that } v_{\partial K} = 0 \ \forall K \in \mathcal{T}_H\};$$

the enriched space  $V_{H,h}$ , simply denoted by  $V_H$  in what follows, is

$$(3.2) \quad V_H \equiv V_{H,h} := V_P + V_{MB},$$

and the RFMB method reads

$$(3.3) \quad \begin{cases} \text{Find } u_H \in V_H \text{ such that} \\ a(u_H, v) = \langle f, v \rangle \quad \forall v \in V_H. \end{cases}$$

Upscaling methodologies which deal with *patches* of elements have been introduced by many authors (see the review [12]).

In order to implement (3.3) by a computer algorithm we can follow, with suitable modifications, the two ways drawn in sections 2.1–2.2. Everything is clearer in the *domain decomposition* framework.

We consider first the approach of section 2.1. Now, unlike in (2.6), the sum in (3.2) is not direct, in the sense that  $V_P \cap V_{MB} \neq \emptyset$ . Therefore one possibility is to replace the splitting (2.8)–(2.9) with the classical Schwarz *alternating* algorithm on the *overlapping* spaces  $V_P$  and  $V_{MB}$  (see, e.g., [3, 10]): from  $u_H^n \in V_H$ , we compute the next iterate  $u_H^{n+1} \in V_H$  by

$$(3.4) \quad \begin{cases} \text{Find } w_{MB} \in V_{MB} \text{ such that} \\ a(w_{MB} + u_H^n, v) = \langle f, v \rangle \quad \forall v \in V_{MB}; \\ \text{Set } u_H^{n+1/2} := w_{MB} + u_H^n; \end{cases}$$

$$(3.5) \quad \begin{cases} \text{Find } w_P \in V_P \text{ such that} \\ a(w_P + u_H^{n+1/2}, v) = \langle f, v \rangle \quad \forall v \in V_P; \\ \text{Set } u_H^{n+1} := w_P + u_H^{n+1/2}. \end{cases}$$

PROPOSITION 3.1. *The error reduction of (3.4)–(3.5) verifies*

$$\frac{\|u_H - u_H^{n+1}\|_{H^1}}{\|u_H - u_H^n\|_{H^1}} \leq C < 1$$

uniformly w.r.t.  $H$ ,  $h$ , and, since (2.1)–(2.2), uniformly w.r.t.  $\mathbf{a}$ .

*Proof.* We can split any  $w_H \in V_H$  into  $w_P + w_{MB}$ , where  $w_P \in V_P$ ,  $w_{MB} \in V_{MB}$ , and  $\|w_P\|_{H^1} + \|w_{MB}\|_{H^1} \leq C\|w_H\|_{H^1}$ . For example, take  $w_P = w_H$  on any macroelement boundary  $\partial K$ , and extend  $w_P$  inside  $K$  following Lemma 4.2, whence (also using Lemma 4.1)

$$\begin{aligned} |w_P|_{H^1(\Omega)}^2 &= \sum_{K \in \mathcal{T}_H} |w_P|_{H^1(K)}^2 \leq C \sum_{K \in \mathcal{T}_H} |w_P|_{H^{1/2}(\partial K)}^2 \\ (3.6) \qquad &\leq C \sum_{K \in \mathcal{T}_H} |w_H|_{H^{1/2}(\partial K)}^2 \\ &\leq C \sum_{K \in \mathcal{T}_H} |w_H|_{H^1(K)}^2 = C|w_H|_{H^1(\Omega)}^2. \end{aligned}$$

The error reduction estimates follows from (3.6) thanks to [10, Theorem 15].  $\square$

More sophisticated iterative procedures (e.g., by means of the conjugate gradient acceleration) may be used instead of (3.4)–(3.5) as well (we refer to [10, section 2]).

The second way for dealing with the RFMB formulation (3.3) is related to the idea of condensing the degrees of freedom related to the macrobubbles; this is, with minor modifications, the idea of section 2.2. Indeed one can introduce the counterpart of  $V_F$  and  $V_{\mathcal{L}}$ , with  $V_{MB}$  in place of  $V_B$  in (2.16) and (2.17), and obtain a coarse formulation which is the counterpart of (2.18) (see [8] for details). This can also be viewed as a *nonoverlapping* domain decomposition approach: we decompose (3.3) into the Steklov–Poincaré problem on the interface  $\bigcup \partial K$  and the problem on the interior of the macroelements  $K$ . When the macroelements  $K$  are much larger than the elements  $T$ , then the actual construction of  $V_H$  could be not competitive; indeed the Steklov–Poincaré problem is not solved directly in a domain decomposition formulation, whereas it is treated by an iterative procedure, after suitable preconditioning (see [10] for details).

**4. Error estimates.** This section is devoted to a priori error analysis of the RFMB method (3.3); the theory also covers the RFB method (2.7), which is the case  $\mathcal{T}_H \equiv \mathcal{T}_h$ .

First, we state some lemmas. Recall that  $\omega$  denotes a generic regular subset of the domain  $\Omega$ , either one-dimensional or two-dimensional. We set  $\tilde{L}^2(\omega) \in L^2(\omega)$  and  $\tilde{H}^1(\omega) \in H^1(\omega)$  (endowed with the norms  $\|\cdot\|_{L^2(\omega)}$  and  $|\cdot|_{H^1(\omega)}$ , respectively) as the subspaces of zero mean value functions. We denote by  $\bar{\Pi}_\omega$  the mean value of  $v$  on  $\omega$ , while the operator  $\tilde{\Pi}_\omega v$  denotes the  $L^2(\omega)$ -projection onto  $\tilde{L}^2(\omega)$ , i.e.,  $\tilde{\Pi}_\omega v := v - \bar{\Pi}_\omega v$ .

We recall that Sobolev spaces of fractional order may be obtained by interpolation between integer order Sobolev spaces; in particular,  $H^{1/2}(\omega) = (L^2(\omega), H^1(\omega))_{(1/2,2)}$ , following the usual notation for the real interpolation method (see [19]). We also introduce  $\tilde{H}^{1/2}(\omega) = (\tilde{L}^2(\omega), \tilde{H}^1(\omega))_{(1/2,2)}$ . Thanks to [19, section 1.17.1, Theorem 2],  $\tilde{H}^{1/2}(\omega) \equiv H^{1/2}(\omega) \cap \tilde{L}^2(\omega)$ , with norms equivalence  $\|v\|_{\tilde{H}^{1/2}(\omega)} \simeq \|v\|_{H^{1/2}(\omega)}$ . Then the norm  $\|\cdot\|_{\tilde{H}^{1/2}(\omega)}$  on  $\tilde{H}^{1/2}(\omega)$  can be extended as a seminorm  $|\cdot|_{H^{1/2}(\omega)}$  to  $H^{1/2}(\omega)$  just setting  $|v|_{H^{1/2}(\omega)} := \|\tilde{\Pi}_\omega v\|_{\tilde{H}^{1/2}(\omega)}$ , yielding  $\|v\|_{H^{1/2}(\omega)} \simeq \|v\|_{L^2(\omega)} + |v|_{H^{1/2}(\omega)}$ .

The following lemma is just a minor variation of the well-known result for traces of  $H^1$  functions.

LEMMA 4.1. *Let  $\omega$  be an open subset of  $\Omega$  with Lipschitz boundary  $\partial\omega$ ; therefore*

$$(4.1) \quad |v|_{H^{1/2}(\partial\omega)} \leq C|v|_{H^1(\omega)} \quad \forall v \in H^1(\omega);$$

conversely, we can extend any  $v \in H^{1/2}(\partial\omega)$  to  $v \in H^1(\omega)$  such that

$$(4.2) \quad |v|_{H^1(\omega)} \leq C|v|_{H^{1/2}(\partial\omega)}.$$

Further, if  $\omega$  is a shape regular (triangular or quadrilateral) element, the constant  $C$  in (4.1) and (4.2) does not depend on the diameter of  $\omega$  (i.e., it depends on the minimum angle).

*Proof.* Consider first  $v \in H^1(\omega)$ ; the classical result, applied to  $\tilde{\Pi}_\omega v$ , is

$$\|\tilde{\Pi}_\omega v\|_{H^{1/2}(\partial\omega)} \leq C\|\tilde{\Pi}_\omega v\|_{H^1(\omega)};$$

therefore (4.1) follows from our definition of  $|\cdot|_{H^{1/2}(\partial\omega)}$  (since  $\tilde{\Pi}_{\partial\omega}\tilde{\Pi}_\omega v = \tilde{\Pi}_{\partial\omega}v$ ) and from  $\|\tilde{\Pi}_\omega v\|_{L^2(\omega)} \leq C|v|_{H^1(\omega)}$ .

Given  $v \in H^{1/2}(\partial\omega)$ , let  $w$  be the harmonic lifting of  $\tilde{\Pi}_{\partial\omega}v$ , which verifies

$$(4.3) \quad \|w\|_{H^1(\omega)} \leq C\|\tilde{\Pi}_{\partial\omega}v\|_{H^{1/2}(\partial\omega)};$$

therefore  $v = w + \bar{\Pi}_{\partial\omega}v$  is the extension inside  $\omega$  we are looking for, and (4.2) follows easily from (4.3) and

$$|v|_{H^1(\omega)} = |w|_{H^1(\omega)} \leq \|w\|_{H^1(\omega)}.$$

Finally, the uniformity of the constant  $C$  in (4.1)–(4.2) for shape regular elements follows easily from a scaling argument.  $\square$

A similar result holds true for finite element discrete functions; in particular, we have the following lemma.

LEMMA 4.2. *Consider a partition  $\mathcal{T}$  of  $\omega$ . Let  $\mathcal{T}$  be a shape regular, admissible, and quasi-uniform partition of mesh-size  $h_{\mathcal{T}}$ . Let  $V_{\mathcal{T}}(\omega) \subset H^1(\omega)$  be a standard finite element space related to  $\mathcal{T}$ , and let  $V_{\mathcal{T}}(\partial\omega) \subset H^{1/2}(\partial\omega)$  be the space of restrictions to  $\partial\omega$  of functions belonging to  $V_{\mathcal{T}}(\omega)$ . Given  $v \in V_{\partial\mathcal{T}}(\omega)$ , we can extend it to a  $v \in V_{\mathcal{T}}(\omega)$  such that (4.2) holds true. The constant  $C$  therein is independent on  $h_{\mathcal{T}}$  and, when  $\omega$  is a shape regular triangle or quadrilateral, on its diameter.*

*Proof.* Take  $v \in V_{\partial\mathcal{T}}(\omega)$ ; let  $w \in H^1(\omega)$  be the harmonic lifting of  $\tilde{\Pi}_{\partial\omega}v \in V_{\partial\mathcal{T}}(\omega)$ , which verifies

$$(4.4) \quad \|w\|_{H^{1+\sigma}(\omega)} \leq C(\sigma)\|\tilde{\Pi}_{\partial\omega}v\|_{H^{1/2+\sigma}(\partial\omega)}$$

for  $0 \leq \sigma < 1/2$ , thanks to the convexity of  $\omega$ . Recall the fractional order inverse inequality

$$(4.5) \quad \|\tilde{\Pi}_{\partial\omega}v\|_{H^{1/2+\sigma}(\partial\omega)} \leq Ch_{\mathcal{T}}^{-\sigma}\|\tilde{\Pi}_{\partial\omega}v\|_{H^{1/2}(\partial\omega)},$$

which is obtained by interpolation on the classical inverse inequalities. We also need the nodal interpolant  $w_I \in V_{\mathcal{T}}(\omega)$  of  $w$ , which satisfies the error estimate

$$(4.6) \quad \|w - w_I\|_{H^1(\omega)} \leq C(\sigma)h_{\mathcal{T}}^{\sigma}\|w\|_{H^{1+\sigma}(\omega)},$$

where now  $0 < \sigma < 1/2$ . Fixing, for example,  $\sigma = 1/4$ , using (4.6), (4.4), and then (4.5), we get

$$(4.7) \quad \|w - w_I\|_{H^1(\omega)} \leq C \|\tilde{\Pi}_{\partial\omega} v\|_{H^{1/2}(\partial\omega)};$$

using again (4.4), this time for  $\sigma = 0$ , we end with

$$(4.8) \quad \|w_I\|_{H^1(\omega)} \leq C \|\tilde{\Pi}_{\partial\omega} v\|_{H^{1/2}(\partial\omega)}.$$

We can therefore extend  $v$  as  $w_I + \bar{\Pi}_{\partial\omega} v$  and (4.2) follows from (4.8); the uniformity w.r.t. the diameter of  $\omega$  is proven by usual scaling arguments.  $\square$

Using previous lemmas, we can now derive error estimates for the RFMB method; the main idea of the analysis is stated in the next proposition.

PROPOSITION 4.3. *The numerical solution  $u_H$  of the RFMB method (3.3) verifies*

$$(4.9) \quad \|u - u_H\|_{H^1(\Omega)}^2 \simeq \inf_{v \in V_P} \sum_{K \in \mathcal{T}_H} |u - v|_{H^{1/2}(\partial K)}^2.$$

*Proof.* Recall (2.3), so the classical argument gives

$$\|u - u_H\|_{H^1(\Omega)}^2 \simeq \inf_{v \in V_H} \|u - v\|_{H^1(\Omega)}^2.$$

As a consequence of Lemma 4.1, for any  $v \in V_H$  we have

$$(4.10) \quad |u - v|_{H^1(\Omega)}^2 = \sum_{K \in \mathcal{T}_H} |u - v|_{H^1(K)}^2 \geq C \sum_{K \in \mathcal{T}_H} |u - v|_{H^{1/2}(\partial K)}^2,$$

and for any  $v \in V_P$  there is a suitable  $w \in V_{MB}$  yielding

$$(4.11) \quad |u - (v + w)|_{H^1(\Omega)}^2 = \sum_{K \in \mathcal{T}_H} |u - (v + w)|_{H^1(K)}^2 \leq C \sum_{K \in \mathcal{T}_H} |u - v|_{H^{1/2}(\partial K)}^2;$$

eventually, taking the infimum w.r.t.  $v$  in both (4.10) and (4.11) and using the Poincaré inequality, we obtain

$$\inf_{v \in V_H} \|u - v\|_{H^1(\Omega)}^2 \simeq \inf_{v \in V_P} \sum_{K \in \mathcal{T}_H} |u - v|_{H^{1/2}(\partial K)}^2. \quad \square$$

According to Proposition 4.3, the accuracy of the RFMB method depends on how well we can approximate the exact solution  $u$  by piecewise polynomials on the skeleton  $\bigcup_{K \in \mathcal{T}_H} \partial K$ .

In the classical framework the best approximation error for  $u$ —the right-hand side of (4.9), in our particular case—is evaluated thanks to higher order regularity assumptions on  $u$ . But, in our case,  $\|u\|_{H^s(\Omega)}$  depends on the small scales of  $\mathbf{a}$  for any  $s > 1$ . Standard Sobolev spaces are therefore nonadequate here; a deeper investigation on the structure of  $u$  is out of the aim of this paper.

Therefore, following [15], we restrict the theoretical analysis to periodic coefficients:

$$(4.12) \quad \begin{aligned} a_{ij}(\mathbf{x}) &\equiv a_{ij,\varepsilon}(\mathbf{x}) := \alpha_{ij}(\mathbf{x}/\varepsilon), \\ &\text{where } \mathbf{y} \mapsto \alpha_{ij}(\mathbf{y}) \text{ are smooth} \\ &\text{and 1-periodic w.r.t. } x_1 \text{ and } x_2. \end{aligned}$$

When  $\varepsilon$  is large compared to  $h$ , the classical analysis for the plain Galerkin method, which extends straightforwardly to the RFMB method, gives a satisfactory error estimate.

**THEOREM 4.4.** *Assuming (4.12), the numerical solution  $u_H$  of the RFMB method (3.3) verifies*

$$(4.13) \quad \|u - u_H\|_{H^1(\Omega)} \leq Ch\varepsilon^{-1} \|f\|_{L^2(\Omega)}.$$

*Proof.* Under our assumptions we have  $H^2$ -regularity for  $u$ , and in particular  $\|u\|_{H^2(\Omega)} \leq C\varepsilon^{-1} \|f\|_{L^2(\Omega)}$  (see [15, (4.4)]); therefore, by means of usual estimates for the nodal interpolant approximation, we get (4.13).  $\square$

The most interesting case is when  $\varepsilon$  is smaller than or comparable to the mesh-size. Because of (4.12), we acquire knowledge on the structure of  $u$  by means of the *homogenization* theory. In particular, from [17] we get the following first order approximation:

$$(4.14) \quad \|u - u_0 - \varepsilon(u_1 - \theta)\|_{H^1(\Omega)} \leq C\varepsilon \|u_0\|_{H^2(\Omega)},$$

where  $u_0$ ,  $u_1$ , and  $\theta$  represent the *homogenized solution*, the *first order correction*, and the *boundary correction*, respectively. The homogenized solution  $u_0$  does not depend on  $\varepsilon$  (while  $u_1$  and  $\theta$  do) and is given by

$$(4.15) \quad \begin{cases} -\nabla \cdot (\mathbf{a}^* \cdot \nabla u_0) = f & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega, \end{cases}$$

where the constant homogenized coefficients  $\mathbf{a}^* \equiv \{a_{ij}^*\}$  are given by

$$a_{ij}^* = \int_{[0,1]^2} \left( \alpha_{ij}(\mathbf{y}) + \sum_k \alpha_{ik}(\mathbf{y}) \frac{\partial \chi_j(\mathbf{y})}{\partial y_k} \right) d\mathbf{y}$$

and  $\chi_i$  are the zero mean value periodic (on  $[0, 1]^2$ ) solutions of the equations

$$-\nabla_{\mathbf{y}} \cdot (\boldsymbol{\alpha} \cdot \nabla_{\mathbf{y}} \chi_j(\mathbf{y})) = \sum_i \frac{\partial \alpha_{ij}(\mathbf{y})}{\partial y_i}.$$

It is worth noting the close relation between the  $\mathbf{a}^*$  defined here and its numerical counterpart (see section 2.1) related to the upscaling effect of the RFB method. Eventually we have

$$(4.16) \quad u_1(\mathbf{x}) = \sum_i \chi_i \left( \frac{\mathbf{x}}{\varepsilon} \right) \frac{\partial u_0}{\partial x_i}(\mathbf{x}),$$

while  $\theta$  is the solution of

$$(4.17) \quad \begin{cases} -\nabla \cdot (\mathbf{a} \cdot \nabla \theta) = 0 & \text{in } \Omega, \\ \theta = u_1 & \text{on } \partial\Omega. \end{cases}$$

Our main result is stated in the next theorem.

**THEOREM 4.5.** *Assuming (4.12), the numerical solution  $u_H$  of the RFMB method (3.3) verifies*

$$(4.18) \quad \|u - u_H\|_{H^1(\Omega)} \leq C(h + \varepsilon + \varepsilon^2 + \varepsilon^{1/2} H^{-1/2} + \varepsilon H^{-1}) \|f\|_{L^2}.$$

*Proof.* Using (4.9) and the first order approximation (4.14) we can split the error into four terms:

$$\begin{aligned}
\|u - u_H\|_{H^1(\Omega)}^2 &\leq C \sum_{K \in \mathcal{T}_H} |u - u_0 - \varepsilon(u_1 - \theta)|_{H^{1/2}(\partial K)}^2 \\
&\quad + C \sum_{K \in \mathcal{T}_H} |u_0 - u_0^I|_{H^{1/2}(\partial K)}^2 \\
(4.19) \quad &\quad + C \sum_{K \in \mathcal{T}_H} |\varepsilon u_1|_{H^{1/2}(\partial K)}^2 \\
&\quad + C \sum_{K \in \mathcal{T}_H} |\varepsilon \theta|_{H^{1/2}(\partial K)}^2 \\
&= I + II + III + IV,
\end{aligned}$$

where  $u_0^I$  denotes the nodal interpolant for  $u_0$ . One key point of the proof is that in *III* the function  $\varepsilon u_1$  (which takes into account the oscillations of  $u$ ) is evaluated on the skeleton  $\bigcup \partial K$ . For the other terms (i.e., *I*, *II*, and *IV*) we shall come back to norms on the whole  $\Omega$  by means of (4.1), instead. Therefore, using also (4.14), we get

$$I \leq C \|u - u_0 - \varepsilon(u_1 - \theta)\|_{H^1(\Omega)}^2 \leq C \varepsilon^2 \|u_0\|_{H^2(\Omega)}^2,$$

while the usual properties of the nodal interpolant yield

$$II \leq C \|u_0 - u_0^I\|_{H^1(\Omega)}^2 \leq Ch^2 \|u_0\|_{H^2(\Omega)}^2.$$

The estimates for *III* and *IV* require a preliminary result. Denote by  $\widehat{K}$  the *reference* macroelement (either the reference square  $[0, 1]^2$  or the reference triangle  $\{(\hat{x}_1, \hat{x}_2) | 0 < \hat{x}_1 < 1, 0 < \hat{x}_2 < 1, \hat{x}_1 + \hat{x}_2 < 1\}$ ) and by  $\hat{v}$  and  $\hat{w}$  two generic smooth functions on  $\widehat{K}$ . From

$$(4.20) \quad \|\hat{w}\hat{v}\|_{L^2(\partial \widehat{K})} \leq \|\hat{w}\|_{L^\infty(\partial \widehat{K})} \|\hat{v}\|_{L^2(\partial \widehat{K})},$$

$$(4.21) \quad \|\hat{w}\hat{v}\|_{H^1(\partial \widehat{K})} \leq C \|\hat{w}\|_{W^{1,\infty}(\partial \widehat{K})} \|\hat{v}\|_{H^1(\partial \widehat{K})}$$

we obtain by interpolation and using the properties of traces

$$\begin{aligned}
(4.22) \quad \|\hat{w}\hat{v}\|_{H^{1/2}(\partial \widehat{K})} &\leq C \|\hat{w}\|_{L^\infty(\partial \widehat{K})}^{1/2} \|\hat{w}\|_{W^{1,\infty}(\partial \widehat{K})}^{1/2} \|\hat{v}\|_{H^{1/2}(\partial \widehat{K})} \\
&\leq C \|\hat{w}\|_{L^\infty(\partial \widehat{K})}^{1/2} \|\hat{w}\|_{W^{1,\infty}(\partial \widehat{K})}^{1/2} \|\hat{v}\|_{H^1(\widehat{K})}.
\end{aligned}$$

From there and using the usual scale argument we get the estimate for  $v$  and  $w$  defined on a generic  $K \in \mathcal{T}_H$  (recall  $H \leq 1$ ):

$$\begin{aligned}
(4.23) \quad |wv|_{H^{1/2}(\partial K)} &\leq C (\|w\|_{L^\infty(\partial K)} \\
&\quad + H^{1/2} \|w\|_{L^\infty(\partial K)}^{1/2} \|\nabla w\|_{L^\infty(\partial K)}^{1/2}) H^{-1} \|v\|_{H^1(K)} \\
&\leq C (H^{-1} \|w\|_{L^\infty(\partial K)} \\
&\quad + H^{-1/2} \|w\|_{L^\infty(\partial K)}^{1/2} \|\nabla w\|_{L^\infty(\partial K)}^{1/2}) \|v\|_{H^1(K)}.
\end{aligned}$$

We use now (4.23) for  $u_1 = \chi(\cdot/\varepsilon) \cdot \nabla u_0$  (as defined in (4.16), where  $\chi \equiv (\chi_1, \chi_2)$ ):

$$\begin{aligned}
(4.24) \quad |u_1|_{H^{1/2}(\partial K)} &\leq C (H^{-1} \|\chi(\cdot/\varepsilon)\|_{L^\infty(\partial K)} \\
&\quad + H^{-1/2} \|\chi(\cdot/\varepsilon)\|_{L^\infty(\partial K)}^{1/2} \|\nabla \chi(\cdot/\varepsilon)\|_{L^\infty(\partial K)}^{1/2}) \|\nabla u_0\|_{H^1(K)} \\
&\leq C (H^{-1} + \varepsilon^{-1/2} H^{-1/2}) \|u_0\|_{H^2(K)},
\end{aligned}$$

where in the last step we also used  $\|\chi(\cdot/\varepsilon)\|_{L^\infty} \leq C$  and  $\|\nabla\chi(\cdot/\varepsilon)\|_{L^\infty(\partial K)} \leq C\varepsilon^{-1}$ , as  $\chi \in C^1([0, 1]^2)$ . This gives

$$III \leq C(\varepsilon H^{-1} + \varepsilon^2 H^{-2})\|u_0\|_{H^2(\Omega)}^2.$$

Because of (4.17),  $\theta$  verifies  $\|\theta\|_{H^1(\Omega)}^2 \leq C\|u_1\|_{H^{1/2}(\partial\Omega)}^2$ ; reasoning as before we also obtain  $\|u_1\|_{H^{1/2}(\partial\Omega)}^2 \leq C(1 + \varepsilon^{-1})\|u_0\|_{H^2(\Omega)}^2$  and we end up with

$$IV = C \sum_{K \in \mathcal{T}_H} |\varepsilon\theta|_{H^{1/2}(\partial K)}^2 \leq C\|\varepsilon\theta\|_{H^1(\Omega)}^2 \leq C(\varepsilon + \varepsilon^2)\|u_0\|_{H^2(\Omega)}^2.$$

Recalling  $\|u_0\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}$ , which follows from the elliptic regularity property of (4.15), we get (4.18).  $\square$

The final error estimate for  $\varepsilon$ ,  $H$ , and  $h$  varying is obtained by taking (4.13) and (4.18) together. Notice that when  $\varepsilon \geq h^{2/3}H^{1/3}$  (4.13) is sharper; for  $\varepsilon \leq h^{2/3}H^{1/3}$  (4.18) is sharper, and, in this case, (4.18) simplifies to  $\|u - u_H\|_{H^1(\Omega)} \leq C(h + \varepsilon^{1/2}H^{-1/2})\|f\|_{L^2}$ ; so in general we have

$$(4.25) \quad \|u - u_H\|_{H^1(\Omega)} \leq C \min\{h\varepsilon^{-1}, h + \varepsilon^{1/2}H^{-1/2}\}\|f\|_{L^2}.$$

Our theoretical analysis shows that the RFB method (which corresponds to  $\mathcal{T}_H \equiv \mathcal{T}_h$ ) is affected by the *resonance error*: for  $\varepsilon \simeq h = H$  there is a loss of accuracy. RFMB behaves a little bit better, since the right-hand side of (4.25) is (at most) of order  $(h/H)^{1/3}$  uniformly w.r.t.  $\varepsilon$  (for what we have seen above, the largest  $\min\{h\varepsilon^{-1}, h + \varepsilon^{1/2}H^{-1/2}\}$  in (4.25) is for  $\varepsilon \simeq h^{2/3}H^{1/3}$ ). The situation is even better if the domain  $\Omega$  is, for example, a square, and the source term  $f$  is more regular, as stated below.

**THEOREM 4.6.** *Assume (4.12); if  $C_1H^3 \leq h \leq C_2H$ , any internal angle of  $\Omega$  is smaller than  $\pi/(s+1)$ , and  $f \in H^s$  for some  $s > 1/2$ , then the numerical solution  $u_H$  of the RFMB method (3.3) verifies*

$$(4.26) \quad \|u - u_H\|_{H^1(\Omega)} \leq Ch^{1/2}H^{-1/2}.$$

*Proof.* We consider here the case  $h \leq \varepsilon \leq (hH)^{1/2}$ , because otherwise (4.26) just follows from (4.25). We reason as in the proof of Theorem 4.5 and split the error as

$$(4.27) \quad \begin{aligned} \|u - u_H\|_{H^1(\Omega)}^2 &\leq C \sum_{K \in \mathcal{T}_H} |u - u_0 - \varepsilon(u_1 - \theta)|_{H^{1/2}(\partial K)}^2 \\ &\quad + C \sum_{K \in \mathcal{T}_H} |u_0 - u_0^I|_{H^{1/2}(\partial K)}^2 \\ &\quad + C \sum_{K \in \mathcal{T}_H} |\varepsilon(u_1 - u_1^I)|_{H^{1/2}(\partial K)}^2 \\ &\quad + C \sum_{K \in \mathcal{T}_H} |\varepsilon\theta|_{H^{1/2}(\partial K)}^2 \\ &= I + II + III + IV; \end{aligned}$$

the only difference between (4.19) and (4.27) is in term *III*, which here contains  $u_1^I$ , the nodal interpolant of  $u_1$  on  $\bigcup \partial K$ . The classical result states  $|u_1 - u_1^I|_{H^{1/2}(\partial K)} \leq Ch^{1/2}|u_1|_{H^1(\partial K)}$ ; actually  $|u_1|_{H^1(\partial K)}$  can be estimated as in the proof of Theorem 4.5:

from (4.21), since on the reference macroelement  $\|\hat{v}\|_{H^1(\partial\mathbf{K})} \leq \|\hat{v}\|_{H^{1+s}(\mathbf{K})}$ , using scaling arguments, we obtain

$$H^{1/2}\|wv\|_{H^1(\partial K)} \leq C(H^{-1}\|w\|_{L^\infty(\partial K)} + \|\nabla w\|_{L^\infty(\partial K)})\|v\|_{H^{1+s}(K)};$$

when applied to  $u_1 = \chi(\cdot/\varepsilon) \cdot \nabla u_0$ , the last inequality gives

$$\|u_1\|_{H^1(\partial K)} \leq C(H^{-3/2} + \varepsilon^{-1}H^{-1/2})\|u_0\|_{H^{2+s}(K)};$$

therefore

$$\begin{aligned} (4.28) \quad III &\leq C \sum_{K \in \mathcal{T}_H} h(\varepsilon^2 H^{-3} + H^{-1})\|u_0\|_{H^{2+s}(K)}^2 \\ &\leq ChH^{-1}\|u_0\|_{H^{2+s}(\Omega)}^2 \\ &\leq ChH^{-1}\|f\|_{H^s(\Omega)}^2; \end{aligned}$$

in the last step we used  $\|u_0\|_{H^{2+s}(\Omega)} \leq C\|f\|_{H^s(\Omega)}$  (see [2]). By our condition on  $h$  and  $\varepsilon$ , we also have

$$I + II + IV \leq ChH^{-1}\|f\|_{L^2(\Omega)},$$

and (4.26) is proved.  $\square$

**5. Numerical tests.** In order to implement a computer algorithm for the formulation proposed in sections 2 and 3 we first need to introduce a fine-scale discretization of the spaces  $V_B$  or  $V_{MB}$ . For those purposes consider a third admissible, shape regular, and quasi-uniform partition  $\mathcal{T}_{h^\sharp}$  of mesh-size  $h^\sharp$  ( $h^\sharp \leq h$ ); we assume  $\mathcal{T}_{h^\sharp}$  to be a subpartition of the previous ones  $\mathcal{T}_h$  and  $\mathcal{T}_H$ . Define the space  $V^\sharp$  of functions which are continuous and piecewise linear (or bilinear) on  $\mathcal{T}_{h^\sharp}$ , and the spaces  $V_B^\sharp := V_B \cap V^\sharp$  and  $V_{MB}^\sharp := V_{MB} \cap V^\sharp$ ; accordingly, define  $V_h^\sharp \equiv V_h^\sharp(\mathcal{T}_h, \mathcal{T}_{h^\sharp}) := V_P \oplus V_B^\sharp \equiv V_h \cap V^\sharp$  and  $V_H^\sharp \equiv V_H^\sharp(\mathcal{T}_H, \mathcal{T}_h, \mathcal{T}_{h^\sharp}) := V_P + V_{MB}^\sharp \equiv V_H \cap V^\sharp$ . Using the spaces  $V_B^\sharp$  and  $V_{MB}^\sharp$  in place of  $V_B$  and  $V_{MB}$ , respectively, we have from sections 2 and 3 fully discrete formulations and we can derive in the usual way the numerical algorithms. In what follows we analyze the new error term due to the fine-scale discretizations of bubble spaces. Let us denote by  $u^\sharp \in V^\sharp$  the plain Galerkin approximation of  $u$  on the fine mesh  $\mathcal{T}_{h^\sharp}$ ,

$$a(u^\sharp, v) = \langle f, v \rangle \quad \forall v \in V^\sharp,$$

and denote by  $u_H^\sharp$  the numerical approximation given by the fully discrete RFMB scheme, i.e.,

$$a(u_H^\sharp, v) = \langle f, v \rangle \quad \forall v \in V_H^\sharp.$$

We can prove that

$$(5.1) \quad \|u^\sharp - u_H^\sharp\|_{H^1(\Omega)}^2 \simeq \inf_{v \in V_P} \sum_{K \in \mathcal{T}_H} |u^\sharp - v|_{H^{1/2}(\partial K)}^2.$$

Indeed (5.1) is analogous to (4.9) when we take the space  $V^\sharp$  instead of  $H_0^1(\Omega)$ : one can reason as in Proposition 4.3, using now Lemma 4.2. Therefore we have, using the

triangle inequality and (4.1),

$$\begin{aligned} \|u^\sharp - u_H^\sharp\|_{H^1(\Omega)}^2 &\simeq \inf_{v \in V_P} \sum_{K \in \mathcal{T}_H} |u^\sharp - v|_{H^{1/2}(\partial K)}^2 \\ &\leq C \left( \sum_{K \in \mathcal{T}_H} |u^\sharp - u|_{H^{1/2}(\partial K)}^2 + \inf_{v \in V_P} \sum_{K \in \mathcal{T}_H} |u - v|_{H^{1/2}(\partial K)}^2 \right) \\ &\leq C \left( |u^\sharp - u|_{H^1(\Omega)}^2 + \inf_{v \in V_P} \sum_{K \in \mathcal{T}_H} |u - v|_{H^{1/2}(\partial K)}^2 \right) \end{aligned}$$

and, eventually, we have for the fully discrete method the error estimate

$$\|u - u_H^\sharp\|_{H^1(\Omega)}^2 \leq C_1 |u^\sharp - u|_{H^1(\Omega)}^2 + C_2 \inf_{v \in V_P} \sum_{K \in \mathcal{T}_H} |u - v|_{H^{1/2}(\partial K)}^2 = I + II.$$

We have decoupled the numerical error into the term  $I$ , due to the approximation of local problems (depending on  $h^\sharp$ ) and the term  $II$ , due to the method itself (depending on  $h$  and  $H$ ). Note that  $I$  is the numerical error for the plain Galerkin approximation on the global fine mesh  $\mathcal{T}_{h^\sharp}$ . In case of periodic coefficients (assumption (4.12)),  $II$  has been analyzed in section 4; the new term  $I$  verifies  $I \leq h^\sharp \varepsilon^{-1} \|f\|_{L^2(\Omega)}$ , thanks to the classical argument. This also tell us what accuracy is needed when solving local problems: taking, for example,  $h^\sharp \leq \min\{h, \varepsilon h + \varepsilon^{3/2} H^{-1/2}\}$ , then the fully discrete method verifies the analogous of estimates (4.25).

In what follows, we test the methods on different model problems on the unit square  $\Omega := [0, 1]^2$ . In all the cases we take isotropic and diagonal  $\mathbf{a}$  (i.e.,  $a_{11} = a_{22}$  and  $a_{12} = a_{21} = 0$ ). In the first case we have periodic coefficients

$$(5.2) \quad a_{11} = a_{22} = \frac{9}{2} \sin\left(\frac{\pi x_1}{\varepsilon}\right) \cos\left(\frac{\pi x_2}{\varepsilon}\right) + \frac{11}{2},$$

and the wavelength  $\varepsilon$  is 0.05; the coefficients take values between 1 and 10 (see the contour plot in Figure 1). We construct a uniform fine mesh  $\mathcal{T}_{h^\sharp}$  of  $512 \times 512 \times 2 = 524288$  triangles and 263169 vertexes; the mesh is fine enough to compute accurately the plain Galerkin approximation  $u^\sharp$  of the exact solution for the model problem; the  $u^\sharp$  is plotted in Figure 2. We test the RFB method on a uniform coarse grid  $\mathcal{T}_h$  of  $2^i \times 2^i$  quadrilateral elements for  $i = 2, 3, 4, \dots, 7$ . For what concerns the RFMB method, we also introduce a uniform grid  $\mathcal{T}_H$  of quadrilaterals, and we take each macroelement  $K \in \mathcal{T}_H$  containing  $2 \times 2 = 4$  elements (referred to as  $RFMB_4$ ) or  $4 \times 4 = 16$  elements (referred to as  $RFMB_{16}$ ). The  $H^1$  relative error (when compared to  $u^\sharp$ ) is shown in Figure 3. We also plot the errors for the plain Galerkin method and the MsFEM on the same coarse meshes  $\mathcal{T}_h$ . We recall that we are considering here the MsFEM without the oversampling technique for its similarity to the RFB method.

For the second test we still have the coefficients of (5.2), but now we let  $\varepsilon$  vary from  $1/2$  up to  $1/128$ , and we compute the numerical errors of the methods on a fixed mesh  $\mathcal{T}_h$  of  $16 \times 16$  quadrilateral elements. The relative errors are plotted in Figure 4.

Then we consider nonperiodic coefficients that are given by the expression

$$(5.3) \quad a_{11} = a_{22} = 1 + \sum_{i=1}^4 \varphi(\gamma_i),$$

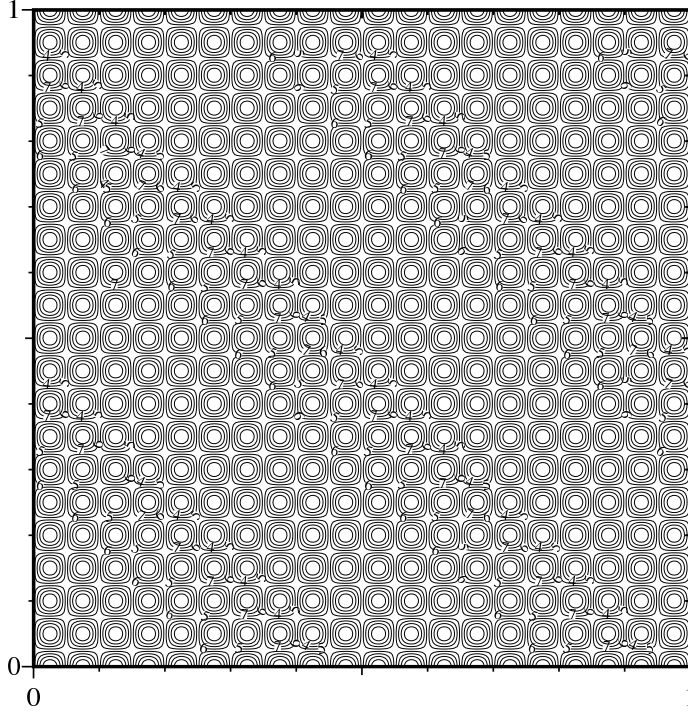


FIG. 1. Contour plot of the periodic coefficients  $a_{11} = a_{22}$ ; levels 2, 3, ..., 9 are shown.

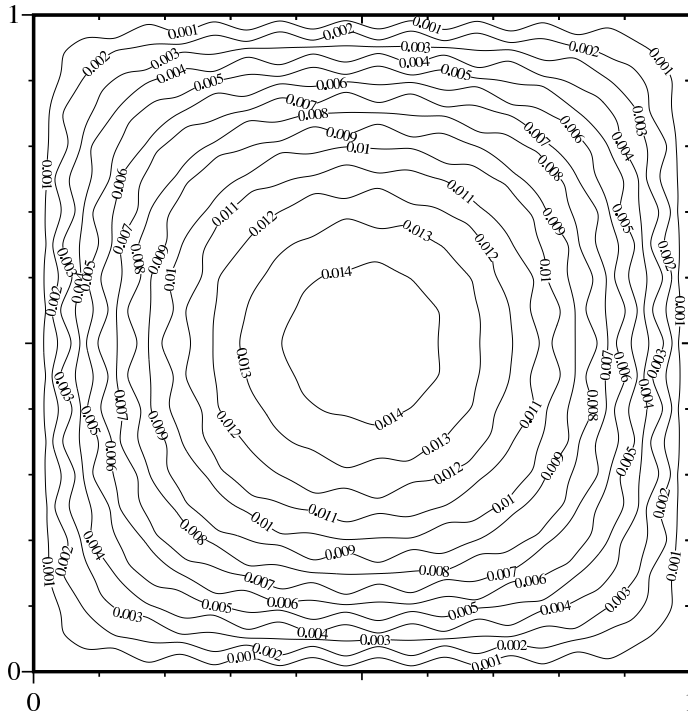
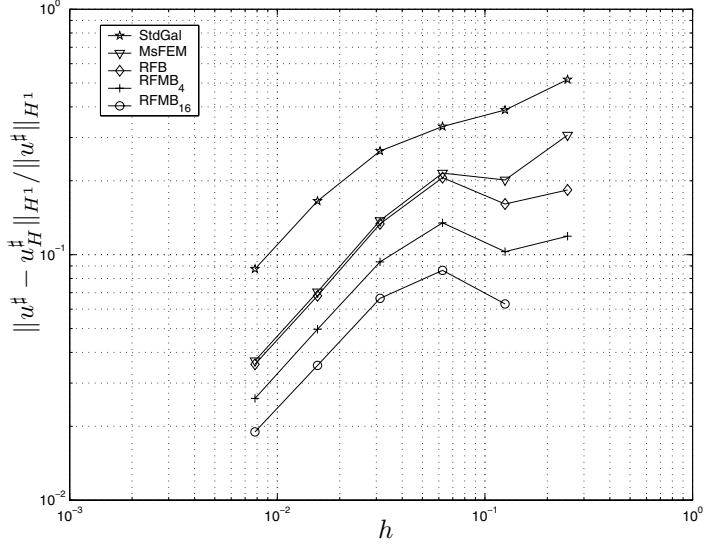
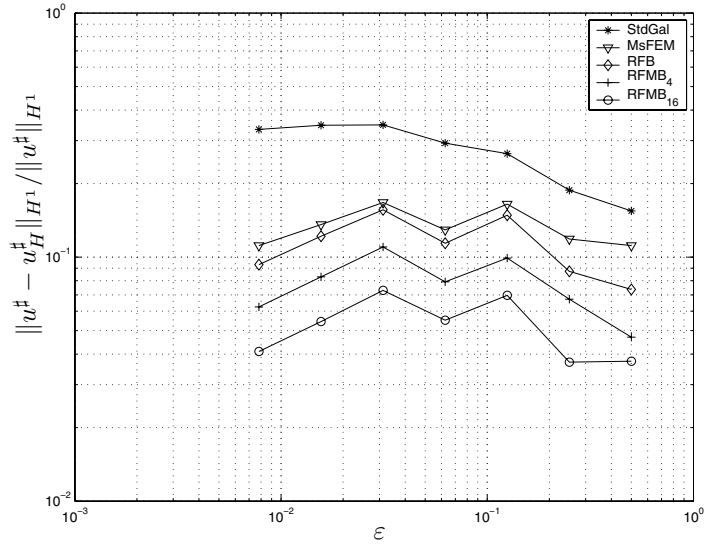


FIG. 2. Contour plot of the solution  $u$  for periodic coefficients; levels  $1 \times 10^{-3}, 2 \times 10^{-4}, \dots, 14 \times 10^{-4}$  are shown.

FIG. 3. *Relative  $H^1$ -error vs.  $h$  for periodic coefficients.*FIG. 4. *Relative  $H^1$ -error vs.  $\varepsilon$  for periodic coefficients.*

where  $\varphi(\gamma_i) := (30 - 25\gamma_i) (\cos(10\pi\gamma_i^2))^2$ ,  $\gamma_1 := \sin(\pi\delta_1)$ ,  $\gamma_2 := (\cos(\pi\delta_2))^2$ ,  $\gamma_3 := (\cos(2\pi\delta_3))^2$ ,  $\gamma_4 := \sin(2\pi\delta_4)$ , and  $\delta_i \equiv \delta_i(\mathbf{x}) := 2^{-1/2}\|\mathbf{x} - \mathbf{v}_i\|$ , with  $\mathbf{v}_i, i = 1, \dots, 4$ , being the four vertexes of  $\Omega$ . Here the smallest length scale is of order  $10^{-2}$ , while the ratio between the maximum and the minimum is  $10^2$  (see the plot of  $a_{11} = a_{22}$  in Figure 5). The fine grid approximation is plotted in Figure 6. We test the numerical methods on the same grids proposed above (i.e.,  $\mathcal{T}_h$  of  $2^i \times 2^i$  quadrilateral elements for  $i = 2, 3, 4, \dots, 7$ ); the numerical errors are plotted in Figure 7.

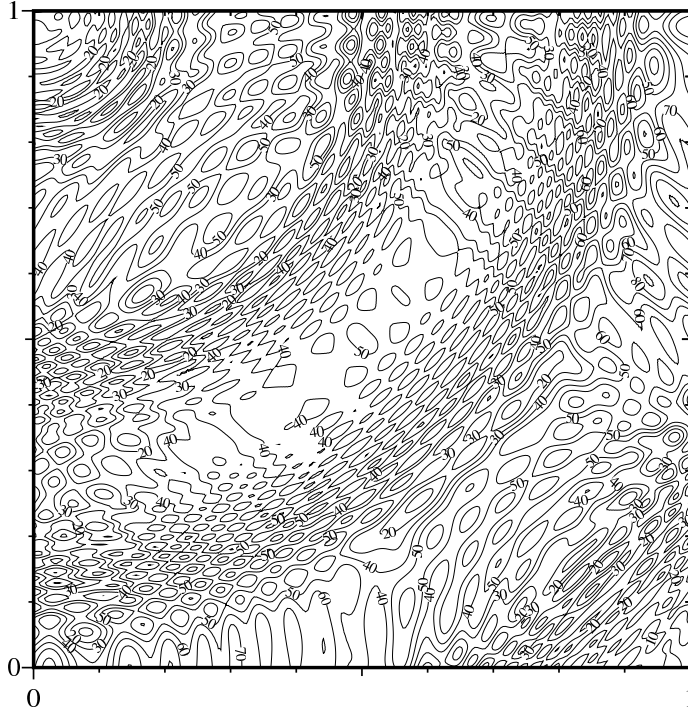


FIG. 5. Contour plot of the nonperiodic coefficients  $a_{11} = a_{22}$ ; levels 10, 20,  $\dots$ , 100 are shown.

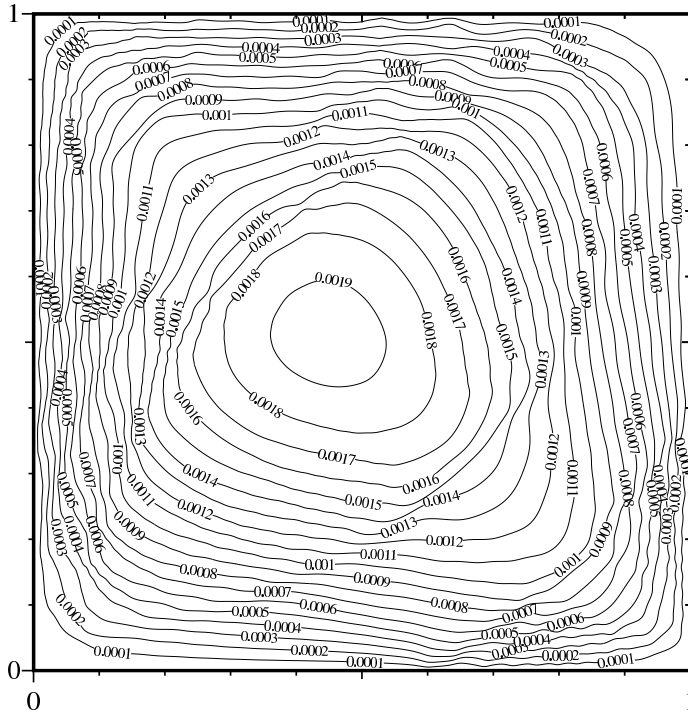


FIG. 6. Contour plot of the solution  $u$  for nonperiodic coefficients; levels  $1 \times 10^{-4}$ ,  $2 \times 10^{-4}$ ,  $\dots$ ,  $19 \times 10^{-4}$  are shown.

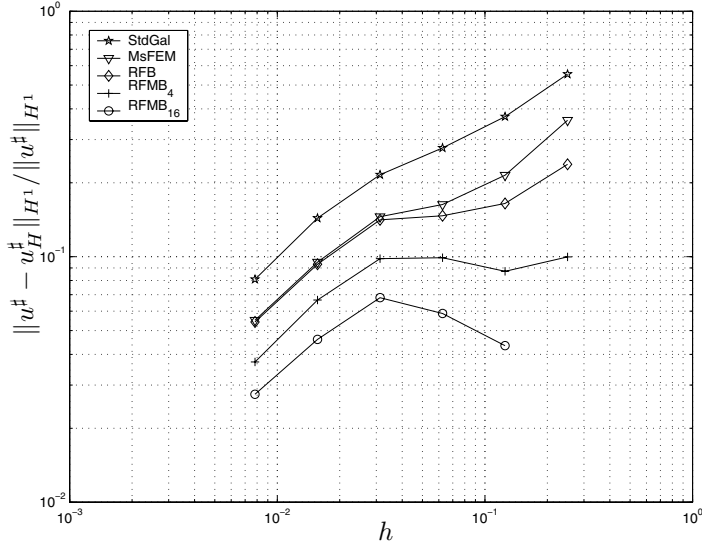


FIG. 7. Relative  $H^1$ -error for nonperiodic coefficients.

Let us discuss first the numerical results for the case of periodic coefficients presented in Figure 3. We clearly see that the plain Galerkin method gives the worst accuracy on all the meshes: the error starts decreasing with the expected first order rate only when the mesh-size is smaller than the scale  $\varepsilon$ . For the other methods, we also see the first order error decreasing for small  $h$ , but the actual error is smaller, in particular for larger  $h$ . All those methods suffer some resonance error, as we see that the maximum error is reached when  $h \approx 0.05$ , which is indeed the length scale of the coefficient oscillations. We see that RFB and MsFEM give almost the same accuracy when  $h$  is small, while on the coarser meshes RFB is more accurate. The RFMB method gives the same qualitative behavior of RFB, but the actual numerical error is smaller. From the practical point of view, if we are using the RFB approach on a coarse mesh and we want to increase the accuracy without using a very small  $h$ , we can use larger macrobubbles.

Figure 4 confirms the better behavior of the RFB and RFMB approaches, compared to the plain Galerkin approach, especially for small  $\varepsilon$ .

The same behavior of the methods is seen in the nonperiodic coefficients test (Figure 7), even though now the resonance error is less located because the coefficients and the solution itself contain many length scales. As before, we can reach a better accuracy on coarse grids enlarging the bubbles.

**6. Conclusion and extensions.** In this paper we considered the RFB method applied to problem (1.1). We also proposed an extension of the method (i.e., RFMB) based on macrobubbles (which live inside many elements). Following [15], we developed the theoretical analysis assuming periodic coefficients. Both the theoretical analysis and the numerical tests confirmed the validity of the methods.

The methods proposed and analyzed in the paper decouple the global problems into local problems on fine grids and a global problem on a coarse grid. In principle, this has two advantages compared to the plain Galerkin method (on a whole fine grid): it easily allows either for a parallel implementation or for upscaling procedures;

those aspects have not been discussed in the paper and deserve further investigations. In particular we have seen that larger macrobubbles (for RFMB) give more accuracy, though at the price of having larger local problems to solve. A following paper (in preparation) will investigate more convenient numerical algorithms based on this technique, also discussing the benefits in terms of CPU time.

In the paper we considered the macroelements  $K \in \mathcal{T}_H$  to be either triangles or quadrangles; this assumption can be relaxed with minor modifications in the analysis: we just need Lemmas 4.1 and 4.2 for the macroelements, and indeed general polygons satisfying suitable shape regularity conditions are allowed as well.

## REFERENCES

- [1] C. BAIocchi, F. BREZZI, AND L. P. FRANCA, *Virtual bubbles and Galerkin-least-squares type methods (Ga.L.S.)*, Comput. Methods Appl. Mech. Engrg., 105 (1993), pp. 125–141.
- [2] S. BERTOLUZZA, *Mesh Desing for the Subdomain Solvers Arising in Non Conforming Steklov-Poincaré Discretizations*, manuscript, 2002.
- [3] J. H. BRAMBLE, J. E. PASCIAK, J. P. WANG, AND J. XU, *Convergence estimates for product iterative methods with applications to domain decomposition*, Math. Comp., 57 (1991), pp. 1–21.
- [4] F. BREZZI, M. O. BRISTEAU, L. P. FRANCA, M. MALLEt, AND G. ROGÉ, *A relationship between stabilized finite element methods and the Galerkin method with bubble functions*, Comput. Methods Appl. Mech. Engrg., 96 (1992), pp. 117–129.
- [5] F. BREZZI, L. P. FRANCA, T. J. R. HUGHES, AND A. RUSSO,  $b = \int g$ , Comput. Methods Appl. Mech. Engrg., 145 (1997), pp. 329–339.
- [6] F. BREZZI, D. MARINI, AND A. RUSSO, *Applications of the pseudo residual-free bubbles to the stabilization of convection-diffusion problems*, Comput. Methods Appl. Mech. Engrg., 166 (1998), pp. 51–63.
- [7] F. BREZZI, D. MARINI, AND E. SÜLI, *Residual-free bubbles for advection-diffusion problems: The general error analysis*, Numer. Math., 85 (2000), pp. 31–47.
- [8] F. BREZZI AND L. D. MARINI, *Augmented spaces, two-level methods, and stabilizing subgrids*, Internat. J. Numer. Methods Fluids, 40 (2002), pp.31–46.
- [9] F. BREZZI AND A. RUSSO, *Choosing bubbles for advection-diffusion problems*, Math. Models Methods Appl. Sci., 4 (1994), pp. 571–587.
- [10] T. F. CHAN AND T. P. MATHEW, *Domain Decomposition Algorithms*, Acta Numer., 3, Cambridge University Press, Cambridge, UK, 1994, pp. 61–143.
- [11] Y. R. EFENDIEV, T. Y. HOU, AND X.-H. WU, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 888–910.
- [12] C. L. FARMER, *Upscaling: A review*, in Proceedings of the Institute of Computational Fluid Dynamics Conference on Numerical Methods for Fluid Dynamics, Oxford, UK, 2001.
- [13] L. P. FRANCA AND A. RUSSO, *Approximation of the Stokes problem by residual-free macro bubbles*, East-West J. Numer. Math., 4 (1996), pp. 265–278.
- [14] L. P. FRANCA AND A. RUSSO, *Deriving upwinding, mass lumping and selective reduced integration by residual-free bubbles*, Appl. Math. Lett., 9 (1996), pp. 83–88.
- [15] T. Y. HOU, X.-H. WU, AND Z. CAI, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913–943.
- [16] J.-L. LIONS AND E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications. Vol. I*, Springer-Verlag, New York, 1972.
- [17] S. MOSKOW AND M. VOGELIUS, *First-order corrections to the homogenized eigenvalues of a periodic composite medium. A convergence proof*, Proc. Roy. Soc. Edinburgh Sect. A, 127 (1997), pp. 1263–1299.
- [18] G. SANGALLI, *Global and local error analysis for the residual-free bubbles method applied to advection-dominated problems*, SIAM J. Numer. Anal., 38 (2000), pp. 1496–1522.
- [19] H. TRIEBEL, *Interpolation Theory, Function Spaces, Differential Operators*, 2nd ed., Johann Ambrosius Barth, Heidelberg, 1995.