A GENERALIZED FINITE ELEMENT METHOD FOR LINEAR THERMOELASTICITY

Axel Målqvist[1,*] and Anna Persson[1]

Abstract. We propose and analyze a generalized finite element method designed for linear quasistatic thermoelastic systems with spatial multiscale coefficients. The method is based on the local orthogonal decomposition technique introduced by Målqvist and Peterseim (*Math. Comp.* **83** (2014) 2583–2603). We prove convergence of optimal order, independent of the derivatives of the coefficients, in the spatial H^1-norm. The theoretical results are confirmed by numerical examples.

1. Introduction

In many applications the expansion and contraction of a material exposed to temperature changes are of great importance. To model this phenomenon a system consisting of an elasticity equation describing the displacement coupled with an equation for the temperature is used, see, *e.g.*, [6]. The full system consists of a hyperbolic elasticity equation coupled with a parabolic equation for the temperature, see [9] for a comprehensive treatment of this formulation. If the inertia effects are negligible, the hyperbolic term in the elasticity equation can be removed. This leads to an elliptic-parabolic system, often referred to as *quasistatic*. This formulation is discussed in, for instance [24, 27]. In some settings it is justified to also remove the parabolic term, which leads to an elliptic-elliptic system, see, *e.g.*, [24, 27]. Since the thermoelastic problem is formally equivalent to the system describing poroelasticity, several papers on this equation are also relevant, see, *e.g.*, [5, 26].

In this paper we study the quasistatic case. Existence and uniqueness of a solution to this system are discussed in [24] within the framework of linear degenerate evolution equations in Hilbert spaces. It is also shown that this system is essentially of parabolic type. Existence and uniqueness are also treated in [27] (only two-dimensional problems) and in [23, 25] some results on the thermoelastic contact problem are presented. The classical finite element method for the thermoelastic system is analyzed in [11, 27], where convergence rates of optimal order are derived for problems with solution in H^2 or higher.

When the elastic medium of interest is strongly heterogeneous, like composite materials, the coefficients are highly varying and oscillating. Commonly, such coefficients are said to have *multiscale* features. For these

[1] Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Göteborg, Sweden. peanna@chalmers.se

problems classical polynomial finite elements, as in [11,27], fail to approximate the solution well unless the mesh width resolves the data variations. This is due to the fact that *a priori* bounds of the error depend on (at least) the spatial H^2-norm of the solution. Since this norm depends on the derivative of the diffusion coefficient, it is of order ϵ^{-1} if the coefficient oscillates with frequency ϵ^{-1}. To overcome this difficulty, several numerical methods have been proposed, see for instance [3,4,15,16,18].

In this paper we suggest a generalized finite element method based on the techniques introduced in [18], often referred to as *local orthogonal decomposition*. This method builds on ideas from the variational multiscale method [15,16], where the solution space is split into a coarse and a fine part. The coarse space is modified such that the basis functions contain information from the diffusion coefficient and have support on small patches. With this approach the basis functions have good approximation properties locally. In [18] the technique is applied to elliptic problems with an arbitrary positive and bounded diffusion coefficient. One of the main advantages is that no assumptions on scale separation or periodicity of the coefficient are needed. Recently, this technique has been applied to several other problems, for instance, semilinear elliptic equations [13], boundary value problems [12], eigenvalue problems [19], linear and semilinear parabolic equations [17], and the linear wave equation [1].

The method we propose in this paper uses generalized finite element spaces similar to those used in [14,18], together with a correction building on the ideas in [12,16]. We prove convergence of optimal order that does not depend on the derivatives of the coefficients. We emphasize that by avoiding these derivatives, the *a priori* bound does not contain any constant of order ϵ^{-1}, although coefficients are highly varying.

In Section 2 we formulate the problem of interest, in Section 3 we first recall the classical finite element method for thermoelasticity and then we define the new generalized finite element method. In Section 4 we perform a localization of the basis functions and in Section 5 we analyze the error. Finally, in Section 6 we present some numerical results.

2. Problem formulation

Let $\Omega \subseteq \mathbb{R}^d$, $d = 2, 3$, be a polygonal/polyhedral domain describing the reference configuration of an elastic body. For a given time $T > 0$ we let $u : [0, T] \times \Omega \to \mathbb{R}^d$ denote the displacement field and $\theta : [0, T] \times \Omega \to \mathbb{R}$ the temperature. To impose Dirichlet and Neumann boundary conditions, we let Γ_D^u and Γ_N^u denote two disjoint segments of the boundary such that $\Gamma := \partial\Omega = \Gamma_D^u \cup \Gamma_N^u$. In addition, we assume $\text{meas}(\Gamma_D^u) > 0$. The segments Γ_D^θ and Γ_N^θ are defined similarly.

We use $(\cdot, \cdot)$ to denote the inner product in $L_2(\Omega)$ and $\|\cdot\|$ for the corresponding norm. Let $H^1(\Omega)$ denote the classical Sobolev space with norm $\|v\|_{H^1(\Omega)}^2 = \|v\|^2 + \|\nabla v\|^2$ and let $H^{-1}(\Omega)$ denote the dual space to H^1. Furthermore, we adopt the notation $L_p([0, T]; X)$ for the Bochner space with the norm

$$\|v\|_{L_p([0,T];X)} = \left(\int_0^T \|v\|_X^p \, \mathrm{dt} \right)^{1/p}, \quad 1 \leq p < \infty,$$

$$\|v\|_{L_\infty([0,T];X)} = \operatorname*{ess\,sup}_{0 \leq t \leq T} \|v\|_X,$$

where X is a Banach space equipped with the norm $\|\cdot\|_X$. The notation $v \in H^1([0, T]; X)$ is used to denote $v, \dot{v} \in L_2([0, T]; X)$. The dependence on the interval $[0, T]$ and the domain Ω is frequently suppressed and we write, for instance, $L_2(L_2)$ for $L_2([0, T]; L_2(\Omega))$. We also define the following subspaces of H^1

$$V^1 := \{v \in (H^1(\Omega))^d : v = 0 \text{ on } \Gamma_D^u\}, \quad V^2 := \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma_D^\theta\}.$$

Under the assumption that the displacement gradients are small, the (linearized) strain tensor is given by:

$$\varepsilon(u) = \frac{1}{2}(\nabla u + \nabla u^\mathsf{T}).$$

Assuming further that the material is isotropic, Hooke's law gives the (total) stress tensor, see *e.g.* [23] and the references therein,

$$\bar{\sigma} = 2\mu\varepsilon(u) + \lambda(\nabla \cdot u)I - \alpha\theta I,$$

where I is the d-dimensional identity matrix, α is the thermal expansion coefficient, and μ and λ are the so called Lamé coefficients given by:

$$\mu = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)},$$

where E denotes Young's elastic modulus and ν denotes Poisson's ratio. The materials of interest are strongly heterogeneous which implies that α, μ, and λ are rapidly varying in space.

The linear quasistatic thermoelastic problem takes the form

$$-\nabla \cdot (2\mu\varepsilon(u) + \lambda\nabla \cdot uI - \alpha\theta I) = f, \quad \text{in } (0,T] \times \Omega, \tag{2.1}$$

$$\dot{\theta} - \nabla \cdot \kappa\nabla\theta + \alpha\nabla \cdot \dot{u} = g, \quad \text{in } (0,T] \times \Omega, \tag{2.2}$$

$$u = 0, \quad \text{in } (0,T] \times \Gamma_D^u, \tag{2.3}$$

$$\bar{\sigma} \cdot n = 0, \quad \text{in } (0,T] \times \Gamma_N^u. \tag{2.4}$$

$$\theta = 0, \quad \text{on } (0,T] \times \Gamma_D^\theta, \tag{2.5}$$

$$\kappa\nabla\theta \cdot n = 0, \quad \text{on } (0,T] \times \Gamma_N^\theta. \tag{2.6}$$

$$\theta(0) = \theta_0, \quad \text{in } \Omega, \tag{2.7}$$

where κ is the heat conductivity parameter, which is assumed to be rapidly varying in space.

Remark 2.1. For simplicity we have assumed homogeneous boundary data (2.3)–(2.6). However, using techniques similar to the ones used in [12, 14] the analysis in this paper can be extended to non-homogeneous situations.

Assumption 2.2. We make the following assumptions on the data

(A1) $\kappa \in L_\infty(\Omega, \mathbb{R}^{d \times d})$, symmetric,

$$0 < \kappa_1 := \underset{x\in\Omega}{\operatorname{ess\,inf}} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\kappa(x)v \cdot v}{v \cdot v}, \quad \infty > \kappa_2 := \underset{x\in\Omega}{\operatorname{ess\,sup}} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\kappa(x)v \cdot v}{v \cdot v},$$

(A2) $\mu, \lambda, \alpha \in L_\infty(\Omega, \mathbb{R})$, and

$$0 < \mu_1 := \underset{x\in\Omega}{\operatorname{ess\,inf}} \mu(x) \le \underset{x\in\Omega}{\operatorname{ess\,sup}} \mu(x) =: \mu_2 < \infty.$$

Similarly, the constants $\lambda_1, \lambda_2, \alpha_1$, and α_2 are used to denote the corresponding upper and lower bounds for λ and α.

(A3) $f, \dot{f} \in L_\infty(L_2), \ddot{f} \in L_\infty(H^{-1}), g \in L_\infty(L_2), \dot{g} \in L_\infty(H^{-1})$, and $\theta_0 \in V^2$.

To pose a variational form we multiply the equations (2.1) and (2.2) with test functions from V^1 and V^2 and using Green's formula together with the boundary conditions (2.3)–(2.6) we arrive at the following weak formulation [11]. Find $u(t, \cdot) \in V^1$ and $\theta(t, \cdot) \in V^2$, such that,

$$(\sigma(u) : \varepsilon(v_1)) - (\alpha\theta, \nabla \cdot v_1) = (f, v_1), \quad \forall v_1 \in V^1, \tag{2.8}$$

$$(\dot{\theta}, v_2) + (\kappa\nabla\theta, \nabla v_2) + (\alpha\nabla \cdot \dot{u}, v_2) = (g, v_2), \quad \forall v_2 \in V^2, \tag{2.9}$$

and the initial value $\theta(0, \cdot) = \theta_0$ is satisfied. Here we use σ to denote the effective stress tensor $\sigma(u) := 2\mu\varepsilon(u) + \lambda(\nabla \cdot u)I$ and : to denote the Frobenius inner product of matrices. Using Korn's (first) inequality we have the following bounds, see, *e.g.*, [7,8],

$$c_\sigma \|v_1\|_{H^1}^2 \leq (\sigma(v_1) : \epsilon(v_1)) \leq C_\sigma \|v_1\|_{H^1}^2, \quad \forall v_1 \in V^1 \tag{2.10}$$

where c_σ (resp. C_σ) depends on μ_1 (resp. μ_2 and λ_2). Similarly, there are constants c_κ (resp. C_κ) depending on the bound κ_1 (resp. κ_2) such that

$$c_\kappa \|v_2\|_{H^1}^2 \leq (\kappa\nabla v_2, \nabla v_2) \leq C_\kappa \|v_2\|_{H^1}^2, \quad \forall v_2 \in V^2. \tag{2.11}$$

Furthermore, we use the following notation for the energy norms induced by the bilinear forms

$$\|v_1\|_\sigma^2 := (\sigma(v_1) : \varepsilon(v_1)), \ v_1 \in V^1, \quad \|v_2\|_\kappa^2 := (\kappa\nabla v_2\nabla v_2), \ v_2 \in V^2.$$

Existence and uniqueness of a solution to (2.8) and (2.9) have been proved in [24,27]. There are also some papers on the solution to contact problems, see [2,25].

Theorem 2.3. *Assume that* (A1)–(A3) *hold and that $\partial\Omega$ is sufficiently smooth. Then there exist u and θ such that $u \in L_2(V^1)$, $\nabla \cdot \dot{u} \in L_2(H^{-1})$, $\theta \in L_2(V^2)$, and $\dot{\theta} \in L_2(H^{-1})$ satisfying* (2.8) *and* (2.9) *and the initial condition $\theta(0, \cdot) = \theta_0$.*

Remark 2.4. We remark that the equations (2.1)–(2.7) also describe a poroelastic system and, hence, the method proposed in this paper also applies to problems in poroelasticity. In this case θ denotes the fluid pressure, κ the permeability and viscosity of the fluid.

3. NUMERICAL APPROXIMATION

In this section we first recall some properties of the classical finite element method for (2.8) and (2.9). In Section 3.2 we propose a new numerical method built on the ideas from [18]. The localization of this method is treated in Section 4.

3.1. Classical finite element

First, we need to define appropriate finite element spaces. For this purpose we let $\{\mathcal{T}_h\}_{h>0}$ be a family of shape regular triangulations of Ω with the mesh size $h_K := \operatorname{diam}(K)$, for $K \in \mathcal{T}_h$. Furthermore, we denote the largest diameter in the triangulation by $h := \max_{K\in\mathcal{T}_h} h_K$. We now define the classical piecewise affine finite element spaces

$$S_h^1 = \{v \in (C(\bar\Omega))^d : v_i|_K \text{ is a polynomial of degree} \leq 1, \ 1 \leq i \leq d, \ \forall K \in \mathcal{T}_h\},$$
$$S_h^2 = \{v \in C(\bar\Omega) : v|_K \text{ is a polynomial of degree} \leq 1, \forall K \in \mathcal{T}_h\},$$

and $V_h^1 = S_h^1 \cap V^1$, $V_h^2 = S_h^2 \cap V^2$.

For the discretization in time we consider, for simplicity, a uniform time step τ such that $t_n = n\tau$ for $n \in \{0, 1, \ldots, N\}$ and $N\tau = T$. The classical finite element method with a backward Euler scheme in time reads; for $n \in \{1, \ldots, N\}$ find $u_h^n \in V_h^1$ and $\theta_h^n \in V_h^2$, such that

$$(\sigma(u_h^n) : \varepsilon(v_1)) - (\alpha\theta_h^n, \nabla \cdot v_1) = (f^n, v_1), \quad \forall v_1 \in V_h^1, \tag{3.1}$$
$$(\bar\partial_t\theta_h^n, v_2) + (\kappa\nabla\theta_h^n, \nabla v_2) + (\alpha\nabla \cdot \bar\partial_t u_h^n, v_2) = (g^n, v_2), \quad \forall v_2 \in V_h^2, \tag{3.2}$$

where $\bar\partial_t\theta_h^n := (\theta_h^n - \theta_h^{n-1})/\tau$ and similarly for $\bar\partial_t u_h^n$. The right hand sides are evaluated at time t_n, that is, $f^n := f(t_n)$ and $g^n := g(t_n)$. Given initial data u_h^0 and θ_h^0 the system (3.1) and (3.2) is well posed [11]. We assume

that $\theta_h^0 \in V_h^1$ is a suitable approximation of θ_0. For u_h^0 we note that $u(0)$ is uniquely determined by (2.8) at $t = 0$, that is, $u(0)$ fulfills the equation

$$(\sigma(u(0)) : \varepsilon(v_1)) - (\alpha\theta^0, \nabla \cdot v_1) = (f^0, v_1), \quad \forall v_1 \in V^1,$$

and we thus define $u_h^0 \in V_h^1$ to be the solution to

$$(\sigma(u_h^0) : \varepsilon(v_1)) - (\alpha\theta_h^0, \nabla \cdot v_1) = (f^0, v_1), \quad \forall v_1 \in V_h^1. \tag{3.3}$$

The following theorem is a consequence of ([11], Thm. 3.1). The convergence rate is optimal for the two first norms. However, it is not optimal for the L_2-norm $\|\theta^n - \theta_h^n\|$. In [11] this is avoided by using second order continuous piecewise polynomials for the displacement (P2-P1 elements). It is, however, noted that the problem is still stable using P1-P1 elements. In this paper we use P1-P1 elements and derive error bounds in the $L_\infty(H^1)$-norm, of optimal order, for both the displacement and the temperature.

Theorem 3.1. *Let u and θ be the solution to (2.8) and (2.9) and $\{u_h^n\}_{n=1}^N$ and $\{\theta_h^n\}_{n=1}^N$ be the solution to (3.1) and (3.2). For $n \in \{1,\ldots,N\}$ we have*

$$\|u^n - u_h^n\|_{H^1} + \left(\sum_{m=1}^n \tau\|\theta^m - \theta_h^m\|_{H^1}^2\right)^{1/2} + \|\theta^n - \theta_h^n\| \le C_{\epsilon^{-1}}h + C\tau,$$

where $C_{\epsilon^{-1}}$ is of order ϵ^{-1} if the material varies on a scale of size ϵ.

Note that the constant $C_{\epsilon^{-1}}$ involved in this error bound contains derivatives of the coefficients. Hence, convergence only takes place when the mesh size h is sufficiently small ($h < \epsilon$). Throughout this paper, it is assumed that h is small enough and V_h^1 and V_h^2 are referred to as reference spaces for the solution. Similarly, u_h^n and θ_h^n are referred to as reference solutions. In Section 5 this solution is compared with the generalized finite element solution. We emphasize that the generalized finite element solution is computed in spaces of lower dimension and hence not as computationally expensive. It should also be noted that there, of course, is an initial cost of generating the generalized finite element space itself, but this space can be reused in each time step.

Remark 3.2. The classical linear elasticity equation can suffer from (Poisson) locking effects when using continuous piecewise linear polynomials (P1 elements) and ν is close to $1/2$. In [14], the GFEM that this paper is based on is applied to linear elasticity equations and it is shown that the Poisson locking effects is reduced. It is reasonable to believe that this property is transferred to the method presented here, but we shall, to simplify the analysis, refrain from investigating this. Instead we assume that the Lamé coefficients are of moderate size, such that the reference solution is a good approximation.

Furthermore, the coupled time-dependent problem can suffer from another type of locking, occurring if $\dot{\theta}$ is neglected in (2.2) and P1 elements are used in both spaces (P1-P1). The locking produces artificial oscillations in the numerical approximation of the temperature (or pressure) for early time steps. However, when $\dot{\theta}$ is not neglected, numerical observations indicate that this locking effect does not occur. See [21, 22] and references therein for a further discussion. Hence, we shall consider a P1-P1 discretization for the reference solution in this paper.

In the following theorem we prove some regularity results for the finite element solution.

Theorem 3.3. *Let $\{u_h^n\}_{n=1}^N$ and $\{\theta_h^n\}_{n=1}^N$ be the solution to (3.1) and (3.2). Then the following bound holds*

$$\left(\sum_{j=1}^n \tau\|\bar{\partial}_t u_h^j\|_{H^1}^2\right)^{1/2} + \left(\sum_{j=1}^n \tau\|\bar{\partial}_t \theta_h^j\|^2\right)^{1/2} + \|\theta_h^n\|_{H^1} \le C(\|g\|_{L_\infty(L_2)} + \|\dot{f}\|_{L_\infty(H^{-1})} + \|\theta_h^0\|_{H^1}). \tag{3.4}$$

If $\theta_h^0 = 0$, then for $n \in \{1, \ldots, N\}$

$$\|\bar{\partial}_t u_h^n\|_{H^1} + \|\bar{\partial}_t \theta_h^n\| + \left(\sum_{j=1}^n \tau \|\bar{\partial}_t \theta_h^j\|_{H^1}^2 \right)^{1/2}$$
$$\leq C \left(\|g\|_{L_\infty(L_2)} + \|\dot{g}\|_{L_\infty(H^{-1})} + \|\dot{f}\|_{L_\infty(H^{-1})} + \|\ddot{f}\|_{L_\infty(H^{-1})} \right). \tag{3.5}$$

If $f = 0$ and $g = 0$, then for $n \in \{1, \ldots, N\}$

$$\|\bar{\partial}_t u_h^n\|_{H^1} + \|\bar{\partial}_t \theta_h^n\| + t_n^{1/2} \|\bar{\partial}_t \theta_h^n\|_{H^1} \leq C t_n^{-1/2} \|\theta_h^0\|_{H^1}. \tag{3.6}$$

Proof. From (3.1) and (3.2) and the initial data (3.3) we deduce that the following relation must hold for $n \geq 1$

$$(\sigma(\bar{\partial}_t u_h^n) : \varepsilon(v_1)) - (\alpha \bar{\partial}_t \theta_h^n, \nabla \cdot v_1) = (\bar{\partial}_t f^n, v_1), \quad \forall v_1 \in V_h^1, \tag{3.7}$$
$$(\bar{\partial}_t \theta_h^n, v_2) + (\kappa \nabla \theta_h^n, \nabla v_2) + (\alpha \nabla \cdot \bar{\partial}_t u_h^n, v_2) = (g^n, v_2), \quad \forall v_2 \in V_h^2. \tag{3.8}$$

By choosing $v_1 = \bar{\partial}_t u_h^n$ and $v_2 = \bar{\partial}_t \theta_h^n$ and adding the resulting equations we have

$$\frac{c_\sigma}{2} \|\bar{\partial}_t u_h^n\|_{H^1}^2 + \frac{1}{2} \|\bar{\partial}_t \theta_h^n\|^2 + (\kappa \nabla \theta_h^n, \nabla \bar{\partial}_t \theta_h^n) \leq C(\|g^n\|^2 + \|\bar{\partial}_t f^n\|_{H^{-1}}^2). \tag{3.9}$$

Note that the coupling terms cancel. By using Cauchy–Schwarz and Young's inequality we can bound

$$\tau(\kappa \nabla \theta_h^n, \nabla \bar{\partial}_t \theta_h^n) = \|\kappa^{1/2} \nabla \theta_h^n\|^2 - (\kappa \nabla \theta_h^n, \nabla \theta_h^{n-1}) \geq \frac{1}{2} \|\theta_h^n\|_\kappa^2 - \frac{1}{2} \|\theta_h^{n-1}\|_\kappa^2.$$

Multiplying (3.9) by τ, summing over n, and using (2.11) gives

$$\sum_{j=1}^n \tau \|\bar{\partial}_t u_h^j\|_{H^1}^2 + \sum_{j=1}^n \tau \|\bar{\partial}_t \theta_h^j\|^2 + \|\theta_h^n\|_{H^1}^2 \leq C \sum_{j=1}^n \tau(\|g^j\|^2 + \|\bar{\partial}_t f^j\|_{H^{-1}}^2) + C \|\theta_h^0\|_{H^1},$$

which is bounded by the right hand side in (3.4).

For the bound (3.5) we note that the following relation must hold for $n \geq 2$

$$(\sigma(\bar{\partial}_t u_h^n) : \varepsilon(v_1)) - (\alpha \bar{\partial}_t \theta_h^n, \nabla \cdot v_1) = (\bar{\partial}_t f^n, v_1), \quad \forall v_1 \in V_h^1, \tag{3.10}$$
$$(\bar{\partial}_t^2 \theta_h^n, v_2) + (\kappa \nabla \bar{\partial}_t \theta_h^n, \nabla v_2) + (\alpha \nabla \cdot \bar{\partial}_t^2 u_h^n, v_2) = (\bar{\partial}_t g^n, v_2), \quad \forall v_2 \in V_h^2. \tag{3.11}$$

Now choose $v_1 = \bar{\partial}_t^2 u_h^n$ and $v_2 = \bar{\partial}_t \theta_h^n$ and add the resulting equations to get

$$(\sigma(\bar{\partial}_t u_h^n) : \varepsilon(\bar{\partial}_t^2 u_h^n)) + (\bar{\partial}_t^2 \theta_h^n, \bar{\partial}_t \theta_h^n) + (\kappa \nabla \bar{\partial}_t \theta_h^n, \nabla \bar{\partial}_t \theta_h^n) = (\bar{\partial}_t f^n, \bar{\partial}_t^2 u_h^n) + (\bar{\partial}_t g^n, \bar{\partial}_t \theta_h^n).$$

Multiplying by τ and using Cauchy–Schwarz and Young's inequality gives

$$\frac{1}{2} \|\bar{\partial}_t u_h^n\|_\sigma^2 + \frac{1}{2} \|\bar{\partial}_t \theta_h^n\|^2 + C\tau \|\bar{\partial}_t \theta_h^n\|_{H^1}^2 \leq \frac{1}{2} \|\bar{\partial}_t \theta_h^{n-1}\|^2 + \frac{1}{2} \|\bar{\partial}_t u_h^{n-1}\|_\sigma^2 + \tau(\bar{\partial}_t f^n, \bar{\partial}_t^2 u_h^n) + C\|\bar{\partial}_t g^n\|_{H^{-1}}^2.$$

Summing over n and using (2.10) now gives

$$\|\bar{\partial}_t u_h^n\|_{H^1}^2 + \|\bar{\partial}_t \theta_h^n\|^2 + \sum_{j=2}^n \tau \|\bar{\partial}_t \theta_h^j\|_{H^1}^2 \leq C \left(\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \|\bar{\partial}_t \theta_h^1\|^2 + \sum_{j=2}^n \tau \left((\bar{\partial}_t f^j, \bar{\partial}_t^2 u_h^j) + \|\bar{\partial}_t g^j\|_{H^{-1}}^2 \right) \right).$$

Here we use summation by parts to get

$$\sum_{j=2}^{n} \tau(\bar{\partial}_t f^j, \bar{\partial}_t^2 u_h^j) = (\bar{\partial}_t f^n, \bar{\partial}_t u_h^n) - (\bar{\partial}_t f^1, \bar{\partial}_t u_h^1) - \sum_{j=2}^{n} \tau(\bar{\partial}_t^2 f^j, \bar{\partial}_t u_h^{j-1})$$

$$\leq C \left(\max_{1 \leq j \leq n} \|\bar{\partial}_t f^j\|_{H^{-1}} + \sum_{j=2}^{n} \tau \|\bar{\partial}_t^2 f^j\|_{H^{-1}} \right) \max_{1 \leq j \leq n} \|\bar{\partial}_t u_h^j\|_{H^1},$$

and $\max_{1 \leq j \leq n} \|\bar{\partial}_t u_h^j\|_{H^1}$ can now be kicked to the left hand side.

To estimate $\bar{\partial}_t \theta_h^1$ and $\bar{\partial}_t u_h^1$ we choose $v_1 = \bar{\partial}_t u_h^1$ and $v_2 = \bar{\partial}_t \theta_h^1$ in (3.7) and (3.8) for $n = 1$. We thus have, since $\theta_h^0 = 0$,

$$\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \|\bar{\partial}_t \theta_h^1\|^2 + \frac{1}{\tau}\|\theta_h^1\|_{H^1}^2 \leq C(\|\bar{\partial}_t f^1\|_{H^{-1}}^2 + \|g^1\|^2).$$

The observation that $\frac{1}{\tau}\|\theta_h^1\|_{H^1}^2 = \tau\|\bar{\partial}_t \theta_h^1\|_{H^1}^2$ completes the bound (3.5).

Now assume $f = 0$ and $g = 0$ and note that the following holds for $n \geq 2$,

$$(\sigma(\bar{\partial}_t^2 u_h^n) : \varepsilon(v_1)) - (\alpha \bar{\partial}_t^2 \theta_h^n, \nabla \cdot v_1) = 0, \quad \forall v_1 \in V_h^1,$$
$$(\bar{\partial}_t^2 \theta_h^n, v_2) + (\kappa \nabla \bar{\partial}_t \theta_h^n, \nabla v_2) + (\alpha \nabla \cdot \bar{\partial}_t^2 u_h^n, v_2) = 0, \quad \forall v_2 \in V_h^2.$$

Choosing $v_1 = \bar{\partial}_t^2 u_h^n$, $v_2 = \bar{\partial}_t^2 \theta_h^n$ and adding the resulting equations gives

$$(\sigma(\bar{\partial}_t^2 u_h^n) : \varepsilon(\bar{\partial}_t^2 u_h^n)) + (\bar{\partial}_t^2 \theta_h^n, \bar{\partial}_t^2 \theta_h^n) + (\kappa \nabla \bar{\partial}_t \theta_h^n, \nabla \bar{\partial}_t^2 \theta_h^n) = 0,$$

where, again, the coupling terms cancel. The two first terms on the left hand side are positive and can thus be ignored. Multiplying by τ and t_n^2 gives after using Cauchy–Schwarz and Young's inequality

$$t_n^2 \|\bar{\partial}_t \theta_h^n\|_\kappa^2 \leq t_{n-1}^2 \|\bar{\partial}_t \theta_h^{n-1}\|_\kappa^2 + (t_n^2 - t_{n-1}^2)\|\bar{\partial}_t^2 \theta_h^{n-1}\|_\kappa^2.$$

Note that $t_n^2 - t_{n-1}^2 \leq 3\tau t_{n-1}$, where we use that $t_n \leq 2t_{n-1}$ if $n \geq 2$. Summing over n now gives

$$t_n^2 \|\bar{\partial}_t \theta_h^n\|_\kappa^2 \leq t_1^2 \|\bar{\partial}_t \theta_h^1\|_\kappa^2 + 3 \sum_{j=2}^{n} \tau t_{j-1} \|\bar{\partial}_t \theta_h^{j-1}\|_\kappa^2.$$

To bound the last sum we choose $v_1 = \bar{\partial}_t^2 u_h^n$, $v_2 = \bar{\partial}_t \theta_h^n$ in (3.10) and (3.11), now with $f = 0$ and $g = 0$. Adding the resulting equations gives

$$(\bar{\partial}_t^2 \theta_h^n, \bar{\partial}_t \theta_h^n) + (\kappa \nabla \bar{\partial}_t \theta_h^n, \nabla \bar{\partial}_t \theta_h^n) + (\sigma(\bar{\partial}_t u_h^n) : \varepsilon(\bar{\partial}_t^2 u_h^n)) = 0,$$

Multiplying by τ and t_n gives after using Cauchy–Schwarz inequality

$$\frac{t_n}{2}\|\bar{\partial}_t u_h^n\|_\sigma^2 + \frac{t_n}{2}\|\bar{\partial}_t \theta_h^n\|^2 + c_\kappa \tau t_n \|\bar{\partial}_t \theta_h^n\|_{H^1}^2 \leq \frac{t_{n-1}}{2}\|\bar{\partial}_t u_h^{n-1}\|_\sigma^2 + \frac{t_{n-1}}{2}\|\bar{\partial}_t \theta_h^{n-1}\|^2 + \frac{\tau}{2}\|\bar{\partial}_t u_h^{n-1}\|_\sigma^2 + \frac{\tau}{2}\|\bar{\partial}_t \theta_h^{n-1}\|^2.$$

Summing over n and using (2.10) thus gives

$$\frac{c_\sigma t_n}{2}\|\bar{\partial}_t u_h^n\|_{H^1}^2 + \frac{t_n}{2}\|\bar{\partial}_t \theta_h^n\|^2 + \sum_{j=2}^{n} \tau t_j \|\bar{\partial}_t \theta_h^j\|_{H^1}^2 \leq \frac{C_\sigma t_1}{2}\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \frac{t_1}{2}\|\bar{\partial}_t \theta_h^1\|^2$$

$$+ C \sum_{j=2}^{n} \tau \left(\|\bar{\partial}_t u_h^{j-1}\|_{H^1}^2 + \|\bar{\partial}_t \theta_h^{j-1}\|^2 \right).$$

To bound the last sum in this estimate we choose $v_1 = \bar{\partial}_t u_h^n$, $v_2 = \bar{\partial}_t \theta_h^n$ in (3.7)–(3.8) and multiply by τ to get

$$c_\sigma \tau \|\bar{\partial}_t u_h^n\|_{H^1}^2 + \tau \|\bar{\partial}_t \theta_h^n\|^2 + \frac{1}{2}\|\theta_h^n\|_\kappa^2 \leq \frac{1}{2}\|\theta_h^{n-1}\|_\kappa^2.$$

Summing over n and using (2.11) gives

$$C \sum_{j=1}^n \tau \left(\|\bar{\partial}_t \theta_h^j\|^2 + \|\bar{\partial}_t u_h^j\|_{H^1}^2 \right) + \frac{c_\kappa}{2}\|\theta_h^n\|_{H^1}^2 \leq \frac{C_\kappa}{2}\|\theta_h^0\|_{H^1}^2. \tag{3.12}$$

It remains to bound $t_1^2\|\bar{\partial}_t \theta_h^1\|_{H^1}^2$, $t_1\|\bar{\partial}_t \theta_h^1\|^2$, and $t_1\|\bar{\partial}_t u_h^1\|_{H^1}$. For this purpose we recall that $t_1 = \tau$ and use (3.12) for $n = 1$ to get

$$t_1\|\bar{\partial}_t u_h^1\|_{H^1} + t_1\|\bar{\partial}_t \theta_h^1\|^2 + t_1^2\|\bar{\partial}_t \theta_h^1\|_{H^1}^2 \leq C(\tau(\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \|\bar{\partial}_t \theta_h^1\|^2) + \|\theta_h^1\|_{H^1}^2 + \|\theta_h^0\|_{H^1}^2) \leq C\|\theta_h^0\|_{H^1}^2.$$

Finally, we have that

$$t_n\|\bar{\partial}_t u_h^n\|_{H^1}^2 + t_n\|\bar{\partial}_t \theta_h^n\|^2 \leq C\|\theta_h^0\|_{H^1}^2, \quad t_n^2\|\bar{\partial}_t \theta_h^n\|_{H^1}^2 \leq C\|\theta_h^0\|_{H^1}^2,$$

and thus (3.6) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

3.2. Generalized finite element

In this section we shall derive a generalized finite element method. First we define V_H^1 and V_H^2 analogously to V_h^1 and V_h^2, but with a larger mesh size $H > h$. In addition, we assume that the family of triangulations $\{\mathcal{T}_H\}_{H>h}$ is quasi-uniform and that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$ such that $V_H^1 \subseteq V_h^1$ and $V_H^2 \subseteq V_h^2$. Furthermore, we use the notation $\mathcal{N} = \mathcal{N}^1 \times \mathcal{N}^2$ to denote the free nodes in $V_H^1 \times V_H^2$. The aim is now to define a new (multiscale) space with the same dimension as $V_H^1 \times V_H^2$, but with better approximation properties. For this purpose we define an interpolation operator $I_H = (I_H^1, I_H^2) : V_h^1 \times V_h^2 \to V_H^1 \times V_H^2$ with the property that $I_H \circ I_H = I_H$ and for all $v = (v_1, v_2) \in V_h^1 \times V_h^2$

$$H_K^{-1}\|v - I_H v\|_{L_2(K)} + \|\nabla I_H v\|_{L_2(K)} \leq C_I \|\nabla v\|_{L_2(\omega_K)}, \quad \forall K \in \mathcal{T}_H, \tag{3.13}$$

where

$$\omega_K := \text{int } \{\hat{K} \in \mathcal{T}_H : \hat{K} \cap K \neq \emptyset\}.$$

Note that I_H^1 is vector-valued. Since the mesh is assumed to be shape regular, the estimates in (3.13) are also global, *i.e.*,

$$H^{-1}\|v - I_H v\| + \|\nabla I_H v\| \leq C\|\nabla v\|, \tag{3.14}$$

where C is a constant depending on the shape regularity parameter, $\gamma > 0$;

$$\gamma := \max_{K \in \mathcal{T}_H} \gamma_K, \text{ with } \gamma_K := \frac{\text{diam } B_K}{\text{diam } K}, \text{ for } K \in \mathcal{T}_H, \tag{3.15}$$

where B_K is the largest ball contained in K.

One example of an interpolation that satisfies the above assumptions is $I_H^i = E_H^i \circ \Pi_H^i$, $i = 1, 2$. Here Π_H^i denotes the piecewise L_2-projection onto $P_1(\mathcal{T}_H)$ $((P_1(\mathcal{T}_H))^d$ if $i = 1)$, the space of functions that are affine on each triangle $K \in \mathcal{T}_H$. Furthermore, E_H^1 is an averaging operator mapping $(P_1(\mathcal{T}_H))^d$ into V_H^1, by (coordinate wise)

$$(E_H^{1,j}(v))(z) = \frac{1}{\text{card}\{K \in \mathcal{T}_H : z \in K\}} \sum_{K \in \mathcal{T}_H : z \in K} v_j|_K(z), \quad 1 \leq j \leq d,$$

where $z \in \mathcal{N}^1$. E_H^2 mapping $P_1(\mathcal{T}_H)$ to V_H^2 is defined similarly. For a further discussion on this interpolation and other available options we refer to [20].

Let us now define the kernels of I_H^1 and I_H^2

$$V_{\mathrm{f}}^1 := \{v \in V_h^1 : I_H^1 v = 0\}, \quad V_{\mathrm{f}}^2 := \{v \in V_h^2 : I_H^2 v = 0\}$$

The kernels are fine scale spaces in the sense that they contain all features that are not captured by the (coarse) finite element spaces V_H^1 and V_H^2. Note that the interpolation leads to the splits $V_h^1 = V_H^1 \oplus V_{\mathrm{f}}^1$ and $V_h^2 = V_H^2 \oplus V_{\mathrm{f}}^2$, meaning that any function $v_1 \in V_h^1$ can be uniquely decomposed as $v_1 = v_{1,H} + v_{1,\mathrm{f}}$, with $v_{1,H} \in V_H^1$ and $v_{1,\mathrm{f}} \in V_{\mathrm{f}}^1$, and similarly for $v_2 \in V_h^2$.

Now, we introduce a Ritz projection onto the fine scale spaces. For this we use the bilinear forms associated with the diffusion in (2.8) and (2.9). The projection of interest is thus $R_{\mathrm{f}} : V_h^1 \times V_h^2 \to V_{\mathrm{f}}^1 \times V_{\mathrm{f}}^2$, such that for all $(v_1, v_2) \in V_h^1 \times V_h^2$, $R_{\mathrm{f}}(v_1, v_2) = (R_{\mathrm{f}}^1 v_1, R_{\mathrm{f}}^2 v_2)$ fulfills

$$(\sigma(v_1 - R_{\mathrm{f}}^1 v_1) : \varepsilon(w_1)) = 0, \quad \forall w_1 \in V_{\mathrm{f}}^1, \tag{3.16}$$

$$(\kappa \nabla(v_2 - R_{\mathrm{f}}^2 v_2), \nabla w_2) = 0, \quad \forall w_2 \in V_{\mathrm{f}}^2. \tag{3.17}$$

Note that this is an uncoupled system and R_{f}^1 and R_{f}^2 are classical Ritz projections.

For any $(v_1, v_2) \in V_h^1 \times V_h^2$ we have, due to the splits of the spaces V_h^1 and V_h^2 above, that

$$v_1 - R_{\mathrm{f}}^1 v_1 = v_{1,H} - R_{\mathrm{f}}^1 v_{1,H}, \quad v_2 - R_{\mathrm{f}}^2 v_2 = v_{2,H} - R_{\mathrm{f}}^2 v_{2,H}.$$

Using this we define the multiscale spaces

$$V_{\mathrm{ms}}^1 := \{v - R_{\mathrm{f}}^1 v : v \in V_H^1\}, \quad V_{\mathrm{ms}}^2 := \{v - R_{\mathrm{f}}^2 v : v \in V_H^2\}. \tag{3.18}$$

Clearly $V_{\mathrm{ms}}^1 \times V_{\mathrm{ms}}^2$ has the same dimension as $V_H^1 \times V_H^2$. Indeed, with $\lambda_x^{1,i}$ denoting a basis function in V_H^1 with a hat function at the ith position at node x, $1 \leq i \leq d$, and λ_y^2 the hat function in V_H^2 at node y, such that

$$V_H^1 \times V_H^2 = \mathrm{span}\{(\lambda_x^{1,i}, 0), (0, \lambda_y^2) : 1 \leq i \leq d, \, (x, y) \in \mathcal{N}\},$$

a basis for $V_{\mathrm{ms}}^1 \times V_{\mathrm{ms}}^2$ is given by:

$$\{(\lambda_x^{1,i} - R_{\mathrm{f}}^1 \lambda_x^{1,i}, 0), (0, \lambda_y^2 - R_{\mathrm{f}}^2 \lambda_y^2) : 1 \leq i \leq d, \, (x, y) \in \mathcal{N}\}. \tag{3.19}$$

Finally, we also note that the splits $V_h^1 = V_{\mathrm{ms}}^1 \oplus V_{\mathrm{f}}^1$ and $V_h^2 = V_{\mathrm{ms}}^2 \oplus V_{\mathrm{f}}^2$ hold, which fulfill the following orthogonality relations

$$(\sigma(v_1) : \varepsilon(w_1)) = 0, \quad \forall v_1 \in V_{\mathrm{ms}}^1, \, w_1 \in V_{\mathrm{f}}^1, \tag{3.20}$$

$$(\kappa \nabla v_2, \nabla w_2) = 0, \quad \forall v_2 \in V_{\mathrm{ms}}^2, \, w_2 \in V_{\mathrm{f}}^2. \tag{3.21}$$

3.2.1. Stationary problem

For the error analysis in Section 5 it is convenient to define the Ritz projection onto the multiscale space using the bilinear form given by the stationary version of (2.8) and (2.9). We thus define $R_{\mathrm{ms}} : V_h^1 \times V_h^2 \to V_{\mathrm{ms}}^1 \times V_{\mathrm{ms}}^2$, such that for all $(v_1, v_2) \in V_h^1 \times V_h^2$, $R_{\mathrm{ms}}(v_1, v_2) = (R_{\mathrm{ms}}^1(v_1, v_2), R_{\mathrm{ms}}^2 v_2)$ fulfills

$$(\sigma(v_1 - R_{\mathrm{ms}}^1(v_1, v_2)) : \varepsilon(w_1)) - (\alpha(v_2 - R_{\mathrm{ms}}^2 v_2), \nabla \cdot w_1) = 0, \quad \forall w_1 \in V_{\mathrm{ms}}^1, \tag{3.22}$$

$$(\kappa \nabla(v_2 - R_{\mathrm{ms}}^2 v_2), \nabla w_2) = 0, \quad \forall w_2 \in V_{\mathrm{ms}}^2. \tag{3.23}$$

Note that we must have $R_{\mathrm{ms}}^2 = I - R_{\mathrm{f}}^2$, but $R_{\mathrm{ms}}^1 \neq I - R_{\mathrm{f}}^1$ in general.

The Ritz projection in (3.22) and (3.23) is upper triangular. Hence, when solving for $R_{\mathrm{ms}}^1(v_1, v_2)$ the term $(\alpha R_{\mathrm{ms}}^2 v_2, \nabla \cdot w_1)$ in (3.22) is known. Since this term has multiscale features and appears on the right hand side,

we impose a correction on $R^1_{\mathrm{ms}}(v_1, v_2)$ inspired by the ideas in [12] and [16]. The correction is defined as the element $\tilde{R}_{\mathrm{f}} v_2 \in V^1_{\mathrm{f}}$, which fulfills

$$(\sigma(\tilde{R}_{\mathrm{f}} v_2) : \varepsilon(w_1)) = (\alpha R^2_{\mathrm{ms}} v_2, \nabla \cdot w_1), \quad \forall w_1 \in V^1_{\mathrm{f}}, \tag{3.24}$$

and we define $\tilde{R}^1_{\mathrm{ms}}(v_1, v_2) = R^1_{\mathrm{ms}}(v_1, v_2) + \tilde{R}_{\mathrm{f}} v_2$.

Note that the Ritz projections are stable in the sense that

$$\|\tilde{R}^1_{\mathrm{ms}}(v_1, v_2)\|_{H^1} \le C(\|v_1\|_{H^1} + \|v_2\|_{H^1}), \quad \|R^2_{\mathrm{ms}} v_2\|_{H^1} \le C\|v_2\|_{H^1}. \tag{3.25}$$

Remark 3.4. The problem to find $\tilde{R}_{\mathrm{f}} v_2$ is posed in the entire fine scale space and is thus computationally expensive to solve. The aim is to localize these computations to smaller patches of coarse elements, see Section 4.

To derive error bounds for this projection we define two operators $\mathcal{A}_1 : V^1_h \times V^2_h \to V^1_h$ and $\mathcal{A}_2 : V^2_h \to V^2_h$ such that for all $(v_1, v_2) \in V^1_h \times V^2_h$ we have

$$(\mathcal{A}_1(v_1, v_2), w_1) = (\sigma(v_1) : \varepsilon(w_1)) - (\alpha v_2, \nabla \cdot w_1), \quad \forall w_1 \in V^1_h, \tag{3.26}$$

$$(\mathcal{A}_2 v_2, w_2) = (\kappa \nabla v_2, \nabla w_2), \quad \forall w_2 \in V^2_h. \tag{3.27}$$

Lemma 3.5. *For all* $(v_1, v_2) \in V^1_h \times V^2_h$ *it holds that*

$$\|v_1 - \tilde{R}^1_{\mathrm{ms}}(v_1, v_2)\|_{H^1} \le C(H\|\mathcal{A}_1(v_1, v_2)\| + \|v_2 - R^2_{\mathrm{ms}} v_2\|) \tag{3.28}$$

$$\le CH(\|\mathcal{A}_1(v_1, v_2)\| + \|v_2\|_{H^1}),$$

$$\|v_2 - R^2_{\mathrm{ms}} v_2\|_{H^1} \le CH\|\mathcal{A}_2 v_2\|. \tag{3.29}$$

Proof. It follows from [18] that (3.29) holds, since (3.23) is an elliptic equation of Poisson type. Using an Aubin–Nitsche duality argument as in, *e.g.*, [17], we can derive the following estimate in the L_2-norm

$$\|v_2 - R^2_{\mathrm{ms}} v_2\| \le CH\|v_2 - R^2_{\mathrm{ms}} v_2\|_{H^1} \le CH\|v_2\|_{H^1},$$

which proves the second inequality in (3.28).

It remains to bound $\|v_1 - \tilde{R}^1_{\mathrm{ms}}(v_1, v_2)\|_{H^1}$. Recall that any $v \in V^1_h$ can be decomposed as

$$v = v - R^1_{\mathrm{f}} v + R^1_{\mathrm{f}} v = (I - R^1_{\mathrm{f}})v + R^1_{\mathrm{f}} v,$$

where $(I - R^1_{\mathrm{f}})v \in V^1_{\mathrm{ms}}$. Using the orthogonality (3.20) and that $(\sigma(\cdot) : \varepsilon(\cdot))$ is a symmetric bilinear form we get

$$(\sigma(\tilde{R}^1_{\mathrm{ms}}(v_1, v_2)) : \varepsilon(v)) = (\sigma(R^1_{\mathrm{ms}}(v_1, v_2) + \tilde{R}_{\mathrm{f}} v_2) : \varepsilon((I - R^1_{\mathrm{f}})v + R^1_{\mathrm{f}} v))$$

$$= (\sigma(R^1_{\mathrm{ms}}(v_1, v_2)) : \varepsilon((I - R^1_{\mathrm{f}})v)) + (\sigma(\tilde{R}_{\mathrm{f}} v_2) : \varepsilon(R^1_{\mathrm{f}} v)).$$

Due to (3.22) and (3.24) we thus have

$$(\sigma(R^1_{\mathrm{ms}}(v_1, v_2)) : \varepsilon((I - R^1_{\mathrm{f}})v)) + (\sigma(\tilde{R}_{\mathrm{f}} v_2) : \varepsilon(R^1_{\mathrm{f}} v))$$

$$= (\sigma(v_1) : \varepsilon((I - R^1_{\mathrm{f}})v)) - (\alpha(v_2 - R^2_{\mathrm{ms}} v_2), \nabla \cdot (I - R^1_{\mathrm{f}})v) + (\alpha R^2_{\mathrm{ms}} v_2, \nabla \cdot R^1_{\mathrm{f}} v)$$

$$= (\mathcal{A}_1(v_1, v_2), (I - R^1_{\mathrm{f}})v) + (\alpha R^2_{\mathrm{ms}} v_2, \nabla \cdot v).$$

Define $e := v_1 - \tilde{R}^1_{\mathrm{ms}}(v_1, v_2)$. Using the above relation together with (3.26) we get the bound

$$c_\sigma \|e\|^2_{H^1} \le (\sigma(e) : \varepsilon(e)) = (\sigma(v_1) : \varepsilon(e)) - (\mathcal{A}_1(v_1, v_2), (I - R^1_{\mathrm{f}})e) - (\alpha R^2_{\mathrm{ms}} v_2, \nabla \cdot e)$$

$$= (\mathcal{A}_1(v_1, v_2), R^1_{\mathrm{f}} e) + (\alpha(v_2 - R^2_{\mathrm{ms}} v_2), \nabla \cdot e)$$

$$\le \|\mathcal{A}_1(v_1, v_2)\| \|R^1_{\mathrm{f}} e\| + C\|v_2 - R^2_{\mathrm{ms}} v_2\| \|e\|_{H^1}.$$

Since $R^1_{\mathrm{f}} e \in V^1_{\mathrm{f}}$ we have due to (3.13)

$$\|R^1_{\mathrm{f}} e\| = \|R^1_{\mathrm{f}} e - I^1_H R^1_{\mathrm{f}} e\| \le CH\|R^1_{\mathrm{f}} e\|_{H^1} \le CH\|e\|_{H^1},$$

where we have used the stability $\|R^1_{\mathrm{f}} v\|_{H^1} \le C\|v\|_{H^1}$ for $v \in V^1_h$. The first inequality in (3.28) now follows. $\square$

Remark 3.6. Without the correction $\tilde{R}_{\mathrm{f}}$ the error bound (3.28) would depend on the derivatives of α,

$$\|v_1 - R_{\mathrm{ms}}^1(v_1, v_2)\|_{H^1} \leq C_{\alpha'}(H\|\mathcal{A}_1(v_1, v_2)\| + \|v_2 - R_{\mathrm{ms}}^2 v_2\|),$$

where α' is large if α has multiscale features.

3.2.2. Time-dependent problem

A generalized finite element method with a backward Euler discretization in time is now defined by replacing V_h^1 with V_{ms}^1 and V_h^2 with V_{ms}^2 in (3.1) and (3.2) and adding a correction similar to (3.24). The method thus reads; for $n \in \{1, \ldots, N\}$ find $\tilde{u}_{\mathrm{ms}}^n = u_{\mathrm{ms}}^n + u_{\mathrm{f}}^n$, with $u_{\mathrm{ms}}^n \in V_{\mathrm{ms}}^1$, $u_{\mathrm{f}}^n \in V_{\mathrm{f}}^1$, and $\theta_{\mathrm{ms}}^n \in V_{\mathrm{ms}}^2$, such that

$$(\sigma(\tilde{u}_{\mathrm{ms}}^n) : \varepsilon(v_1)) - (\alpha\theta_{\mathrm{ms}}^n, \nabla \cdot v_1) = (f^n, v_1), \quad \forall v_1 \in V_{\mathrm{ms}}^1, \tag{3.30}$$

$$(\bar{\partial}_t\theta_{\mathrm{ms}}^n, v_2) + (\kappa\nabla\theta_{\mathrm{ms}}^n, \nabla v_2) + (\alpha\nabla \cdot \bar{\partial}_t\tilde{u}_{\mathrm{ms}}^n, v_2) = (g^n, v_2), \quad \forall v_2 \in V_{\mathrm{ms}}^2, \tag{3.31}$$

$$(\sigma(u_{\mathrm{f}}^n) : \varepsilon(w_1)) - (\alpha\theta_{\mathrm{ms}}^n, \nabla \cdot w_1) = 0, \quad \forall w_1 \in V_{\mathrm{f}}^1. \tag{3.32}$$

where $\theta_{\mathrm{ms}}^0 = R_{\mathrm{ms}}^2\theta_h^0$. Furthermore, we define $\tilde{u}_{\mathrm{ms}}^0 := u_{\mathrm{ms}}^0 + u_{\mathrm{f}}^0$, where $u_{\mathrm{f}}^0 \in V_{\mathrm{f}}^1$ is defined by (3.32) for $n = 0$ and $u_{\mathrm{ms}}^0 \in V_{\mathrm{ms}}^1$, such that

$$(\sigma(\tilde{u}_{\mathrm{ms}}^0) : \varepsilon(v_1)) - (\alpha\theta_{\mathrm{ms}}^0, \nabla \cdot v_1) = (f^0, v_1), \quad \forall v_1 \in V_{\mathrm{ms}}^1. \tag{3.33}$$

Lemma 3.7. *The problem* (3.30) *and* (3.31) *is well-posed.*

Proof. Given u_{ms}^{n-1}, $\theta_{\mathrm{ms}}^{n-1}$, and u_{f}^{n-1}, equations (3.30)–(3.32) yields a square system. Hence, it is sufficient to prove that the solution is unique. Let $v_1 = u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1}$ in (3.30) and $v_2 = \tau\theta_{\mathrm{ms}}^n$ in (3.31) and add the resulting equations to get

$$(\sigma(u_{\mathrm{ms}}^n) : \varepsilon(u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1})) + (\sigma(u_{\mathrm{f}}^n) : \varepsilon(u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1})) + \tau(\bar{\partial}_t\theta_{\mathrm{ms}}^n, \theta_{\mathrm{ms}}^n) + c_\kappa\tau\|\theta_{\mathrm{ms}}^n\|_{H^1}^2 + (\alpha\nabla \cdot (u_{\mathrm{f}}^n - u_{\mathrm{f}}^{n-1}), \theta_{\mathrm{ms}}^n)$$
$$\leq (f^n, u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1}) + \tau(g^n, \theta_{\mathrm{ms}}^n).$$

Using the orthogonality (3.20) and (3.32) this simplifies to

$$(\sigma(u_{\mathrm{ms}}^n) : \varepsilon(u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1})) + \tau(\bar{\partial}_t\theta_{\mathrm{ms}}^n, \theta_{\mathrm{ms}}^n) + c_\kappa\tau\|\theta_{\mathrm{ms}}^n\|_{H^1}^2 + c_\sigma\|u_{\mathrm{f}}^n\|_{H^1}^2$$
$$\leq (f^n, u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1}) + \tau(g^n, \theta_{\mathrm{ms}}^n) + (\sigma(u_{\mathrm{f}}^n) : \varepsilon(u_{\mathrm{f}}^{n-1})).$$

Now, using that $(\sigma(\cdot) : \varepsilon(\cdot))$ is a symmetric bilinear form we get the following identity

$$(\sigma(v) : \varepsilon(v - w)) = \frac{1}{2}(\sigma(v) : \varepsilon(v)) + \frac{1}{2}(\sigma(v - w) : \varepsilon(v - w)) - \frac{1}{2}(\sigma(w) : \varepsilon(w)), \tag{3.34}$$

and using Cauchy–Schwarz and Young's inequality we derive

$$(f^n, u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1}) \leq C\|f^n\|_{H^{-1}} + \frac{1}{2}(\sigma(u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1}) : \varepsilon(u_{\mathrm{ms}}^n - u_{\mathrm{ms}}^{n-1})).$$

This, together with the estimate $\tau(\bar{\partial}_t\theta_{\mathrm{ms}}^n, \theta_{\mathrm{ms}}^n) \geq \frac{1}{2}\|\theta_{\mathrm{ms}}^n\|^2 - \frac{1}{2}\|\theta_{\mathrm{ms}}^{n-1}\|^2$ and (2.10), leads to

$$\frac{c_\sigma}{2}\|u_{\mathrm{ms}}^n\|_{H^1}^2 + \frac{1}{4}\|\theta_{\mathrm{ms}}^n\|^2 + c_\kappa\tau\|\theta_{\mathrm{ms}}^n\|_{H^1}^2 + \frac{c_\sigma}{2}\|u_{\mathrm{f}}^n\|_{H^1}^2$$
$$\leq C(\|f^n\|_{H^{-1}}^2 + \tau\|g^n\|^2 + \|\theta_{\mathrm{ms}}^{n-1}\|^2 + \|u_{\mathrm{ms}}^{n-1}\|_{H^1}^2 + \|u_{\mathrm{f}}^{n-1}\|_{H^1}^2).$$

Hence, a unique solution exists. $\qquad\square$

4. Localization

In this section we show how to truncate the basis functions, which is motivated by the exponential decay of (3.19). We consider a localization inspired by the one proposed in [12], which is performed by restricting the fine scale space to patches of coarse elements defined by the following; for $K \in \mathcal{T}_H$

$$\omega_0(K) := \text{int } K,$$
$$\omega_k(K) := \text{int } \left(\cup \{ \hat{K} \in \mathcal{T}_H : \hat{K} \cap \overline{\omega_{k-1}(K)} \neq \emptyset \} \right), \quad k = 1, 2, \dots$$

Now let $V_{\mathrm{f}}^1(\omega_k(K)) := \{ v \in V_{\mathrm{f}}^1 : v(z) = 0 \text{ on } (\overline{\Omega} \setminus \Gamma_N^u) \setminus \omega_k(K) \}$ be the restriction of V_{f}^1 to the patch $\omega_k(T)$. We define $V_{\mathrm{f}}^2(\omega_k(K))$ similarly.

The localized fine scale space can now be used to approximate the fine scale part of the basis functions in (3.19), which significantly reduces the computational cost for these problems. Let $(\cdot, \cdot)_\omega$ denote the L_2 inner product over a subdomain $\omega \subseteq \Omega$ and define the local Ritz projection $R_{\mathrm{f},k}^K : V_h^1 \times V_h^2 \to V_{\mathrm{f}}^1(\omega_k(K)) \times V_{\mathrm{f}}^2(\omega_k(K))$ such that for all $(v_1, v_2) \in V_h^1 \times V_h^2$, $R_{\mathrm{f},k}^K(v_1, v_2) = (R_{\mathrm{f},k}^{K,1} v_1, R_{\mathrm{f},k}^{K,2} v_1)$ fulfills

$$(\sigma(R_{\mathrm{f},k}^{K,1} v_1) : \varepsilon(w_1))_{\omega_k(K)} = (\sigma(v_1) : \varepsilon(w_1))_K, \quad \forall w_1 \in V_{\mathrm{f}}^1(\omega_k(K)), \tag{4.1}$$

$$(\kappa \nabla(R_{\mathrm{f},k}^{K,2} v_2), \nabla w_2)_{\omega_k(K)} = (\kappa \nabla v_2, \nabla w_2)_K, \quad \forall w_2 \in V_{\mathrm{f}}^2(\omega_k(K)). \tag{4.2}$$

Note that if we replace $\omega_k(K)$ with Ω in (4.1) and (4.2) and denote the resulting projection

$$R_{\mathrm{f}}^K(v_1, v_2) = (R_{\mathrm{f}}^{K,1} v_1, R_{\mathrm{f}}^{K,2} v_2),$$

then for all $(v_1, v_2) \in V_h^1 \times V_h^2$ we have

$$R_{\mathrm{f}}(v_1, v_2) = \sum_{K \in \mathcal{T}_H} R_{\mathrm{f}}^K(v_1, v_2) = \sum_{K \in \mathcal{T}_H} (R_{\mathrm{f}}^{K,1} v_1, R_{\mathrm{f}}^{K,2} v_2).$$

Motivated by this we now define the localized fine scale projection as

$$R_{\mathrm{f},k}(v_1, v_2) := \sum_{K \in \mathcal{T}_H} R_{\mathrm{f},k}^K(v_1, v_2) = \sum_{K \in \mathcal{T}_H} (R_{\mathrm{f},k}^{K,1} v_1, R_{\mathrm{f},k}^{K,2} v_2), \tag{4.3}$$

and the localized multiscale spaces

$$V_{\mathrm{ms},k}^1 := \{ v_1 - R_{\mathrm{f},k}^1 v_1 : v_1 \in V_H^1 \}, \quad V_{\mathrm{ms},k}^2 := \{ v_2 - R_{\mathrm{f},k}^2 v_2 : v_2 \in V_H^2 \}, \tag{4.4}$$

with the corresponding localized basis

$$\{ (\lambda_x^{1,i} - R_{\mathrm{f},k}^1 \lambda_x^{1,i}, 0), (0, \lambda_y^2 - R_{\mathrm{f},k}^2 \lambda_y^2) : 1 \leq i \leq d, (x, y) \in \mathcal{N} \}. \tag{4.5}$$

4.1. Stationary problem

In this section we define a localized version of the stationary problem (3.22) and (3.23). Let $R_{\mathrm{ms},k} : V_h^1 \times V_h^2 \to V_{\mathrm{ms},k}^1 \times V_{\mathrm{ms},k}^2$, such that for all $(v_1, v_2) \in V_h^1 \times V_h^2$, $R_{\mathrm{ms},k}(v_1, v_2) = (R_{\mathrm{ms},k}^1(v_1, v_2), R_{\mathrm{ms},k}^2 v_2)$. The method now reads; find

$$\tilde{R}_{\mathrm{ms},k}^1(v_1, v_2) = R_{\mathrm{ms},k}^1(v_1, v_2) + \sum_{K \in \mathcal{T}_H} \tilde{R}_{\mathrm{f},k}^K v_2, \quad \text{where } \tilde{R}_{\mathrm{f},k}^K v_2 \in V_{\mathrm{f}}^1(\omega_k(K)),$$

and $R^2_{\mathrm{ms},k}v_2$ such that

$$(\sigma(v_1 - \tilde{R}^1_{\mathrm{ms},k}(v_1,v_2)) : \varepsilon(w_1)) - (\alpha(v_2 - R^2_{\mathrm{ms},k}v_2), \nabla \cdot w_1) = 0, \quad \forall w_1 \in V^1_{\mathrm{ms},k}, \tag{4.6}$$

$$(\kappa\nabla(v_2 - R^2_{\mathrm{ms},k}v_2), \nabla w_2) = 0, \quad \forall w_2 \in V^2_{\mathrm{ms},k}. \tag{4.7}$$

$$(\sigma(\tilde{R}^K_{\mathrm{f},k}v_2) : \varepsilon(w)) - (\alpha R^2_{\mathrm{ms},k}v_2, \nabla \cdot w)_K = 0, \quad \forall w \in V^1_{\mathrm{f}}(w_k(K)). \tag{4.8}$$

Note that the Ritz projection is stable in the sense that

$$\|\tilde{R}^1_{\mathrm{ms},k}(v_1,v_2)\|_{H^1} \le C(\|v_1\|_{H^1} + \|v_2\|_{H^1}), \quad \|R^2_{\mathrm{ms},k}v_2\|_{H^1} \le C\|v_2\|_{H^1}. \tag{4.9}$$

The following two lemmas give a bound on the error introduced by the localization.

Lemma 4.1. *For all* $(v_1,v_2) \in V^1_h \times V^2_h$, *there exists* $\xi \in (0,1)$, *such that*

$$\|R^1_{\mathrm{f},k}v_1 - R^1_{\mathrm{f}}v_1\|^2_{H^1} \le Ck^d\xi^{2k} \sum_{K \in \mathcal{T}_H} \|R^{K,1}_{\mathrm{f}}v_1\|^2_{H^1}, \tag{4.10}$$

$$\|R^2_{\mathrm{f},k}v_2 - R^2_{\mathrm{f}}v_2\|^2_{H^1} \le Ck^d\xi^{2k} \sum_{K \in \mathcal{T}_H} \|R^{K,2}_{\mathrm{f}}v_2\|^2_{H^1}, \tag{4.11}$$

$$\|\tilde{R}_{\mathrm{f},k}v_2 - \tilde{R}_{\mathrm{f}}v_2\|^2_{H^1} \le Ck^d\xi^{2k} \sum_{K \in \mathcal{T}_H} \|\tilde{R}^K_{\mathrm{f}}v_2\|^2_{H^1}. \tag{4.12}$$

The bounds (4.10) and (4.11) are direct results from [14] and [12,18] respectively, while (4.12) follows from [14] by a slight modification of the right hand side. We omit the proof here.

The next lemma gives a bound for the localized Ritz projection.

Lemma 4.2. *For all* $(v_1,v_2) \in V^1_h \times V^2_h$ *there exist* $\xi \in (0,1)$ *such that*

$$\|v_1 - \tilde{R}^1_{\mathrm{ms},k}(v_1,v_2)\|_{H^1} \le C(H + k^{d/2}\xi^k)(\|\mathcal{A}_1(v_1,v_2)\| + \|v_2\|_{H^1}), \tag{4.13}$$

$$\|v_2 - R^2_{\mathrm{ms},k}v_2\|_{H^1} \le C(H + k^{d/2}\xi^k)\|\mathcal{A}_2 v_2\|. \tag{4.14}$$

Proof. It follows from [12] that (4.14) holds. To prove (4.13) we let $v_H \in V^1_H$ and $v_{H,k} \in V^1_H$ be elements such that

$$R^1_{\mathrm{ms}}(v_1,v_2) = v_H - R^1_{\mathrm{f}}v_H, \quad R^1_{\mathrm{ms},k}(v_1,v_2) = v_{H,k} - R^1_{\mathrm{f},k}v_{H,k}.$$

Define $e := v_1 - \tilde{R}^1_{\mathrm{ms},k}(v_1,v_2)$. From (4.6) and (4.7) we get have the following identity for any $z \in V^1_{\mathrm{ms},k}$

$$(\sigma(e) : \varepsilon(e)) - (\alpha(v_2 - R^2_{\mathrm{ms},k}v_2), \nabla \cdot e) = (\sigma(e) : \varepsilon(v_1 - z - \tilde{R}_{\mathrm{f},k}v_1)) - (\alpha(v_2 - R^2_{\mathrm{ms},k}v_2), \nabla \cdot (v_1 - z - \tilde{R}_{\mathrm{f},k}v_2)).$$

Using this with $z = v_H - R^1_{\mathrm{f},k}v_H \in V^1_{\mathrm{ms},k}$ we get

$$\begin{aligned} c_\sigma\|e\|^2_{H^1} \le (\sigma(e) : \varepsilon(e)) &= (\sigma(e) : \varepsilon(v_1 - v_H - R^1_{\mathrm{f},k}v_H - \tilde{R}_{\mathrm{f}}v_1)) \\ &\quad - (\alpha(v_2 - R^2_{\mathrm{ms},k}v_2), \nabla \cdot (v_1 - v_H - R^1_{\mathrm{f},k}v_H - \tilde{R}_{\mathrm{f},k}v_2)) \\ &\quad + (\alpha(v_2 - R^2_{\mathrm{ms},k}v_2), \nabla \cdot e). \end{aligned}$$

Now, using Cauchy–Schwarz and Young's inequality we get

$$\|e\|^2_{H^1} \le C(\|v_1 - v_H - R^1_{\mathrm{f},k}v_H - \tilde{R}_{\mathrm{f},k}v_2\|^2_{H^1} + \|v_2 - R^2_{\mathrm{ms},k}v_2\|^2),$$

where the last term is bounded in (4.14). For the first term we get

$$\|v_1 - v_H - R^1_{\mathrm{f},k}v_H - \tilde{R}_{\mathrm{f},k}v_2\|_{H^1} \leq \|v_1 - (v_H - R^1_{\mathrm{f}}v_H + \tilde{R}_{\mathrm{f}}v_2)\|_{H^1} + \|R^1_{\mathrm{f}}v_H - R^1_{\mathrm{f},k}v_H\|_{H^1} + \|\tilde{R}_{\mathrm{f}}v_2 - \tilde{R}_{\mathrm{f},k}v_2\|_{H^1}$$
$$\leq \|v_1 - \tilde{R}^1_{\mathrm{ms}}(v_1, v_2)\|_{H^1} + \|R^1_{\mathrm{f}}v_H - R^1_{\mathrm{f},k}v_H\|_{H^1} + \|\tilde{R}_{\mathrm{f}}v_2 - \tilde{R}_{\mathrm{f},k}v_2\|_{H^1},$$

where the first term on the right hand side is bounded in Lemma 3.5. For the second term we use Lemma 4.1 to get

$$\|R^1_{\mathrm{f}}v_H - R^1_{\mathrm{f},k}v_H\|^2_{H^1} \leq Ck^d\xi^{2k}\sum_{K\in\mathcal{T}_H}\|R^{K,1}_{\mathrm{f}}v_H\|^2_{H^1} \leq Ck^d\xi^{2k}\sum_{K\in\mathcal{T}_H}\|v_H\|^2_{H^1(K)}$$
$$= Ck^d\xi^{2k}\|v_H\|^2_{H^1} = Ck^d\xi^{2k}\|I_H(v_H - R^1_{\mathrm{f}}v_H)\|^2_{H^1}$$
$$= Ck^d\xi^{2k}\|I_H R^1_{\mathrm{ms}}(v_1, v_2)\|^2_{H^1} \leq Ck^d\xi^{2k}\|R^1_{\mathrm{ms}}(v_1, v_2)\|^2_{H^1}.$$

We can bound this further by using (3.25) and (3.26), such that

$$\|R^1_{\mathrm{ms}}(v_1, v_2)\|_{H^1} \leq C(\|v_1\|_{H^1} + \|v_2\|_{H^1}) \leq C(\|\mathcal{A}_1(v_1, v_2)\| + \|v_2\|_{H^1}).$$

Similar arguments, using Lemma 4.1 and (4.8), prove

$$\|\tilde{R}_{\mathrm{f}}v_2 - \tilde{R}_{\mathrm{f},k}v_2\|_{H^1} \leq Ck^{d/2}\xi^k\|v_2\|_{H^1},$$

and (4.13) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 4.3. To preserve linear convergence, the localization parameter k should be chosen such that $k = c\log(H^{-1})$ for some constant c. With this choice of k we get $k^{d/2}\xi^k \sim H$ and we get linear convergence in Lemma 4.2.

Remark 4.4. It is possible to consider different patch sizes k_u, k_θ, and k_{f} in (4.1), (4.2), and (4.8), corresponding to the displacement u, the temperature θ, and the additional correction, respectively. This would give us three different values of ξ, say ξ_u, ξ_θ, and ξ_{f}, in Lemma 4.1. To achieve linear convergence, the patch sizes should be chosen such that $k_u^{d/2}\xi_u^{k_u}, k_\theta^{d/2}\xi_\theta^{k_\theta}, k_{\mathrm{f}}^{d/2}\xi_{\mathrm{f}}^{k_{\mathrm{f}}} \sim H$, see Remark 4.3. This may be advantageous if one equation requires significantly larger patches. However, we shall, for simplicity, consider the same patch size k in all instances.

We note that the orthogonality relation (3.20) does not hold when V^1_{ms} is replaced by $V^1_{\mathrm{ms},k}$. However, we have that $V^1_{\mathrm{ms},k}$ and V^1_{f} are almost orthogonal in the sense that

$$(\sigma(v) : \varepsilon(w)) \leq Ck^{d/2}\xi^k\|v\|_{H^1}\|w\|_{H^1}, \quad \forall v \in V^1_{\mathrm{ms},k}, w \in V^1_{\mathrm{f}}. \tag{4.15}$$

To prove this, note that $v = v_{H,k} - R^1_{\mathrm{f},k}v_{H,k}$ for some $v_{H,k} \in V^1_H$, and

$$(\sigma(v) : \varepsilon(w)) = (\sigma(v_{H,k} - R^1_{\mathrm{f}}v_{H,k}) : \varepsilon(w)) + (\sigma(R^1_{\mathrm{f}}v_{H,k} - R^1_{\mathrm{f},k}v_{H,k}) : \varepsilon(w))$$
$$= (\sigma(R^1_{\mathrm{f}}v_{H,k} - R^1_{\mathrm{f},k}v_{H,k}) : \varepsilon(w)) \leq C_\sigma\|R^1_{\mathrm{f}}v_{H,k} - R^1_{\mathrm{f},k}v_{H,k}\|_{H^1}\|w\|_{H^1},$$

where we have used that $v_{H,k} - R^1_{\mathrm{f}}v_{H,k} \in V^1_{\mathrm{ms}}$ and the orthogonality (3.20). Due to Lemma 4.1 we now have

$$\|R^1_{\mathrm{f}}v_{H,k} - R^1_{\mathrm{f},k}v_{H,k}\|^2_{H^1} \leq Ck^d\xi^{2k}\sum_{K\in\mathcal{T}_H}\|R^{K,1}_{\mathrm{f}}v_{H,k}\|^2_{H^1} \leq Ck^d\xi^{2k}\sum_{K\in\mathcal{T}_H}\|v_{H,k}\|^2_{H^1(K)}$$
$$= Ck^d\xi^{2k}\|v_{H,k}\|^2_{H^1} = Ck^d\xi^{2k}\|I_H(v_{H,k} - R^1_{\mathrm{f},k}v_{H,k})\|^2_{H^1}$$
$$= Ck^d\xi^{2k}\|I_H v\|^2_{H^1} \leq Ck^d\xi^{2k}\|v\|^2_{H^1},$$

and (4.15) follows.

4.2. Time-dependent problem

A localized version of (3.30)–(3.32) is now defined by replacing V_{ms}^1 with $V_{\text{ms},k}^1$ and V_{ms}^2 with $V_{\text{ms},k}^2$. The method thus reads; for $n \in \{1, \ldots, N\}$ find

$$\tilde{u}_{\text{ms},k}^n = u_{\text{ms},k}^n + \sum_{K \in \mathcal{T}_H} u_{\text{f},k}^{n,K}, \text{ with } u_{\text{ms},k}^n \in V_{\text{ms},k}^1, \ u_{\text{f},k}^{n,K} \in V_{\text{f}}^1(\omega_k(K)),$$

and $\theta_{\text{ms},k}^n \in V_{\text{ms},k}^2$, such that

$$(\sigma(\tilde{u}_{\text{ms},k}^n) : \varepsilon(v_1)) - (\alpha\theta_{\text{ms},k}^n, \nabla \cdot v_1) = (f^n, v_1), \quad \forall v_1 \in V_{\text{ms},k}^1, \tag{4.16}$$

$$(\bar{\partial}_t\theta_{\text{ms},k}^n, v_2) + (\kappa\nabla\theta_{\text{ms},k}^n, \nabla v_2) + (\alpha\nabla \cdot \bar{\partial}_t\tilde{u}_{\text{ms},k}^n, v_2) = (g^n, v_2), \quad \forall v_2 \in V_{\text{ms},k}^2, \tag{4.17}$$

$$(\sigma(u_{\text{f},k}^{n,K}) : \varepsilon(w_1)) - (\alpha\theta_{\text{ms},k}^n, \nabla \cdot w_1)_K = 0, \qquad \forall w_1 \in V_{\text{f}}^1(\omega_k(K)), \tag{4.18}$$

where $\theta_{\text{ms},k}^0 = R_{\text{ms},k}^2\theta_h^0$. Furthermore, we define $\tilde{u}_{\text{ms},k}^0 = u_{\text{ms},k}^0 + \sum_{K \in \mathcal{T}_H} u_{\text{f},k}^{0,K}$, where $u_{\text{f},k}^{0,K} \in V_{\text{f}}^1(\omega_k(K))$ is defined by (4.18) for $n = 0$ and $u_{\text{ms},k}^0 \in V_{\text{ms}}^1$ such that

$$(\sigma(\tilde{u}_{\text{ms},k}^0) : \varepsilon(v_1)) - (\alpha\theta_{\text{ms},k}^0, \nabla \cdot v_1) = (f^0, v_1), \quad \forall v_1 \in V_{\text{ms},k}^1. \tag{4.19}$$

We also define $u_{\text{f},k}^n := \sum_{K \in \mathcal{T}_H} u_{\text{f},k}^{n,K}$. Note that for u_{f}^n we have due to (3.32)

$$(\sigma(u_{\text{f}}^n) : \varepsilon(w_1)) - (\alpha\theta_{\text{ms}}^n, \nabla \cdot w_1) = 0, \quad \forall w_1 \in V_{\text{f}}^1.$$

For the localized version $u_{\text{f},k}^n$ this relation is not true. Instead, we prove the following lemma.

Lemma 4.5. *For $w_1 \in V_{\text{f}}^1$, it holds that*

$$|(\sigma(u_{\text{f},k}^n) : \varepsilon(w_1)) - (\alpha\theta_{\text{ms},k}^n, \nabla \cdot w_1)| \leq Ck^{d/2}\xi^k\|\theta_{\text{ms},k}^n\|\|w_1\|_{H^1}.$$

Proof. Note that from (4.18) we have

$$(\sigma(u_{\text{f},k}^{n,K}) : \varepsilon(w_1)) - (\alpha\theta_{\text{ms},k}^n, \nabla \cdot w_1)_K = 0, \quad \forall w_1 \in V_{\text{f}}^1(\omega_k(K)). \tag{4.20}$$

This equation can be viewed as the localization of the following problem. Find $z_{\text{f}}^n \in V_{\text{f}}^1$, such that

$$(\sigma(z_{\text{f}}^n) : \varepsilon(w_1)) - (\alpha\theta_{\text{ms},k}^n, \nabla \cdot w_1) = 0, \quad \forall w_1 \in V_{\text{f}}^1. \tag{4.21}$$

Now ([14], Lem. 4.4) gives the bound

$$\|z_{\text{f}}^n - u_{\text{f},k}^n\|_{H^1}^2 \leq Ck^d\xi^{2k} \sum_{K \in \mathcal{T}_H} \|z_{\text{f}}^{n,K}\|_{H^1}^2$$

where $z_{\text{f}}^n = \sum_{K \in \mathcal{T}_H} z_{\text{f}}^{n,K}$ such that

$$(\sigma(z_{\text{f}}^{n,K}) : \varepsilon(w_1)) - (\alpha\theta_{\text{ms},k}^n, \nabla \cdot w_1)_K = 0, \quad \forall w_1 \in V_{\text{f}}^1.$$

Using this we derive the bound

$$\|z_{\text{f}}^n - u_{\text{f},k}^n\|_{H^1}^2 \leq Ck^d\xi^{2k} \sum_{K \in \mathcal{T}_H} \|z_{\text{f}}^{n,K}\|_{H^1}^2 \leq Ck^d\xi^{2k} \sum_{K \in \mathcal{T}_H} \|\theta_{\text{ms},k}^n\|_{L_2(K)}^2 = Ck^d\xi^{2k}\|\theta_{\text{ms},k}^n\|^2. \tag{4.22}$$

Now, to prove the lemma we use (4.21) and Cauchy–Schwarz inequality to get

$$|(\sigma(u_{\text{f},k}^n) : \varepsilon(w_1)) - (\alpha\theta_{\text{ms},k}^n, \nabla \cdot w_1)| = |(\sigma(u_{\text{f},k}^n - z_{\text{f}}^n) : \varepsilon(w_1))| \leq C_\sigma\|u_{\text{f},k}^n - z_{\text{f}}^n\|_{H^1}\|w_1\|_{H^1}.$$

Applying (4.22) finishes the proof. $\qquad \square$

The proof can be modified slightly to show the following bound

$$|(\sigma(\bar{\partial}_t u^n_{\mathrm{f},k}) : \varepsilon(w_1)) - (\alpha \bar{\partial}_t \theta^n_{\mathrm{ms},k}, \nabla \cdot w_1)| \leq C k^{d/2} \xi^k \|\bar{\partial}_t \theta^n_{\mathrm{ms},k}\| \|w_1\|_{H^1}. \tag{4.23}$$

Also note that it follows, by choosing $w_1 = u^n_{\mathrm{f},k}$ and $w_1 = \bar{\partial}_t u^n_{\mathrm{f},k}$ respectively, that

$$\|u^n_{\mathrm{f},k}\|_{H^1} \leq C \|\theta^n_{\mathrm{ms},k}\|, \quad \|\bar{\partial}_t u^n_{\mathrm{f},k}\|_{H^1} \leq C \|\bar{\partial}_t \theta^n_{\mathrm{ms},k}\|. \tag{4.24}$$

To prove that (4.16)–(4.18) is well posed, we need the following condition on the size of H.

Assumption 4.6. We make the following assumption on the size of H.

(A4) $H \leq \min\left(\frac{1}{4 C_{\mathrm{co}}}, \frac{c_\sigma}{(C_{\mathrm{co}} + C_{\mathrm{ort}})}\right)$, where C_{co} is the constant in Lemma 4.5 and C_{ort} is the constant in the almost orthogonal property (4.15).

Lemma 4.7. *Assuming* (A4) *the problem* (4.16)–(4.18) *is well-posed.*

Proof. This proof is similar the proof of Lemma 3.7, but we need to account for the lack of orthogonality and the fact that (3.32) is not satisfied.

Given $u^{n-1}_{\mathrm{ms},k}$, $\theta^{n-1}_{\mathrm{ms},k}$, and $u^{n-1}_{\mathrm{f},k} = \sum_K u^{n-1,K}_{\mathrm{f},k}$, the equations (4.16)–(4.18) yields a square system, so it is sufficient to prove that the solution is unique. Choosing $v_1 = u^n_{\mathrm{ms},k} - u^{n-1}_{\mathrm{ms},k}$ in (4.16) and $v_2 = \tau \theta^n_{\mathrm{ms},k}$ in (4.17) and adding the resulting equations we get

$$(\sigma(u^n_{\mathrm{ms},k}) : \varepsilon(u^n_{\mathrm{ms},k} - u^{n-1}_{\mathrm{ms},k})) + (\sigma(u^n_{\mathrm{f},k}) : \varepsilon(u^n_{\mathrm{ms},k} - u^{n-1}_{\mathrm{ms},k})) + \tau(\bar{\partial}_t \theta^n_{\mathrm{ms},k}, \theta^n_{\mathrm{ms},k})$$
$$+ c_\kappa \tau \|\theta^n_{\mathrm{ms},k}\|^2_{H^1} + (\alpha \nabla \cdot (u^n_{\mathrm{f},k} - u^{n-1}_{\mathrm{f},k}), \theta^n_{\mathrm{ms},k})$$
$$\leq (f^n, u^n_{\mathrm{ms},k} - u^{n-1}_{\mathrm{ms},k}) + \tau(g^n, \theta^n_{\mathrm{ms},k}).$$

Now, using (3.34) and

$$(f^n, u^n_{\mathrm{ms},k} - u^{n-1}_{\mathrm{ms},k}) \leq C \|f^n\|_{H^{-1}} + \frac{1}{2}(\sigma(u^n_{\mathrm{ms},k} - u^{n-1}_{\mathrm{ms},k}) : \varepsilon(u^n_{\mathrm{ms},k} - u^{n-1}_{\mathrm{ms},k})).$$

together with the estimate $\tau(\bar{\partial}_t \theta^n_{\mathrm{ms},k}, \theta^n_{\mathrm{ms},k}) \geq \frac{1}{2}\|\theta^n_{\mathrm{ms},k}\|^2 - \frac{1}{2}\|\theta^{n-1}_{\mathrm{ms},k}\|^2$, gives

$$\frac{c_\sigma}{2}\|u^n_{\mathrm{ms},k}\|^2_{H^1} + \frac{1}{4}\|\theta^n_{\mathrm{ms},k}\|^2 + c_\kappa \tau \|\theta^n_{\mathrm{ms},k}\|^2_{H^1} + (\sigma(u^n_{\mathrm{f},k}) : \varepsilon(u^n_{\mathrm{ms},k})) + (\alpha \nabla \cdot u^n_{\mathrm{f},k}, \theta^n_{\mathrm{ms},k})$$

$$\leq C \|f^n\|^2_{H^{-1}} + \frac{\tau}{2}\|g^n\|^2 + \frac{C_\sigma}{2}\|u^{n-1}_{\mathrm{ms},k}\|^2_{H^1} + \frac{1}{2}\|\theta^{n-1}_{\mathrm{ms},k}\|^2$$
$$+ (\sigma(\tilde{u}^n_{\mathrm{f},k}) : \varepsilon(u^{n-1}_{\mathrm{ms},k})) + (\alpha \nabla \cdot u^{n-1}_{\mathrm{f},k}, \theta^n_{\mathrm{ms},k}).$$

Using Lemma 4.5 we have

$$(\alpha \nabla \cdot u^n_{\mathrm{f},k}, \theta^n_{\mathrm{ms},k}) = (\alpha \theta^n_{\mathrm{ms},k}, \nabla \cdot u^n_{\mathrm{f},k}) - (\sigma(u^n_{\mathrm{f},k}) : \varepsilon(u^n_{\mathrm{f},k})) + (\sigma(u^n_{\mathrm{f},k}) : \varepsilon(u^n_{\mathrm{f},k}))$$
$$\geq -|(\alpha \theta^n_{\mathrm{ms},k}, \nabla \cdot u^n_{\mathrm{f},k}) - (\sigma(u^n_{\mathrm{f},k}) : \varepsilon(u^n_{\mathrm{f},k}))| + c_\sigma \|u^n_{\mathrm{f},k}\|^2_{H^1}$$
$$\geq -C_{\mathrm{co}} k^{d/2} \xi^k \|u^n_{\mathrm{f},k}\|_{H^1} \|\theta^n_{\mathrm{ms},k}\| + c_\sigma \|u^n_{\mathrm{f},k}\|^2_{H^1},$$

and the almost orthogonal property (4.15) gives

$$|(\sigma(u^n_{\mathrm{f},k}) : \varepsilon(u^n_{\mathrm{ms},k}))| \geq -C_{\mathrm{ort}} k^{d/2} \xi^k \|u^n_{\mathrm{f},k}\|_{H^1} \|u^n_{\mathrm{ms},k}\|_{H^1}.$$

Now, using that k should be chosen such that linear convergence is obtained, see Remark 4.3, that is $k^{d/2} \xi^k \sim H$, we conclude after using Young's inequality that

$$\left(\frac{c_\sigma}{2} - \frac{C_{\mathrm{ort}} H}{2}\right) \|u^n_{\mathrm{ms},k}\|^2_{H^1} + \left(\frac{1}{8} - \frac{C_{\mathrm{co}} H}{2}\right) \|\theta^n_{\mathrm{ms},k}\|^2 + c_\kappa \|\theta^n_{\mathrm{ms},k}\|^2_{H^1} + \left(c_\sigma - \frac{(C_{\mathrm{co}} + C_{\mathrm{ort}}) H}{2}\right) \|u^n_{\mathrm{f},k}\|^2_{H^1}$$

$$\leq C\left(\|f^n\|^2_{H^{-1}} + \tau \|g^n\|^2 + \|u^{n-1}_{\mathrm{ms},k}\|^2_{H^1} + \|\theta^{n-1}_{\mathrm{ms},k}\|^2 + \|u^{n-1}_{\mathrm{f},k}\|^2_{H^1}\right),$$

where Assumption 4.6 guarantees that the coefficients are positive. Hence, a unique solution exists. $\qquad \square$

5. Error analysis

In this section we analyze the error of the generalized finite element method. The results are based on Assumption 4.6. In the analysis we utilize the following property, which is similar to Lemma 4.5.

Lemma 5.1. *Let* $\tilde{e}_{f,k}^n := \tilde{R}_{f,k}\theta_h^n - u_{f,k}^n$ *and* $\eta_\theta^n := R_{ms,k}^2\theta_h^n - \theta_{ms,k}^n$. *Then, for* $w_1 \in V_f^1$, *it holds that*

$$|(\sigma(\tilde{e}_{f,k}^n) : \varepsilon(w_1)) - (\alpha\eta_\theta^n, \nabla \cdot w_1)| \leq Ck^{d/2}\xi^k\|\eta_\theta^n\|\|w_1\|_{H^1}.$$

Proof. The proof is similar to the proof of Lemma 4.5. We omit the details. □

This can be modified slightly to show the following bound

$$|(\sigma(\bar{\partial}_t\tilde{e}_{f,k}^n) : \varepsilon(w_1)) - (\alpha\bar{\partial}_t\eta_\theta^n, \nabla \cdot w_1)| \leq Ck^{d/2}\xi^k\|\bar{\partial}_t\eta_\theta^n\|\|w_1\|_{H^1}. \tag{5.1}$$

Also note that it follows, by choosing $w_1 = \tilde{e}_{f,k}^n$ and $w_1 = \bar{\partial}_t\tilde{e}_{f,k}^n$ respectively, that

$$\|\tilde{e}_{f,k}^n\|_{H^1} \leq C\|\eta_\theta^n\|, \quad \|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1} \leq C\|\bar{\partial}_t\eta_\theta^n\|. \tag{5.2}$$

Theorem 5.2. *Assume that* (A4) *holds. Let* $\{u_h^n\}_{n=1}^N$ *and* $\{\theta_h^n\}_{n=1}^N$ *be the solution to* (3.1) *and* (3.2) *and* $\{\tilde{u}_{ms,k}^n\}_{n=1}^N$ *and* $\{\theta_{ms,k}^n\}_{n=1}^N$ *the solution to* (4.16)–(4.18). *For* $n \in \{1, \ldots, N\}$ *we have*

$$\|u_h^n - \tilde{u}_{ms,k}^n\|_{H^1} + \|\theta_h^n - \theta_{ms,k}^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\Big(\|g\|_{L_\infty(L_2)} + \|\dot{g}\|_{L_\infty(H^{-1})} + \|f\|_{L_\infty(L_2)}$$
$$+ \|\dot{f}\|_{L_\infty(L_2)} + \|\ddot{f}\|_{L_\infty(H^{-1})} + t_n^{-1/2}\|\theta_h^0\|_{H^1}\Big).$$

The proof of Theorem 5.2 is based on two lemmas.

Lemma 5.3. *Assume that* $\theta_h^0 = 0$ *and* (A4) *holds. Let* $\{u_h^n\}_{n=1}^N$ *and* $\{\theta_h^n\}_{n=1}^N$ *be the solution to* (3.1) *and* (3.2) *and* $\{\tilde{u}_{ms,k}^n\}_{n=1}^N$ *and* $\{\theta_{ms,k}^n\}_{n=1}^N$ *the solution to* (4.16)–(4.18). *For* $n \in \{1, \ldots, N\}$ *we have*

$$\|u_h^n - \tilde{u}_{ms,k}^n\|_{H^1} + \|\theta_h^n - \theta_{ms,k}^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\Big(\|g\|_{L_\infty(L_2)} + \|\dot{g}\|_{L_\infty(H^{-1})} + \|f\|_{L_\infty(L_2)}$$
$$+ \|\dot{f}\|_{L_\infty(L_2)} + \|\ddot{f}\|_{L_\infty(H^{-1})}\Big).$$

Proof. We divide the error into the terms

$$u_h^n - \tilde{u}_{ms,k}^n = u_h^n - \tilde{R}_{ms,k}^1(u_h^n, \theta_h^n) + \tilde{R}_{ms,k}^1(u_h^n, \theta_h^n) - \tilde{u}_{ms,k}^n =: \tilde{\rho}_u^n + \tilde{\eta}_u^n,$$
$$\theta_h^n - \theta_{ms,k}^n = \theta_h^n - R_{ms,k}^2\theta_h^n + R_{ms,k}^2\theta_h^n - \theta_{ms,k}^n =: \rho_\theta^n + \eta_\theta^n.$$

We also adopt the following notation

$$\tilde{e}_{f,k}^n := \tilde{R}_{f,k}\theta_h^n - u_{f,k}^n, \quad \eta_u^n := \tilde{\eta}_u^n - \tilde{e}_{f,k}^n = R_{ms,k}^1(u_h^n, \theta_h^n) - u_{ms,k}^n.$$

From (3.2) it follows that

$$(\kappa\nabla\theta_h^n, \nabla v_2) = (g^n - \bar{\partial}_t\theta_h^n - \nabla \cdot \bar{\partial}_t u_h^n, v_2), \quad \forall v_2 \in V_h^2,$$

so by Lemma 4.2 we have the bound

$$\|\rho_\theta^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\|P_h^2 g^n - \bar{\partial}_t\theta_h^n - \nabla \cdot \bar{\partial}_t u_h^n\|,$$

where we have used that $\eta_\theta^0 = 0$. Furthermore, we note that if $\theta_h^0 = 0$, then $\tilde{R}_{f,k}\theta_h^0 = 0$ and $u_{f,k}^0 = 0$. Hence, $\tilde{e}_{f,k}^0 = 0$. From (4.19) and (3.3) we have, if $\theta_h^0 = \theta_{\mathrm{ms},k}^0 = 0$, for $v_1 \in V_{\mathrm{ms},k}^1$,

$$(\sigma(u_{\mathrm{ms},k}^0) : \varepsilon(v_1)) = (f^0, v_1) = (\sigma(u_h^0) : \varepsilon(v_1)) = (\sigma(R_{\mathrm{ms},k}^1(u_h^0, 0)) : \varepsilon(v_1)),$$

so also $\eta_u^0 = 0$.

To bound $\bar{\partial}_t \rho_\theta^n$ and $\alpha\nabla \cdot \bar{\partial}_t \tilde{\rho}_u^n$ we note that due to (3.1) and (3.3), $\bar{\partial}_t u_h^n$ and $\bar{\partial}_t \theta_h^n$ satisfy the equation

$$(\sigma(\bar{\partial}_t u_h^n) : \varepsilon(v_1)) - (\alpha\bar{\partial}_t \theta_h^n, \nabla \cdot v_1) = (\bar{\partial}_t f^n, v_1), \quad \forall v_1 \in V_h^1.$$

Hence, by Lemma 4.2 and the Aubin–Nitsche duality argument we have

$$\|\bar{\partial}_t \rho_\theta^n\|_{H^{-1}} \leq \|\bar{\partial}_t \rho_\theta^n\| \leq C(H + k^{d/2}\xi^k)\|\bar{\partial}_t \rho_\theta^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\|\bar{\partial}_t \theta_h^n\|_{H^1}, \tag{5.8}$$

and for $\bar{\partial}_t \tilde{\rho}_u^n$ we get

$$\|\alpha\nabla \cdot \bar{\partial}_t \tilde{\rho}_u^n\|_{H^{-1}} \leq \alpha_2\|\nabla \cdot \bar{\partial}_t \tilde{\rho}_u^n\| \leq C\|\bar{\partial}_t \tilde{\rho}_u^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)(\|\bar{\partial}_t f^n\| + \|\bar{\partial}_t \theta_h^n\|_{H^1}). \tag{5.9}$$

Thus, using (2.10), we arrive at the following bound

$$\sum_{j=1}^n \tau\|\eta_\theta^j\|_{H^1}^2 + \|\eta_u^n\|_{H^1}^2 + \|\tilde{e}_{f,k}^n\|_{H^1}^2 + \|\eta_\theta^n\|^2 \leq C(H + k^{d/2}\xi^k)^2 \sum_{j=1}^n \tau\left(\|\bar{\partial}_t \theta_h^j\|_{H^1}^2 + \|\bar{\partial}_t f^j\|^2\right)$$

$$+ Ck^d\xi^{2k}\sum_{j=1}^n \tau(\|\bar{\partial}_t \eta_u^j\|_{H^1}^2 + \|\bar{\partial}_t \tilde{e}_{f,k}^j\|_{H^1}^2), \tag{5.10}$$

where we apply Theorem 3.3 to the first sum on the right hand side. If we can find an upper bound on $\sum_{j=1}^n \tau(\|\bar{\partial}_t \eta_u^j\|_{H^1}^2 + \|\bar{\partial}_t \tilde{e}_{f,k}^j\|^2)$, then (5.10) gives a bound for $\|\tilde{\eta}_u^n\|_{H^1} \leq \|\eta_u^n\|_{H^1} + \|\tilde{e}_{f,k}^n\|_{H^1}$. This is done next, and we bound $\|\eta_\theta^n\|_{H^1}$ at the same time. For this purpose, we choose $v_2 = \bar{\partial}_t \eta_\theta^n$ in (5.4) and note that it follows from (5.3) that

$$(\sigma(\bar{\partial}_t \tilde{\eta}_u^n) : \varepsilon(\bar{\partial}_t \eta_u^n)) - (\alpha\bar{\partial}_t \eta_\theta^n, \nabla \cdot \bar{\partial}_t \eta_u^n) = 0. \tag{5.11}$$

This also holds for $n = 1$ since $\eta_\theta^0 = 0$ and $\tilde{\eta}_u^0 = 0$. Thus, by adding the resulting equations, we have

$$c_\sigma\|\bar{\partial}_t \eta_u^n\|_{H^1}^2 + \|\bar{\partial}_t \eta_\theta^n\|^2 + (\kappa\nabla\eta_\theta^n, \nabla\bar{\partial}_t \eta_\theta^n)$$
$$= (-\bar{\partial}_t \rho^n, \bar{\partial}_t \eta_\theta^n) - (\sigma(\bar{\partial}_t \tilde{e}_{f,k}^n) : \varepsilon(\bar{\partial}_t \eta_u^n)) - (\alpha\nabla \cdot \bar{\partial}_t \tilde{e}_{f,k}^n, \bar{\partial}_t \eta_\theta^n)$$
$$\leq \|\bar{\partial}_t \rho^n\|\|\bar{\partial}_t \eta_\theta^n\| + C_{\mathrm{ort}}k^{d/2}\xi^k\|\bar{\partial}_t \tilde{e}_{f,k}^n\|_{H^1}\|\bar{\partial}_t \eta_u^n\|_{H^1} - (\alpha\nabla \cdot \bar{\partial}_t \tilde{e}_{f,k}^n, \bar{\partial}_t \eta_\theta^n)$$

where we have used (4.15). For the last term we use Lemma 5.1 to achieve

$$-(\alpha\nabla \cdot \bar{\partial}_t \tilde{e}_{f,k}^n, \bar{\partial}_t \eta_\theta^n) \leq C_{\mathrm{co}}k^{d/2}\xi^k\|\bar{\partial}_t \tilde{e}_{f,k}^n\|_{H^1}\|\bar{\partial}_t \eta_\theta^n\| - (\sigma(\bar{\partial}_t \tilde{e}_{f,k}^n) : \varepsilon(\bar{\partial}_t \tilde{e}_{f,k}^n)).$$

Thus, we have

$$c_\sigma(\|\bar{\partial}_t \eta_u^n\|_{H^1}^2 + \|\bar{\partial}_t \tilde{e}_{f,k}^n\|_{H^1}^2) + \|\bar{\partial}_t \eta_\theta^n\|^2 + (\kappa\nabla\eta_\theta^n, \nabla\bar{\partial}_t \eta_\theta^n)$$
$$\leq \|\bar{\partial}_t \rho^n\|\|\bar{\partial}_t \eta_\theta^n\| + C_{\mathrm{ort}}k^{d/2}\xi^k\|\bar{\partial}_t \tilde{e}_{f,k}^n\|_{H^1}\|\bar{\partial}_t \eta_u^n\|_{H^1} + C_{\mathrm{co}}k^{d/2}\xi^k\|\bar{\partial}_t \tilde{e}_{f,k}^n\|_{H^1}\|\bar{\partial}_t \eta_\theta^n\|,$$

and using Young's inequality we deduce

$$\left(c_\sigma - \frac{C_{\mathrm{ort}}k^{d/2}\xi^k}{2}\right)\|\bar{\partial}_t \eta_u^n\|_{H^1}^2 + \left(c_\sigma - \frac{(C_{\mathrm{ort}} + C_{\mathrm{co}})k^{d/2}\xi^k}{2}\right)\|\bar{\partial}_t \tilde{e}_{f,k}^n\|_{H^1}^2)$$
$$+ \left(\frac{1}{2} - \frac{C_{\mathrm{co}}k^{d/2}\xi^k}{2}\right)\|\bar{\partial}_t \eta_\theta^n\|^2 + (\kappa\nabla\eta_\theta^n, \nabla\bar{\partial}_t \eta_\theta^n) \leq C\|\bar{\partial}_t \rho^n\|^2,$$

where Assumption 4.6 guarantees that the coefficients are positive. Multiplying by τ, using that

$$\tau(\kappa\nabla\eta_\theta^n, \nabla\bar{\partial}_t\eta_\theta^n) \geq \frac{1}{2}(\|\eta_\theta^n\|_\kappa^2 - \|\eta_\theta^{n-1}\|_\kappa^2),$$

and summing over n we derive

$$\sum_{j=1}^n \tau(\|\bar{\partial}_t\eta_u^j\|_{H^1}^2 + \|\bar{\partial}_t\tilde{e}_{f,k}^j\|_{H^1}^2 + \|\bar{\partial}_t\eta_\theta^j\|^2) + \|\eta_\theta^n\|_{H^1}^2$$

$$\leq C\sum_{j=1}^n \tau\|\bar{\partial}_t\rho^j\|^2 \leq C(H + k^{d/2}\xi^k)\sum_{j=1}^n \tau(\|\bar{\partial}_t f^j\|^2 + \|\bar{\partial}_t\theta_h^j\|_{H^1}^2),$$

where we have used that $\eta_\theta^0 = 0$, the bound (2.11), and (5.8) and (5.9). We can now apply Theorem 3.3. Thus, the lemma follows for $\|\theta_h^n - \theta_{\mathrm{ms},k}^n\|_{H^1}$. Moreover, this bounds the last terms in (5.10), which completes the proof. $\qquad\square$

Lemma 5.4. *Assume that $f = 0$ and $g = 0$, and that (A4) holds. Let $\{u_h^n\}_{n=1}^N$ and $\{\theta_h^n\}_{n=1}^N$ be the solution to (3.1) and (3.2) and $\{\tilde{u}_{\mathrm{ms},k}^n\}_{n=1}^N$ and $\{\theta_{\mathrm{ms},k}^n\}_{n=1}^N$ be the solution to (4.16)–(4.18). For $n \in \{1,\dots,N\}$ we have*

$$\|u_h^n - \tilde{u}_{\mathrm{ms},k}^n\|_{H^1} + t_n^{1/2}\|\theta_h^n - \theta_{\mathrm{ms},k}^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\|\theta_h^0\|_{H^1}. \tag{5.12}$$

Proof. As in the proof of Lemma 5.3 we split the error into two parts

$$u_h^n - \tilde{u}_{\mathrm{ms},k}^n = \tilde{\rho}_u^n + \tilde{\eta}_u^n, \quad \theta_h^n - \theta_{\mathrm{ms},k}^n = \rho_\theta^n + \eta_\theta^n,$$

where Lemma 4.2 and Theorem 3.3 gives

$$\|\rho_\theta^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\| - \bar{\partial}_t\theta_h^n - \nabla\cdot\bar{\partial}_t u_h^n\| \leq C(H + k^{d/2}\xi^k)t_n^{-1/2}\|\theta_h^0\|_{H^1},$$
$$\|\tilde{\rho}_u^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\|\theta_h^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\|\theta_h^0\|_{H^1}.$$

Now, note that (5.4) and (5.11) holds also when $f = 0$ and $g = 0$. In particular, (5.11) holds also for $n = 1$ due to the definition of $u_{\mathrm{ms},k}^0$ and u_h^0 in (4.19) and (3.3) respectively. By choosing $v_2 = \bar{\partial}_t\eta_\theta^n$ and adding the resulting equations we derive

$$c_\sigma\|\bar{\partial}_t\eta_u^n\|_{H^1}^2 + \|\bar{\partial}_t\eta_\theta^n\|^2 + (\kappa\nabla\eta_\theta^n, \nabla\bar{\partial}_t\eta_\theta^n) + (\sigma(\bar{\partial}_t\tilde{e}_{f,k}^n) : \epsilon(\bar{\partial}_t\eta_u^n)) + (\alpha\nabla\cdot\bar{\partial}_t\tilde{e}_{f,k}^n, \bar{\partial}_t\eta_\theta^n) \leq \|\bar{\partial}_t\rho^n\|\|\bar{\partial}_t\eta_\theta^n\|.$$

Recall $\rho^n = \rho_\theta^n + \alpha\nabla\cdot\tilde{\rho}_u^n$. As in the proof of Lemma 5.3 we get from Lemma 5.2

$$(\alpha\nabla\cdot\bar{\partial}_t\tilde{e}_{f,k}^n, \bar{\partial}_t\eta_\theta^n) \geq -C_{\mathrm{co}}k^{d/2}\xi^k\|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1}\|\bar{\partial}_t\eta_\theta^n\| + (\sigma(\bar{\partial}_t\tilde{e}_{f,k}^n) : \varepsilon(\bar{\partial}_t\tilde{e}_{f,k}^n)),$$

and from (4.15)

$$(\sigma(\tilde{e}_{f,k}^n) : \varepsilon(\bar{\partial}_t\eta_u^n)) \geq -C_{\mathrm{ort}}k^{d/2}\xi^k\|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1}\|\bar{\partial}_t\eta_u^n\|_{H^1}.$$

Hence, we have

$$\left(c_\sigma - \frac{C_{\mathrm{ort}}k^{d/2}\xi^k}{2}\right)\|\bar{\partial}_t\eta_u^n\|_{H^1}^2 + \left(c_\sigma - \frac{(C_{\mathrm{ort}} + C_{\mathrm{co}})k^{d/2}\xi^k}{2}\right)\|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1}^2$$

$$+ \left(\frac{1}{2} - \frac{C_{\mathrm{co}}k^{d/2}\xi^k}{2}\right)\|\bar{\partial}_t\eta_\theta^n\|^2 + (\kappa\nabla\eta_\theta^n, \nabla\bar{\partial}_t\eta_\theta^n) \leq \|\bar{\partial}_t\rho^n\|^2,$$

and Assumption 4.6 guarantees that the coefficients are positive. Multiplying by τt_n^2, using that

$$\tau(\kappa\nabla\eta_\theta^n, \nabla\bar{\partial}_t\eta_\theta^n) \geq \frac{1}{2}(\|\eta_\theta^n\|_\kappa^2 - \|\eta_\theta^{n-1}\|_\kappa^2)$$

and $t_n^2 - t_{n-1}^2 \leq 3\tau t_{n-1}$, for $n \geq 2$, now give

$$C\tau t_n^2(\|\bar{\partial}_t\eta_u^n\|_{H^1}^2 + \|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1}^2 + \|\bar{\partial}_t\eta_\theta^n\|^2) + \frac{t_n^2}{2}\|\eta_\theta^n\|_\kappa^2 - \frac{t_{n-1}^2}{2}\|\eta_\theta^{n-1}\|_\kappa^2 \leq C\tau t_n^2\|\bar{\partial}_t\rho^n\|^2 + C\tau t_{n-1}\|\eta_\theta^{n-1}\|_\kappa^2.$$

Note that this inequality also holds for $n = 1$, since $\eta_\theta^0 = 0$ (recall $\theta_{\mathrm{ms},k}^0 = R_{\mathrm{ms},k}^2\theta_h^0$). Summing over n gives and using (2.11)

$$C\sum_{j=1}^n \tau t_j^2(\|\bar{\partial}_t\eta_u^j\|_{H^1}^2 + \|\bar{\partial}_t\tilde{e}_{f,k}^j\|_{H^1}^2 + \|\bar{\partial}_t\eta_\theta^j\|^2) + c_\kappa t_n^2\|\eta_\theta^n\|_{H^1}^2 \leq C\sum_{j=1}^n \tau t_j^2\|\bar{\partial}_t\rho^j\|^2 + C\sum_{j=1}^{n-1} \tau t_j\|\eta_\theta^j\|_{H^1}^2, \quad (5.13)$$

and since $f^n = 0$ and $g^n = 0$, Lemma 4.2 and the Aubin–Nitsche trick as in (5.8) together with Theorem 3.3 give

$$\|\bar{\partial}_t\rho^n\| \leq \|\bar{\partial}_t\rho_\theta^n\| + \alpha_2\|\bar{\partial}_t\rho_u^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)(\|\bar{\partial}_t\theta_h^n\|_{H^1} + \|\nabla\cdot\bar{\partial}_t u_h^n\|) \leq C(H + k^{d/2}\xi^k)t_n^{-1}\|\theta_h^0\|_{H^1}. \quad (5.14)$$

To bound the last sum on the right hand side in (5.13) we choose $v_1 = \bar{\partial}_t\eta_u^n$ and $v_2 = \eta_\theta^n$ in (5.4) and (5.3) and add the resulting equations. This gives

$$(\sigma(\eta_u^n):\varepsilon(\bar{\partial}_t\eta_u^n)) + (\bar{\partial}_t\eta_\theta^n, \eta_\theta^n) + (\kappa\nabla\eta_\theta^n, \nabla\eta_\theta^n) = (-\bar{\partial}_t\rho^n, \eta_\theta^n) - (\sigma(\tilde{e}_{f,k}^n):\varepsilon(\bar{\partial}_t\eta_u^n)) - (\alpha\nabla\cdot\bar{\partial}_t\tilde{e}_{f,k}^n, \eta_\theta^n),$$

where the use of (5.6) and (5.7) gives

$$(\sigma(\eta_u^n):\varepsilon(\bar{\partial}_t\eta_u^n)) + (\sigma(\tilde{e}_{f,k}^n):\varepsilon(\bar{\partial}_t\tilde{e}_{f,k}^n)) + (\bar{\partial}_t\eta_\theta^n, \eta_\theta^n) + (\kappa\nabla\eta_\theta^n, \nabla\eta_\theta^n)$$
$$\leq \|\bar{\partial}_t\rho^n\|\|\eta_\theta^n\| + Ck^{d/2}\xi^k\|\eta_\theta^n\|(\|\bar{\partial}_t\eta_u^n\|_{H^1} + \|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1}).$$

Multiplying by τt_n and using that $t_n - t_{n-1} = \tau$ we get

$$C\tau t_n\|\eta_\theta^n\|_{H^1}^2 + \frac{t_n}{2}(\|\eta_u^n\|_\sigma^2 + \|\tilde{e}_{f,k}^n\|_\sigma^2) - \frac{t_{n-1}}{2}(\|\eta_u^{n-1}\|_\sigma^2 + \|\tilde{e}_{f,k}^{n-1}\|_\sigma^2) + \frac{t_n}{2}\|\eta_\theta^n\|^2 - \frac{t_{n-1}}{2}\|\eta_\theta^{n-1}\|^2$$
$$\leq Ct_n\tau(\|\bar{\partial}_t\rho^n\|\|\eta_\theta^n\| + k^{d/2}\xi^k\|\eta_\theta^n\|(\|\bar{\partial}_t\eta_u^n\|_{H^1} + \|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1})) + C\tau(\|\eta_u^{n-1}\|_\sigma^2 + \|\tilde{e}_{f,k}^{n-1}\|_\sigma^2 + \|\eta_\theta^{n-1}\|^2)$$
$$\leq Ct_n^2\tau\|\bar{\partial}_t\rho^n\|^2 + C_y t_n^2\tau k^d\xi^{2k}(\|\bar{\partial}_t\eta_u^n\|_{H^1}^2 + \|\bar{\partial}_t\tilde{e}_{f,k}^n\|_{H^1}^2) + C\tau(\|\tilde{\eta}_u^{n-1}\|_\sigma^2 + \|\tilde{e}_{f,k}^{n-1}\|_\sigma^2 + \|\eta_\theta^{n-1}\|^2 + \|\eta_\theta^n\|^2),$$

where we have used Young's (weighted) inequality on the form, $\tau t_n ab \leq \tau t_n^2 a^2 + \tau b^2/4$, in the last step. For the second term we have used the inequality with an additional C_y, i.e. $\tau t_n ab \leq C_y\tau t_n^2 a^2 + (4C_y)^{-1}\tau b^2$. Note that C_y can be made arbitrarily small. Summing over n and using (2.10) now gives

$$C\sum_{j=1}^n \tau t_j\|\eta_\theta^j\|_{H^1}^2 + \frac{c_\sigma t_n}{2}(\|\eta_u^n\|_{H^1}^2 + \|\tilde{e}_{f,k}^n\|_{H^1}^2) + \frac{t_n}{2}\|\eta_\theta^n\|^2$$
$$\leq C\sum_{j=1}^n \tau t_j^2\|\bar{\partial}_t\rho^j\|^2 + C_y k^d\xi^{2k}\sum_{j=1}^n \tau t_j^2(\|\bar{\partial}_t\eta_u^j\|_{H^1}^2 + \|\bar{\partial}_t\tilde{e}_{f,k}^j\|_{H^1}^2))$$
$$+ C\sum_{j=0}^n \tau(\|\eta_u^j\|_{H^1}^2 + \|\tilde{e}_{f,k}^j\|_{H^1}^2 + \|\eta_\theta^j\|^2). \quad (5.15)$$

We can now use (5.13) to deduce

$$\sum_{j=1}^{n} \tau t_j^2 (\|\bar{\partial}_t \eta_u^j\|_{H^1}^2 + \|\bar{\partial}_t \tilde{e}_{f,k}^j\|_{H^1}^2) \leq C \sum_{j=1}^{n} \tau t_j^2 \|\bar{\partial}_t \rho^j\|^2 + C \sum_{j=1}^{n-1} \tau t_j \|\eta_\theta^j\|_{H^1}^2.$$

Using this in (5.15) gives

$$C \sum_{j=1}^{n} \tau t_j \|\eta_\theta^j\|_{H^1}^2 + \frac{c_\sigma t_n}{2} (\|\eta_u^n\|_{H^1}^2 + \|\tilde{e}_{f,k}^n\|_{H^1}^2) + \frac{t_n}{2} \|\eta_\theta^n\|^2 \leq C \sum_{j=1}^{n} \tau t_j^2 \|\bar{\partial}_t \rho^j\|^2 + C_y k^d \xi^{2k} \sum_{j=1}^{n} \tau t_j \|\eta_\theta^j\|_{H^1}^2$$

$$+ C \sum_{j=0}^{n} \tau (\|\eta_u^j\|_{H^1}^2 + \|\tilde{e}_{f,k}^j\|_{H^1}^2 + \|\eta_\theta^j\|^2). \quad (5.16)$$

Since C_y now can be made arbitrarily small the term $C_y k^d \xi^{2k} \sum_{j=1}^{n} \tau t_j \|\eta_\theta^j\|_{H^1}^2$ can be moved to the left hand side. To estimate the last sum on the right hand side in (5.16) we multiply (5.4) by τ and sum over n to get

$$(\eta_\theta^n - \eta_\theta^0, v_2) + (\kappa \nabla \sum_{j=1}^{n} \tau \eta_\theta^j, \nabla v_2) + (\alpha \nabla \cdot \tilde{\eta}_u^n - \tilde{\eta}_u^0, v_2) = (-\rho^n + \rho^0, v_2), \quad (5.17)$$

where we note that $\eta_\theta^0 = 0$ and $\tilde{\eta}_u^0 = 0$. By choosing $v_1 = \eta_u^n$ in (5.3) and $v_2 = \eta_\theta^n$ in (5.17) and adding the resulting equations we get

$$c_\sigma \|\eta_u^n\|_{H^1}^2 + \|\eta_\theta^n\|^2 + (\kappa \sum_{j=1}^{n} \tau \nabla \eta_\theta^j, \nabla \eta_\theta^n) \leq \| - \rho^n + \rho^0 \| \|\eta_\theta^n\| - (\sigma(\tilde{e}_{f,k}^n) : \varepsilon(\eta_u^n)) - (\alpha \nabla \cdot \tilde{e}_{f,k}^n, \eta_\theta^n)$$

$$\leq \| - \rho^n + \rho^0 \| \|\eta_\theta^n\| + C_{\text{ort}} k^{d/2} \xi^k \|\tilde{e}_{f,k}^n\|_{H^1} \|\eta_u^n\|_{H^1}$$

$$+ C_{\text{co}} k^{d/2} \xi^k \|\tilde{e}_{f,k}^n\|_{H^1} \|\eta_\theta^n\| - c_\sigma \|\tilde{e}_{f,k}^n\|_{H^1}^2,$$

where we have used the almost orthogonal property (4.15) and Lemma 4.5. We conclude that

$$\left(c_\sigma - \frac{C_{\text{ort}} k^{d/2} \xi^k}{2} \right) \|\eta_u^n\|_{H^1}^2 + \left(c_\sigma - \frac{(C_{\text{ort}} + C_{\text{co}}) k^{d/2} \xi^k}{2} \right) \|\tilde{e}_{f,k}^n\|_{H^1}^2$$

$$+ \left(\frac{1}{2} - \frac{C_{\text{co}} k^{d/2} \xi^k}{2} \right) \|\eta_\theta^n\|^2 + \left(\kappa \sum_{j=1}^{n} \tau \nabla \eta_\theta^j, \nabla \eta_\theta^n \right) \leq C \| - \rho^n + \rho^0 \|^2, \quad (5.18)$$

and Assumption 4.6 guarantees positive coefficients. Now, note that we have the bound

$$\left(\kappa \sum_{j=1}^{n} \tau \nabla \eta_\theta^j, \nabla \eta_\theta^n \right) = \left(\kappa \sum_{j=1}^{n} \tau \nabla \eta_\theta^j, \bar{\partial}_t \left(\sum_{j=1}^{n} \tau \nabla \eta_\theta^j \right) \right) \geq \frac{1}{2\tau} \left(\left\| \sum_{j=1}^{n} \tau \eta_\theta^j \right\|_\kappa^2 - \left\| \sum_{j=1}^{n-1} \tau \eta_\theta^j \right\|_\kappa^2 \right),$$

with the convention that $\sum_{j=1}^{0} \tau \eta_\theta^j = 0$. Multiplying (5.18) by τ, summing over n, and using (2.11) thus gives

$$\sum_{j=1}^{n} \tau (\|\eta_u^j\|_{H^1}^2 + \|\tilde{e}_{f,k}^j\|_{H^1}^2 + \|\eta_\theta^j\|^2) + \frac{c_\kappa}{2} \left\| \sum_{j=1}^{n} \tau \eta_\theta^j \right\|_{H^1}^2 \leq C \sum_{j=1}^{n} \tau \| - \rho^j + \rho^0 \|^2$$

$$\leq C(H + k^{d/2} \xi^k)^2 \sum_{j=1}^{n} \tau \|\theta_h^0\|_{H^1}^2$$

$$\leq C(H + k^{d/2} \xi^k)^2 t_n \|\theta_h^0\|_{H^1}^2. \quad (5.19)$$

Here we have used the Aubin–Nitsche duality argument, Lemma 4.2 and Theorem 3.3 to deduce

$$\|\rho^n\| \leq \|\rho_\theta^n\| + C\|\rho_u^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)(\|\rho_\theta^n\|_{H^1} + \|\theta_h^n\|_{H^1})$$
$$\leq C(H + k^{d/2}\xi^k)\|\theta_h^n\|_{H^1} \leq C(H + k^{d/2}\xi^k)\|\theta_h^0\|_{H^1}, \quad n \geq 0.$$

Combining (5.13), (5.14), (5.16), and (5.19) we get

$$t_n^2\|\eta_\theta^n\|_{H^1}^2 + t_n\|\eta_u^n\|_{H^1}^2 + t_n\|\tilde{e}_{\mathrm{f},k}^n\|_{H^1}^2 \leq C(H + k^{d/2}\xi^k)^2 t_n\|\theta_h^0\|_{H^1}^2,$$

which completes the proof. $\qquad\square$

Proof of Theorem 5.2. Since the problem is linear we can split the solution

$$u_h^n = \bar{u}_h^n + \hat{u}_h^n, \quad \theta_h^n = \bar{\theta}_h^n + \hat{\theta}_h^n,$$

where $\bar{u}_h^n$ and $\bar{\theta}_h^n$ solves (3.1) and (3.2) with $f = 0$ and $g = 0$ and $\hat{u}_h^n$ and $\hat{\theta}_h^n$ solves (3.1) and (3.2) with $\theta_h^0 = 0$. The theorem now follows by applying Lemmas 5.3 and 5.4. $\qquad\square$

6. Numerical examples

In this section we perform two numerical examples. For a discussion on how to implement the type of generalized finite element efficiently described in this paper we refer to [10].

The first numerical example models a composite material which is preheated to a fix temperature and at time $t_0 = 0$ the piece is subject to a cool-down.

The domain is set to be the unit square $\Omega = [0,1] \times [0,1]$ and we assume that the temperature has a homogeneous Dirichlet boundary condition, that is $\Gamma_D^\theta = \partial\Omega$ and $\Gamma_N^\theta = \emptyset$. For the displacement we assume the bottom boundary to be fix and for the remaining part of the boundary we prescribe a homogeneous Neumann condition, that is $\Gamma_D^u = [0,1] \times 0$ and $\Gamma_N^u = \partial\Omega \setminus \Gamma_D^u$.

The composite is assumed to be built up according to Figure 1. The white part in the figure denotes a background material and the black parts an insulated material. The black squares are of size $2^{-5} \times 2^{-5}$. We assume that the Lamé coefficients μ and λ take the values μ_1 and λ_1 on the insulated material, and μ_2 and λ_2 on the background material. In this experiment we have set $\mu_1/\mu_2 = 10$ and $\lambda_1/\lambda_2 = 50$. Similarly, using subscript 1 for the insulated material and subscript 2 for the background material, we set $\alpha_1/\alpha_2 = 10$ and $\kappa = \kappa_i \cdot I$, for $i = 1, 2$, where I is the 2-dimensional identity matrix and $\kappa_1/\kappa_2 = 10$. Furthermore, we have chosen to set $f = [0,0]^\mathsf{T}$ (no external body forces) and $g = -10$.
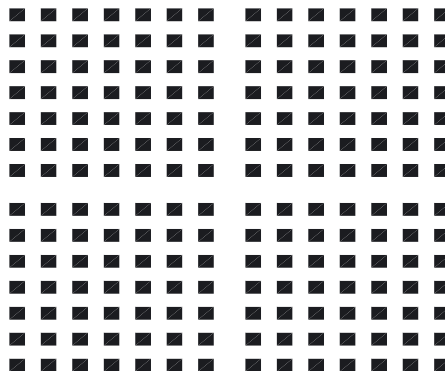

FIGURE 1. Composite material on the unit square. One black square is of size $2^{-5} \times 2^{-5}$.

The initial data must be zero on the boundary Γ_D^θ, so we have chosen to put $\theta^0 = 500x(1-x)y(1-y)$ and θ_h^0 to the L_2-projection of θ^0 to V_h^2. For the generalized finite element solution we have chosen $\theta_{\mathrm{ms},k}^0 = R_{\mathrm{ms},k}^2 \theta_h^0$ and $\tilde{u}_{\mathrm{ms},k}^0$ is given by (4.19).

The domain is discretized using a uniform triangulation. The reference solution is computed on a mesh of $h = \sqrt{2} \times 2^{-6}$ which resolves the fine parts (the black squares) in the material. The generalized finite element method (GFEM) in (4.16)–(4.18) is computed for five decreasing values of the mesh size, namely, $H = \sqrt{2} \times 2^{-1}, \sqrt{2} \times 2^{-2}, \ldots, \sqrt{2} \times 2^{-5}$, with the patch sizes $k = 1, 1, 2, 2, 3$. For comparison, we also compute the corresponding classical finite element (FEM) solution on the coarse meshes using continuous piecewise affine polynomials for both spaces (P1-P1). The solutions satisfies (3.1) and (3.2) with h replaced by H and are denoted u_H^n and θ_H^n respectively for $n = 1, \ldots, N$. When computing these solutions we have evaluated the integrals exactly to avoid quadrature errors.

We have chosen to set $T = 1$ and $\tau = 0.05$ for all values of H and for the reference solution. The solutions are compared at the time point N.

Note that the implementation of the corrections $u_{\mathrm{f},k}^{n,K}$ in (4.18) given by

$$(\sigma(u_{\mathrm{f},k}^{n,K}) : \varepsilon(w_1)) - (\alpha \theta_{\mathrm{ms},k}^n, \nabla \cdot w_1)_K = 0, \quad \forall w_1 \in V_{\mathrm{f}}^1(\omega_k(K)),$$

should *not* be computed explicitly at each time step. It is more efficient to compute x_y^K, given by

$$(\sigma(x_y^K) : \varepsilon(w_1)) - (\alpha(\lambda_y^2 - R_{\mathrm{f},k}^2 \lambda_y^2), \nabla \cdot w_1)_K = 0, \quad \forall w_1 \in V_{\mathrm{f}}^1(\omega_k(K)),$$

where $\{(\cdot, y) \in \mathcal{N} : \lambda_y^2 - R_{\mathrm{f},k}^2 \lambda_y^2\}$ is the basis for $V_{\mathrm{ms},k}^2$. Now, since $\theta_{\mathrm{ms},k}^n = \sum_y \beta_y^n(\lambda_y^2 - R_{\mathrm{f},k}^2 \lambda_y^2)$, we have the identity

$$u_{\mathrm{f},k}^n = \sum_K u_{\mathrm{f},k}^{n,K} = \sum_K \sum_y \beta_y^n x_y^K.$$

With this approach, we only need to compute x_y^K once before solving for the system (4.16) and (4.17) for $n = 1, \ldots, N$.

The relative errors in the H^1-seminorm $\|\nabla \cdot\|$ are shown in Figure 2. The left graph shows the relative errors for the displacement, $\|\nabla(\tilde{u}_{\mathrm{ms},k}^N - u_h^N)\|/\|\nabla u_h^N\|$ and $\|\nabla(u_H^N - u_h^N)\|/\|\nabla u_h^N\|$. The right graph shows the error

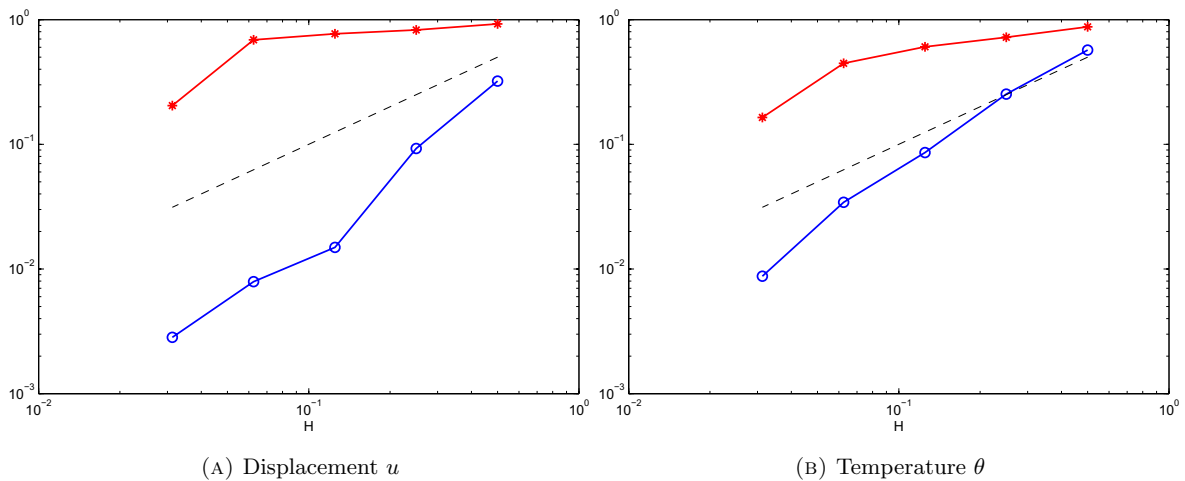(A) Displacement u (B) Temperature θ

FIGURE 2. Relative errors using GFEM (*blue* ∘) and P1-P1 FEM (*red* ∗) for the linear thermoelasticity problem plotted against the mesh size H. The dashed line is H. (color online)
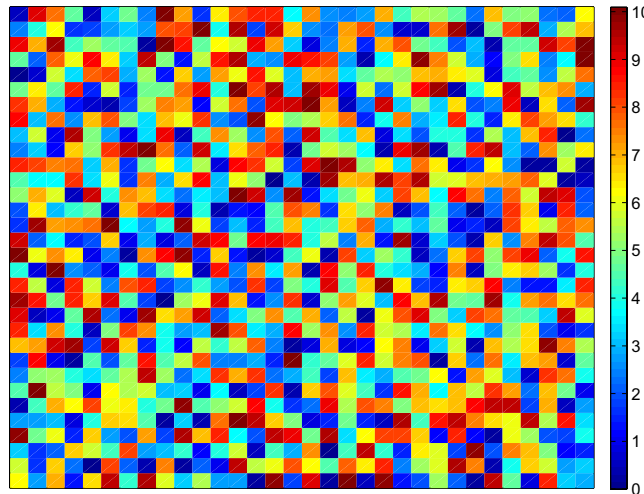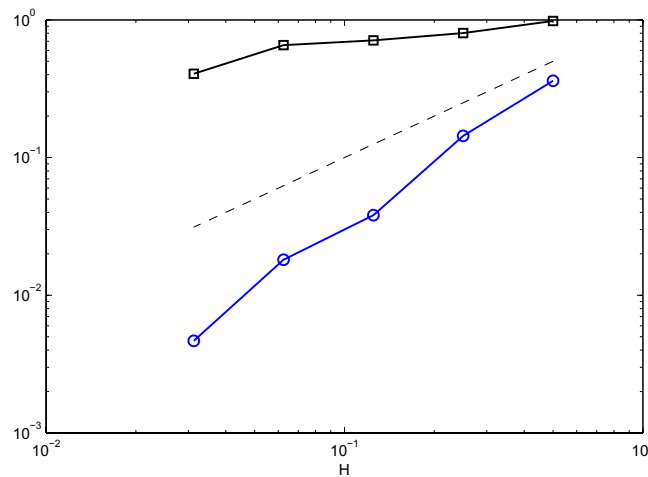
FIGURE 3. A plot of the coefficient α.

FIGURE 4. Relative errors for the displacement u using GFEM with correction for α (*blue* $\circ$) and GFEM without correction for α (*black* $\square$) for the linear thermoelasticity problem plotted against the mesh size H. The dashed line is H. (color online)

for the temperature $\|\nabla(\theta_{\mathrm{ms},k}^N - \theta_h^N)\|/\|\nabla\theta_h^N\|$ and $\|\nabla(\theta_H^N - \theta_h^N)\|/\|\nabla\theta_h^N\|$. As expected the generalized finite element shows convergence of optimal order and outperforms the classical finite element.

The second example shows the importance of the additional correction (4.18), which is designed to handle multiscale behavior in the coefficient α. The computational domain, the spatial and the time discretization, and the patch sizes remain the same as in the first example. However, we let $\Gamma_D = \partial\Omega$ and $\Gamma_N = \emptyset$ in this case.

To test the influence of α we let the other coefficients be constants, $\mu = \lambda = 1$ and $\kappa = I$, where the I is the 2-dimensional identity matrix. The coefficient α takes values between 0.1 and 10 according to Figure 3. The boxes are of size $2^{-5} \times 2^{-5}$ and, hence, the reference mesh of size $h = \sqrt{2} \cdot 2^{-6}$ is sufficiently small to resolve the variations in α.

A. MÅLQVIST AND A. PERSSON

The initial data is set to $\theta^0 = x(1-x)y(1-y)$ and θ_h^0 is the L_2-projection of θ^0 onto V_h^2. For the generalized finite element solution we have chosen $\theta_{\mathrm{ms},k}^0 = R_{\mathrm{ms},k}^2 \theta_h^0$ and $\tilde{u}_{\mathrm{ms},k}^0$ is given by (4.19), as in our first example. Furthermore, we have chosen to set $f = [1\ 1]^\intercal$ and $g = 10$.

The generalized finite element method (GFEM) in (4.16)–(4.18) is computed for the five decreasing values of the mesh size used in the first example. For comparison, we compute the generalized finite element without the additional correction on $u_{\mathrm{ms},k}^n$. In this case the system (4.16)–(4.18) simplifies to

$$(\sigma(u_{\mathrm{ms},k}^n) : \varepsilon(v_1)) - (\alpha\theta_{\mathrm{ms},k}^n, \nabla \cdot v_1) = (f^n, v_1), \quad \forall v_1 \in V_{\mathrm{ms},k}^1,$$
$$(\bar{\partial}_t \theta_{\mathrm{ms},k}^n, v_2) + (\kappa\nabla\theta_{\mathrm{ms},k}^n, \nabla v_2) + (\alpha\nabla \cdot \bar{\partial}_t u_{\mathrm{ms},k}^n, v_2) = (g^n, v_2), \quad \forall v_2 \in V_{\mathrm{ms},k}^2.$$

The relative errors in the H^1-seminorm are shown in Figure 2. The graph shows the errors for the displacement with correction for α, $\|\nabla(\tilde{u}_{\mathrm{ms},k}^N - u_h^N)\|/\|\nabla u_h^N\|$ and the error without correction for α $\|\nabla(u_{\mathrm{ms},k}^N - u_h^N)\|/\|\nabla u_h^N\|$. As expected the GFEM with correction for α shows convergence of optimal order and outperforms the GFEM without correction for α. This is due to the fact that, in the case when the correction for α is neglected, the constant in (4.13) (and hence also the constant in Thm. 5.2) depends on the variations in α (see Rem. 3.6).

References

[1] A. Abdulle and P. Henning, Localized orthogonal decomposition method for the wave equation with a continuum of scales. *Math. Comp.* **86** (2017) 549–587.

[2] K.T. Andrews, P. Shi, M. Shillor and S. Wright, Thermoelastic contact with Barber's heat exchange condition. *Appl. Math. Optim.* **28** (1993) 11–48.

[3] I. Babuška and R. Lipton, Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. *Multiscale Model. Simul.* **9** (2011) 373–406.

[4] I. Babuška and J.E. Osborn, Generalized finite element methods: their performance and their relation to mixed methods. *SIAM J. Numer. Anal.* **20** (1983) 510–536.

[5] M.A. Biot, General theory of three-dimensional consolidation. *J. Appl. Phys.* **18** (1941) 155–164.

[6] M.A. Biot, Thermoelasticity and irreversible thermodynamics. *J. Appl. Phys.* **27** (1956) 240–253.

[7] S.C. Brenner and R.L. Scott, The mathematical theory of finite element methods. Vol. 15 of *Texts in Applied Mathematics*, 3rd edn. Springer, New York (2008).

[8] P.G. Ciarlet, Mathematical elasticity. Vol. I. Three-dimensional elasticity. Vol. 20 of *Studies in Mathematics and its Applications.* North-Holland Publishing Co., Amsterdam (1988).

[9] C.M. Dafermos, On the existence and the asymptotic stability of solutions to the equations of linear thermoelasticity. *Arch. Ration. Mech. Anal.* **29** (1968) 241–271.

[10] Ch. Engwer, P. Henning, A. Målqvist and D. Peterseim, Efficient implementation of the localized orthogonal decomposition method. Preprint arXiv:1602.01658 (2016).

[11] A. Ern and S. Meunier, *A posteriori* error analysis of Euler-Galerkin approximations to coupled elliptic-parabolic problems. *ESAIM: M2AN* **43** (2009) 353–375.

[12] P. Henning and A. Målqvist, Localized orthogonal decomposition techniques for boundary value problems. *SIAM J. Sci. Comput.* **36** (2014) A1609–A1634.

[13] P. Henning, A. Målqvist and D. Peterseim, A localized orthogonal decomposition method for semi-linear elliptic problems. *ESAIM: M2AN* **48** (2014) 1331–1349.

[14] P. Henning and A. Persson, A multiscale method for linear elasticity reducing Poisson locking. *Comput. Methods Appl. Mech. Engrg.* **310** (2016) 156–171.

[15] T.J.R. Hughes, G.R. Feijóo, L. Mazzei and J.-B. Quincy. The variational multiscale method – a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.* **166** (1998) 3–24.

[16] M.G. Larson and A. Målqvist, Adaptive variational multiscale methods based on a posteriori error estimation: energy norm estimates for elliptic problems. *Comput. Methods Appl. Mech. Engrg.* **196** (2007) 2313–2324.

[17] A. Målqvist and A. Persson, Multiscale techniques for parabolic equations. Preprint arXiv:1504.08140 (2015).

[18] A. Målqvist and D. Peterseim, Localization of elliptic multiscale problems. *Math. Comp.* **83** (2014) 2583–2603.

[19] A. Målqvist and D. Peterseim, Computation of eigenvalues by numerical upscaling. *Numer. Math.* **130** (2015) 337–361.

[20] D. Peterseim, Variational multiscale stabilization and the exponential decay of fine-scale correctors. In vol. 114 of *Lect. Notes Comput. Sci. Engrg.* Springer (2016) 343–369.

[21] P.J. Phillips and M.F. Wheeler, Overcoming the problem of locking in linear elasticity and poroelasticity: an heuristic approach. *Comput. Geosci.* **13** (2009) 5–12.

[22] P.J. Phillips and M.F. Wheeler, A coupling of mixed and discontinuous Galerkin finite-element methods for poroelasticity. *Comput. Geosci.* **12** (2008) 417–435.

[23] P. Shi and M. Shillor, Existence of a solution to the N-dimensional problem of thermoelastic contact. *Commun. Partial Differ. Eq.* **17** (1992) 1597–1618.

[24] R.E. Showalter, Diffusion in poro-elastic media. *J. Math. Anal. Appl.* **251** (2000) 310–340.

[25] X. Xu, The N-dimensional quasistatic problem of thermoelastic contact with Barber's heat exchange conditions. *Adv. Math. Sci. Appl.* **6** (1996) 559–587.

[26] A. Ženíšek, The existence and uniqueness theorem in Biot's consolidation theory. *Apl. Mat.* **29** (1984) 194–211.

[27] A. Ženíšek, Finite element methods for coupled thermoelasticity and coupled consolidation of clay. *RAIRO Anal. Numér.* **18** (1984) 183–205.