ESAIM: M2AN 55 (2021) 1375–1404
https://doi.org/10.1051/m2an/2021023

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

A GENERALIZED FINITE ELEMENT METHOD FOR THE STRONGLY DAMPED WAVE EQUATION WITH RAPIDLY VARYING DATA

Per Ljung[1,*], Axel Målqvist[1] and Anna Persson[2]

Abstract. We propose a generalized finite element method for the strongly damped wave equation with highly varying coefficients. The proposed method is based on the localized orthogonal decomposition introduced in Målqvist and Peterseim [*Math. Comp.* **83** (2014) 2583–2603], and is designed to handle independent variations in both the damping and the wave propagation speed respectively. The method does so by automatically correcting for the damping in the transient phase and for the propagation speed in the steady state phase. Convergence of optimal order is proven in $L_2(H^1)$-norm, independent of the derivatives of the coefficients. We present numerical examples that confirm the theoretical findings.

Mathematics Subject Classification. 35K10, 65M60.

Received November 5, 2020. Accepted May 17, 2021.

1. Introduction

This paper is devoted to the study of numerical solutions to the strongly damped wave equation with highly varying coefficients. The equation takes the general form

$$\ddot{u} - \nabla \cdot (A\nabla\dot{u} + B\nabla u) = f, \tag{1.1}$$

on a bounded domain Ω. Here, A and B represent the system's damping and wave propagation respectively, f denotes the source term, and the solution u is a displacement function. This equation commonly appears in the modelling of viscoelastic materials, where the strong damping $-\nabla \cdot A\nabla\dot{u}$ arises due to the stress being represented as the sum of an elastic part and a viscous part [6, 13]. Viscoelastic materials have several applications in engineering, including noise dampening, vibration isolation, and shock absorption (see [20] for more applications). In particular, in multiscale applications, such as modelling of porous media or composite materials, A and B are both rapidly varying.

There has been much recent work regarding strongly damped wave equations. For instance, well-posedness of the problem is discussed in [7, 19, 21], asymptotic behavior in [3, 8, 30, 34] solution blowup in [2, 12], and decay estimates in [18]. In particular, FEM for the strongly damped wave equation has been analyzed in [24]

Keywords and phrases. Strongly damped wave equation, multiscale, localized orthogonal decomposition, finite element method, reduced basis method.

[1] Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden.

[2] Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden.

*Corresponding author: perlj@chalmers.se

using the Ritz–Volterra projection, and [23] uses the classical Ritz-projection in the homogeneous case with Rayleigh damping. In these papers, convergence of optimal order is shown. However, in the case of piecewise linear polynomials, the convergence relies on at least H^2-regularity in space. Consequently, since the H^2-norm depends on the derivatives of the coefficients, the error is bounded by $\|u\|_{H^2} \sim \max(\varepsilon_A^{-1}, \varepsilon_B^{-1})$ where ε_A and ε_B denote the scales at which A and B vary respectively. The convergence order is thus only valid when the mesh width h fulfills $h < \min(\varepsilon_A, \varepsilon_B)$. In other words, we require a mesh that is fine enough to resolve the variations of A and B, which becomes computationally challenging. This type of difficulty is common for equations with rapidly varying data, an issue for which several numerical methods have been developed (see *e.g.* [4, 5, 22, 28]). None of these methods are however applicable to the strongly damped wave equation, where two different multiscale coefficients have to be dealt with. In this paper, we propose a novel multiscale method based on the localized orthogonal decomposition (LOD) method.

The LOD method is based on the variational multiscale method presented in [17]. It was first introduced in [29], and has since then been further developed and analyzed for several types of problems (see *e.g.* [1, 15, 16, 25, 27]). In particular, Målqvist and Peterseim [26] studies the LOD method for quadratic eigenvalue problems, which correspond to time-periodic wave equations with weak damping. The main idea of the method is based on a decomposition of the solution space into a coarse and a fine part. The decomposition is done by defining an interpolant that maps functions from an infinite dimensional space into a finite dimensional FE-space. In this way, the kernel of the interpolant captures the finescale features that the coarse FE-space misses, and hence defines the finescale space. Subsequently, one may use the orthogonal complement to this finescale space with respect to a problem-dependent Ritz-projection as a modified FE-space. In the case of time-dependent problems, the LOD method performs particularly well in the sense that the modified FE-space only needs to be computed once, and can then be re-used in each time step.

Multiscale methods, as the localized orthogonal decomposition, are usually designed to handle problems with a single multiscale coefficient. In this sense, the strongly damped wave equation is different, as an extra coefficient appears due to the strong damping. Hence, one of the main challenges for the novel method is how to incorporate the finescale behavior of both coefficients in the computation. Nevertheless, it should be noted that existing multiscale methods are applicable for some special cases of this equation. An example is the case of Rayleigh damping where the coefficients are proportional to each other. Other examples are the steady state case, the transient phase in which the solution evolves rapidly in time, as well as the case of weak damping where no spatial derivatives are present on the damping term.

In this paper we present a generalized finite element method (GFEM), with a backward Euler time stepping for solving the strongly damped wave equation. The method uses both the damping and diffusion coefficients to construct a generalized finite element space, similar to those in *e.g.* [25, 29]. The solution is then evaluated in this space, but to account for the time dependence, an additional correction is added to it. However, this correction is evaluated on the fine scale, and thus expensive to compute. To overcome this issue, we prove spatial exponential decay for the corrections so that we can restrict the problems to patches in a similar manner as for the modified basis functions in [29]. The effect of the proposed method is that the multiscale basis compensates for the damping early on in the simulation when it is dominant and then gradually starts to compensate for the wave propagation which is dominant at steady state. This is done seamlessly and automatically by the method. Furthermore, we prove optimal order convergence in $L_2(H^1)$-norm for this method. Following this, we show that it is sufficient to compute the finescale corrections for only a few time steps by applying reduced basis (RB) techniques. For related work on RB methods, see *e.g.* [9, 10, 14], and for an introduction to the topic we refer to [32].

The outline of the paper is as follows: In Section 2 we present the weak formulation and classical FEM for the strongly damped wave equation, along with necessary assumptions. Section 3 is devoted to the generalized finite element method and its localization procedure. In Section 4 error estimates for the method are proven. Section 5 covers the details of the RB approach, and finally in Section 6 we illustrate numerical examples that confirm the theory derived in this paper.

2. Weak formulation and classical FEM

We consider the wave equation with strong damping of the following form

$$\ddot{u} - \nabla \cdot (A\nabla \dot{u} + B\nabla u) = f, \quad \text{in } \Omega \times (0,T], \tag{2.1}$$
$$u = 0, \quad \text{on } \Gamma \times (0,T], \tag{2.2}$$
$$u(0) = u_0, \quad \text{in } \Omega, \tag{2.3}$$
$$\dot{u}(0) = v_0 \quad \text{in } \Omega, \tag{2.4}$$

where $T > 0$ and Ω is a polygonal (or polyhedral) domain in $\mathbb{R}^d$, $d = 2, 3$, and $\Gamma := \partial\Omega$. The coefficients A and B describe the damping and propagation speed respectively, and f denotes the source function of the system. We assume $A = A(x)$, $B = B(x)$ and $f = f(x,t)$, *i.e.* the multiscale coefficients are independent of time.

Denote by $H_0^1(\Omega)$ the classical Sobolev space with norm

$$\|v\|_{H^1(\Omega)}^2 = \|v\|_{L_2(\Omega)}^2 + \|\nabla v\|_{L_2(\Omega)}^2$$

whose functions vanish on Γ. Moreover, let $L_p(0,T;\mathcal{B})$ be the Bochner space with norm

$$\|v\|_{L_p(0,T;\mathcal{B})} = \left(\int_0^T \|v\|_{\mathcal{B}}^p \,\mathrm{d}t \right)^{1/p}, \quad p \in [1,\infty),$$
$$\|v\|_{L_\infty(0,T;\mathcal{B})} = \operatorname*{ess\,sup}_{t\in[0,T]} \|v\|_{\mathcal{B}},$$

where $\mathcal{B}$ is a Banach space with norm $\|\cdot\|_{\mathcal{B}}$. In this paper, the following assumptions are made on the data.

Assumption 2.1. *The damping and propagation coefficients $A, B \in L_\infty(\Omega, \mathbb{R}^{d\times d})$ are symmetric and satisfy*

$$0 < \alpha_- := \operatorname*{ess\,inf}_{x\in\Omega} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{A(x)v\cdot v}{v\cdot v} < \operatorname*{ess\,sup}_{x\in\Omega} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{A(x)v\cdot v}{v\cdot v} =: \alpha_+ < \infty,$$
$$0 < \beta_- := \operatorname*{ess\,inf}_{x\in\Omega} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{B(x)v\cdot v}{v\cdot v} < \operatorname*{ess\,sup}_{x\in\Omega} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{B(x)v\cdot v}{v\cdot v} =: \beta_+ < \infty.$$

In addition, we assume that $f \in L_\infty([0,T];L_2(\Omega))$ and $\dot{f} \in L_2([0,T];L_2(\Omega))$.

For the spatial discretization, let $\{\mathcal{T}_h\}_{h>0}$ denote a family of shape regular elements that form a partition of the domain Ω. For an element $K \in \mathcal{T}_h$, let the corresponding mesh size be defined as $h_K := \operatorname{diam}(K)$, and denote the largest diameter of the partition by $h := \max_{K\in\mathcal{T}_h} h_K$. We now define the classical FE-space using continuous piecewise linear polynomials as

$$S_h := \{v \in \mathcal{C}(\bar{\Omega}) : v\big|_\Gamma = 0,\ v\big|_K \text{ is a polynomial of partial degree} \le 1,\ \forall K \in \mathcal{T}_h\},$$

and let $V_h = S_h \cap H_0^1(\Omega)$. The semi-discrete FEM becomes: find $u_h : [0,T] \to V_h$ such that

$$(\ddot{u}_h, v) + a(\dot{u}_h, v) + b(u_h, v) = (f, v), \quad \forall v \in V_h,\ t \in [0,T], \tag{2.5}$$

with initial values $u_h(0) = u_h^0$ and $\dot{u}_h(0) = v_h^0$ where $u_h^0, v_h^0 \in V_h$ are appropriate approximations of u_0 and v_0 respectively. Here $(\cdot,\cdot)$ denotes the usual L_2-inner product, $a(\cdot,\cdot) = (A\nabla\cdot, \nabla\cdot)$, and $b(\cdot,\cdot) = (B\nabla\cdot, \nabla\cdot)$.

For the temporal discretization, let $0 = t_0 < t_1 < \ldots < t_N = T$ be a uniform partition with time step $t_n - t_{n-1} = \tau$. We apply a backward Euler scheme to get the fully discrete system: find $u_h^n \in V_h$ such that

$$(\bar{\partial}_t^2 u_h^n, v) + a(\bar{\partial}_t u_h^n, v) + b(u_h^n, v) = (f^n, v), \quad \forall v \in V_h, \tag{2.6}$$

for $n \geq 2$. Here, the discrete derivative is defined as $\bar{\partial}_t u_h^n = (u_h^n - u_h^{n-1})/\tau$. The first initial value is given by $u_h^0 \in V_h$. The second initial value u_h^1 should be an approximation of $u(t_1)$ and could be chosen as $u_h^1 = u_h^0 + \tau v_h^0$.

For results on regularity and error estimates for the FEM solution of the strongly damped wave equation, we refer to [23]. Moreover, existence and uniqueness of a solution to (2.6) is guaranteed by Lax–Milgram.

In the analysis, we use the notations $\|\cdot\|_a^2 := a(\cdot,\cdot)$, $\|\cdot\|_b^2 := b(\cdot,\cdot)$, as well as $|||\cdot|||^2 = \tilde{a}(\cdot,\cdot) := a(\cdot,\cdot)+\tau b(\cdot,\cdot)$, and the fact that these are equivalent with the H^1-norm. That is, there exist positive constants $C_a, C_b, C_{\tilde{a}}, c_a, c_b, c_{\tilde{a}} \in \mathbb{R}$, such that

$$\begin{aligned} c_a\|v\|_{H^1}^2 \leq \|v\|_a^2 \leq C_a\|v\|_{H^1}^2, \quad \forall v \in H^1(\Omega), \\ c_b\|v\|_{H^1}^2 \leq \|v\|_b^2 \leq C_b\|v\|_{H^1}^2, \quad \forall v \in H^1(\Omega), \\ c_{\tilde{a}}\|v\|_{H^1}^2 \leq |||v|||^2 \leq C_{\tilde{a}}\|v\|_{H^1}^2, \quad \forall v \in H^1(\Omega). \end{aligned} \tag{2.7}$$

Theorem 2.2. *The solution u_h^n to (2.6) satisfies the following bounds*

$$\|\bar{\partial}_t u_h^n\|_{L_2}^2 + \sum_{j=2}^{n} \tau\|\bar{\partial}_t u_h^j\|_{H^1}^2 + \|u_h^n\|_{H^1}^2 \leq C\sum_{j=2}^{n}\tau\|f^j\|_{H^{-1}}^2 + C\left(\|\bar{\partial}_t u_h^1\|_{L_2}^2 + \|u_h^1\|_{H^1}^2\right), \quad \text{for } n \geq 2, \tag{2.8}$$

$$\sum_{j=2}^{n} \tau\|\bar{\partial}_t^2 u_h^j\|_{L_2}^2 + \|\bar{\partial}_t u_h^n\|_{H^1}^2 \leq C\sum_{j=2}^{n}\tau\|f^j\|_{L_2}^2 + C\left(\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \|u_h^1\|_{H^1}^2\right), \quad \text{for } n \geq 2. \tag{2.9}$$

Proof. To prove (2.8), choose $v = \tau\bar{\partial}_t u_h^n$ in (2.6) to get

$$\tau\left(\bar{\partial}_t^2 u_h^n, \bar{\partial}_t u_h^n\right) + \tau\|\bar{\partial}_t u_h^n\|_a^2 + \tau b\left(u_h^n, \bar{\partial}_t u_h^n\right) = \tau\left(f^n, \bar{\partial}_t u_h^n\right). \tag{2.10}$$

Due to Cauchy–Schwarz and Young's inequality we have the following lower bound

$$\tau(\bar{\partial}_t^2 u_h^n, \bar{\partial}_t u_h^n) = \|\bar{\partial}_t u_h^n\|_{L_2}^2 - \left(\bar{\partial}_t u_h^{n-1}, \bar{\partial}_t u_h^n\right) \geq \frac{1}{2}\|\bar{\partial}_t u_h^n\|_{L_2}^2 - \frac{1}{2}\|\bar{\partial}_t u_h^{n-1}\|_{L_2}^2,$$

and similarly

$$\tau b(u_h^n, \bar{\partial}_t u_h^n) \geq \frac{1}{2}\|u_h^n\|_b^2 - \frac{1}{2}\|u_h^{n-1}\|_b^2.$$

Similar bounds will be used repeatedly throughout the paper. Summing (2.10) over n gives

$$\frac{1}{2}\|\bar{\partial}_t u_h^n\|_{L_2}^2 - \frac{1}{2}\|\bar{\partial}_t u_h^1\|_{L_2}^2 + \sum_{j=2}^{n}\tau\|\bar{\partial}_t u_h^j\|_a^2 + \frac{1}{2}\|u_h^n\|_b^2 - \frac{1}{2}\|u_h^1\|_b^2 \leq \sum_{j=2}^{n}\tau\|f^j\|_{H^{-1}}\|\bar{\partial}_t u_h^j\|_{H^1}.$$

Using the equivalence of the norms (2.7), Cauchy–Schwarz and Young's (weighted) inequality to subtract $\sum_{j=2}^{n}\tau\|\bar{\partial}_t u_h^j\|_{H^1}^2$ from both sides, we get exactly (2.8).

The proof of (2.9) is similar. We choose $v = \tau\bar{\partial}_t^2 u_h^n$ in (2.6) and sum over n to get

$$\sum_{j=2}^{n}\tau\|\bar{\partial}_t^2 u_h^j\|_{L_2}^2 + \frac{1}{2}\|\bar{\partial}_t u_h^n\|_a^2 - \frac{1}{2}\|\bar{\partial}_t u_h^1\|_a^2 + \sum_{j=2}^{n}\tau b\left(u_h^j, \bar{\partial}_t^2 u_h^j\right) \leq \sum_{j=2}^{n}\tau\|f^j\|_{L_2}\|\bar{\partial}_t^2 u_h^j\|_{L_2}.$$

For the sum involving the bilinear form $b(\cdot,\cdot)$ we use summation by parts to get

$$\sum_{j=2}^{n}\tau b\left(u_h^j, \bar{\partial}_t^2 u_h^j\right) = \sum_{j=3}^{n} -\tau b\left(\bar{\partial}_t u_h^j, \bar{\partial}_t u_h^{j-1}\right) - b\left(u_h^2, \bar{\partial}_t u_h^1\right) + b\left(u_h^n, \bar{\partial}_t u_h^n\right).$$

Using (2.8), the equivalence of the norms (2.7), and Young's weighted inequality we have

$$\begin{aligned}
&\left|\sum_{j=3}^{n}\tau b\left(\bar\partial_t u_h^j,\bar\partial_t u_h^{j-1}\right)+b\left(u_h^2,\bar\partial_t u_h^1\right)-b\left(u_h^n,\bar\partial_t u_h^n\right)\right|\\
&\qquad\le C\sum_{j=2}^{n}\tau\|\bar\partial_t u_h^j\|_{H^1}^2+C\left(\|u_h^2\|_{H^1}^2+\|\bar\partial_t u_h^1\|_{H^1}^2\right)+C\|u_h^n\|_{H^1}^2+C_\epsilon\|\bar\partial_t u_h^n\|_a^2\\
&\qquad\le C\sum_{j=2}^{n}\tau\|f^j\|_{H^{-1}}^2+C\left(\|\bar\partial_t u_h^1\|_{H^1}^2+\|u_h^1\|_{H^1}^2\right)+C_\epsilon\|\bar\partial_t u_h^n\|_a^2.
\end{aligned}$$

Since C_ϵ can be made arbitrarily small, it can be kicked to the left hand side. Using that $\|f^j\|_{H^{-1}}^2\le C\|f^j\|_{L_2}^2$ we deduce (2.9). □

3. Generalized finite element method

This section is dedicated to the development of a multiscale method based on the framework of the standard LOD. First of all, we introduce some notation for the discretization. Let V_H be a FE-space defined analogously to V_h in previous section, but with larger mesh size $H>h$. Moreover, we assume that the corresponding family of partitions $\{\mathcal{T}_H\}_{H>h}$ is, in addition to shape-regular, also quasi-uniform. Denote by $\mathcal{N}$ the set of interior nodes of V_H and by λ_x the standard hat function for $x\in\mathcal{N}$, such that $V_H=\operatorname{span}(\{\lambda_x\}_{x\in\mathcal{N}})$. Finally, we make the assumption that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$, such that $V_H\subseteq V_h$.

3.1. Ideal method

To define a generalized finite element method for our problem, we aim to construct a multiscale space V_{ms} of the same dimension as V_H, but with better approximation properties. For the construction of such a multiscale space, let $I_H:V_h\to V_H$ be an interpolation operator that has the projection property $I_H=I_H\circ I_H$ and satisfies

$$H_K^{-1}\|v-I_Hv\|_{L_2(K)}+\|\nabla I_Hv\|_{L_2(K)}\le C_I\|\nabla v\|_{L_2(N(K))},\quad\forall K\in\mathcal{T}_H,\ v\in V_h,\tag{3.1}$$

where $N(K):=\{K'\in\mathcal{T}_H:\overline{K'}\cap\overline{K}\neq\emptyset\}$. Furthermore, for a shape-regular and quasi-uniform partition, the estimate (3.1) can be summed into the global estimate

$$H^{-1}\|v-I_H\|_{L_2(\Omega)}+\|\nabla I_Hv\|_{L_2(\Omega)}\le C_\gamma\|\nabla v\|_{L_2(\Omega)},$$

where C_γ depends on the interpolation constant C_I and the shape regularity parameter defined as

$$\gamma:=\max_{K\in\mathcal{T}_H}\gamma_K,\quad\text{where}\quad\gamma_K=\frac{\operatorname{diam}(B_K)}{\operatorname{diam}(K)}.$$

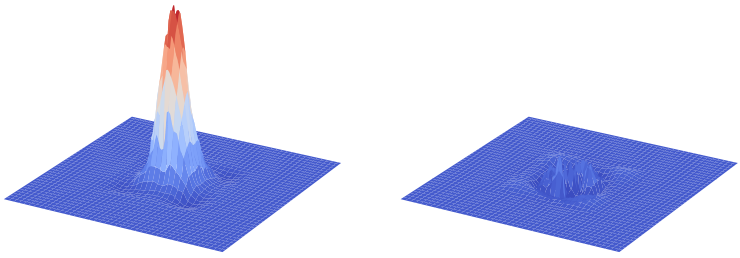
Here B_K denotes the largest ball inside K. A commonly used example of such an interpolant is $I_H=E_H\circ\Pi_H$, where Π_H is the piecewise L_2-projection onto $P_1(\mathcal{T}_H)$, the space of functions that are affine on each triangle $K\in\mathcal{T}_H$, and $E_H:P_1(\mathcal{T}_H)\to V_H$ is an averaging operator that, to each free node $x\in\mathcal{N}$, assigns the arithmetic mean of corresponding function values on intersecting elements, *i.e.*

$$(E_H(v))(x)=\frac{1}{\operatorname{card}\{K\in\mathcal{T}_H:x\in\overline{K}\}}\sum_{K\in\mathcal{T}_H:x\in\overline{K}}v\big|_K(x).$$

For more discussion regarding possible choices of interpolants, see *e.g.* [11] or [31].

Let the space V_{f} be defined by the kernel of the interpolant, *i.e.*

$$V_{\text{f}}=\ker(I_H)=\{v\in V_h:I_Hv=0\}.$$



(A) (B)

FIGURE 1. The modified basis function $\lambda_x - \phi_x$ and the Ritz-projected hat function ϕ_x. (A) $\lambda_x - \phi_x$. (B) ϕ_x.

That is, V_{f} is a finescale space in the sense that it captures the features that are excluded from the coarse FE-space. This consequently leads to the decomposition

$$V_h = V_H \oplus V_{\mathrm{f}},$$

such that every function $v \in V_h$ has a unique decomposition $v = v_H + v_{\mathrm{f}}$, where $v_H \in V_H$ and $v_{\mathrm{f}} \in V_{\mathrm{f}}$.

In the case of the LOD method for the standard wave equation (see [1]), one considers a Ritz-projection based solely on the B-coefficient to construct a multiscale space. Instead, the goal is to define a multiscale space based on the inner product $a(\cdot,\cdot) + \tau b(\cdot,\cdot)$ (for a fixed τ) and add additional correction to account for the time-dependency. This particular choice of scalar product comes from the backward Euler time-stepping formulation and both simplifies the analysis and is more natural in the implementation. Another possibility is to choose $a(\cdot,\cdot)$ as scalar product. For $v \in V_H$, we consider the Ritz-projection $R_{\mathrm{f}} : V_H \to V_{\mathrm{f}}$ defined by

$$a(R_{\mathrm{f}} v, w) + \tau b(R_{\mathrm{f}} v, w) = a(v, w) + \tau b(v, w), \quad \forall w \in V_{\mathrm{f}}.$$

Using this projection, we may define the multiscale space $V_{\mathrm{ms}} := V_H - R_{\mathrm{f}} V_H$ such that

$$V_h = V_{\mathrm{ms}} \oplus V_{\mathrm{f}}, \quad \text{and} \quad a(v_{\mathrm{ms}}, v_{\mathrm{f}}) + \tau b(v_{\mathrm{ms}}, v_{\mathrm{f}}) = 0. \tag{3.2}$$

Note that $\dim(V_{\mathrm{ms}}) = \dim(V_H)$, and hence we can view V_{ms} as a modified coarse space that contains finescale information of A and B. Next, we may use the Ritz-projection to define the basis functions for the space V_{ms}. For $x \in \mathcal{N}$, denote by $\phi_x := R_{\mathrm{f}} \lambda_x \in V_{\mathrm{f}}$ the solution to the (global) corrector problem

$$a(\phi_x, w) + \tau b(\phi_x, w) = a(\lambda_x, w) + \tau b(\lambda_x, w), \quad \forall w \in V_{\mathrm{f}}. \tag{3.3}$$

We can now construct our basis for V_{ms} as $\{\lambda_x - \phi_x\}_{x \in \mathcal{N}}$ which includes the behavior of the coefficients. For an illustration of the Ritz-projected hat function, as well as the modified basis function for V_{ms}, see Figure 1.

We may now formulate our ideal (but impractical) method. Since the solution space can be decomposed as $V_h = V_{\mathrm{ms}} \oplus V_{\mathrm{f}}$, the idea is to solve a coarse scale problem in V_{ms}, and then add additional correction from a problem on the fine scale. The method reads: find $u_{\mathrm{lod}}^n = v^n + w^n$, where $v^n \in V_{\mathrm{ms}}$ and $w^n \in V_{\mathrm{f}}$ such that

$$\tau\left(\bar{\partial}_t^2 v^n, z\right) + a\left(v^n, z\right) + \tau b\left(v^n, z\right) = \tau\left(f^n, z\right) + a\left(u_{\mathrm{lod}}^{n-1}, z\right), \quad \forall z \in V_{\mathrm{ms}}, \tag{3.4}$$

$$a\left(w^n, z\right) + \tau b\left(w^n, z\right) = a\left(u_{\mathrm{lod}}^{n-1}, z\right), \quad \forall z \in V_{\mathrm{f}}, \tag{3.5}$$

for $n \geq 2$ with initial data $u_{\mathrm{lod}}^0 = u_h^0 \in V_{\mathrm{ms}}$ and $u_{\mathrm{lod}}^1 = u_h^1 \in V_{\mathrm{ms}}$. The initial data is chosen in V_{ms} to simplify the implementation of the finescale correctors. We further discuss this choice in Section 4.4.

Remark 3.1. Note that in (3.5), we do not take the source function nor the second derivative into account. This is because we can subtract an interpolant within the L_2-product, so that the corresponding error converges at the same order as the method itself. Moreover, the v^n-part and w^n-part have been excluded from the bilinear form $a(\cdot,\cdot)+\tau b(\cdot,\cdot)$ in (3.4) and (3.5) respectively, due to the orthogonality between V_{ms} and V_{f}.

Note that the multiscale space V_{ms} is created using (3.3) with small τ. Thus, the A-coefficient dominates the system for short times. Moreover, we note from (3.5) that for N large enough, we reach a steady state so that $w^N \approx w^{N-1}$ and $v^N \approx v^{N-1}$. We get for $z \in V_{\text{f}}$

$$a\left(w^N, z\right) + \tau b\left(w^N, z\right) \approx a\left(u_{\text{lod}}^N, z\right) = a\left(v^N, z\right) + a\left(w^N, z\right) = -\tau b\left(v^N, z\right) + a\left(w^N, z\right),$$

due to the orthogonality. Hence, by rearranging terms we have that

$$b\left(v^N, z\right) + b\left(w^N, z\right) = b\left(u_{\text{lod}}^N, z\right) \approx 0,$$

which shows that the solution converges to a state where it is orthogonal with respect to B.

3.2. Localized method

The method we have considered so far is based on the global projection (3.3) onto the finescale space V_{f}, which results in a large linear system that is expensive to solve. Moreover, the basis correctors yield a global support that makes the linear system (3.4) not sparse, but dense. Hence, we wish to localize the computations onto coarse grid patches in order to yield a sparse matrix system.

To localize the corrector problem, we first introduce the patches to which the support of each basis function is to be restricted. For $\omega \subset \Omega$, let $N(\omega) := \{K \in \mathcal{T}_H : \overline{K} \cap \overline{\omega} \neq \emptyset\}$, and define a patch $N^k(\omega)$ of size k as

$$\begin{aligned} N^1(\omega) &:= N(\omega), \\ N^k(\omega) &:= N(N^{k-1}(\omega)), \quad \text{for } k \geq 2. \end{aligned}$$

Given these coarse grid patches, we may restrict the finescale space V_{f} to them by defining

$$V_{\text{f},k}^{\omega} := \{v \in V_{\text{f}} : \text{supp}(v) \subseteq N^k(\omega)\},$$

for a subdomain $\omega \subset \Omega$. In particular, we will commonly use $\omega = T \in \mathcal{T}_H$ and $\omega = x \in \mathcal{N}$.

Next, define the element restricted Ritz-projection R_{f}^T such that $R_{\text{f}}^T v \in V_{\text{f}}$ is the solution to the system

$$a\left(R_{\text{f}}^T v, z\right) + \tau b\left(R_{\text{f}}^T v, z\right) = \int_T (A + \tau B)\, \nabla v \cdot \nabla z \, \mathrm{d}x, \quad \forall z \in V_{\text{f}}.$$

Note that we may construct the global Ritz-projection as the sum

$$R_{\text{f}} v = \sum_{T \in \mathcal{T}_H} R_{\text{f}}^T v.$$

For $k \in \mathbb{N}$, we may restrict the projection to a patch by letting $R_{\text{f},k}^T : V_H \to V_{\text{f},k}^T$ be such that $R_{\text{f},k}^T v \in V_{\text{f},k}^T$ solves

$$a\left(R_{\text{f},k}^T v, z\right) + \tau b\left(R_{\text{f},k}^T v, z\right) = \int_T (A + \tau B)\, \nabla v \cdot \nabla z \, \mathrm{d}x, \quad \forall z \in V_{\text{f},k}^T.$$

By summation we yield the corresponding global version as

$$R_{\text{f},k} v = \sum_{T \in \mathcal{T}_H} R_{\text{f},k}^T v.$$

Finally, we may construct a localized multiscale space as $V_{\text{ms},k} := V_H - R_{\text{f},k} V_H$, spanned by $\{\lambda_x - R_{\text{f},k}\lambda_x\}_{x \in \mathcal{N}}$.

In order to justify the act of localization, it is required that a corrector ϕ_x vanishes rapidly outside an area of its corresponding node x. Indeed, the following theorem (see [27], Thm. 4.1) shows that the corrector ϕ_x satisfies an exponential decay away from its node, making the localization procedure viable.

Theorem 3.2. *There exists a constant $c \geq (8C_I\gamma(2+C_I))^{-1}$, that only depends on the mesh constant γ, such that for any $T \in \mathcal{T}_H$ and any $v \in H_0^1(\Omega)$ the solution $\phi \in V_{\mathrm{f}}$ of the variational problem*

$$\tilde{a}(\phi, w) = \int_T \left(\tilde{A}\nabla v\right) \cdot \nabla w \,\mathrm{d}x, \quad \forall w \in V_{\mathrm{f}}$$

satisfies

$$\|\tilde{A}^{1/2}\nabla\phi\|_{L_2(\Omega\setminus N^k(T))} \leq \sqrt{2}\exp\left(-c\tfrac{\alpha_- + \tau\beta_-}{\alpha_+ + \tau\beta_+}k\right)\|\tilde{A}^{1/2}\nabla v\|_{L_2(T)}, \quad \forall k \in \mathbb{N},$$

where $\tilde{A} = A + \tau B$.

With the space $V_{\mathrm{ms},k}$ defined, we are able to localize the computations on the coarse scale system in (3.4) by replacing the multiscale space by its localized counterpart. It remains to localize the computations of the finescale system in (3.5), which equivalently can be written as

$$a\left(\bar{\partial}_t w^n, z\right) + b\left(w^n, z\right) = \frac{1}{\tau}a\left(v^{n-1}, z\right).$$

We replace the right hand side by its localized version $v_k^{n-1} \in V_{\mathrm{ms},k}$ and note that $v_k^{n-1} = \sum_{x\in\mathcal{N}} \alpha_x^{n-1}\left(\lambda_x - R_{\mathrm{f},k}\lambda_x\right)$. Thus, we seek our localized finescale solution as $w_k^n = \sum_{x\in\mathcal{N}} w_{k,x}^n$, where $w_{k,x}^n \in V_{\mathrm{f},k}^x$ solves

$$a\left(\bar{\partial}_t w_{k,x}^n, z\right) + b\left(w_{k,x}^n, z\right) = \frac{1}{\tau}a\left(\alpha_x^{n-1}\left(\lambda_x - R_{\mathrm{f},k}\lambda_x\right), z\right), \quad \forall z \in V_{\mathrm{f},k}^x, \tag{3.6}$$

so that the computation of this equation is localized to a patch surrounding the node $x \in \mathcal{N}$. We introduce the functions $\xi_{k,x}^l \in V_{\mathrm{f},k}^x$ as solution to the parabolic equation

$$a\left(\bar{\partial}_t \xi_{k,x}^l, z\right) + b\left(\xi_{k,x}^l, z\right) = a\left(\frac{1}{\tau}\chi_1(l)\left(\lambda_x - R_{\mathrm{f},k}\lambda_x\right), z\right), \quad \forall z \in V_{\mathrm{f},k}^x, \tag{3.7}$$

for $l = 1, 2, \ldots, N$ with initial value $\xi_{k,x}^0 = 0$, and where $\chi_1(l)$ is an indicator function that equals 1 when $l = 1$ and 0 otherwise. We claim that $w_{k,x}^n = \sum_{l=1}^n \alpha_x^{n-l}\xi_{k,x}^l$ is the solution to (3.6). This follows as for all $z \in V_{\mathrm{f},k}^x$

$$\begin{aligned}
a\left(\bar{\partial}_t w_{k,x}^n, z\right) + b\left(w_{k,x}^n, z\right) &= a\left(\bar{\partial}_t \sum_{l=1}^n \alpha_x^{n-l}\xi_{k,x}^l, z\right) + b\left(\sum_{l=1}^n \alpha_x^{n-l}\xi_{k,x}^l, z\right) \\
&= \sum_{l=2}^n \alpha_x^{n-l}\left(a\left(\bar{\partial}_t\xi_{k,x}^l, z\right) + b\left(\xi_{k,x}^l, z\right)\right) + \alpha_x^{n-1}\left(a\left(\bar{\partial}_t\xi_{k,x}^1, z\right) + b\left(\xi_{k,x}^1, z\right)\right) \\
&= 0 + a\left(\alpha_x^{n-1}\left(\lambda_x - R_{\mathrm{f},k}\lambda_x\right), z\right).
\end{aligned}$$

With the localized computations established, the GFEM reads: find $u_{\mathrm{lod},k}^n = v_k^n + w_k^n$, where $v_k^n = \sum_{x\in\mathcal{N}} \alpha_x^n\left(\lambda_x - R_{\mathrm{f},k}\lambda_x\right) \in V_{\mathrm{ms},k}$ solves

$$\tau\left(\bar{\partial}_t^2 v_k^n, z\right) + a\left(v_k^n, z\right) + \tau b\left(v_k^n, z\right) = \tau\left(f^n, z\right) + a\left(u_{\mathrm{lod},k}^{n-1}, z\right), \quad \forall z \in V_{\mathrm{ms},k}, \tag{3.8}$$

and $w_k^n = \sum_{x\in\mathcal{N}}\sum_{l=1}^n \alpha_x^{n-l}\xi_{k,x}^l$, where $\xi_{k,x}^l \in V_{\mathrm{f},k}^x$ solves (3.7).

To justify the fact that we localize the finescale equation, we require a result similar to that of Theorem 3.2, but for the functions $\{\xi_x^l\}_{l=1}^N$. We finish this section about localization by proving that these functions satisfy the exponential decay required for the localization procedure to be viable.

Theorem 3.3. *For any node $x \in \mathcal{N}$, let $\xi_x^n \in V_{\mathrm{f}}$ be the solution to*

$$a\left(\bar{\partial}_t \xi_x^n, z\right) + b\left(\xi_x^n, z\right) = a\left(\frac{1}{\tau}\chi_1(n)\left(\lambda_x - R_{\mathrm{f}}\lambda_x\right), z\right), \quad \forall z \in V_{\mathrm{f}},$$

with initial value $\xi_x^0 = 0$. Then there exist constants $c > 0$ and $C > 0$ such that for any $k \geq 1$

$$\|\xi_x^n\|_{H^1(\Omega\backslash N^k(x))} \leq Ce^{-ck}\|\lambda_x\|_{H^1},$$

for sufficiently small time step τ.

Proof. Assume $k \geq 5$. First, we analyze the problem for the first time step, which when multiplied by τ can be written as

$$a\left(\xi_x^1, z\right) + \tau b\left(\xi_x^1, z\right) = a\left(\lambda_x - \phi_x, z\right), \quad \forall z \in V_{\mathrm{f}}, \tag{3.9}$$

where $\phi_x = R_{\mathrm{f}}\lambda_x$. We denote $\tilde{a} = a + \tau b$ such that $\tilde{a}\left(\phi_x, z\right) = \tilde{a}\left(\lambda_x, z\right)$ for all $z \in V_{\mathrm{f}}$. Furthermore we use the energy norm $|||\cdot||| := \sqrt{\tilde{a}\left(\cdot,\cdot\right)}$, and by $|||\cdot|||_D$ we denote the restriction of the norm onto a domain D. As seen in the proof of Theorem 4.1 in [27], the result in Theorem 3.2 can be written as

$$|||\phi_x|||_{\Omega\backslash N^k(x)} \leq C_\phi \mu^{\lfloor k/4 \rfloor}|||\lambda_x|||,$$

for some $\mu < 1$. Moreover we define the cut-off function $\eta_k \in V_H$ by

$$\eta_k := \begin{cases} 1, & \text{in } \Omega\backslash N^{k+1}(x), \\ 0, & \text{in } N^k(x), \end{cases}$$

for $x \in \mathcal{N}$. Now let $\nu = \eta_{k-3}$. Then we have that

$$\begin{aligned} \operatorname{supp}(\nu) &= \Omega\backslash N^{k-3}(x), \\ \operatorname{supp}(\nabla\nu) &= N^{k-2}(x)\backslash N^{k-3}(x). \end{aligned}$$

With this setting, we note that

$$\begin{aligned} |||\xi_x^1|||_{\Omega\backslash N^k}^2 &\leq \int_\Omega \nu\tilde{A}\nabla\xi_x^1 \cdot \nabla\xi_x^1 \,\mathrm{d}x = \int_\Omega \tilde{A}\nabla\xi_x^1 \cdot \nabla\left(\nu\xi_x^1\right) \,\mathrm{d}x - \int_\Omega \tilde{A}\nabla\xi_x^1 \cdot \xi_x^1\nabla\nu \,\mathrm{d}x \\ &\leq \underbrace{\left|\int_\Omega \tilde{A}\nabla\xi_x^1 \cdot \nabla\left(1 - I_H\right)\left(\nu\xi_x^1\right) \,\mathrm{d}x\right|}_{=:M_1} + \underbrace{\left|\int_\Omega \tilde{A}\nabla\xi_x^1 \cdot \nabla I_H\left(\nu\xi_x^1\right) \,\mathrm{d}x\right|}_{=:M_2} + \underbrace{\left|\int_\Omega \tilde{A}\nabla\xi_x^1 \cdot \xi_x^1\nabla\nu \,\mathrm{d}x\right|}_{=:M_3}, \end{aligned}$$

where we have denoted $\tilde{A} = A + \tau B$. We now proceed to estimate the terms M_1, M_2 and M_3 separately. For M_1, we use the problem (3.9) with $z = \left(1 - I_H\right)\left(\nu\xi_x^1\right) \in V_{\mathrm{f}}$ to get

$$M_1 = \left|\int_\Omega A\nabla\left(\lambda_x - \phi_x\right) \cdot \nabla\left(1 - I_H\right)\left(\nu\xi_x^1\right) \,\mathrm{d}x\right| = \left|\tau\int_{\Omega\backslash N^{k-3}} B\nabla\left(\phi_x\right) \cdot \nabla\left(1 - I_H\right)\left(\nu\xi_x^1\right) \,\mathrm{d}x\right|,$$

where we have used the $\tilde{a}$-orthogonality between V_{ms} and V_{f}, that the integral is zero on $\operatorname{supp}(\lambda_x)$, and that the support of the remaining integrand is $\Omega\backslash N^{k-4}$. Thus, we get that

$$M_1 \leq \tau\frac{\beta_+}{\alpha_-}|||\phi_x|||_{\Omega\backslash N^{k-4}}\left|||\xi_x^1\right|||_{\Omega\backslash N^{k-4}} \leq \tau\frac{\beta_+}{\alpha_-}C_\phi\mu^{\lfloor\frac{k-4}{4}\rfloor}|||\lambda_x|||\left|||\xi_x^1\right|||_{\Omega\backslash N^{k-4}}.$$

Moreover, by similar calculations as in the proof of Theorem 4.1 in [27], from M_2 and M_3 we get

$$M_2, M_3 \leq \tilde{C}\left|\left|\left|\xi_x^1\right|\right|\right|_{N^k\setminus N^{k-4}}^2,$$

for a constant $\tilde{C} > 0$. In total, for $\varepsilon \in (0,1)$, we find that

$$\begin{aligned}\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^k}^2 &\leq \tau\frac{\beta_+}{\alpha_-}C_\phi\mu^{\lfloor\frac{k-4}{4}\rfloor}\left|\left|\left|\lambda_x\right|\right|\right|\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^{k-4}} + \tilde{C}\left|\left|\left|\xi_x^1\right|\right|\right|_{N^k\setminus N^{k-4}}^2\\ &\leq \frac{\left(\beta_+C_\phi\right)^2}{\alpha_-^2\varepsilon}\tau^2\mu^{2\lfloor\frac{k-4}{4}\rfloor}\left|\left|\left|\lambda_x\right|\right|\right|^2 + \varepsilon\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^{k-4}}^2 + \tilde{C}\left(\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^{k-4}}^2 - \left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^k}^2\right).\end{aligned}$$

Let $\delta := \left(\varepsilon+\tilde{C}\right)\left(1+\tilde{C}\right)^{-1} < 1$, and set $\kappa = \max(\delta,\mu) < 1$. Then, by rearranging the terms we get the inequality

$$\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^k}^2 \leq \frac{\left(\beta_+C_\phi\right)^2}{\alpha_-^2\varepsilon\left(1+\tilde{C}\right)}\tau^2\kappa^{2\lfloor\frac{k-4}{4}\rfloor}\left|\left|\left|\lambda_x\right|\right|\right|^2 + \kappa\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^{k-4}}^2.$$

Repeating the estimate, we end up with

$$\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^k}^2 \leq \kappa^{\lfloor k/4\rfloor}\left|\left|\left|\xi_x^1\right|\right|\right|_\Omega^2 + \frac{\left(\beta_+C_\phi\right)^2}{\alpha_-^2\varepsilon\left(1+\tilde{C}\right)}\left|\left|\left|\lambda_x\right|\right|\right|^2\sum_{i=0}^{\lfloor k/4\rfloor-1}\tau^2\kappa^i\kappa^{2\lfloor\frac{k-4-4i}{4}\rfloor}.$$

We proceed by estimating $\left|\left|\left|\xi_x^1\right|\right|\right|_\Omega$. By choosing $z = \xi_x^1$ in (3.9) we get

$$\left|\left|\left|\xi_x^1\right|\right|\right|^2 \leq \left|\left|\left|\lambda_x-\phi_x\right|\right|\right|\left|\left|\left|\xi_x^1\right|\right|\right| \leq \left|\left|\left|\lambda_x\right|\right|\right|\left|\left|\left|\xi_x^1\right|\right|\right|,$$

since

$$\left|\left|\left|\lambda_x-\phi_x\right|\right|\right|^2 = \tilde{a}\left(\lambda_x-\phi_x,\lambda_x-\phi_x\right) \leq \left|\left|\left|\lambda_x-\phi_x\right|\right|\right|\left|\left|\left|\lambda_x\right|\right|\right|.$$

Moreover, for $i = 0,1,2,\ldots,\lfloor k/4\rfloor-1$, we note that

$$\kappa^{i+2\lfloor\frac{k-4-4i}{4}\rfloor} \leq \kappa^{\lfloor k/4\rfloor-1+2\lfloor\frac{k-4-4(k/4-1)}{4}\rfloor} = \kappa^{\lfloor k/4\rfloor-1} \tag{3.10}$$

so in total we have the estimate

$$\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^k} \leq \sqrt{1+C_0\tau^2}\kappa^{\frac{1}{2}\lfloor k/4\rfloor}\left|\left|\left|\lambda_x\right|\right|\right|, \quad \text{with} \quad C_0 = \frac{\left(\beta_+C_\phi\right)^2\kappa^{-1}}{\alpha_-^2\varepsilon\left(1+\tilde{C}\right)}\left(\lfloor k/4\rfloor-1\right).$$

Recall that this is for the first time step. In next time step, we consider the problem

$$a\left(\xi_x^2,z\right) + \tau b\left(\xi_x^2,z\right) = a\left(\xi_x^1,z\right), \quad \forall z \in V_{\mathrm{f}}.$$

As for the first time step, we split the estimate into the similar integrals M_1, M_2, and M_3, and get

$$M_1 \leq \tau\frac{\beta_+}{\alpha_-}\left|\left|\left|\xi_x^1\right|\right|\right|_{\Omega\setminus N^{k-4}}\left|\left|\left|\xi_x^2\right|\right|\right|_{\Omega\setminus N^{k-4}} \leq \tau\frac{\beta_+}{\alpha_-}\sqrt{1+C_0\tau^2}\kappa^{\frac{1}{2}\lfloor\frac{k-4}{4}\rfloor}\left|\left|\left|\lambda_x\right|\right|\right|\left|\left|\left|\xi_x^2\right|\right|\right|_{\Omega\setminus N^{k-4}},$$

while M_2 and M_3 remain the same. In total, we get the estimate

$$\left(1+\tilde{C}\right)\left|\left|\left|\xi_x^2\right|\right|\right|_{\Omega\setminus N^k}^2 \leq \frac{\beta_+^2}{\alpha_-^2\varepsilon}\tau^2\left(1+C_0\tau^2\right)\kappa^{\lfloor\frac{k-4}{4}\rfloor}\left|\left|\left|\lambda_x\right|\right|\right|^2 + \left(\varepsilon+\tilde{C}\right)\left|\left|\left|\xi_x^2\right|\right|\right|_{\Omega\setminus N^{k-4}}^2.$$

Once again, by letting $\delta = \left(\varepsilon + \tilde{C}\right) / \left(1 + \tilde{C}\right)$ and since $\delta \leq \kappa$, we get

$$\begin{aligned}\left|\left|\left|\xi_x^2\right|\right|\right|_{\Omega\setminus N^k}^2 &\leq \frac{\beta_+^2}{\alpha_-^2 \varepsilon \left(1+\tilde{C}\right)} \tau^2 \left(1 + C_0\tau^2\right) \kappa^{\left\lfloor \frac{k-4}{4} \right\rfloor} \left|\left|\left|\lambda_x\right|\right|\right|^2 + \kappa \left|\left|\left|\xi_x^2\right|\right|\right|_{\Omega\setminus N^{k-4}}^2 \\ &\leq \kappa^{\lfloor k/4 \rfloor} \left|\left|\left|\lambda_x\right|\right|\right|^2 + \frac{\beta_+^2}{\alpha_-^2 \varepsilon \left(1+\tilde{C}\right)} \left(1 + C_0\tau^2\right) \sum_{i=0}^{\lfloor k/4 \rfloor - 1} \tau^2 \kappa^i \kappa^{\left\lfloor \frac{k-4-4i}{4} \right\rfloor}.\end{aligned}$$

Once again we use (3.10) to conclude that

$$\left|\left|\left|\xi_x^2\right|\right|\right|_{\Omega\setminus N^k} \leq \sqrt{1 + C_1\tau^2\left(1 + C_0\tau^2\right)} \kappa^{\frac{1}{2}\lfloor k/4 \rfloor} \left|\left|\left|\lambda_x\right|\right|\right| = \sqrt{1 + C_1\tau^2 + C_1\tau^2 C_0\tau^2} \kappa^{\frac{1}{2}\lfloor k/4 \rfloor} \left|\left|\left|\lambda_x\right|\right|\right|,$$

where

$$C_1 = \frac{\beta_+^2 \kappa^{-1}}{\alpha_-^2 \varepsilon \left(1+\tilde{C}\right)} \left(\lfloor k/4 \rfloor - 1\right).$$

Inductively, we get for arbitrary time step n the estimate

$$\left|\left|\left|\xi_x^n\right|\right|\right|_{\Omega\setminus N^k} \leq \kappa^{\frac{1}{2}\lfloor k/4 \rfloor} \left|\left|\left|\lambda_x\right|\right|\right| \sqrt{\sum_{i=0}^{n-1} \left(C_1\tau^2\right)^i + \left(C_1\tau^2\right)^n C_0\tau^2}. \tag{3.11}$$

Since $\kappa^{\frac{1}{2}\lfloor k/4 \rfloor} \leq \kappa^{\frac{1}{2}(k/4-1)} = \kappa^{-\frac{1}{2}} e^{-\frac{1}{8}\log(1/\kappa)k}$, and since the energy norm is equivalent to the H^1-norm, the theorem holds for $k \geq 5$. We show that the estimate (3.11) is still valid for $k \leq 4$. For the n:th time step, let $z = \xi_x^n$ in (3.9). This yields the estimate

$$\left|\left|\left|\xi_x^n\right|\right|\right| \leq \left|\left|\left|\xi_x^{n-1}\right|\right|\right| \leq \ldots \leq \left|\left|\left|\xi_x^1\right|\right|\right| \leq \left|\left|\left|\lambda_x\right|\right|\right|.$$

Furthermore if $k < 4$ we have that $\kappa^{\frac{1}{2}\lfloor k/4 \rfloor} = 1$, and under the assumption that $C_1\tau^2 < 1$, we have that

$$\sum_{i=0}^{n-1} \left(C_1\tau^2\right)^i = \frac{\left(C_1\tau^2\right)^n - 1}{C_1\tau^2 - 1} > 1,$$

which shows that the estimate holds for $k < 4$. If $k = 4$ we can bound $\kappa^{\frac{1}{2}\lfloor k/4 \rfloor} \leq \kappa^{\frac{1}{2}\lfloor k/5 \rfloor}$ and repeat the same argument. □

Remark 3.4. Note that the sum that appears in (3.11) converges to

$$\sum_{i=0}^{n-1} \left(C_1\tau^2\right)^i \xrightarrow[n\to\infty]{} \frac{1}{1 - C_1\tau^2},$$

which means that the total constant in (3.11) behaves nicely for sufficiently small time steps. More specifically, for time steps

$$\tau \leq \sqrt{\frac{\alpha_-^2 \varepsilon \kappa \left(1+\tilde{C}\right)}{\beta_+^2 \left(\lfloor k/4 \rfloor - 1\right)}}.$$

4. Error estimates

In this section we derive error estimates of the ideal method (3.4) and (3.5). The additional error due to localization can be controlled in terms of the localization parameter k. This is further discussed in Remark 4.9. We begin by considering an auxiliary problem.

4.1. Auxiliary problem

The auxiliary problem is defined as the standard variational formulation for the strongly damped wave equation, but we exclude the second order time derivative. Moreover, we let the starting time $t = t_0$ be general and set the time discretization to $t = t_0 < t_1 < \ldots < t_N = T$. Note that N here might be different from the discretization of the fully discrete equation (2.6), but the time step τ, and thus the multiscale space, are the same. The auxiliary problem is to find $Z_h^n \in V_h$ for $n = 1, \ldots, N$, such that

$$a\left(\bar{\partial}_t Z_h^n, v\right) + b\left(Z_h^n, v\right) = \left(f^n, v\right), \quad \forall v \in V_h, \tag{4.1}$$

with initial value $Z_h^0 \in V_{\text{ms}}$. Equivalently, multiply (4.1) by τ and we may consider

$$a\left(Z_h^n, v\right) + \tau b\left(Z_h^n, v\right) = \tau\left(f^n, v\right) + a\left(Z_h^{n-1}, v\right), \quad \forall v \in V_h. \tag{4.2}$$

Existence of a solution to this problem is guaranteed by Lax–Milgram. For simplicity, we make the assumption that the initial data for the damped wave equation (2.6) is already in the multiscale space V_{ms}, such that

$$u_h^0 = u_{\text{lod}}^0 \in V_{\text{ms}}, \quad u_h^1 = u_{\text{lod}}^1 \in V_{\text{ms}}.$$

For general initial data we refer to Section 4.4 below. Furthermore, to limit the technical details in the proof we have chosen to analyze the error in the $L_2\left(H^1\right)$-norm instead of the pointwise (in time) H^1-norm.

The solution space can be decomposed as $V_h = V_{\text{ms}} \oplus V_{\text{f}}$, such that the solution can be written as $Z_h^n = v^n + w^n$ where $v^n \in V_{\text{ms}}$ and $w^n \in V_{\text{f}}$. If we insert this into the system in (4.2) and consider test functions $z \in V_{\text{ms}}$, the left hand side becomes

$$a\left(Z_h^n, z\right) + \tau b\left(Z_h^n, z\right) = a\left(v^n, z\right) + \tau b\left(v^n, z\right),$$

where we have used the orthogonality between V_{ms} and V_{f} with respect to $a\left(\cdot,\cdot\right) + \tau b\left(\cdot,\cdot\right)$. Likewise, if test functions $z \in V_{\text{f}}$ are considered, the left hand side becomes

$$a\left(Z_h^n, z\right) + \tau b\left(Z_h^n, z\right) = a\left(w^n, z\right) + \tau b\left(w^n, z\right).$$

With these findings, we define the approximation to the auxiliary problem as to find $Z_{\text{lod}}^n = v^n + w^n$, where $v^n \in V_{\text{ms}}$ and $w^n \in V_{\text{f}}$ such that

$$a\left(v^n, z\right) + \tau b\left(v^n, z\right) = \tau\left(f^n, z\right) + a\left(Z_{\text{lod}}^{n-1}, z\right), \quad \forall z \in V_{\text{ms}}, \tag{4.3}$$

$$a\left(w^n, z\right) + \tau b\left(w^n, z\right) = a\left(Z_{\text{lod}}^{n-1}, z\right), \quad \forall z \in V_{\text{f}}, \tag{4.4}$$

with initial data $Z_{\text{lod}}^0 \in V_{\text{ms}}$. Note that if $f = 0$, then $Z_h^n = Z_{\text{lod}}^n$ for every n, meaning that the method reproduces Z_h^n exactly. For the auxiliary problem, we prove the following error estimates.

Theorem 4.1. *Let Z_h^n be the solution to (4.1) and Z_{lod}^n the solution to (4.3) and (4.4). Assume that $Z_{\text{lod}}^0 - Z_h^0 = 0$, and $f^n \in L_2\left(\Omega\right)$, for $n \geq 0$, then the error is bounded by*

$$\|Z_h^n - Z_{\text{lod}}^n\|_{H^1} \leq CH \sum_{j=1}^{n} \tau \|f^j\|_{L_2}. \tag{4.5}$$

In addition,

$$\sum_{j=1}^{n}\tau\|Z_h^j-Z_{\mathrm{lod}}^j\|_{L_2}^2\le CH^2\sum_{j=1}^{n}\tau\|f^j\|_{L_2}^2, \tag{4.6}$$

and if $f^n=\bar\partial_t g^n$, *for some* $\{g^n\}_{n=0}^N$ *such that* $g^n\in V_h$, *then*

$$\sum_{j=1}^{n}\tau\|Z_h^j-Z_{\mathrm{lod}}^j\|_{L_2}^2\le CH^2\left(\sum_{j=1}^{n}\tau\|g^j\|_{L_2}^2+\|g^0\|_{L_2}^2\right), \tag{4.7}$$

where C *does not depend on the variations in* A *or* B.

Proof. Since $Z_h^n\in V_h$ there are $\bar v^n\in V_{\mathrm{ms}}$ and $\bar w^n\in V_{\mathrm{f}}$ such that $Z_h^n=\bar v^n+\bar w^n$. Let $e^n=Z_h^n-Z_{\mathrm{lod}}^n$, and consider

$$\begin{aligned}|||e^n|||^2&:=a\left(e^n,e^n\right)+\tau b\left(e^n,e^n\right)\\&=\tau\left(f^n,e^n\right)+a\left(Z_h^{n-1},e^n\right)-a\left(v^n,e^n\right)-\tau b\left(v^n,e^n\right)-a\left(w^n,e^n\right)-\tau b\left(w^n,e^n\right).\end{aligned}$$

For $v^n\in V_{\mathrm{ms}}$ we have due to the orthogonality and (4.3)

$$\begin{aligned}a\left(v^n,e^n\right)+\tau b\left(v^n,e^n\right)&=a\left(v^n,\bar v^n-v^n\right)+\tau b\left(v^n,\bar v^n-v^n\right)\\&=\tau\left(f^n,\bar v^n-v^n\right)+a\left(Z_{\mathrm{lod}}^{n-1},\bar v^n-v^n\right).\end{aligned}$$

Similarly, for $w^n\in V_{\mathrm{f}}$ we use the orthogonality and (4.4) to get

$$a\left(w^n,e^n\right)+\tau b\left(w^n,e^n\right)=a\left(Z_{\mathrm{lod}}^{n-1},\bar w^n-w^n\right).$$

Hence,

$$\begin{aligned}|||e^n|||^2&=\tau\left(f^n,e^n\right)+a\left(Z_h^{n-1},e^n\right)-\tau\left(f^n,\bar v^n-v^n\right)-a\left(Z_{\mathrm{lod}}^{n-1},\bar v^n-v^n\right)-a\left(Z_{\mathrm{lod}}^{n-1},\bar w^n-w^n\right)\\&=\tau\left(f^n,\bar w^n-w^n\right)+a\left(Z_h^{n-1}-Z_{\mathrm{lod}}^{n-1},e^n\right).\end{aligned}$$

The first term can be bounded by using the interpolation operator I_H

$$\begin{aligned}\tau|\left(f^n,\bar w^n-w^n\right)|&\le\tau\|f^n\|_{L_2}\|\bar w^n-w^n-I_H\left(\bar w^n-w^n\right)\|_{L_2}\le CH\tau\|f^n\|_{L_2}\|\bar w^n-w^n\|_{H^1}\\&\le CH\tau\|f^n\|_{L_2}\|e^n\|_{H^1}\le CH\tau\|f^n\|_{L_2}|||e^n|||.\end{aligned}$$

For the second term we note that $Z_h^{n-1}-Z_{\mathrm{lod}}^{n-1}=e^{n-1}$ so that

$$|||e^n|||\le CH\tau\|f^n\|_{L_2}+\left|\left|\left|e^{n-1}\right|\right|\right|.$$

Using this bound repeatedly and $e^0=0$ we get

$$|||e^n|||\le CH\sum_{j=1}^{n}\tau\|f^j\|_{L_2}.$$

This concludes the proof since $\|e^n\|_{H^1}\le C|||e^n|||$.

To prove the remaining bounds in L_2-norm, we define the forward difference operator $\tilde\partial_t x^n=\left(x^{n+1}-x^n\right)/\tau$ and consider the dual problem: find $x_h^j\in V_h$ for $j=n-1,\dots,0$, such that $x_h^n=0$ and

$$a\left(-\tilde\partial_t x_h^j,z\right)+b\left(x_h^j,z\right)=\left(e^{j+1},z\right),\quad\forall z\in V_h. \tag{4.8}$$

Note that this problem moves backwards in time. By choosing $z = x_h^j$ in (4.8) and performing a classical energy argument, we deduce

$$\|x_h^j\|_{H^1}^2 + \sum_{k=j}^{n} \tau \|x_h^k\|_{H^1}^2 \leq C \sum_{k=j+1}^{n} \tau \|e^k\|_{L_2}^2. \tag{4.9}$$

Similarly, by choosing $z = -\tilde{\partial}_t x_h^j$, we achieve

$$\|x_h^j\|_{H^1}^2 + \sum_{k=j}^{n-1} \tau \|\tilde{\partial}_t x_h^k\|_{H^1}^2 \leq C \sum_{k=j+1}^{n} \tau \|e^k\|_{L_2}^2. \tag{4.10}$$

Now, use (4.8) to get

$$\sum_{j=1}^{n} \tau \|e^j\|_{L_2}^2 = \sum_{j=1}^{n} \tau a\left(-\tilde{\partial}_t x_h^{j-1}, e^j\right) + \tau b\left(x_h^{j-1}, e^j\right).$$

Summation by parts gives

$$\sum_{j=1}^{n} \tau \|e^j\|_{L_2}^2 = \sum_{j=1}^{n} \tau a\left(-\tilde{\partial}_t x_h^{j-1}, e^j\right) + \tau b\left(x_h^{j-1}, e^j\right) = \sum_{j=1}^{n} \tau a\left(x_h^{j-1}, \bar{\partial}_t e^j\right) + \tau b\left(x_h^{j-1}, e^j\right), \tag{4.11}$$

where we have used $x^n = e^0 = 0$. Furthermore, we use the equations (4.1) and (4.3), and the orthogonality in (3.2), to show that the following Galerkin orthogonality holds for $z_{\text{ms}} \in V_{\text{ms}}$

$$\begin{aligned} a\left(\bar{\partial}_t e^j, z_{\text{ms}}\right) + b\left(e^j, z_{\text{ms}}\right) &= a\left(\bar{\partial}_t Z_h^j, z_{\text{ms}}\right) + b\left(Z_h^j, z_{\text{ms}}\right) - \frac{1}{\tau} a\left(v^j, z_{\text{ms}}\right) - b\left(v^j, z_{\text{ms}}\right) + \frac{1}{\tau} a\left(Z_{\text{lod}}^{j-1}, z_{\text{ms}}\right) \\ &= \left(f^j, z_{\text{ms}}\right) - \left(f^j, z_{\text{ms}}\right) = 0. \end{aligned} \tag{4.12}$$

Let $x_h^j = x_{\text{ms}}^j + x_{\text{f}}^j$, for some $x_{\text{ms}}^j \in V_{\text{ms}}$, $x_{\text{f}}^j \in V_{\text{f}}$. Using the orthogonality (4.12) and the equations (4.4) and (4.1) we deduce

$$\begin{aligned} \sum_{j=1}^{n} \tau a\left(x_h^{j-1}, \bar{\partial}_t e^j\right) + \tau b\left(x_h^{j-1}, e^j\right) &= \sum_{j=1}^{n} \tau a\left(x_{\text{f}}^{j-1}, \bar{\partial}_t e^j\right) + \tau b\left(x_{\text{f}}^{j-1}, e^j\right) = \sum_{j=1}^{n} \tau a\left(x_{\text{f}}^{j-1}, \bar{\partial}_t Z_h^j\right) + \tau b\left(x_{\text{f}}^{j-1}, Z_h^j\right) \\ &= \sum_{j=1}^{n} \tau \left(x_{\text{f}}^{j-1}, f^j\right). \end{aligned}$$

If $f^j \in L_2(\Omega)$, then we may subtract $I_H x_{\text{f}}^{j-1} = 0$ and use (3.1) to achieve

$$\sum_{j=1}^{n} \tau \left(x_{\text{f}}^{j-1}, f^j\right) \leq CH \sum_{j=1}^{n} \tau \|x_{\text{f}}^{j-1}\|_{H^1} \|f^j\|_{L_2} \leq CH \left(\sum_{j=1}^{n} \tau \|x_{\text{f}}^{j-1}\|_{H^1}^2\right)^{1/2} \left(\sum_{j=1}^{n} \tau \|f^j\|_{L_2}^2\right)^{1/2}.$$

Note that $\left|\left|\left|x_{\text{f}}^{j-1}\right|\right|\right|^2 + \left|\left|\left|x_{\text{ms}}^{j-1}\right|\right|\right|^2 \leq \left|\left|\left|x_h^{j-1}\right|\right|\right|^2$. Hence the energy estimate (4.9) can now be used to achieve (4.6). If $f^j = \bar{\partial}_t g^j$ one may use summation by parts to achieve

$$\begin{aligned} \sum_{j=1}^{n} \tau \|e^j\|_{L_2}^2 = \sum_{j=1}^{n} \tau \left(x_{\text{f}}^{j-1}, \bar{\partial}_t g^j\right) &\leq \sum_{j=1}^{n} \tau \left(-\tilde{\partial}_t x_{\text{f}}^{j-1}, g^j\right) - \left(x_{\text{f}}^0, g^0\right) \\ &\leq CH \sum_{j=1}^{n} \tau \|\tilde{\partial}_t x_{\text{f}}^{j-1}\|_{H^1} \|g^j\|_{L_2} + CH \|x_{\text{f}}^0\|_{H^1} \|g^0\|_{L_2}, \end{aligned}$$

where we have used $x_{\text{f}}^n = x_h^n = 0$. Using (4.10) we conclude (4.7). □

Remark 4.2. The bound in (4.6) is not of optimal order, but it is useful in the error analysis.

The next lemma gives error estimates for the discrete time derivative of the error. In the analysis of the (full) damped wave equation we use $g = \bar{\partial}_t u_h$, see Lemma 4.5. If the initial data is nonzero we expect $\bar{\partial}_t g^1$ below to be of order t_1^{-1} in L_2-norm. A detailed explanation of this is given below. Hence, we have a blow up close to zero due to low regularity of the initial data. Therefore, we need to multiply the error by t_j. This is similar to the parabolic case for nonsmooth initial data see, *e.g.* [33].

Lemma 4.3. *Let Z_h^n be the solution to (4.1) and Z_{lod}^n the solution to (4.3) and (4.4). Assume $Z_{\mathrm{lod}}^0 - Z_h^0 = 0$. If $\bar{\partial}_t f^n \in L_2(\Omega)$, for $n \geq 1$, then*

$$\sum_{j=2}^{n} \tau \|\bar{\partial}_t \left(Z_h^j - Z_{\mathrm{lod}}^j \right)\|_{L_2}^2 \leq CH^2 \left(\sum_{j=2}^{n} \tau \|\bar{\partial}_t f^j\|_{L_2}^2 + \|f^1\|_{L_2}^2 \right) \tag{4.13}$$

and if $f^n = \bar{\partial}_t g^n$, for some $\{g^n\}_{n=0}^N$, such that $g^n \in V_h$, then

$$\sum_{j=2}^{n} \tau \|\bar{\partial}_t \left(Z_h^j - Z_{\mathrm{lod}}^j \right)\|_{L_2}^2 \leq CH^2 \left(\sum_{j=2}^{n} \tau \|\bar{\partial}_t g^j\|_{L_2}^2 + \|\bar{\partial}_t g^1\|_{L_2}^2 \right) \tag{4.14}$$

and, in addition, the following bound holds

$$\sum_{j=2}^{n} \tau t_j^2 \|\bar{\partial}_t \left(Z_h^j - Z_{\mathrm{lod}}^j \right)\|_{L_2}^2 \leq CH^2 \left(\sum_{j=2}^{n} \tau \|\bar{\partial}_t g^j\|_{L_2}^2 + t_1^2 \|\bar{\partial}_t g^1\|_{L_2}^2 \right), \tag{4.15}$$

where C does not depend on the variations in A or B.

Proof. The proof of (4.13) is similar to (4.6). Let $e^j = Z_h^j - Z_{\mathrm{lod}}^j$ and define the dual problem

$$a\left(-\tilde{\partial}_t x_h^j, z\right) + b\left(x_h^j, z\right) = \left(\bar{\partial}_t e^{j+1}, z\right), \quad \forall z \in V_h, \ j = n-1, \ldots, 0, \tag{4.16}$$

with $x_h^n = 0$. Choosing $z = \bar{\partial}_t e^{j+1}$ and performing summation by parts we deduce

$$\begin{aligned} \sum_{j=2}^{n} \tau \|\bar{\partial}_t e^j\|_{L_2}^2 &= \sum_{j=2}^{n} \tau a\left(-\tilde{\partial}_t x_h^{j-1}, \bar{\partial}_t e^j\right) + \tau b\left(x_h^{j-1}, \bar{\partial}_t e^j\right) \\ &= \sum_{j=2}^{n} \tau a\left(x_h^{j-1}, \bar{\partial}_t^2 e^j\right) + \tau b\left(x_h^{j-1}, \bar{\partial}_t e^j\right) + a\left(x_h^1, \bar{\partial}_t e^1\right), \end{aligned} \tag{4.17}$$

where we used that $x_h^n = 0$. Following the same argument as for (4.7), but with a difference quotient, we arrive at

$$\sum_{j=2}^{n} \|\bar{\partial}_t e^j\|_{L_2}^2 = \sum_{j=2}^{n} \tau a\left(x_h^{j-1}, \bar{\partial}_t^2 e^j\right) + \tau b\left(x_h^{j-1}, \bar{\partial}_t e^j\right) + a\left(x_h^1, \bar{\partial}_t e^1\right) = \sum_{j=2}^{n} \tau \left(x_{\mathrm{f}}^{j-1}, \bar{\partial}_t f^j\right) + a\left(x_h^1, \bar{\partial}_t e^1\right).$$

Using $e^0 = 0$, we deduce

$$a\left(x_h^1, \bar{\partial}_t e^1\right) = \frac{1}{\tau} a\left(x_h^1, e^1\right) \leq \frac{C}{\tau} H \|x_h^1\|_{H^1} \tau \|f^1\|_{L_2} \leq CH \|x_h^1\|_{H^1} \|f^1\|_{L_2},$$

and with $\bar{\partial}_t f^j \in L_2(\Omega)$ we get

$$\sum_{j=2}^{n} \|\bar{\partial}_t e^j\|_{L_2}^2 \le CH \sum_{j=2}^{n} \tau \|x_{\mathrm{f}}^{j-1}\|_{H^1} \|\bar{\partial}_t f^j\|_{L_2} + CH \|x_h^1\|_{H^1} \|f^1\|_{L_2},$$

and (4.13) follows by using an energy estimate of x_h^j similar to (4.9), but with $\bar{\partial}_t e^j$ on the right hand side.

If $f^j = \bar{\partial}_t g^j$ we proceed as for (4.7) to achieve

$$\sum_{j=2}^{n} \tau \left(x_{\mathrm{f}}^{j-1}, \bar{\partial}_t f^j \right) \le CH \sum_{j=2}^{n} \tau \|\tilde{\partial}_t x_{\mathrm{f}}^{j-1}\|_{H^1} \|\bar{\partial}_t g^j\|_{L_2} + CH \|x_{\mathrm{f}}^1\|_{H^1} \|\bar{\partial}_t g^1\|_{L_2}$$

and (4.14) follows by using energy estimates similar to (4.10).

For (4.15) we consider the dual problem

$$a\left(-\tilde{\partial}_t x_h^j, z\right) + b\left(x_h^j, z\right) = \left(t_{j+1} \bar{\partial}_t e^{j+1}, z\right), \quad \forall z \in V_h, \ j = n-1, \dots, 0. \tag{4.18}$$

A simple energy estimate shows

$$\|x_h^j\|_{H^1}^2 + \sum_{k=j}^{n-1} \tau \|\tilde{\partial}_t x_h^k\|_{H^1}^2 \le C \sum_{k=j}^{n-1} \tau t_{k+1}^2 \|\bar{\partial}_t e^{k+1}\|_{L_2}^2, \quad j = 0, \dots, n-1. \tag{4.19}$$

Now choosing $z = t_{j+1} \bar{\partial}_t e^{j+1}$ in (4.18) and performing summation by parts gives

$$\begin{aligned} \sum_{j=2}^{n} \tau t_j^2 \|\bar{\partial}_t e^j\|_{L_2}^2 &= \sum_{j=2}^{n} \tau a\left(-\tilde{\partial}_t x_h^{j-1}, t_j \bar{\partial}_t e^j\right) + \tau b\left(x_h^{j-1}, t_j \bar{\partial}_t e^j\right) \\ &= \sum_{j=2}^{n} \Big(\tau a\left(x_h^{j-1}, t_j \bar{\partial}_t^2 e^j\right) + \tau b\left(x_h^{j-1}, t_j \bar{\partial}_t e^j\right) \\ &\quad + a\left(x_h^{j-1}, (t_j - t_{j-1}) \bar{\partial}_t e^{j-1}\right) \Big) + a\left(x_h^1, t_1 \bar{\partial}_t e^1\right). \end{aligned} \tag{4.20}$$

The first two terms of the sum can be handled similarly to (4.14),

$$\sum_{j=2}^{n} \tau a\left(x_h^{j-1}, t_j \bar{\partial}_t^2 e^j\right) + \tau b\left(x_h^{j-1}, t_j \bar{\partial}_t e^j\right) = \sum_{j=2}^{n} \tau \left(x_{\mathrm{f}}^{j-1}, t_j \bar{\partial}_t f^j\right).$$

Now, using summation by parts we achieve

$$\begin{aligned} \sum_{j=2}^{n} \tau \left(x_{\mathrm{f}}^{j-1}, t_j \bar{\partial}_t f^j\right) &= \sum_{j=2}^{n} \tau \left(\left(-\tilde{\partial}_t x_{\mathrm{f}}^{j-1}, t_j f^j\right) - \left(x_{\mathrm{f}}^j, f^j\right) \right) - \left(x_{\mathrm{f}}^1, t_2 f^1\right) \\ &\le CH \left(\sum_{j=2}^{n} \tau \left(t_j \|\tilde{\partial}_t x_{\mathrm{f}}^{j-1}\|_{H^1} \|f^j\|_{L_2} + \|x_{\mathrm{f}}^j\|_{H^1} \|f^j\|_{L_2} \right) + t_2 \|x_{\mathrm{f}}^1\|_{H^1} \|f^1\|_{L_2} \right) \end{aligned}$$

where we can use (4.19). Note that in the first term we can use the (crude) bound $t_j^2 \le t_n^2$ and let the constant C depend on T. Moreover, we use that $t_2 \le 2t_1$. We get

$$\sum_{j=2}^{n} \tau \left(x_{\mathrm{f}}^{j-1}, t_j \bar{\partial}_t f^j\right) \le CH \left(\sum_{j=1}^{n} \tau t_j^2 \|\bar{\partial}_t e^j\|_{L_2}^2 \right)^{1/2} \left(\left(\sum_{j=2}^{n} \tau \|f^j\|_{L_2}^2 \right)^{1/2} + t_1 \|f^1\|_{L_2} \right).$$

For the third term in (4.20), we use $t_j - t_{j-1} = \tau$ and once again perform summation by parts to get

$$\sum_{j=2}^{n} \tau a\left(x_h^{j-1}, \bar{\partial}_t e^{j-1}\right) = \sum_{j=2}^{n} \tau a\left(-\tilde{\partial}_t x_h^{j-1}, e^{j-1}\right),$$

where we have used $x_h^n = e^0 = 0$. Combining (4.19) and (4.5) we get

$$\begin{aligned}\sum_{j=2}^{n} \tau a\left(-\tilde{\partial}_t x_h^{j-1}, e^{j-1}\right) &\leq C \max_{j=1,\dots,n} \|e^j\|_{H^1} \left(\sum_{j=2}^{n} \tau\right)^{1/2} \left(\sum_{j=2}^{n} \tau \|\tilde{\partial}_t x_h^{j-1}\|_{H^1}^2\right)^{1/2} \\ &\leq CH \sum_{j=1}^{n} \tau \|f^j\|_{L_2} \left(\sum_{j=1}^{n} \tau t_j^2 \|\bar{\partial}_t e^j\|_{L_2}^2\right)^{1/2}.\end{aligned}$$

For the last term in (4.20) we use (4.19) and (4.5) for $n = 1$ to achieve

$$a\left(x_h^1, t_1 \bar{\partial}_t e^1\right) = a\left(x_h^1, e^1\right) \leq CH \left(\sum_{j=1}^{n} \tau t_j^2 \|\bar{\partial}_t e^j\|_{L_2}^2\right)^{1/2} t_1 \|f^1\|_{L_2},$$

and (4.15) follows by letting $f^j = \bar{\partial}_t g^j$. □

4.2. The damped wave equation

For the error analysis of the full damped wave equation we shall make use of the projection corresponding to the auxiliary problem. For $u_h^n \in V_h$, let $X^n = X_v^n + X_w^n \in V_h$ with $X_v^n \in V_{\text{ms}}$ and $X_w^n \in V_{\text{f}}$ such that

$$a\left(X_v^n - u_h^n, z\right) + \tau b\left(X_v^n - u_h^n, z\right) = a\left(X^{n-1} - u_h^{n-1}, z\right), \quad \forall z \in V_{\text{ms}}, \tag{4.21}$$

$$a\left(X_w^n, z\right) + \tau b\left(X_w^n, z\right) = a\left(X^{n-1}, z\right), \quad \forall z \in V_{\text{f}}. \tag{4.22}$$

Note that since u_h^n solves (2.6), the system (4.21) and (4.22) is equivalent to

$$a\left(X_v^n, z\right) + \tau b\left(X_v^n, z\right) = \tau\left(f^n - \bar{\partial}_t^2 u_h^n, z\right) + a\left(X^{n-1}, z\right), \quad \forall z \in V_{\text{ms}}, \tag{4.23}$$

$$a\left(X_w^n, z\right) + \tau b\left(X_w^n, z\right) = a\left(X^{n-1}, z\right), \quad \forall z \in V_{\text{f}}. \tag{4.24}$$

That is, we may view u_h^n and X^n as the solution and approximation to the auxiliary problem with source data $f^n - \bar{\partial}_t^2 u_h^n$. We deduce following lemma.

Lemma 4.4. *Let u_h^n be the solution to (2.6) and X^n the solution to (4.21) and (4.22). The error satisfies the following bounds*

$$\|X^n - u_h^n\|_{H^1} \leq CH \sum_{j=2}^{n} \tau \|f^j - \bar{\partial}_t^2 u_h^j\|_{L_2}, \qquad n \geq 2, \tag{4.25}$$

$$\sum_{j=2}^{n} \tau \|X^j - u_h^j\|_{L_2}^2 \leq CH^2 \left(\sum_{j=2}^{n} \tau \left(\|f^j\|_{L_2}^2 + \|\bar{\partial}_t u_h^j\|_{L_2}^2\right) + \|\bar{\partial}_t u_h^1\|_{L_2}^2\right), \qquad n \geq 2, \tag{4.26}$$

where C does not depend on the variations in A or B.

Proof. We let the auxiliary problem (4.1) start at t_1 with initial data u_h^1, such that the error at the initial time is zero, *i.e.* $e^0 = 0$. The bound (4.25) now follows directly from (4.5) with $f^n - \bar{\partial}_t^2 u_h^n$ as right hand side. The second bound (4.26) follows from (4.6) and (4.7) with $f^n \in L_2(\Omega)$ and $g^n = \bar{\partial}_t u_h^{n+1}$. □

In a similar way me may deduce bounds for the (discrete) time derivative of the error. As a direct consequence of Lemma 4.3, we get the following result.

Lemma 4.5. *Let u_h^n be the solution to (2.6) and X^n the solution to (4.21) and (4.22). The following bounds hold*

$$\sum_{j=3}^{n} \tau \|\bar{\partial}_t \left(X^j - u_h^j \right) \|_{L_2}^2 \leq CH^2 \left(\sum_{j=3}^{n} \tau \left(\|\bar{\partial}_t f^j\|_{L_2}^2 + \|\bar{\partial}_t^2 u_h^j\|_{L_2}^2 \right) + \|f^2\|_{L_2}^2 + \|\bar{\partial}_t^2 u_h^2\|_{L_2}^2 \right), \tag{4.27}$$

$$\sum_{j=3}^{n} \tau t_j^2 \|\bar{\partial}_t \left(X^j - u_h^j \right) \|_{L_2}^2 \leq CH^2 \left(\sum_{j=3}^{n} \tau \left(\|\bar{\partial}_t f^j\|_{L_2}^2 + \|\bar{\partial}_t^2 u_h^j\|_{L_2}^2 \right) + t_2^2 \|f^2\|_{L_2}^2 + t_2^2 \|\bar{\partial}_t^2 u_h^2\|_{L_2}^2 \right), \tag{4.28}$$

where C does not depend on the variations in A or B.

Lemma 4.6. *Let u_h^n and u_{lod}^n be the solutions to (2.6) and (3.4), (3.5), respectively. Assume that $u_0 = u_1 = 0$. The error is bounded by*

$$\sum_{j=2}^{n} \tau \|u_{\text{lod}}^j - u_h^j\|_{H^1}^2 \leq CH^2 \left(\sum_{j=1}^{n} \tau \left(\|f^j\|_{L_2}^2 + \|\bar{\partial}_t f^j\|_{L_2}^2 \right) + \max_{j=1,\ldots,n} \|f^j\|_{L_2}^2 \right),$$

for $n \geq 2$, where C does not depend on the variations in A or B.

Proof. Begin by splitting the error into two contributions

$$u_{\text{lod}}^n - u_h^n = u_{\text{lod}}^n - X^n + X^n - u_h^n =: \theta^n + \rho^n,$$

where X^n is the solution to the simplified problem in (4.21) and (4.22). By Lemma 4.4 ρ^n is bounded by

$$\|\rho^n\|_{H^1} \leq CH \sum_{j=2}^{n} \tau \left(\|f^j\|_{L_2} + \|\bar{\partial}_t^2 u_h^j\|_{L_2} \right),$$

and we can now apply the energy bound (2.9). It remains to bound θ^n. Recall that for any $z \in V_h$ we have $z = z_{\text{ms}} + z_{\text{f}}$ for some $z_{\text{ms}} \in V_{\text{ms}}$ and $z_{\text{f}} \in V_{\text{f}}$. Using that $u_{\text{lod}}^n = v^n + w^n$ satisfies (3.4) and the orthogonality (3.2) we get

$$\left(\bar{\partial}_t^2 u_{\text{lod}}^n, z_{\text{ms}}\right) + a\left(\bar{\partial}_t u_{\text{lod}}^n, z_{\text{ms}}\right) + b\left(u_{\text{lod}}^n, z_{\text{ms}}\right) = \left(f^n, z_{\text{ms}}\right) + \left(\bar{\partial}_t^2 w^n, z_{\text{ms}}\right).$$

Similarly, due to (3.5) and the orthogonality,

$$\left(\bar{\partial}_t^2 u_{\text{lod}}^n, z_{\text{f}}\right) + a\left(\bar{\partial}_t u_{\text{lod}}^n, z_{\text{f}}\right) + b\left(u_{\text{lod}}^n, z_{\text{f}}\right) = \left(\bar{\partial}_t^2 u_{\text{lod}}^n, z_{\text{f}}\right).$$

For X^n we use (4.21) and (4.22) and the orthogonality to deduce

$$\left(\bar{\partial}_t^2 X^n, z\right) + a\left(\bar{\partial}_t X^n, z\right) + b\left(X^n, z\right) = \left(\bar{\partial}_t^2 X^n, z\right) + \left(f^n - \bar{\partial}_t^2 u_h^n, z_{\text{ms}}\right), \quad z \in V_h,$$

Hence, θ^n satisfies

$$\left(\bar{\partial}_t^2 \theta^n, z\right) + a\left(\bar{\partial}_t \theta^n, z\right) + b\left(\theta^n, z\right) = \left(-\bar{\partial}_t^2 \rho^n, z\right) - \left(\bar{\partial}_t^2 u_h^n, z_{\text{f}}\right) + \left(\bar{\partial}_t^2 u_{\text{lod}}^n, z_{\text{f}}\right) + \left(\bar{\partial}_t^2 w^n, z_{\text{ms}}\right), \quad z \in V_h,$$

with $\theta^0 = \theta^1 = 0$, since $u_{\text{lod}}^0 = u_h^0 = X^0$ and $u_{\text{lod}}^1 = u_h^1 = X^1$. Let $\tilde{\theta}^n = \sum_{j=2}^{n} \tau \theta^j$. Multiplying by τ and summing over n gives

$$\left(\bar{\partial}_t \theta^n, z\right) + a\left(\theta^n, z\right) + b\left(\tilde{\theta}^n, z\right) \leq \left(-\bar{\partial}_t \rho^n, z\right) - \left(\bar{\partial}_t u_h^n - \bar{\partial}_t u_h^1, z_{\text{f}}\right) + \left(\bar{\partial}_t u_{\text{lod}}^n - \bar{\partial}_t u_{\text{lod}}^1, z_{\text{f}}\right) \tag{4.29}$$

$$+\left(\bar{\partial}_t w^n-\bar{\partial}_t w^1,z_{\mathrm{ms}}\right),$$

where we have used that $\theta^1=\theta^0=\rho^1=\rho^0=0$. Using the interpolant I_H we deduce

$$\left(\bar{\partial}_t u_h^n,z_{\mathrm{f}}\right)+\left(\bar{\partial}_t u_{\mathrm{lod}}^n,z_{\mathrm{f}}\right)+\left(\bar{\partial}_t w^n,z_{\mathrm{ms}}\right)\leq CH\left(\|\bar{\partial}_t u_h^n\|_{L_2}+\|\bar{\partial}_t u_{\mathrm{lod}}^n\|_{L_2}\right)\|z\|_{H^1}+CH\|\bar{\partial}_t u_{\mathrm{lod}}^n\|_{H^1}\|z_{\mathrm{ms}}\|_{L_2},$$

for $1\leq n\leq N$. Let $\alpha(n)=\|\bar{\partial}_t u_h^n\|_{L_2}+\|\bar{\partial}_t u_{\mathrm{lod}}^n\|_{H^1}$. Since $\|z_{\mathrm{ms}}\|_{L_2}\leq C\|z\|_{H^1}$ and $\alpha(1)=0$ due to the vanishing initial data, we get

$$\left(\bar{\partial}_t\theta^n,z\right)+a\left(\theta^n,z\right)+b\left(\tilde{\theta}^n,z\right)\leq\left(-\bar{\partial}_t\rho^n,z\right)+CH\alpha(n)\|z\|_{H^1},\ z\in V_h.$$

Now, choose $z=\theta^n=\bar{\partial}_t\tilde{\theta}^n$ in (4.29). We get

$$\frac{1}{2}\|\theta^n\|_{L_2}^2-\frac{1}{2}\|\theta^{n-1}\|_{L_2}^2+\tau\|\theta^n\|_a^2+\frac{1}{2}\|\tilde{\theta}^n\|_b^2-\frac{1}{2}\|\tilde{\theta}^{n-1}\|_b^2\leq\tau\|\bar{\partial}_t\rho^n\|_{L_2}\|\theta^n\|_{L_2}+CH\tau\alpha(n)\|\theta^n\|_{H^1}.$$

Summing over n gives

$$\|\theta^n\|_{L_2}^2+\sum_{j=2}^{n}\tau\|\theta^j\|_{H^1}^2+\|\tilde{\theta}^n\|_{H^1}^2\leq\sum_{j=2}^{n}\tau\|\bar{\partial}_t\rho^j\|_{L_2}\|\theta^j\|_{L_2}+CH\sum_{j=2}^{n}\tau\alpha(j)\|\theta^j\|_{H^1}.$$

Now using that $\|\theta^n\|_{L_2}\leq\|\theta^n\|_{H^1}$ and Young's weighted inequality, θ^j can be kicked back to the left hand side. We deduce

$$\sum_{j=2}^{n}\tau\|\theta^j\|_{H^1}^2\leq C\sum_{j=2}^{n}\tau\|\bar{\partial}_t\rho^j\|_{L_2}^2+CH^2\sum_{j=2}^{n}\tau\alpha(j)^2.$$

Using Lemma 4.5 we have

$$\sum_{j=2}^{n}\tau\|\theta^j\|_{H^1}^2\leq CH^2\left(\sum_{j=2}^{n}\tau\left(\|\bar{\partial}_t f^j\|_{L_2}^2+\|\bar{\partial}_t^2 u_h^j\|_{L_2}^2\right)+\|\bar{\partial}_t^2 u_h^2\|_{L_2}^2\right)+CH^2\sum_{j=2}^{n}\tau\alpha(j)^2.$$

To bound $\|\bar{\partial}_t^2 u_h^2\|_{L_2}^2$, we consider (2.6) for $n=2$ and choose $v=\bar{\partial}_t^2 u_h^2$, which gives

$$\left(\bar{\partial}_t^2 u_h^2,\bar{\partial}_t^2 u_h^2\right)+a\left(\bar{\partial}_t u_h^2,\bar{\partial}_t^2 u_h^2\right)+b\left(u_h^2,\bar{\partial}_t^2 u_h^2\right)=\left(\bar{\partial}_t f^2,\bar{\partial}_t^2 u_h^2\right).$$

Due to the vanishing initial data $\bar{\partial}_t u_h^2=\tau^{-1}u_h^2$ and $\bar{\partial}_t^2 u_h^2=\tau^{-2}u_h^2$. We get

$$\|\bar{\partial}_t^2 u_h^2\|_{L_2}^2+\frac{1}{\tau^3}\|u_h^2\|_a^2+\frac{1}{\tau^2}\|u_h^2\|_b^2=\left(f^2,\bar{\partial}_t^2 u_h^2\right),\tag{4.30}$$

and we deduce

$$\|\bar{\partial}_t^2 u_h^2\|_{L_2}^2\leq C\|f^2\|_{L_2}^2.$$

All terms, except $\sum_{j=2}^{n}\tau\|\bar{\partial}_t u_{\mathrm{lod}}^j\|_{H^1}^2$ that appears in $\sum_{j=2}^{n}\alpha^2(j)$, can now be bounded by using the regularity in Theorem 2.2. To bound $\sum_{j=2}^{n}\tau\|\bar{\partial}_t u_{\mathrm{lod}}^j\|_{H^1}^2$ we choose $z=\bar{\partial}_t v^n$ and $z=\bar{\partial}_t w^n$ in (3.4) and (3.5) respectively. Adding the two equations and using the orthogonality between V_{ms} and V_{f} we achieve

$$\left(\bar{\partial}_t^2 v^n,\bar{\partial}_t v^n\right)+a\left(\bar{\partial}_t u_{\mathrm{lod}}^n,\bar{\partial}_t u_{\mathrm{lod}}^n\right)+b\left(u_{\mathrm{lod}}^n,\bar{\partial}_t u_{\mathrm{lod}}^n\right)=\left(f^n,\bar{\partial}_t v^n\right)\leq C_\epsilon\|f^n\|_{L_2}^2+\epsilon\|\bar{\partial}_t v^n\|_{L_2}^2.$$

Note that $\|\bar{\partial}_t v^n\|_{L_2} \leq C\|\nabla\bar{\partial}_t v^n\|_{L_2} \leq C|||\bar{\partial}_t v^n||| = C|||\bar{\partial}_t u^n_{\mathrm{lod}}||| \leq C\|\bar{\partial}_t u^n_{\mathrm{lod}}\|_a$, so we may choose ϵ small enough such that $|||\bar{\partial}_t u^n_{\mathrm{lod}}|||$ can be kicked to the left hand side. As in the proof of Theorem 2.2 we may now deduce

$$\|\bar{\partial}_t v^n\|^2_{L_2} + \sum_{j=2}^{n} \tau\|\bar{\partial}_t u^j_{\mathrm{lod}}\|^2_{H^1} + \|u^n_{\mathrm{lod}}\|^2_{H^1} \leq C\left(\sum_{j=2}^{n}\|f^j\|^2_{L_2} + \|\bar{\partial}_t u^1_h\|^2_{L_2} + \|u^1_h\|^2_{H^1}\right), \tag{4.31}$$

where we have used that $v^1 = u^1_{\mathrm{lod}} = u^1_h \in V_{\mathrm{ms}}$. However, we have assumed vanishing initial data so these terms disappear. The lemma follows. □

Lemma 4.7. *Let u^n_h and u^n_{lod} be the solutions to (2.6) and (3.4), (3.5), respectively. Assume that $f = 0$. The error is bounded by*

$$\sum_{j=2}^{n} \tau t_j^2\|u^j_{\mathrm{lod}} - u^j_h\|^2_{H^1} \leq CH^2\left(\|\bar{\partial}_t u^1_h\|^2_{H^1} + \|u^1_h\|^2_{H^1} + \|u^0_h\|^2_{H^1}\right), \quad n \geq 2,$$

where C does not depend on the variations in A or B.

Proof. We follow the steps in the proof of Lemma 4.6 to equation (4.29). Note that $\|\rho^n\|_{H^1}$ can be bounded by Lemma 4.4 and the energy bound in (2.9) with $f = 0$.

Now, let $\tilde{\theta}^n = \sum_{j=2}^{n}\tau\theta^j$. Choose $z = \theta^n = \bar{\partial}_t\tilde{\theta}^n$ in (4.29) and multiply by τt_n^2. We get

$$\begin{aligned}\frac{t_n^2}{2}\|\theta^n\|^2_{L_2} - \frac{t_{n-1}^2}{2}\|\theta^{n-1}\|^2_{L_2} + \tau t_n^2\|\theta^n\|^2_a + \frac{t_n^2}{2}\|\tilde{\theta}^n\|^2_b - \frac{t_{n-1}^2}{2}\|\tilde{\theta}^{n-1}\|^2_b \\ \leq \tau t_n^2\|\bar{\partial}_t\rho^n\|_{L_2}\|\theta^n\|_{L_2} + CHt_n^2\tau\left(\alpha(n) + \alpha(1)\right)\|\theta^n\|_{H^1} + \frac{\left(t_n^2 - t_{n-1}^2\right)}{2}\|\theta^{n-1}\|^2_{L_2} + \frac{\left(t_n^2 - t_{n-1}^2\right)}{2}\|\tilde{\theta}^{n-1}\|^2_b.\end{aligned}$$

Summing over n and using $t_n^2 - t_{n-1}^2 \leq 2\tau t_n$ gives

$$\begin{aligned}t_n^2\|\theta^n\|^2_{L_2} + \sum_{j=2}^{n}\tau t_j^2\|\theta^j\|^2_{H^1} + t_n^2\|\tilde{\theta}^n\|^2_{H^1} \leq C\sum_{j=2}^{n}\tau t_j^2\|\bar{\partial}_t\rho^j\|_{L_2}\|\theta^j\|_{L_2} + CH\sum_{j=2}^{n}\tau t_j^2\left(\alpha(j) + \alpha(1)\right)\|\theta^j\|_{H^1} \\ + C\sum_{j=2}^{n}\tau t_j\|\theta^j\|^2_{L_2} + C\sum_{j=2}^{n}\tau t_j\|\tilde{\theta}^j\|^2_b.\end{aligned} \tag{4.32}$$

From the first two sums on the right hand side we can kick $t_j\|\theta^j\|_{L_2} \leq t_j\|\theta^j\|_{H^1}$ and $t_j\|\theta^j\|_{H^1}$ to the left hand side. The remaining two sums needs to be bounded by other energy estimates.

Multiply (4.29) by τ and sum over n to get

$$\begin{aligned}(\theta^n, z) + a\left(\tilde{\theta}^n, z\right) + b\left(\sum_{j=2}^{n}\tau\tilde{\theta}^j, z\right) \leq (\rho^n, z) - \left(u^n_h - u^1_h, z_{\mathrm{f}}\right) + \left(u^n_{\mathrm{lod}} - u^1_{\mathrm{lod}}, z_{\mathrm{f}}\right) + \left(w^n - w^1, z_{\mathrm{ms}}\right) \\ + t_n\left(\left(\bar{\partial}_t u^1_h, z_{\mathrm{f}}\right) - \left(\bar{\partial}_t u^1_{\mathrm{lod}}, z_{\mathrm{f}}\right) - \left(\bar{\partial}_t w^1, z_{\mathrm{ms}}\right)\right).\end{aligned} \tag{4.33}$$

where we have used $\theta^1 = \rho^1 = 0$. As in the proof of Lemma 4.6 we get

$$(u^n_h, z_{\mathrm{f}}) + (u^n_{\mathrm{lod}}, z_{\mathrm{f}}) + (w^n, z_{\mathrm{ms}}) \leq CH\left(\|u^n_h\|_{L_2} + \|u^n_{\mathrm{lod}}\|_{L_2}\right)\|z\|_{H^1} + CH\|u^n_{\mathrm{lod}}\|_{H^1}\|z_{\mathrm{ms}}\|_{L_2},$$

for $1 \leq n \leq N$. Let $\beta(n) = \|u^n_h\|_{L_2} + \|u^n_{\mathrm{lod}}\|_{H^1}$. Choose $z = \tilde{\theta}^n = \bar{\partial}_t\sum_{j=1}^{n}\tau\tilde{\theta}^j$. Similar to above energy estimates, we get

$$\|\tilde{\theta}^n\|^2_{L_2} + \sum_{j=2}^{n}\tau\|\tilde{\theta}^j\|^2_a + \|\sum_{j=2}^{n}\tau\tilde{\theta}^j\|^2_b \leq C\sum_{j=2}^{n}\tau\|\rho^j\|^2_{L_2} + CH^2\sum_{j=2}^{n}\tau\left(\beta(j) + \beta(1) + \alpha(1)\right)^2. \tag{4.34}$$

Since $\sum_{j=2}^{n}\tau t_j\|\tilde{\theta}^j\|_b^2 \le C(t_n)\sum_{j=2}^{n}\tau\|\tilde{\theta}^j\|_a^2$ we may use (4.34) in (4.32). This gives

$$t_n^2\|\theta^n\|_{L_2}^2+\sum_{j=2}^{n}\tau t_j^2\|\theta^j\|_{H^1}^2+t_n^2\|\tilde{\theta}^n\|_{H^1}^2 \le C\sum_{j=2}^{n}\tau\left(t_j^2\|\bar{\partial}_t\rho^j\|_{L_2}^2+\|\rho^j\|_{L_2}^2\right)+C\sum_{j=2}^{n}\tau t_j\|\theta^j\|_{L_2}^2 \tag{4.35}$$
$$+CH^2\sum_{j=2}^{n}\tau\left(t_j^2\left(\alpha(j)+\alpha(1)\right)^2+\left(\beta(j)+\beta(1)+\alpha(1)\right)^2\right).$$

It remains to bound $C\sum_{j=2}^{n}\tau t_j\|\theta^j\|_{L_2}^2$. For this purpose, choose $z=\theta^n=\bar{\partial}_t\tilde{\theta}^n$ in (4.33). Multiply by $t_n\tau$ and sum over n to achieve

$$\sum_{j=2}^{n}\tau t_j\|\theta^j\|_{L_2}^2+t_n\|\tilde{\theta}^n\|_a^2+\sum_{j=2}^{n}t_j\tau b\left(\sum_{k=2}^{j}\tau\tilde{\theta}^k,\bar{\partial}_t\tilde{\theta}^j\right) \le C\sum_{j=1}^{n}\tau t_j\|\rho^j\|_{L_2}^2+C\sum_{j=2}^{n}\tau\|\tilde{\theta}^j\|_a^2 \tag{4.36}$$
$$+CH\sum_{j=2}^{n}\tau t_j\left(\beta(j)+\beta(1)+\alpha(1)\right)\|\theta^j\|_{H^1}.$$

Note that $\|\theta^j\|_{H^1}$ is the last sum in only present in the right hand side. The second term on the right hand side is bounded by (4.34). For the term involving the bilinear form $b(\cdot,\cdot)$ we use summation by parts to get

$$-\sum_{j=2}^{n}\tau b\left(t_j\sum_{k=2}^{j}\tau\tilde{\theta}^k,\bar{\partial}_t\sum_{k=2}^{j}\tau\theta^k\right) \le \sum_{j=2}^{n}\tau b\left(t_j\tilde{\theta}^j+\sum_{k=2}^{j}\tau\tilde{\theta}^k,\tilde{\theta}^{j-1}\right)-b\left(t_n\sum_{j=2}^{n}\tau\tilde{\theta}^j,\tilde{\theta}^n\right)$$
$$\le C\sum_{j=2}^{n}\tau t_j\|\tilde{\theta}^j\|_b^2+C\|\sum_{j=2}^{n}\tau\tilde{\theta}^j\|_b^2+C_\epsilon t_n^2\|\tilde{\theta}^n\|_{H^1}^2.$$

Here the constant C_ϵ can be made arbitrarily small due to Young's weighted inequality. The first two terms can be bounded by (4.34). Thus, (4.36) becomes

$$\sum_{j=2}^{n}\tau t_j\|\theta^j\|_{L_2}^2+t_n\|\tilde{\theta}^n\|_a^2 \le C\sum_{j=1}^{n}\tau\left(t_j\|\rho^j\|_{L_2}^2+\|\rho^j\|_{L_2}^2\right)+CH\sum_{j=2}^{n}\tau t_j\left(\beta(j)+\beta(1)+\alpha(1)\right)\|\theta^j\|_{H^1}$$
$$+CH^2\sum_{j=2}^{n}\tau\left(\beta(j)+\beta(1)+\alpha(1)\right)^2+C_\epsilon t_n^2\|\tilde{\theta}^n\|_{H^1}^2.$$

Using this in (4.35) we arrive at

$$t_n^2\|\theta^n\|_{L_2}^2+\sum_{j=2}^{n}\tau t_j^2\|\theta^j\|_{H^1}^2+t_n^2\|\tilde{\theta}^n\|_{H^1}^2 \le C\sum_{j=2}^{n}\tau\left(t_j^2\|\bar{\partial}_t\rho^j\|_{L_2}^2+t_j\|\rho^j\|_{L_2}^2\right)$$
$$+CH^2\sum_{j=2}^{n}\tau\left(t_j^2\left(\alpha(j)+\alpha(1)\right)^2+\left(\beta(j)+\beta(1)+\alpha(1)\right)^2\right) \tag{4.37}$$
$$+CH\sum_{j=2}^{n}\tau t_j\left(\beta(j)+\beta(1)+\alpha(1)\right)\|\theta^j\|_{H^1}+C_\epsilon t_n^2\|\tilde{\theta}^n\|_{H^1}^2.$$

Using Lemmas 4.5 and 4.4 with $f=0$ we deduce for the first two terms in (4.37)

$$C\sum_{j=2}^{n}\tau\left(t_j^2\|\bar{\partial}_t\rho^j\|_{L_2}^2+t_j\|\rho^j\|_{L_2}^2\right) \le CH^2\left(\sum_{j=2}^{n}\tau\|\bar{\partial}_t^2u_h^j\|_{L_2}^2+t_2^2\|\bar{\partial}_t^2u_h^2\|_{L_2}^2+\|\bar{\partial}_tu_h^1\|_{L_2}^2\right),$$

where we can use (2.9) for $n = 2$ and $f = 0$ to bound $\bar{\partial}_t^2 u_h^2$. We get

$$t_2^2 \|\bar{\partial}_t^2 u_h^2\|_{L_2}^2 \leq C\tau \left(\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \|u_h^1\|_{H^1}^2 \right).$$

For the remaining terms in (4.37) we note that $t_j \|\theta^j\|_{H^1}$ now may be kicked to left hand side using Cauchy–Schwarz and Young's weighted inequality. The term involving C_ϵ can also be moved to the left hand side. All terms involving $\alpha(j)$ and $\beta(j)$ can be bounded by (2.8) and (4.31). This finishes the proof after using the regularity in Theorem 2.2 with $f = 0$. □

4.3. Error bound for the ideal method

We get the final result by combining the two previous lemmas.

Corollary 4.8. *Let u_h^n and u_{lod}^n be the solutions to (2.6) and (3.4), (3.5), respectively. The solutions can be split into $u_h^n = u_{h,1}^n + u_{h,2}^n$ and $u_{\text{lod}}^n = u_{\text{lod},1}^n + u_{\text{lod},2}^n$, where the first part has vanishing initial data, and the second part a vanishing right hand side. The error is bounded by*

$$\sum_{j=2}^{n} \tau \|u_{h,1}^j - u_{\text{lod},1}^j\|_{H^1}^2 \leq CH^2 \left(\sum_{j=1}^{n} \tau \left(\|f^j\|_{L_2}^2 + \|\bar{\partial}_t f^j\|_{L_2}^2 \right) + \max_{j=1,\ldots,n} \|f^j\|_{L_2}^2 \right), \quad \text{for } n \geq 2$$

and

$$\sum_{j=2}^{n} \tau t_j^2 \|u_{h,2}^j - u_{\text{lod},2}^j\|_{H^1}^2 \leq CH^2 \left(\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \|u_h^1\|_{H^1}^2 + \|u_h^0\|_{H^1}^2 \right), \quad \text{for } n \geq 2.$$

Proof. This is a direct consequence of Lemmas 4.6 and 4.7 together with the fact that the problem is linear so the error can be split into two contributions satisfying the conditions of each lemma. □

Remark 4.9. The result from Corollary 4.8 is derived for the ideal method presented in (3.4) and (3.5). The GFEM in (3.7) and (3.8) will yield yet another error from the localization procedure. However, due to the exponential decay in Theorems 3.2 and 3.3, it holds for the choice $k \approx |\log(H)|$ that the perturbation from the ideal method is of higher order and the derived result in Corollary 4.8 is still valid. For the details regarding the error from the localization procedure, we refer to [29].

4.4. Initial data

For general initial data $u_h^0, u_h^1 \in V_h$ we consider the projections $R_{\text{ms}} u_h^0$ and $R_{\text{ms}} u_h^1$, where $R_{\text{ms}} = I - R_{\text{f}}$ is the Ritz-projection onto V_{ms}. Let v be the difference between two solutions to the damped wave equation with the different initial data. From (2.8) it follows that

$$\|v\|_{H^1}^2 \leq C \left(\|\bar{\partial}_t \left(u_h^1 - R_{\text{ms}} u_h^1 \right)\|_{L_2}^2 + \|u_h^1 - R_{\text{ms}} u_h^1\|_b^2 \right),$$

where we have chosen to keep the b-norm. For the first term we may use the interpolant I_H to achieve H. For the second term use

$$\|u_h^1 - R_{\text{ms}} u_h^1\|_b^2 \leq \frac{\beta_+}{\alpha_- + \tau\beta_-} \left(\|u_h^1 - R_{\text{ms}} u_h^1\|_a^2 + \tau \|u_h^1 - R_{\text{ms}} u_h^1\|_b^2 \right). \tag{4.38}$$

If the initial data fulfills the following condition for some $g \in L_2(\Omega)$

$$a\left(u_h^1, v\right) + \tau b\left(u_h^1, v\right) = (g, v), \quad \forall v \in V_h, \tag{4.39}$$

then we may deduce

$$\|u_h^1 - R_{\text{ms}} u_h^1\|_a^2 + \tau \|u_h^1 - R_{\text{ms}} u_h^1\|_b^2 = \left(g, u_h^1 - R_{\text{ms}} u_h^1\right) \leq CH \|g\|_{L_2} \|u_h^1 - R_{\text{ms}} u_h^1\|_{H^1}.$$

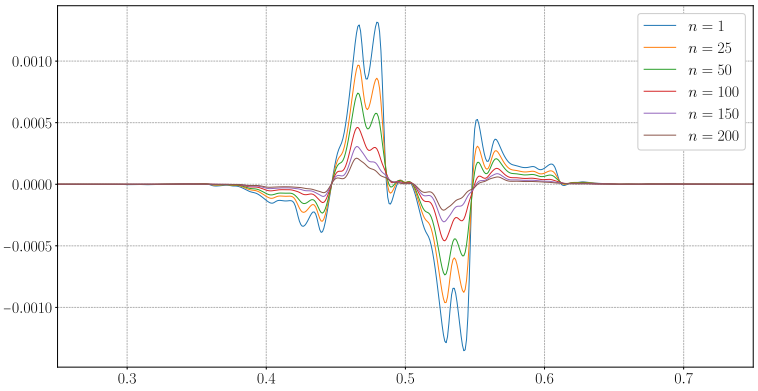


FIGURE 2. The behavior of the correction functions $\xi_{k,x}^n$ with increasing n. The time step is $\tau = 0.01$ and k is here chosen so that the support covers the entire interval.

Hence, the error introduced by the projection of the initial data is of order H. The condition (4.39) appears when applying the LOD method to classical wave equations, see [1], where it is referred to as "well prepared data". We note in our case that if B is small compared to A, that is if the damping is strong, then the constant in (4.38) is small. In some sense, this means that the condition in (4.39) is of "less importance", which is consistent with the fact that strong damping reduces the impact of the initial data over time.

5. Reduced basis method

The GFEM as it is currently stated requires us to solve the system in (3.7) for each coarse node in each time step, *i.e.* N number of times. We will alter the method by applying a reduced basis method, such that it will suffice to find the solutions for $M < N$ time steps, and compute the remaining in a significantly cheaper and efficient way.

First of all, we note how the system (3.7) that $\xi_{k,x}^n$ solves resembles a parabolic type equation with no source term. That is, the solution will decay exponentially until it is completely vanished. An example of how $\xi_{k,x}^n$ vanish with increasing n can be seen in Figure 2, where the coefficients are given as

$$A(x) = \left(2 - \sin\left(\tfrac{2\pi x}{\varepsilon_A}\right)\right)^{-1} \quad \text{and} \quad B(x) = \left(2 - \cos\left(\tfrac{2\pi x}{\varepsilon_B}\right)\right)^{-1},$$

with $\varepsilon_A = 2^{-4}$ and $\varepsilon_B = 2^{-6}$.

In Figure 2 it is also seen how the solutions decay with a similar shape through all time steps. This gives the idea that it is possible to only evaluate the solutions for a few time steps, and utilize these solutions to find the remaining ones. This idea can be further investigated by storing the solutions $\{\xi_{k,x}^n\}_{n=1}^N$ and analyzing the corresponding singular values. The singular values are plotted and seen in Figure 3. It is seen how the values decrease rapidly, and that most of the values lie on machine precision level. In practice, this means that the information in $\{\xi_{k,x}^n\}_{n=1}^N$ can be extracted from only a few $\xi_{k,x}^n$'s. We use this property to decrease the computational complexity by means of a reduced basis method. We remark that singular value decomposition is not used for the method itself, but is merely used as a tool to analyze the possibility of applying reduced basis methods.

The main idea behind reduced basis methods is to find an approximate solution in a low-dimensional space $V_{M,k,x}^{\mathrm{RB}}$, which is created using a number of already computed solutions. More precisely, to construct a basis for this space, one first computes M solutions $\{\xi_{k,x}^m\}_{m=1}^M$, where $M < N$. By orthonormalizing these solutions using *e.g.* Gram–Schmidt orthonormalization, we yield a set of vectors $\{\zeta_{k,x}^m\}_{m=1}^M$, called the reduced basis.

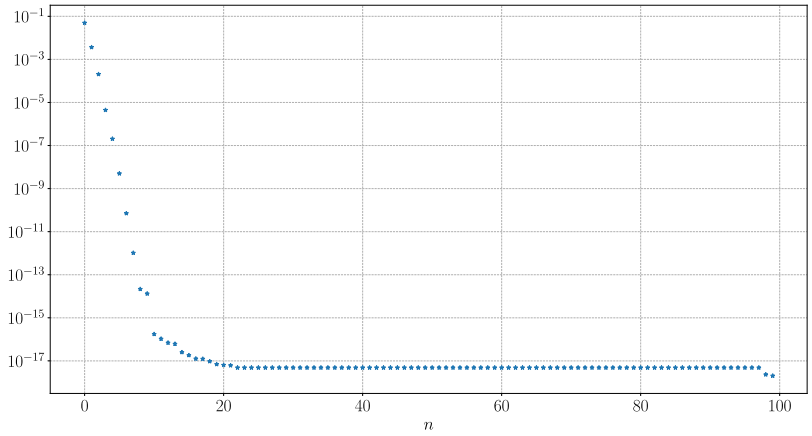


FIGURE 3. The singular values obtained when performing a singular value decomposition of the matrix created by storing the finescale corrections $\{\xi_{k,x}^n\}_{n=1}^N$ with $N = 100$.

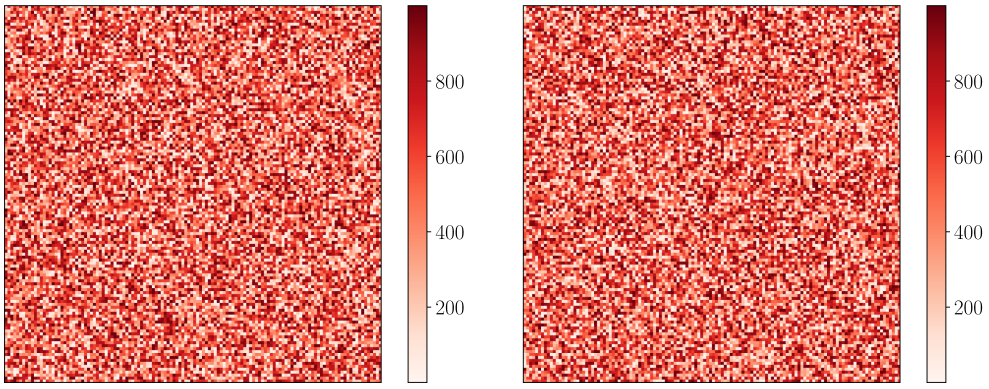
Consequently, the reduced basis space becomes $V_{M,k,x}^{\mathrm{RB}} = \operatorname{span}\left(\{\zeta_{k,x}^m\}_{m=1}^M\right)$ for each node $x \in \mathcal{N}$. With this space created, the procedure of finding $\{\xi_{k,x}^n\}_{n=1}^N$ is now reduced to finding $\{\xi_{k,x}^n\}_{n=1}^M$, and then approximate the remaining solutions by $\{\xi_{k,x}^{n,\mathrm{rb}}\}_{n=M+1}^N \subset V_{M,k,x}^{\mathrm{RB}}$. More precisely, for $l = M+1, M+2, \ldots, N$ we seek $\xi_{k,x}^{l,\mathrm{rb}} \in V_{M,k,x}^{\mathrm{RB}}$ such that

$$a\left(\bar{\partial}_t \xi_{k,x}^{l,\mathrm{rb}}, z\right) + b\left(\xi_{k,x}^{l,\mathrm{rb}}, z\right) = a\left(\frac{1}{\tau}\chi_1(l)\left(\lambda_x - R_{\mathrm{f},k}\lambda_x\right), z\right), \quad \forall z \in V_{M,k,x}^{\mathrm{RB}}. \tag{5.1}$$

The matrix system to solve for a solution in $V_{M,k,x}^{\mathrm{RB}}$ is of dimension $M \times M$, so when M is chosen small, the last $N - M$ solutions are significantly cheaper to compute, which solves the issue of computing N problems on the finescale space.

When constructing the reduced basis $\{\zeta_{k,x}^m\}_{m=1}^M$, it is important to be aware of the fact that the solution corrections $\{\xi_{k,x}^n\}_{n=1}^N$ all show very similar behavior. In practice, this implies that many of the $\xi_{k,x}^n$'s are linearly dependent, hence causing floating point errors to become of significant size in the RB-space $V_{M,k,x}^{\mathrm{RB}}$. To work around this issue, one may include a relative tolerance level that removes a vector from the basis if it is too close to being linearly dependent to one of the previously orthonormalized vectors. Note that this may imply that we get $\tilde{M} < M$ basis vectors in our RB-space instead of M. One may moreover use this tolerance level as a criterion for the amount of solutions, M, to pre-compute. That is, once the first vector is removed from the orthonormalization process, then the RB-space contains sufficient information and no more solutions need to be added.

In total, the novel method first requires that we solve N_H number of systems on the localized fine scale in order to construct the multiscale space $V_{\mathrm{ms},k}$. Moreover, we require to solve a localized fine system N_H times for M time steps to create the RB-space $V_{M,k,x}^{\mathrm{RB}}$ for each coarse node $x \in \mathcal{N}$. By utilizing the RB-space, the remaining $N - M$ finescale corrections are then solved for in an $M \times M$ matrix system, and we yield the sought solution $u_{\mathrm{lod},k}^{N,\mathrm{rb}}$ by computing a matrix system on the coarse grid with the multiscale space $V_{\mathrm{ms},k}$. We summarize the reduced basis approach in Algorithm 1.



(A) (B)

FIGURE 4. The two different coefficients used for the numerical examples. The contrast is $\alpha_+/\alpha_- = \beta_+/\beta_- = 10^4$. (A) $A(x,y)$. (B) $B(x,y)$.

Algorithm 1: Summary how to compute the finescale correctors $\{\xi^n_{k,x}\}_{n=1}^N$ with the RB-approach.

```
Pick M, relative_tol;
for all x ∈ 𝒩 do
    for l = 1, 2, ..., M do
        Solve (3.7) for ξ^l_{k,x} and store;
    Compute reduced basis {ζ^m_{k,x}}_{m=1}^{M̃} = gram_schmidt({ξ^m_{k,x}}_{m=1}^{M}, relative_tol);
    Construct V^RB_{M̃,k,x} = span({ζ^m_{k,x}}_{m=1}^{M̃});
    for l = M + 1, ..., N do
        Solve (5.1) for ξ^{l,rb}_{k,x} and store;
```

6. Numerical examples

In this section we illustrate numerical examples for the novel method. At first, we present numerical examples that confirm the theoretical findings derived in this paper. In addition, we provide a practical example related to seismology where the Marmousi model is used together with the Ricker wavelet as source function.

6.1. Academic example

In this section we present numerical examples that illustrate the performance of the established theory. For all examples, we consider the domain to be the unit square $\Omega = [0,1] \times [0,1]$. The coefficients $A(x,y)$ and $B(x,y)$ used in these examples are generated randomly with values in the interval $[10^{-1}, 10^3]$, and examples of such are seen in Figure 4. Moreover, as initial value for each example we set $u_0 = u_1 = 0$, and the source function is given by $f = 1$.

The first example is used to show how the performance is effected by the localization parameter k. Here, we evaluate the solution on the full grid, u^n_{lod}, and compare it with the localized solution, $u^n_{\text{lod},k}$, as k varies. For the example the time step $\tau = 0.02$ was used and final time was set to $T = 1$. The fine and coarse meshes were set to $h = 2^{-7}$ and $H = 2^{-4}$ respectively, and we let $k = 2, 3, \ldots, 7$. The relative error between the functions can

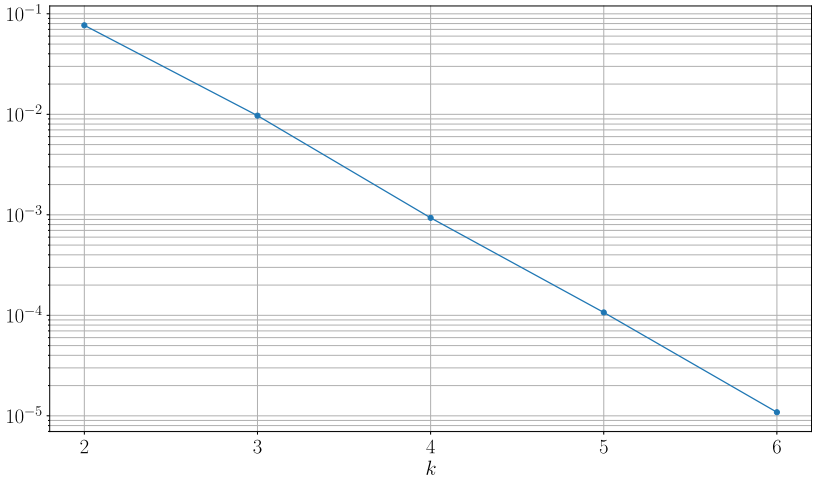


FIGURE 5. Relative H^1-error $\|u^N_{\text{lod}} - u^N_{\text{lod},k}\|_{H^1} / \|u^N_{\text{lod}}\|_{H^1}$ between the non-localized and localized method, plotted against the layer number k.

be seen in Figure 5. Here we can see how the error decays exponentially as k increases, verifying the theoretical findings regarding the localization procedure.

For the second example, the performance of the GFEM in (3.7) and (3.8) depending on the coarse mesh width H is shown. For this example, the fine mesh width is set to $h = 2^{-8}$, and for each coarse mesh width the localization parameter is set to $k = \log_2(1/H)$. Moreover, the time step is set to $\tau = 0.02$ (for the GFEM as well as the reference solution) and the solution is evaluated at $T = 1$. To compute the error, we use a FEM solution on the fine mesh as a reference solution. The error as a function of $1/H$ can be seen in Figure 6. Here it is seen how the error for the novel method decays faster than linearly, confirming the error estimates derived in Section 4. For comparison, Figure 6 also shows the error of the standard FEM solution, as well as the solution using the standard LOD method with correction solely on A and B respectively, *i.e.* corrections based on the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ respectively and without finescale correctors. As expected, the error of these methods stay at a constant level through all coarse grid sizes.

At last, we compute the solution where the system (3.7) is computed using the reduced basis approach. For this example, we let the number of pre-computed solutions M vary, and see how the error between the solutions $u^n_{\text{lod},k}$ and $u^{n,\text{rb}}_{\text{lod},k}$ behaves. In the example we have the fine mesh $h = 2^{-8}$, the coarse mesh $H = 2^{-5}$, the time step $\tau = 0.02$, and the final time $T = 1$. The result can be seen in Figure 7. Here it is seen how the error decreases rapidly with the amount of pre-computed solutions. Note that it is sufficient to compute approximately 10 solutions to yield an error smaller than the discretization error for the main method in Figure 6. This for the case when the number of time steps are $N = 50$. We emphasize that a large increment in time steps does not impact the number of pre-computed solutions M significantly, making the RB-approach relatively more efficient the more time steps that are considered.

6.2. Marmousi model

We finish by demonstrating the novel method (with the reduced basis approach) on a more practical example. A commonly used model problem in seismology is the Marmousi model, which we use to construct our coefficients $A(x,y)$ and $B(x,y)$. This is done by applying midpoint quadrature on the image seen in Figure 8a so that the dimensions work with the fine mesh, which is set to $h = 2^{-8}$. The scales on the coefficients are set to $\alpha_- = 2$, $\alpha_+ = 5$, $\beta_- = 1$ and $\beta_+ = 10$. As earlier, the spatial domain is set to the unit square and the temporal domain

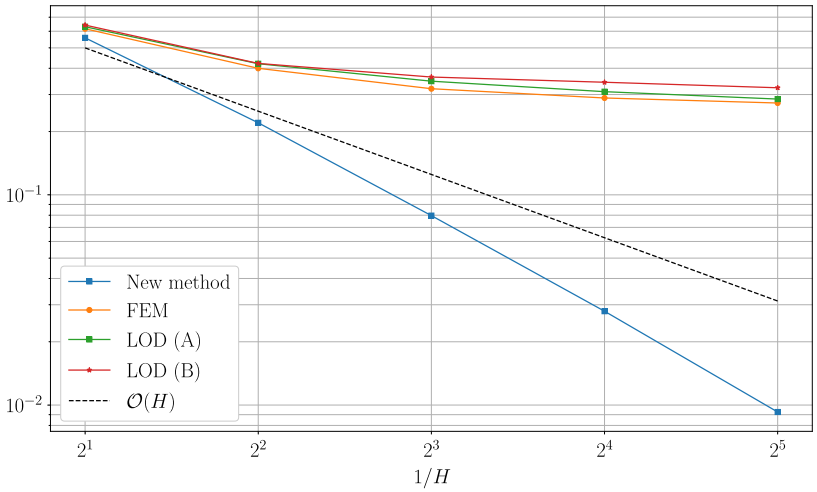


FIGURE 6. Relative H^1-error $\|u^N_{\text{ref}} - u^N_{\text{lod},k}\|_{H^1}/\|u^N_{\text{ref}}\|_{H^1}$ between the reference solution and the approximate solution computed with the proposed method (without reduced basis computations).

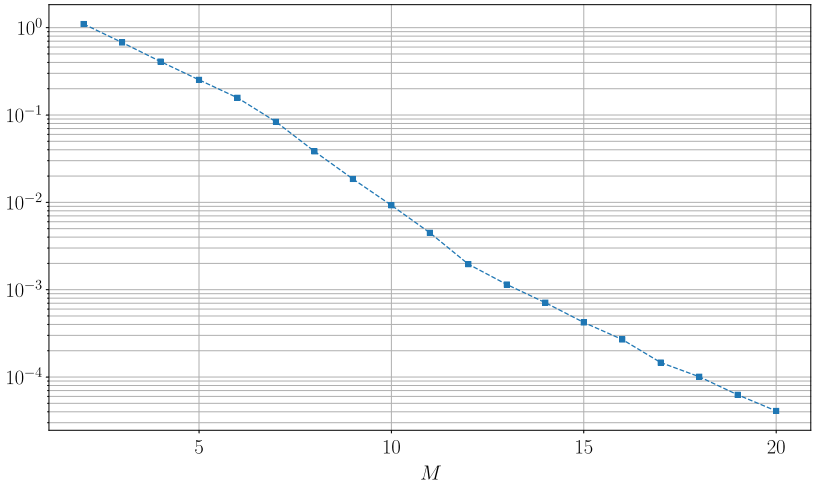
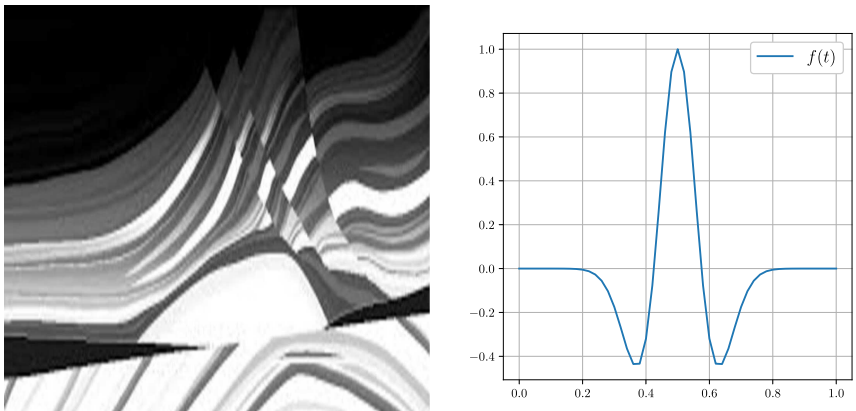


FIGURE 7. Relative H^1-error $\|u^N_{\text{lod},k} - u^{N,\text{rb}}_{\text{lod},k}\|_{H^1}/\|u^N_{\text{lod},k}\|_{H^1}$ between the solution with and without the reduced basis approach, plotted against the number of pre-computed solutions.

to $[0,1]$, discretized with time step $\tau = 0.02$. As source function we use the Ricker wavelet defined as

$$f(x,y,t) = \chi_P(x,y)\left(1 - 2\pi^2\nu^2(t-t')^2\right)e^{-\pi^2\nu^2(t-t')^2}$$

where ν denotes the frequency, t' is the center of the wavelet and $\chi_P(x,y)$ is the indicator function equal to 1 on $P \subset \Omega$ and 0 elsewhere. For our example we let $\nu = 3$, $t' = 0.5$ and $P = [0.5-2h, 0.5+2h] \times [0.5-2h, 0.5+2h]$. The temporal behaviour of the wavelet can be seen in Figure 8b. The solution is computed for coarse mesh sizes $H = 2^{-1}, 2^{-2}, \ldots, 2^{-6}$, with localization parameter $k = \log_2(1/H)$, and for the construction of each reduced basis space $V^{\text{RB}}_{M,k,x}$ we pre-compute 15 vectors, *i.e.* $M = 15$. The error is computed with respect to a reference



(A) (B)

FIGURE 8. *Left*: Marmousi data used to create the coefficients $A(x, y)$ and $B(x, y)$. *Right*: how the Ricker wavelet $f(t)$ varies over time. (A) Marmousi data. (B) Ricker wavelet $f(t)$.

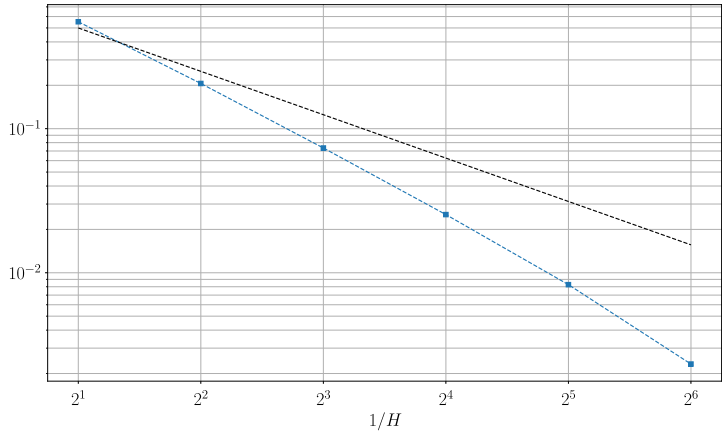


FIGURE 9. Relative H^1-error $\|u^N_{\text{ref}} - u^{N,\text{rb}}_{\text{lod},k}\|_{H^1}/\|u^N_{\text{ref}}\|_{H^1}$ between the full method with reduced basis approach and the reference solution for the practical example using Marmousi data.

solution evaluated on the fine grid, and is displayed in Figure 9. It seen here how the full method converges faster than linearly.

REFERENCES

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

[2] G. Avalos and I. Lasiecka, Optimal blowup rates for the minimal energy null control of the strongly damped abstract wave equation. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2** (2003) 601–616.

[3] J. Azevedo, C. Cuevas and H. Soto, Qualitative theory for strongly damped wave equations. *Math. Methods Appl. Sci.* **40** (2017) 08.

[4] I. Babuska and R. Lipton, Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. *SIAM J. Multiscale Model. Simul.* **9** (2010) 373–406.

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

[6] E. Bonetti, E. Rocca, R. Scala and G. Schimperna, On the strongly damped wave equation with constraint. *Commun. Part. Differ. Equ.* **42** (2017) 1042–1064.

[7] A. Carvalho and J. Cholewa, Local well posedness for strongly damped wave equations with critical nonlinearities. *Bull. Aust. Math. Soc.* **66** (2002) 443–463.

[8] C. Cuevas, C. Lizama and H. Soto, Asymptotic periodicity for strongly damped wave equations. In: Vol. 2013 of *Abstract and Applied Analysis*. Hindawi (2013).

[9] N. Dal Santo, S. Deparis, A. Manzoni and A. Quarteroni, Multi space reduced basis preconditioners for large-scale parametrized PDEs. *SIAM J. Sci. Comput.* **40** (2018) A954–A983.

[10] M. Drohmann, B. Haasdonk and M. Ohlberger, Adaptive reduced basis methods for nonlinear convection-diffusion equations. In: Vol. 4 of *Finite Volumes for Complex Applications VI. Problems & Perspectives. Volume 1, 2*, *Springer Proc. Math.* Springer, Heidelberg (2011) 369–377.

[11] C. Engwer, P. Henning, A. Målqvist and D. Peterseim, Efficient implementation of the localized orthogonal decomposition method. *Comput. Methods Appl. Mech. Eng.* **350** (2019) 123–153.

[12] F. Gazzola and M. Squassina, Global solutions and finite time blow up for damped semilinear wave equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **23** (2006) 185–207.

[13] P.J. Graber and J.L. Shomberg, Attractors for strongly damped wave equations with nonlinear hyperbolic dynamic boundary conditions. *Nonlinearity* **29** (2016) 1171.

[14] B. Haasdonk, M. Ohlberger and G. Rozza, A reduced basis method for evolution schemes with parameter-dependent explicit operators. *Electron. Trans. Numer. Anal.* **32** (2008) 145–161.

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

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

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

[18] R. Ikehata, Decay estimates of solutions for the wave equations with strong damping terms in unbounded domains. *Math. Methods Appl. Sci.* **24** (2001) 659–670.

[19] V. Kalantarov and S. Zelik, A note on a strongly damped wave equation with fast growing nonlinearities. *J. Math. Phys.* **56** (2015) 011501.

[20] P. Kelly, Solid Mechanics Part I: An Introduction to Solid Mechanics. University of Auckland (2019).

[21] A. Khanmamedov, Strongly damped wave equation with exponential nonlinearities. *J. Math. Anal. Appl.* **419** (2014) 663–687.

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

[23] S. Larsson, V. Thomée and L.B. Wahlbin, Finite-element methods for a strongly damped wave equation. *IMA J. Numer. Anal.* **11** (1991) 115–142.

[24] Y. Lin, V. Thomée and L. Wahlbin, *Ritz-Volterra projections to finite element spaces and applications to integro-differential and related equations*. Technical report (Cornell University. Mathematical Sciences Institute). Mathematical Sciences Institute, Cornell University (1989).

[25] A. Målqvist and A. Persson, Multiscale techniques for parabolic equations. *Numer. Math.* **138** (2018) 191–217.

[26] A. Målqvist and D. Peterseim, Generalized finite element methods for quadratic eigenvalue problems. *ESAIM: M2AN* **51** (2017) 147–163.

[27] A. Målqvist and D. Peterseim, Numerical Homogenization beyond Periodicity and Scale Separation. To appear in SIAM Spotlight (2020).

[28] A. Målqvist and D. Peterseim, Computation of eigenvalues by numerical upscaling. *Numer. Math.* **130** (2012) 337–361.

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

[30] P. Massatt, Limiting behavior of strongly damped nonlinear wave equations. *J. Differ. Equ.* **48** (1983) 334–349.

[31] D. Peterseim, Variational multiscale stabilization and the exponential decay of fine-scale correctors. In: Vol. 114 of *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations. Lect. Notes Comput. Sci. Eng.* Springer, [Cham] (2016) 341–367.

[32] A. Quarteroni, A. Manzoni and F. Negri, Reduced Basis Methods for Partial Differential Equations: An Introduction. Springer (2016).

[33] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, 2nd edition. *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin (2006).

[34] G.F. Webb, Existence and asymptotic behavior for a strongly damped nonlinear wave equation. *Can. J. Math.* **32** (1980) 631–643.