A CONTINUOUS WAVELET-GALERKIN METHOD
FOR THE LINEAR WAVE EQUATION

HOANG NGUYEN† AND GIANPAOLO EVANGELISTA†

Abstract. We consider the continuous space-time Galerkin method for the linear second-order
wave equation proposed by French and Peterson in 1996. A bottleneck for this approach is how to
solve the discrete problems effectively. In this paper, we tackle this bottleneck by essentially employ-
ing wavelet bases in space. We show how to decouple the corresponding linear system and we prove
that the resulting subsystems can be uniformly preconditioned by simple diagonal preconditioners,
leading to efficient iterative solutions.

Key words. wave equation, continuous Galerkin, finite element, wavelet, preconditioning

AMS subject classifications. 65T60, 65M60, 65F35

1. Introduction. The continuous Galerkin method, or continuous space-time
Galerkin method, is a variational technique dealing with evolution problems, using
approximation spaces of continuous functions in both space and time. An advantage
of this approach is that methods of any order of accuracy are relatively easy to for-
mulate. For wave propagation problems, this approach is particularly appropriate,
since the discrete problems inherit energy conservation properties from the continu-
ous problems. This is especially useful when dealing with nonlinear wave problems,
see [11] and the references therein.

A bottleneck for the practical computation of variational techniques in space-time,
in particular the continuous Galerkin method for solving the wave equation, is that,
except when the accuracy in time is of the lowest order, the resulting linear system
is, in general, a coupled system, and it is not obvious how to solve it efficiently.
This is not a problem when one invokes the method of lines with spatial Galerkin
discretization, leading e.g. to the celebrated Runge-Kutta Discontinuous Galerkin
schemes, see [4] and the references therein. One drawback of such approaches is that
the Courant-Friedrichs-Levy (CFL) condition has to be taken into account, cf. [14].
Recently, (semi) explicit discontinuous Galerkin methods have been proposed in [9]
and [14]. In those latter approaches, by generating appropriate space-time meshes in
a time-stepping manner, one obtains merely small-sized uncoupled linear subsystems,
while the CFL condition is effectively avoided. However, the mesh generation would
pose some difficulties when the spatial dimension becomes, e.g., larger than two.

In this paper, we reconsider the continuous space-time Galerkin method pro-
posed in [11], tackling the bottleneck mentioned above. In particular, we show how
the linear system can be decoupled and how the resulting subsystems can be uni-
formly preconditioned, meaning that the condition numbers of the preconditioned
subsystems are bounded independently on both spatial and temporal stepsizes.
The decoupling merely relies on decompositions of temporal system matrices, while the
uniform preconditioning mainly relies on the compressibility of operators in wavelet
coordinates and the Riesz properties of wavelet bases (the latter means that wavelets
generate Riesz bases for a scale of Sobolev spaces). Moreover, all of the resulting
preconditioned subsystems are symmetric positive definite (SPD), leading to efficient
iterative solutions in the sense that the computational complexity is of linear order with respect to the number of degrees of freedom. Hence, our method can be viewed as a semi explicit method as well.

Although non-wavelet bases, in combination with preconditioning techniques such as those discussed in [1] and the references therein, might also offer uniformly well conditioned (sub)systems, we have opted for wavelet bases because of the following. Firstly, wavelet bases offer diagonal preconditioners in many applications, thanks to their Riesz properties (see e.g. [6, 5], and [15] for a related problem). It turns out that we are able to obtain diagonal preconditioners for the wave problem also, and our numerical results indicate satisfactory performances. Secondly, sparse wavelet approximations to parabolic problems have recently been proposed in [19], dramatically reducing the number of degrees of freedom while retaining the order of accuracy. Finally, wavelet bases have successfully been used in adaptive approximations to operator equations, such as partial differential equations or (boundary) integral equations, see e.g. [6, 5]. Hence, this paper can also be viewed as an intermediate step toward other wavelet applications to the wave equation, which will be objects of future work. It should be noted that although we focus on nodal bases for the temporal discretization, our approach would apply to other temporal bases as well. Moreover, our approach would also work well with other approaches such as the discontinuous Galerkin methods for second-order hyperbolic problems proposed in [13] or other Galerkin methods for parabolic problems.

The outline of this paper is as follows: In the rest of this section, we specify our notation. In Section 2, we formulate the wave propagation problem and we recall the continuous Galerkin approximation to its solution as well as the discretization error estimates given in [11]. In addition, we give a matrix vector representation of the resulting coupled linear system. In Section 3, we discuss the concept of multiresolution approximation and biorthogonal wavelets. In addition, we prove some spectral properties of the representation of operators in wavelet coordinates, In Section 4, we show how the coupled linear system from Section 2 can be decoupled and preconditioned, employing the results of Section 3. Our numerical results are presented in Section 5.

We begin with some basic notation and definitions that will be used throughout this paper. First of all, the time derivative will be denoted by the dot notation. To make the notation not unnecessarily complicated, we will drop references to the underlying domain or index sets whenever there is no risk of confusion, i.e., we will write $L^2$ for $L^2(\Omega)$ etc.. Unless otherwise stated, $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ will denote some canonical inner product and norm, e.g., the Euclidean inner product, the spectral norm, the $L^2$-norm, etc.. In order to avoid repeated use of generic but unspecified constants, by $C \lesssim D$ we mean that $C$ can be bounded by a multiple of $D$ independently of parameters on which $C$ and $D$ may depend. In cases of interest, we will state the (in)dependence explicitly. Obviously, $C \gtrsim D$ means $D \lesssim C$, and $C \sim D$ means $C \lesssim D$ and $C \gtrsim D$.

We will adopt the following compact notation from the literature (cf. [6]). For $\Sigma$ being a countable collection of functions in some separable Hilbert space $H$, equipped with some inner product $\langle \cdot, \cdot \rangle_H$ and norm $\| \cdot \|_H$, we will formally identify $\Sigma$ with a column vector (of functions in $H$): for $c = (c_\sigma)_{\sigma \in \Sigma}$ being a column vector of scalars, $c^T \Sigma$ will denote the formal series $\sum_{\sigma \in \Sigma} c_\sigma \sigma$; and likewise for $C = (c_{\varsigma, \sigma})_{\varsigma, \sigma \in \Sigma}$ being a matrix, $C \Sigma$ will denote the column vector $(\sum_{\sigma \in \Sigma} c_{\varsigma, \sigma} \sigma)_{\varsigma \in \Sigma}$. Such a collection $\Sigma$ is called a Riesz system (in $H$) if

$$\|c^T \Sigma\|_H \sim \|c\| \quad (c \in l_2(\Sigma)),$$
where $l_2(\Sigma) := \{ \mathbf{c} = (c_\sigma)_{\sigma \in \Sigma} : \| \mathbf{c} \| = \| \mathbf{c} \|_{l_2(\Sigma)} = (\sum_{\sigma \in \Sigma} c_\sigma^2)^{\frac{1}{2}} < \infty \}$. In addition, if such a system is a basis for $H$ then it is called a Riesz basis. Likewise, for $\Sigma$ and $\hat{\Sigma}$ being two countable collections of functions in $H$, $\langle \Sigma, \hat{\Sigma} \rangle_H$ will denote the matrix $(\langle \sigma, \hat{\sigma} \rangle_H)_{\sigma \in \Sigma, \hat{\sigma} \in \hat{\Sigma}}$, and so, for $A$ and $\hat{A}$ being two matrices of appropriate dimensions, 

$$(A \Sigma, \hat{A} \hat{\Sigma})_H = A \langle \Sigma, \hat{\Sigma} \rangle_H \hat{A}^T.$$ 

Finally, for $s \geq 0$, $H^s$ and $H^s_0$ will denote the usual Sobolev spaces on a $d$-dimensional domain $\Omega$, whereas for $s < 0$, $H^s$ and $H^s_0$ will denote the dual of $H^{-s}$ and $H^{-s}_0$ respectively, i.e. $H^s = (H^{-s})'$ and $H^s_0 = (H^{-s}_0)'$, with $(H^0)' = H^0 = H^0_0 = (H^0_0)' = L^2$ being the pivot space. For $H$ being any of the Sobolev spaces above, $\langle \cdot, \cdot \rangle_H$ and $\| \cdot \|_H$ will denote the usual inner product and the usual norm on $H$ respectively.

2. Continuous Galerkin approximation and error estimates. In this section, we slightly simplify the formulation and error estimates in [11] so that they better fit our setting. The coupled linear system resulting from the continuous Galerkin approach will also be given.

2.1. The linear wave equation. Let $\mathcal{R} \ni T > 0$ and let $\Omega$ be an open, bounded $d$-dimensional domain with sufficiently smooth boundary $\partial \Omega$. We consider the following initial boundary value problem: Given $f = f(t,x)$, $U_0 = U_0(x)$ and $V_0 = V_0(x)$, find $U = U(t,x)$ such that:

$$\begin{align*}
\ddot{U} - \Delta U &= f & \text{in } (0,T) \times \Omega, \\
U &= 0 & \text{on } (0,T) \times \partial \Omega, \\
U(0,\cdot) &= U_0 & \text{in } \Omega, \\
\dot{U}(0,\cdot) &= V_0 & \text{in } \Omega,
\end{align*}$$

(2.1)

with $\Delta$ being the Laplacian operator. Note that, with

$$V = \dot{U}, \ Y = \begin{pmatrix} U \\ V \end{pmatrix}, \ Y_0 = \begin{pmatrix} U_0 \\ V_0 \end{pmatrix}, \ F = \begin{pmatrix} 0 \\ f \end{pmatrix}, \ \text{and } A = \begin{bmatrix} 0 & -I \\ -\Delta & 0 \end{bmatrix},$$

(2.1) is equivalent to

$$\begin{align*}
\dot{Y} + AY &= F & \text{in } (0,T) \times \Omega, \\
Y &= 0 & \text{on } (0,T) \times \partial \Omega, \\
Y(0,\cdot) &= Y_0 & \text{in } \Omega.
\end{align*}$$

(2.2)

Under standard regularity assumptions on the given data, results on the existence, uniqueness and regularity of the solution of (2.1) are known. Further, with $A$ being considered as an operator in $H^1_0 \times L^2$ with domain $\text{dom}(A) = (H^2 \cap H^1_0) \times H^1_0$, $A$ generates a strongly continuous semigroup, see e.g. [17]. The latter is essential for the derivation of the results in the next section. In this paper, we tacitly assume that the given data are sufficiently regular so that the corresponding results or formulations are meaningful, cf. [11]. As mentioned in [11], although we confine ourselves to (2.1), the results therein can be easily generalized to the case where $-\Delta$ is replaced by any time-independent, selfadjoint, second-order uniformly elliptic operator. Our results can also be generalized to that case.

2.2. Continuous Galerkin approximation. Continuous Galerkin approximations to the solution of (2.1) are defined in the tensor product of the temporal and spatial approximation spaces as follows. For the spatial discretization, with $j$ and $p$ being non-negative discretization parameters, the approximation space

$$V_j = V_j(p) = \{ v_j = v_j(x) : x \in \Omega \} \subset H^1_0$$
is assumed to satisfy the following direct estimates, or Jackson estimates:
\[
\inf_{v_j \in V_j} \| w - v_j \|_{L^2} \lesssim 2^{-js} \| w \|_{H^s},
\]
(2.3)
for \( w \in H^s \cap H^0 \) and \( 0 \leq s \leq p + 1 \). In addition, with \( P_z : H^0 \to V_j \) being the Galerkin projection from \( H^0 \) onto \( V_j \) defined by
\[
\langle \nabla (w - P_z w), \nabla v_j \rangle_{L^2} = 0 \quad \forall v_j \in V_j,
\]
(2.4)
we assume that
\[
\| w - P_z w \|_{H^m} \lesssim 2^{-j(s-m)} \| w \|_{H^s},
\]
(2.5)
for \( w \in H^s \cap H^0 \), \( m \in \{0, 1\} \) and \( 0 \leq s \leq p + 1 \). For the temporal discretization, with \( P_q(I) \) being the space of polynomials of degree at most \( q \) on some interval \( I \), let
\[
V_h = V_h(q) := \{ v_h = v_h(t) : t \in [0, T], v_h \in C([0, T]), (v_h)_{|I_n} \in P_q(I_n) \}
\]
where \( 0 = t_0 < t_1 < \cdots < t_N = T \) is a partition of \( (0, T) \) into intervals \( I_n := (t_{n-1}, t_n) \) of length \( h_n := t_n - t_{n-1} \), and
\[
h := \max_{n} h_n. \tag{2.6}
\]
Throughout this paper, we will assume that \( p, q \geq 1 \).

Based on the formulation (2.2), with
\[
S_n := \Omega \times I_n, V_{hj} := V_h \otimes V_j, u_0 := P_z(U_0) \text{ and } v_0 := P_z(V_0),
\]
the approximations \( u \) and \( v \) to \( U \) and \( V \) respectively are the solution of the following variational problem: Find \( u, v \in V_{hj} \) such that \( u(0, \cdot) = u_0, v(0, \cdot) = v_0, \) and
\[
\begin{align*}
\langle \dot{u} - v, \sigma \rangle_{L^2(S_n)} &= 0 \quad \forall \sigma \in P_{q-1}(I_n) \otimes V_j, \\
\langle \dot{v}, \varsigma \rangle_{L^2(S_n)} + \langle \nabla u, \nabla \varsigma \rangle_{L^2(S_n)} &= \langle f, \varsigma \rangle_{L^2(S_n)} \quad \forall \varsigma \in P_{q-1}(I_n) \otimes V_j
\end{align*}
\]
(2.7)
for \( n = 1, \ldots, N \). As we will see in Section 4 how \( u_0 \) and \( v_0 \) can be efficiently approximated, we simply assume here and in the following that \( u_0 \) and \( v_0 \) are given. Note that the test functions are of one degree lower in time, namely \((q - 1)\), which takes the continuity of \( u \) and \( v \) at the temporal nodes into account, allowing for \( u \) and \( v \) to be computed successively on each space-time slab \( S_n \) as follows. Let \( u_n := u_{I_n} \) and \( v_n := v_{|I_n} \), then (2.7) is equivalent to the following variational problem: For \( u_0(t, \cdot) := u_0, v_0(t, \cdot) := v_0 \forall t, \) find \( u_n, v_n \in P_q(I_n) \otimes V_j \) for \( n = 1, \ldots, N \) such that:
\[
\begin{align*}
\langle \dot{u}_n - v_n, \sigma \rangle_{L^2(S_n)} &= 0 \quad \forall \sigma \in P_{q-1}(I_n) \otimes V_j, \\
\langle \dot{v}_n, \varsigma \rangle_{L^2(S_n)} + \langle \nabla u_n, \nabla \varsigma \rangle_{L^2(S_n)} &= \langle f, \varsigma \rangle_{L^2(S_n)} \quad \forall \varsigma \in P_{q-1}(I_n) \otimes V_j, \\
& \quad u_n(t_{n-1}, \cdot) = u_{n-1}(t_{n-1}, \cdot), \\
& \quad v_n(t_{n-1}, \cdot) = v_{n-1}(t_{n-1}, \cdot).
\end{align*}
\]
(2.8)

The existence and uniqueness of the solution \( u \) and \( v \) of (2.7) are proven in [11]. In addition, \( u \) and \( v \) satisfy the following (simplified) error estimates:

**Theorem 2.1** ([11]). Let \( U \) and \( V \) be the solution of (2.1), and let \( u \) and \( v \) be their approximations defined in (2.7). For \( t \in [0, T] \), there holds
\[
\begin{align*}
(i) \quad & \| u(t) - U(t) \|_{L^2} \leq h^{q+1} + 2^{-j(p+1)}, \\
(ii) \quad & \| u(t) - U(t) \|_{H^0} \leq h^{q+1} + 2^{-jp}.
\end{align*}
\]
In addition, for $1 \leq n \leq N$, there holds

(iii) $\|u(t_n) - U(t_n)\|_{L_2}^2, \|v(t_n) - V(t_n)\|_{L_2}^2 \lesssim h^{2q + 2^{-j(p+1)}}$,

(iv) $\|u(t_n) - U(t_n)\|_{H^1_0}, \|v(t_n) - V(t_n)\|_{H^1_0} \lesssim h^{2q + 2^{-jp}}$.

Remark 2.2. The statements of Theorem 2.1 are meaningful only when the given data $f, U_0$ and $V_0$, and so the exact solution $U$ and $V$, are sufficiently regular. Further, the unspecified constants therein depend on the given data, the exact solution and the final time $T$ (see [11] for details). In addition, in order to obtain the predicted rates of convergence in an optimal manner, $h_n$ and $j$ should relate to each other in one or other way. For example, the optimal $L^2$ rate of convergence is obtained with either $h_n \approx 2^{-j}$ or $h_n \approx 2^{-2^{-j}}$, as followed from parts (i) and (iii). $\Box$

Remark 2.3. Suppose that $V_j$ is a $C^q$ Lagrange finite element space incorporating continuous piecewise polynomials of degree $p$ with respect to some mesh consisting of regular elements with diameters of order $2^{-j}$. Then it is well-known that, by interpolation between Sobolev spaces, Jackson estimates (2.3) are satisfied. Further, by the regularity of the elliptic problem (2.4), the estimates (2.5) are also satisfied (see e.g. [3, 2]). Further, not only estimates of type (2.3) and (2.5), but also several inverse estimates are satisfied by $P_q(I_n)$, and so by $V_n$, as well (see [11, 10] for details). All of these estimates are essential for the derivation of Theorem 2.1. $\Box$

2.3. Matrix vector representation. Recall that $\langle \cdot, \cdot \rangle$ will denote some canonical inner product (in this section, either $\langle \cdot, \cdot \rangle_{L^2(I_n)}$, or $\langle \cdot, \cdot \rangle_{L^2(\Omega)}$, or $\langle \cdot, \cdot \rangle_{H^1(I_n)}$). For $\hat{q} \geq 0$, let $\Xi_n^\hat{q} = (\xi_{n,0}, \cdots, \xi_{n,\hat{q}})^T$ be a basis for $P_{\hat{q}}(I_n)$, then we write $P_{\hat{q}}(I_n) \otimes V_j \ni w_n = \sum_{i=0}^{\hat{q}} \xi_{n,i} w_{n,i}$ with $w_{n,i} \in V_j$. In addition, let $\Phi_j$ be a basis for $V_j$, then we write $V_j \ni w_{n,i} = w_n^T \Phi_j$, and so, with $w_n := (w_{n,0}, \cdots, w_{n,\hat{q}})^T$,

$$
w_n = w_n^T(\Xi_n^\hat{q} \otimes \Phi_j)
$$

is the representation of $w_n$ with respect to the basis $\Xi_n^\hat{q} \otimes \Phi_j$ for $P_{\hat{q}}(I_n) \otimes V_j$, where $\Xi_n^\hat{q} \otimes \Phi_j = (\xi_{n,0}(\Phi_j)^T, \cdots, \xi_{n,\hat{q}}(\Phi_j)^T)$. Note that, with $I_j$ being the $|\Phi_j| \times |\Phi_j|$ identity matrix where $|\Phi_j|$ is the cardinal of $\Phi_j$ (equal to the dimension of $V_j$), it holds that

$$
w_n^T(\Xi_n^\hat{q} \otimes \Phi_j) = w_n^T[\Xi_n^\hat{q} \otimes I_j]\Phi_j.
$$

(2.9)

In this paper, we write $u_n = u_n^T(\Xi_n^\hat{q} \otimes \Phi_j)$ and $v_n = v_n^T(\Xi_n^\hat{q} \otimes \Phi_j)$. Employing (2.9), (2.8) is represented as

$$
\langle u_n^T(\Xi_n^\hat{q} \otimes \Phi_j) - v_n^T(\Xi_n^\hat{q} \otimes \Phi_j), \Xi_n^{\hat{q}-1} \otimes \Phi_j \rangle = 0,
$$

$$
(\langle v_n^T(\Xi_n^\hat{q} \otimes \Phi_j), \Xi_n^{\hat{q}-1} \otimes \Phi_j \rangle + \langle u_n^T(\Xi_n^\hat{q} \otimes \Phi_j), \Xi_n^{\hat{q}-1} \otimes \nabla \Phi_j \rangle) = f_n^T,
$$

$$
u_n^T[\Xi_n^{\hat{q}}(\nabla (t_{n-1} \otimes I_j)]\Phi_j = u_{n-1}^T(\Phi_j),
$$

$$
v_n^T[\Xi_n^{\hat{q}}(\nabla (t_{n-1} \otimes I_j)]\Phi_j = v_{n-1}^T(\Phi_j),
$$

with the right hand side

$$
f_n^T := (f, \Xi_n^{\hat{q}-1} \otimes \Phi_j),
$$

and

$$
u_0^T \Phi_j := u_0 \text{ and } v_0^T \Phi_j := v_0,
$$

whereas for $n = 2, \cdots, N$

$$
u_n^T := u_{n-1}^T[\Xi_n^{\hat{q}}(t_{n-1} \otimes I_j)] \text{ and } v_n^T := v_{n-1}^T[\Xi_n^{\hat{q}}(t_{n-1} \otimes I_j)].\]
or equivalently
\[
\begin{align*}
    & [(\Xi_n^{-1}, \mathcal{E}_n^f) \otimes (\Phi_j, \Phi_j)] u_n - [(\Xi_n^{-1}, \mathcal{E}_n^f) \otimes (\Phi_j, \Phi_j)] v_n = 0, \\
    & [(\Xi_n^{-1}, \mathcal{E}_n^f) \otimes (\nabla \Phi_j, \nabla \Phi_j)] u_n + [(\Xi_n^{-1}, \mathcal{E}_n^f) \otimes (\Phi_j, \Phi_j)] v_n = f_n, \\
    & [(\Xi_n(t_{n-1}))^T \otimes I] u_n = u_{[n]}, \\
    & [(\Xi_n(t_{n-1}))^T \otimes I] v_n = v_{[n]},
\end{align*}
\]

(2.10)

Note that the dimension of this coupled linear system is \(2(q + 1)|\Phi_j|\). Further, since (2.7) has a unique solution, (2.10) also has a unique solution, independently on the choice of bases.

3. Wavelet approximation and representation. In this section, we briefly discuss the concept of multiresolution approximation, that will be our spatial approximation, and biorthogonal wavelets. In addition, based on the compressibility of operators in wavelet coordinates and the Riesz properties of wavelet bases, we prove some spectral properties of the representation of operators in wavelet coordinates.

3.1. Multiresolution approximation and wavelet bases. In the following, let \(H_s := H_s^2\) for \(0 \leq s \leq 1\) and \(H_s := H^s \cap H_0^1 \subset H^s\) for \(s > 1\), whereas \(H_s := (H_{-s})'\) for \(s < 0\). We recall the following result:

**Theorem 3.1** ([8]). Consider the following two multiresolution analyses
\[
\begin{align*}
    V_0 & \subset V_1 \subset \cdots \subset L^2, \text{ with } \text{clos}_{L^2}(\cup_{j \geq 0} V_j) = L^2 \\
    \tilde{V}_0 & \subset \tilde{V}_1 \subset \cdots \subset L^2, \text{ with } \text{clos}_{L^2}(\cup_{j \geq 0} \tilde{V}_j) = L^2.
\end{align*}
\]

Suppose that

1. \(\exists \) uniformly bounded biorthogonal projectors \(Q_j : L^2 \to L^2\) such that
   \[\text{Im} Q_j = V_j, \text{ Im}(I - Q_j) = \tilde{V}_j^{1\perp}.\]

2. both sequences satisfy Jackson estimates with parameters \(p + 1 > 0, \tilde{p} + 1 > 0\) uniformly in \(j\), i.e.,
   \[
   \inf_{v_j \in V_j} \|v - v_j\|_{L^2} \lesssim 2^{-j(p+1)} \|v\|_{H_{p+1}} , \quad (v \in H_{p+1}),
   \]
   \[
   \inf_{\tilde{v}_j \in \tilde{V}_j} \|\tilde{v} - \tilde{v}_j\|_{L^2} \lesssim 2^{-j(\tilde{p}+1)} \|\tilde{v}\|_{H_{\tilde{p}+1}} , \quad (\tilde{v} \in H_{\tilde{p}+1}).
   \]

3. both sequences satisfy Bernstein estimates with parameters \(0 < \gamma < p + 1, 0 < \tilde{\gamma} < \tilde{p} + 1\) uniformly in \(j\), i.e., for \(s \in [0, \gamma)\) and \(\tilde{s} \in [0, \tilde{\gamma})\), it holds that
   \[\|v_j\|_{H_s} \lesssim 2^{js} \|v_j\|_{L^2} \quad (v_j \in V_j), \]
   \[\|\tilde{v}_j\|_{H_{\tilde{s}}} \lesssim 2^{j\tilde{s}} \|\tilde{v}_j\|_{L^2} \quad (\tilde{v}_j \in \tilde{V}_j).\]

Then, for \(s \in (-\tilde{p} + 1, \gamma)\) and \(\tilde{s} \in (-\tilde{\gamma}, p + 1)\), with \(Q_{-1} := 0\),

**R1**
\[
\begin{align*}
    \|\sum_{j=0}^{\infty} w_j\|^2_{H_s} & \lesssim \sum_{j=0}^{\infty} 4^{js} \|w_j\|^2_{L^2} \quad (w_j \in \text{Im}(Q_j - Q_{j-1})) \\
    \sum_{j=0}^{\infty} 4^{j\tilde{s}} \|Q_j - Q_{j-1}u\|^2_{L^2} & \lesssim \|u\|^2_{H_{\tilde{s}}} \quad (u \in H_{\tilde{s}}).
\end{align*}
\]
For $s \in (-\tilde{\gamma}, \gamma)$, the mappings $(w_j)_j \mapsto \sum_{j=-\infty}^{\infty} w_j$ and $u \mapsto ((Q_j - Q_{j-1})u)_j$, which are bounded in the sense of (R1), are each others inverse. Thus, for $s \in (-\tilde{\gamma}, \gamma)$,

\[ \sum_{j=-\infty}^{\infty} 4^{js} ||(Q_j - Q_{j-1})u||^2_{L^2} (u \in H_s). \]

Analogous results (R1*) and (R2*) are valid at the dual side, i.e., with interchanged roles of $(Q_j, d, \gamma)$ and $(Q^*_j, \tilde{d}, \tilde{\gamma})$.

Now, for $j > 0$ and $I_j$ being some appropriate index set, let $\Psi_j := (\psi_{j,z})_{z \in I_j}$, whose elements are called (biorthogonal) wavelets, be uniform $L^2$-Riesz bases for the detail spaces $W_j := V_j \cap \tilde{V}_{j-1}^\perp = \text{Im}(Q_j - Q_{j-1})$.

Further, let $\Psi_0$ be some arbitrary basis for $W_0 := V_0$. As a direct consequence of Theorem 3.1, we obtain the following. Firstly, our spatial approximation spaces $V_j$ for $j \geq 0$, that are nested closed subspaces of $L^2 = L^2(\Omega)$ incorporating Dirichlet boundary condition (i.e. $(v_j)|_{\partial \Omega} = 0 \forall v_j \in V_j$), can be decomposed as

\[ V_j = V_{j-1} \oplus W_j, \]

with $V_{-1} := \{0\}$. Repeating this two-level decomposition, we obtain the following multilevel decomposition

\[ V_j = \bigoplus_{l=0}^{j} W_l. \]

Secondly, let the following wavelet collection

\[ \Psi^j := \bigcup_{l=0}^{j} \Psi_l = (\psi^j_0, \cdots, \psi^j_j)^T \]

be the corresponding wavelet basis, or multilevel basis, for $V_j$, then the following norm equivalences hold uniformly (in $j$)

\[ \|c^T \Psi^j\|_{H_s}^2 = \sum_{l=0}^{j} 4^{ls} \|c_l\|^2_{L^2} \]  

(3.1)

for $s \in (-\tilde{\gamma}, \gamma)$ and $c = (c^j_0, \cdots, c^j_j)^T$ with $c_l \in \mathbb{R}^{|I_l|}$ and $l = 0, \cdots, j$. In the rest of this paper, $c$ and $d$ will denote vectors of wavelet coefficients, accordingly to this multilevel setting.

It follows from Bernstein estimates that if $\gamma > 1$ then the assumption $V_j \subset H^1_0 = H^1_1$ made in Subsection 2.2 is satisfied. In the next two subsections, we will briefly discuss two types of approximation spaces and their wavelet bases suitable for our purpose. In particular, $\gamma$ will be larger than 1 and the wavelets will have compact supports. The latter implies that an application of $\langle \Psi_j, \Psi_j \rangle_{L^2}$, or $\langle \nabla \Psi_j, \nabla \Psi_j \rangle_{L^2}$, to a vector of appropriate dimension can be carried out within linear complexity. For a general theory of multiresolution approximation and wavelets, see [5, 6].

Remark 3.2. Since (2.3) follows from (3) by interpolation between Sobolev spaces, both (2.3) and (3) are called Jackson estimates. When the spaces $V_j$ contain all polynomials of degree at most $p$, satisfying Dirichlet boundary condition, and
are spanned by compactly supported functions such that the diameters of their supports are of order $2^{-j}$ uniformly, Jackson estimates with parameter $p+1$ are typically valid (cf. Remark 2.3). Likewise, when $V_j$ are spanned by piecewise smooth functions from $C^m(\Omega)$, Bernstein estimates with $\gamma = m + \frac{d}{2}$ are typically valid ($m=1$ means that they are not globally continuous). Further, the norm equivalences (3.1) are also valid in the limit, i.e., $\|c^T\Psi\|^2_{H_j} \approx \sum_{l=0}^{m} d^l \|c_l\|^2_{L^2}$ with $\Psi := \lim_{j \to \infty} \Psi^j = \bigcup_{l=0}^{m} \Psi_l$, and thus the collections $\bigcup_{j=0}^{\infty} 2^{-j} \Psi_l$ are in fact Riesz bases for $H_s$ for $s \in (-\gamma, \gamma)$.

3.2. Finite element wavelets. In the finite element setting, the spaces $V_j$ are nothing else than $C^0$ Lagrange finite element spaces incorporating Dirichlet boundary condition: With $M_0$ being some fixed $d$-simplex (the reference element), we first fix some conforming dyadic refinement of $M_0$ into $2^d$ congruent subsimplices. Now, let $M_0$ be some (initial) collection of $d$-simplices $M$ such that $\bigcup_{M \in M_0} M$ is a conforming triangulation, or mesh, of $\Omega$. Starting from $M_0$, we obtain a sequence of conforming triangulation $(M_j)_j$ of $\bigcup_{M \in M_j} M$ as the triangulation obtained by applying the above fixed dyadic refinement to all $M \in M_j$ (by means of affine bijections). The spaces $V_j$ are the collections of continuous piecewise polynomials of degree $p$, with respect to the triangulations $M_j$, that vanish at $\partial \Omega$. Employing standard finite element techniques, compactly supported biorthogonal wavelet bases for $V_j$, with $\gamma = 3$ and basically any desired $\tilde{p}+1$, can be effectively constructed, see e.g. [8, 16].

Since the spaces $V_j$ consist of continuous piecewise polynomials of degree $p$ on regular elements with diameters of order $2^{-j}$, the Jackson and Bernstein parameter are $p+1$ and $\gamma = \frac{5}{2}$ respectively, as mentioned in the previous subsection. Although this finite element approach has the advantage of inheriting the widely used and well documented finite element techniques, it is not feasible when $d$ becomes larger than three. In this case, one can employ the wavelets discussed in the next subsection.

3.3. Tensor product based spline-wavelets. First, let us consider the simple case when $\Omega$ is the unit $d$-cube $[0,1]^d$. For $d=1$, the spaces $V_j$ are defined through a basis, generated by dyadic dilations and translations of a single scaling function, that is a B-spline of degree $p$ in our case, with suitable adaptations near the boundary so that Dirichlet boundary condition is satisfied. Compacty supported biorthogonal wavelets, with basically any desired $\tilde{p}+1$, can then be constructed following traditional lines. For $d > 1$, wavelets on the $d$-cube are defined by tensor products of wavelets constructed on the interval $[0,1]$, and so the spaces $V_j$ can be defined. Now, for general domains $\Omega$, one might employ domain decomposition techniques as follows: First, $\Omega$ is decomposed into disjoint patches, each of them being some parametric image of the $d$-cube. Then, roughly speaking, wavelets on $\Omega$ are constructed by lifting wavelets on the $d$-cube to the patches and then "gluing" them over the interfaces.

Since the spaces $V_j$ have polynomials reproduction properties, meaning that polynomials of degree $p$ can be reproduced, the Jackson parameter is also $p+1$, equal to that in the finite element case. But, since the scaling functions are splines, the Bernstein parameter is now $\gamma = p + \frac{5}{2}$, larger than that in the finite element case. For more information about tensor product based spline-wavelets, see e.g. [7].

3.4. Operators in wavelet coordinates. For $r \in (-\gamma, \gamma)$, let $L_r : H_r \to H_{-r}$ be a bounded elliptic operator. Here, boundedness means $|\langle v, L_r u \rangle| \lesssim \|v\|_{H_r} \|u\|_{H_r}$, and ellipticity means $\langle u, L_r u \rangle \gtrsim \|u\|^2_{H_r}$. It follows from the boundedness and ellipticity of $L_r$ that

$$\langle \cdot, L_r \cdot \rangle \approx \|\cdot\|^2_{H_r}.$$
In addition, we assume for simplicity that $\mathcal L_r, L^*_r : H_{\gamma} \rightarrow H_{\gamma - 2r}$ are bounded, where $L^*_r$ is the adjoint of $\mathcal L_r$, and thus $\mathcal L_r, L^*_r : H_{\gamma + r} \rightarrow H_{\gamma - r}$ are bounded for any $\hat \gamma = \gamma, \in [0, \gamma - r]$. For a discussion about sufficient conditions for this continuity property, see e.g. [18] and the references therein. In particular, when $\mathcal L_r$ is a differential operator given by its associated bilinear form

$$\langle v, \mathcal L_r u \rangle = \sum_{|\alpha|, |\beta| \leq m} \langle \partial^\alpha v, a_{\alpha \beta} \partial^\beta u \rangle,$$

with $\alpha$ and $\beta$ being multi-indices, and $\partial^\alpha$ and $\partial^\beta$ being the corresponding partial differential operators, then a sufficient condition for this continuity property is that $a_{\alpha \beta} \in C^{(\hat \gamma - r)}(\Omega)$, with $|\cdot|$ being the ceiling function. By Cauchy-Schwarz inequality and the norm equivalences (3.1), we then obtain for any $\hat \gamma \in [0, \min\{\hat \gamma + r, \gamma - r\})$ and $w_l, w_r \in V_l$ and $w_T \in V_T$

$$|\langle w_l, \mathcal L_r w_l \rangle| \leq \|w_l\|_{H_{\gamma + r}} \|\mathcal L_r w_l\|_{H_{\gamma - r}} \lesssim 2^{-\hat \gamma (t - l)} \|w_l\|_{H_r} \|w_l\|_{H_r},$$

and analogously

$$|\langle w_l, \mathcal L_r w_l \rangle| = |\langle L^*_r w_T, w_l \rangle| \lesssim 2^{-\hat \gamma (t' - l')} \|w_T\|_{H_r} \|w_l\|_{H_r},$$

and so

$$|\langle w_T, \mathcal L_r w_T \rangle| \lesssim 2^{-\hat \gamma (t' - l')} \|w_T\|_{H_r} \|w_T\|_{H_r}. \quad (3.2)$$

In fact, (3.2) is even valid for any $\hat \gamma \in [0, \min\{\hat \gamma + r, \gamma - r\})$ as pointed out in [18]. Since $\langle \cdot, \mathcal L_r \cdot \rangle \simeq \|\cdot\|_{H_r}$, it follows from (3.1) that, for all $j$ and $c = (c_j^0, \cdots, c_j^T)$,

$$c^T (\Psi^j, \mathcal L_r \Psi^j) c \simeq c^T D_j^2 c, \quad (3.3)$$

with $D_j$ being the diagonal matrix defined by $D_j = \text{diag}(2^j I^{(|\Psi|)})_{l=0, \cdots, l}$ where $I^{(|\Psi|)}$ denotes the $|\Psi| \times |\Psi|$ identity matrix, i.e., $D_j c = (2^j c_j^0, \cdots, 2^j c_j^T)$. Note that $(\Psi^j, \mathcal L_r \Psi^j)$ is invertible and is the representation of $\mathcal L_r$ with respect to $\Psi^j$. Also, (3.2) can be written as

$$|d_j^T (\Psi^j, \mathcal L_r \Psi^j) c_j| \lesssim 2^{-\hat \gamma (t' - l')} \|d_j\|_2 \|c_j\|,$$  

(4.4)

for any $c_j \in \mathbb{R}^{(\Psi_j)}$, $d_j \in \mathbb{R}^{(\Psi_j)}$ and $\hat \gamma \in [0, \min\{\hat \gamma + r, \gamma - r\})$.

In the following, with $A$ and $B$ being two matrices, we will write $A \approx B$ if $x^T A x \approx x^T B x$ for all $x$ of appropriate dimension. In particular, we write (3.3) as

$$\langle \Psi^j, \mathcal L_r \Psi^j \rangle \approx D_j^2,$$  

(3.5)

which is equivalent to $\langle \Psi^j, \mathcal L_r \Psi^j \rangle^{-1} \approx D_j^{-2r}$. Further, it follows from (3.3) that, for any $d = (d_j^0, \cdots, d_j^T)^T$,

$$\sup_d \frac{|d_j^T D_j^{-r} (\Psi^j, \mathcal L_r \Psi^j) D_j^{-r} c|}{\|d\|} \geq \frac{c^T D_j^{-r} (\Psi^j, \mathcal L_r \Psi^j) D_j^{-r} c}{\|c\|} \approx \|c\|,$$

whereas it follows from the boundedness of $\mathcal L_r$ and (3.1) that

$$|d_j^T D_j^{-r} (\Psi^j, \mathcal L_r \Psi^j) D_j^{-r} c| \leq \|D_j^{-r} d_j^T \Psi^j\|_{H_r} \|D_j^{-r} c^T \Psi^j\|_{H_r} \approx \|d\| \|c\|,$$

and so $|D_j^{-r} (\Psi^j, \mathcal L_r \Psi^j) D_j^{-r} c| \approx \|c\|$ for all $c$, which will be written as

$$\|D_j^{-r} (\Psi^j, \mathcal L_r \Psi^j) D_j^{-r} c\| \approx 1.$$  

(3.6)
LEMMA 3.3. With \( \mathcal{L}_r \) being given as above and \( \hat{r} \in \mathbb{R} \), it holds that

\[
\langle \mathcal{L}_r \Psi^j, \Psi^j \rangle D_j^{2\hat{r}} \langle \Psi^j, \mathcal{L}_r \Psi^j \rangle \approx D^{2\hat{r}} D^{-2\hat{r}} D_j^{2\hat{r}},
\]

provided that \( \gamma \) and \( \bar{\rho} \) are sufficiently large.

Proof. Let \( L_j \) and \( L_{[j]}^{[l]} \) denote \( \langle \Psi^j, \mathcal{L}_r \Psi^j \rangle \) and \( \langle \Psi^j, \mathcal{L}_r \Psi^j \rangle \) for \( l, l' = 0, \ldots, j \) respectively. Further, let \( L_{[j]} \) denote the block diagonal matrix \( \text{diag}(L_{[j]}^{[l]}), l = 0, \ldots, j \), i.e. \( L_{[j]} c = ((L_{[j]}^{[0]} c_0)^T, \ldots, (L_{[j]}^{[j]} c_j)^T)^T \) with \( c = (c_0^T, \ldots, c_j^T)^T \) as before. Note that the counterparts of (3.5) and (3.6) within any level \( l \) are \( L_{[j]}^{[l]} \approx 2^{2rl} I[l] \) and \( \|L_{[j]}^{[l]}\| \approx 2^{2rl} \) respectively, with \( I[l] \) being the \( |\Psi_l| \times |\Psi_l| \) identity matrix as above.

It is sufficient to prove that \( \exists \epsilon \in (0, 1) \) such that

\[
\|D_j^{-\epsilon} [L_{j} - L_{[j]}] c \| \leq \epsilon \|D_j^{-\epsilon} L_{[j]} c \| \tag{3.7}
\]

for all \( c \). Indeed, since \( D_j^{-\epsilon} L_{j} c = D_j^{-\epsilon} L_{[j]} c + D_j^{-\epsilon} [L_{j} - L_{[j]}] c \), having proven (3.7), we infer that

\[
\|D_j^{-\epsilon} L_{j} c \| \leq \|D_j^{-\epsilon} L_{[j]} c \| + \|D_j^{-\epsilon} [L_{j} - L_{[j]}] c \| \leq (1 + \epsilon) \|D_j^{-\epsilon} L_{[j]} c \|
\]

and that

\[
\|D_j^{-\epsilon} L_{j} c \| \geq \|D_j^{-\epsilon} L_{[j]} c \| - \|D_j^{-\epsilon} [L_{j} - L_{[j]}] c \| \geq (1 - \epsilon) \|D_j^{-\epsilon} L_{[j]} c \|
\]

i.e.,

\[
\|D_j^{-\epsilon} L_{j} c \| \approx \|D_j^{-\epsilon} L_{[j]} c \|.
\]

In addition, it follows from \( \|L_{[j]}^{[l]}\| \approx 2^{2rl} \) that

\[
\|D_j^{-\epsilon} L_{[j]} c \| \approx \|D_j^{-\epsilon} D_j^{2\hat{r}} c \|.
\]

The last two relations yield \( \|D_j^{-\epsilon} L_{j} c \| \approx \|D_j^{-\epsilon} D_j^{2\hat{r}} c \| \) for all \( c \), proving our assertion.

Now, we prove (3.7) as follows. Let \( d = (d_0^T, \ldots, d_j^T)^T \), then it holds that

\[
|\langle d, D_j^{-\epsilon} [L_j - L_{[j]}] c \rangle | = |\langle D_j^{-\epsilon} d, [L_j - L_{[j]}] c \rangle | = \sum_{l=0}^{j} \sum_{l' \neq l=0}^{j} |\langle 2^{-\epsilon} d_l, L_{[j]}^{[l']} c_{l'} \rangle |. \tag{3.8}
\]

By (3.4) and \( \|L_{[j]}^{[l]}\| \approx 2^{2rl} \), it holds for any \( \hat{\gamma} = \hat{\gamma}_r \in [0, \min\{\beta + 1 + r, \gamma - r\}] \) that

\[
\begin{align*}
\sum_{l=0}^{j} \sum_{l' \neq l=0}^{j} |\langle 2^{-\epsilon} d_l, L_{[j]}^{[l']} c_{l'} \rangle | & \lesssim \sum_{l=0}^{j} \sum_{l' \neq l=0}^{j} 2^{-\hat{\gamma}|l-l'|} 2^{(\epsilon - r)|l-l'|} \|d_l\| 2^{rl} \|c_l\| \\
& = \sum_{l=0}^{j} \sum_{l' \neq l=0}^{j} 2^{-\hat{\gamma}|l-l'|} 2^{-(\epsilon - r)|l-l'|} \|d_l\| 2^{(2\epsilon - r)|l\|} \|c_l\| \\
& \lesssim \sum_{l=0}^{j} \sum_{l' \neq l=0}^{j} 2^{-(\gamma - r)|l-l'|} 2^{rl} \|d_l\| 2^{-\epsilon} \|L_{[j]}^{[l]} c_{l'}\|.
\end{align*}
\]

Let \( A \in \mathbb{R}^{(j+1) \times (j+1)} \) and \( x, y \in \mathbb{R}^{j+1} \) be defined by \( A_{i,i'} = (1 - \delta_{i,i'}) 2^{-(\gamma - r)|i-i'|} |i-i'| \), and \( x_i = 2^{-\epsilon(r-1)} |l_{[j]}^{[l-1]} c_{l-1}| \) and \( y_i = \|d_{l-1}\| \), then the last double sum is \( y^T A x \). Since \( y^T A x \leq \|A\| \|y\| \|x\| \leq (\|A\|_1, \|A\|_\infty)^2 \|y\| \|x\| \), we infer from (3.9) that
\[ \sum_{l=0}^{j} \sum_{l'i=0}^{j} |\langle 2^{-r} d_{l'}, L_{j,l'}^{r,l} c_{l}' \rangle| \lesssim 2^{-((\gamma_{l'} - |r - r'|)} (\sum_{l'=0}^{j} \|d_{l'}\|^2)^{\frac{1}{2}} (\sum_{l=0}^{j} \|2^{-r} L_{j,l}^{r,l} c_{l}\|^2)^{\frac{1}{2}} \]  

(3.10)

Since \(d\) is arbitrary, it follows from (3.8) and (3.10) that

\[ \|D_{j}^{-r}[L_{j} - L_{j}[\epsilon]]\| \lesssim 2^{-((\gamma_{l'} - |r - r'|)} \|D_{j}^{-r}L_{j}[\epsilon]\|. \]  

(3.11)

yielding (3.7) if \(\gamma\) and \(\hat{p}\) are sufficiently large, since \(\gamma \in [0, \min\{\hat{p} + 1 + r, \gamma - r\}\) arbitrary. \(\Box\)

In addition, let \(\hat{r} \in (-\gamma, \gamma)\), and let \(L_{\hat{r}}^r\) be the corresponding operator satisfying all of the assumptions above with interchanged roles of \(r\) and \(\hat{r}\). Lemma 3.3 is an intermediate step to the following more general result.

**Lemma 3.4.** With \(L_{\hat{r}}, L_{\hat{r}}\) being given as above and \(r, \hat{r} \in \mathbb{R}\) and \(\alpha \geq 0\), it holds that

\[ \langle [L_{\hat{r}}, \Psi^{j}, \Psi^{j}] + \alpha(L_{\hat{r}}, \Psi^{j}, \Psi^{j})\rangle D_{j}^{-2r}([\Psi^{j}, L_{\hat{r}}, \Psi^{j}] + \alpha(\Psi^{j}, L_{\hat{r}}, \Psi^{j})] \approx \rangle \]

\[ \|D_{2}^{2r} + \alpha D_{2}^{2r}\|D_{2}^{-2r}D_{2}^{2r} + \alpha D_{2}^{2r}\| \] independently on \(\alpha\), provided that \(\gamma\) and \(\hat{p}\) are sufficiently large.

**Proof.** The proof is analogous to and employs the proof of Lemma 3.3. Let \(L_{j}, L_{j}[\epsilon]\) and \(L_{j}^{r,l}\) be the stiffness, the block diagonal matrix and the block matrices respectively, defined as in the proof of Lemma 3.3, corresponding to \(L_{\hat{r}}\). Likewise, let \(L_{j}, L_{j}[\epsilon]\) and \(L_{j}^{r,l}\) be those matrices corresponding to \(L_{\hat{r}}\). Note that within any level \(l\) we have \([L_{j}^{r,l} + \alpha L_{j}^{r,l}] \approx (2^{-r} + \alpha 2^{-2r})L_{j}^{r,l}\) independently on \(\alpha\), from which we infer that \([\|L^{r,l}_j + \alpha L^{r,l}_j\| c_l] \approx (2^{-r} + \alpha 2^{-2r})\|c_l\|\) independently on \(\alpha\). Combining this with \([L_{j}^{r,l}] \approx 2^{-r}\| c_l\|\) and \([\alpha L^{r,l}_j] \approx \alpha 2^{-2r}\| c_l\|\), we obtain \([L_{j}^{r,l} + \alpha L_{j}^{r,l}] \approx 2^{-r} + \alpha 2^{-2r}\) independently on \(\alpha\).

It is sufficient to prove that \(\exists \epsilon \in (0, 1)\) being independent on \(\alpha\) such that

\[ \|D_{j}^{-r}[L_{j} + \alpha \tilde{L}_{j}] - [L_{j}[\epsilon] + \alpha \tilde{L}_{j}[\epsilon]]\| \leq \epsilon \|D_{j}^{-r}[L_{j}] + \alpha \tilde{L}_{j}[\epsilon]\| \]  

(3.12)

for all \(c = (c_0^{T}, \ldots, c_{j}^{T})^{T}\). Indeed, having proven (3.12), we infer that

\[ \|D_{j}^{-r}[L_{j} + \alpha \tilde{L}_{j}]\| \approx \|D_{j}^{-r}[L_{j}[\epsilon]] + \alpha \tilde{L}_{j}[\epsilon]\| \] independently on \(\alpha\). In addition, it follows from \([L_{j}^{r,l} + \alpha L_{j}^{r,l}] \approx 2^{-r} + \alpha 2^{-2r}\) that

\[ \|D_{j}^{-r}[L_{j}] + \alpha \tilde{L}_{j}[\epsilon]\| \approx \|D_{j}^{-r}[D_{2}^{2r} + \alpha D_{2}^{2r}]\| \] independently on \(\alpha\). The last two relations yield our assertion.

Now, (3.12) is a consequence of (3.11), \([L_{j}^{r,l}] \approx 2^{-r}\| c_l\|\), \([\alpha L_{j}^{r,l}] \approx \alpha 2^{-2r}\| c_l\|\), and \([L_{j}^{r,l} + \alpha L_{j}^{r,l}] \approx 2^{-r} + \alpha 2^{-2r}\) as follows. For any \(\gamma_{j} \in [0, \min\{\hat{p} + 1 + r, \gamma - r\}\) and \(\gamma_{j} \in [0, \min\{\hat{p} + 1 + r, \gamma - r\}\), it holds that

\[ \|D_{j}^{-r}[L_{j} + \alpha \tilde{L}_{j}] - [L_{j}[\epsilon] + \alpha \tilde{L}_{j}[\epsilon]]\| \lesssim 2^{-\max\{\gamma_{l'} - |r - r'|, \gamma_{l'} - |r - r'|\}} (\|D_{j}^{-r}[L_{j}]\| + \|\alpha D_{j}^{-r}[\tilde{L}_{j}]\|) \]

\[ \lesssim 2^{-\max\{\gamma_{l'} - |r - r'|, \gamma_{l'} - |r - r'|\}} (\|D_{j}^{-r}[L_{j}]\| + \|\alpha D_{j}^{-r}[\tilde{L}_{j}]\|) \]

\[ \lesssim 2^{-\max\{\gamma_{l'} - |r - r'|, \gamma_{l'} - |r - r'|\}} (2 \|D_{j}^{-r}[D_{2}^{2r} + \alpha D_{2}^{2r}]\|) \]

\[ \approx 2^{-\max\{\gamma_{l'} - |r - r'|, \gamma_{l'} - |r - r'|\}} \|D_{j}^{-r}[L_{j}] + \alpha \tilde{L}_{j}[\epsilon]\| \] independently on \(\alpha\), yielding (3.12) if \(\gamma\) and \(\hat{p}\) are sufficiently large. \(\Box\)
4. Decoupling and preconditioning. In this section, we show how the coupled linear system (2.10) can be decoupled and preconditioned by choosing nodal bases, or Lagrange bases, for the temporal approximation and wavelet bases for the spatial approximation. Although we focus on nodal bases, we will briefly discuss how our approach would also work with other temporal bases.

4.1. Temporal discretization. In the following, we will use the notation defined in Section 2. In addition, for $\hat{q} \geq 0$, let $\Xi^q = (\xi_0^q, \cdots, \xi_n^q)^T$ be the nodal basis, consisting of $\hat{q} + 1$ Lagrange polynomials of degree $\hat{q}$, for $P_{\hat{q}}(0, 1)$ with respect to some fixed partition $0 = \tau_0^\hat{q} < \cdots < \tau_{\hat{q}} = 1$ of $(0, 1)$. For $n = 1, \cdots, N$, with $\varphi_n$ being the affine bijections from $(0, 1)$ to $I_n$ defined by $t = \varphi_n(t) = h_n t - \xi_{n-1}$, let $\Xi_n^q$ be the nodal basis for $P_{\hat{q}}(I_n)$ defined by $\Xi_n^q = \Xi^q \circ \varphi_n^{-1} = (\xi_0^q \circ \varphi_n^{-1}, \cdots, \xi_n^q \circ \varphi_n^{-1})^T$.

Since $(\Xi^q(0) = (1, 0, \cdots, 0)^T$, we obtain from the last two subsystems in (2.10)

$$u_{n, 0} = u_{[n]},$$
$$v_{n, 0} = v_{[n]}.$$  \hspace{1cm} (4.1)

Note that, $u_{n, [i]}\Phi_j = u_0$ and $v_{n, [i]}\Phi_j = v_0$ as before whereas now, since $\Xi^q(1) = (0, \cdots, 0, 1)^T$, $u_{n, [i]} = u_{n-1, i}$ and $v_{n, [i]} = v_{n-1, i}$ for $n = 2, \cdots, N$. Further, by the ‘pull-back’ via the affine bijections $\varphi_n$ above, it holds that

$$\langle \Xi^q_n, \Xi^q \rangle = h_n \langle \Xi^q_n, \Xi^q \rangle$$
$$\text{and}$$

$$\langle \Xi^q_n, \Xi^q \rangle = \langle \Xi^q_n, \Xi^q \rangle,$$  \hspace{1cm} (4.2)

with $\Xi^q := ((\xi_0^q), \cdots, (\xi_n^q))^T$. Let $A$ denote the matrix obtained from $\langle \Xi^q, \Xi^q \rangle$ by deleting its first column $a$, i.e.,

$$a := \langle \Xi^q, \xi_0^q \rangle$$
$$\text{and}$$

$$\hat{a} := \langle \Xi^q, (\xi_0^q, \cdots, (\xi_n^q))^T \rangle,$$  \hspace{1cm} (4.3)

and analogously, let

$$\hat{A} := \langle \Xi^q, (\xi_0^q, \cdots, (\xi_n^q))^T \rangle$$
$$\text{and}$$

$$\hat{A} := \langle \Xi^q, (\xi_0^q, \cdots, (\xi_n^q))^T \rangle,$$  \hspace{1cm} (4.4)

then, by employing (4.1) and (4.2), we obtain the following linear system from the first two subsystems in (2.10)

$$\begin{bmatrix}
\hat{A} \otimes M_j & -h_n A \otimes M_j \\
h_n A \otimes S_j & \hat{A} \otimes M_j
\end{bmatrix}
\begin{bmatrix}
u_{0, n} \\
u_{0, n}
\end{bmatrix}
= 
\begin{bmatrix}
\hat{f}_{1, n} \\
\hat{f}_{2, n}
\end{bmatrix},$$  \hspace{1cm} (4.5)

with the remaining unknowns

$$u_{0, n} := (u_{[n]}, \cdots, u_{[n], q})^T$$
$$\text{and}$$

$$v_{0, n} := (v_{[n], 1}, \cdots, v_{[n], q})^T,$$

the spatial mass- and stiffness matrix

$$M_j := \langle \Phi_j, \Phi_j \rangle$$
$$\text{and}$$

$$S_j := \langle \nabla \Phi_j, \nabla \Phi_j \rangle,$$

and the right hand side

$$f_{1, n} := -\hat{a} \otimes M_j u_{[n]} + h_n a \otimes M_j v_{[n]}$$
$$\text{and}$$

$$f_{2, n} := -h_n a \otimes S_j u_{[n]} - \hat{a} \otimes M_j v_{[n]} + f_n.$$

Note that the dimension of the system (4.5) is $2\hat{q}/\hat{q}$. In the next subsection, we will decompose this system into $2\hat{q}$ systems of dimension $|\Phi_j|$ that yield, together with the two identities of dimension $|\Phi_j|$ in (4.1), our approximations to $U$ and $V$. 
4.2. Decoupling. In order to decompose (4.5), we note that, besides $M_j$ and $S_j$, the two $q \times q$ matrices $A$ and $\hat{A}$ are also invertible:

**Lemma 4.1.** For $q = q - 1$, let $\mathfrak{X} = (\xi_{q0}^1, \cdots, \xi_{q0}^q)^T$ be the nodal bases, consisting of $r + 1$ Lagrange polynomials of degree $q$, for $P_q(0,1)$ with respect to some fixed partitions $0 = \tau_{q0}^0 < \cdots < \tau_{q0}^q = 1$ of $(0,1)$ as before. Then the two $q \times q$ matrices $A$ and $\hat{A}$ defined in (4.3) and (4.4) respectively, i.e.,

$$A = (\xi_{q0}^1, \cdots, \xi_{q0}^q)^T \quad \text{and} \quad \hat{A} = (\xi_{q0}^1)^T, (\xi_{q0}^2, \cdots, (\xi_{q0}^q)^T)$$

are invertible.

**Proof.** First, suppose that the rank of $A$ is smaller than $q$. Then $\exists \ x = (x_1, \cdots, x_q)^T \neq 0$ such that $A x = 0$. In other words, with $p_1 := x^T (\xi_{q0}^1, \cdots, \xi_{q0}^q)^T \in P_q(0,1)$, we have $p_1 \neq 0$ and $\langle \mathfrak{X}^{-1}, p_1 \rangle = 0$, or equivalently we have

$$0 \neq p_1 \perp P_{q-1}(0,1)$$

on the one hand. On the other hand, since $p_1(\tau_{q0}^0) = p_1(0) = 0, \exists 0 \neq p_2 \in P_{q-1}(0,1)$ such that $p_1 = \tau p_2$, and so

$$\langle p_1, p_2 \rangle = (\tau, p_2^2) > 0,$$

which gives a contradiction. Thus $A$ is invertible.

Next, note that the collection $\{(\xi_{q0}^1)_\tau, \cdots, (\xi_{q0}^q)_\tau\}$ spans $P_{q-1}(0,1)$. Since $\mathfrak{X}$ is a partition of unity, i.e., $\sum_{q=1}^q (\xi_{q0}^q)_\tau = 1$, we have $\sum_{q=1}^q (\xi_{q0}^q)_\tau = 0$. In particular, we have

$$(\xi_{q0}^q)_\tau = -\sum_{i=1}^q (\xi_{q0}^i)_\tau,$$

from which we conclude that $(\xi_{q0}^1, \cdots, (\xi_{q0}^q)^T)$ is a basis for $P_{q-1}(0,1)$. Thus, since $\mathfrak{X}^{-1}$ is also basis for $P_{q-1}(0,1)$, $\hat{A}$ is invertible. \(\square\)

By this lemma, we employ a modified Schur complement of the system matrix in (4.5) to obtain the following decomposition of (4.5)

$$[\hat{A} \otimes M_j - h_1 A^2 \otimes S_j] \begin{pmatrix} u_{0;n} \\ v_{0;n} \end{pmatrix} = \begin{pmatrix} \tilde{f}_{1;n} \\ \tilde{f}_{2;n} \end{pmatrix},$$

with $\tilde{f}_{1;n}$ and $\tilde{f}_{2;n}$ being certain functions of $f_{1;n}$ and $f_{2;n}$. In other words, we obtain the following $q$ systems of dimension $|\Phi_j|$ for $v_{0;n}$

$$[I \otimes M_j] v_{0;n} = [h_1 A^{-1} A \otimes M_j] u_{0;n} - [h_1 A^{-1} A \otimes I_j] \tilde{f}_{1;n},$$

with $I_j$ denoting the $|\Phi_j|$ by $|\Phi_j|$ identity matrix as before whereas $I = I_{q \times q}$ denoting the $q \times q$ identity matrix, and with $u_{0;n}$ being given by the following coupled system of dimension $q|\Phi_j|

$$[I \otimes M_j + h_1^2 A^{-1} A \otimes S_j] u_{0;n} = \tilde{f}_{2;n}.$$

In the following, we show how (4.7) can be decoupled. Let the real Schur decomposition of $[\hat{A}^{-1} A]^2$ be given by

$$[\hat{A}^{-1} A]^2 = QTQ^T,$$
with $Q$ being an orthogonal matrix and $T$ being a real upper quasi-triangular matrix, by which we mean that $T$ is a sum of a real upper triangular matrix $T_1$ with zeros on the diagonal and a real block diagonal matrix $T_2$ with block size of at most $2 \times 2$. Note that any element $\lambda_i$ on the diagonal of $T_2$ is a real eigenvalue $[A^{-1}A]^2$ and that any $2 \times 2$ matrix $B_i$ on the diagonal of $T_2$ has two complex conjugate eigenvalues $\lambda_i, \pm i\lambda_i, \overline{\lambda_i}$ with $\lambda_i, \overline{\lambda_i} \neq 0$, that are also two complex conjugate eigenvalues of $[A^{-1}A]^2$ (see e.g. [12]). We then obtain the following block upper quasi-triangular system from (4.7)

$$[I \otimes M_j + h^2_n T \otimes S_j] \tilde{u}_{0,n} = [Q^T \otimes I_j] \tilde{f}_{2,n}, \quad (4.9)$$

with $\tilde{u}_{0,n} := [Q^T \otimes I_j] u_{0,n}$, i.e., the solution of (4.7) is given by

$$u_{0,n} = [Q \otimes I_j] \tilde{u}_{0,n}.$$ 

Obviously, solving (4.9), and so (4.7), means successively solving either a system of dimension $|\Phi_j|$ of the form

$$[M_j + \lambda_i h_{n}^2 S_j] x = b \quad (4.10)$$

or a coupled system of dimension $2|\Phi_j|$ of the form

$$[I_{2 \times 2} \otimes M_j + h^2_n B_i \otimes S_j] \xi = b, \quad (4.11)$$

with $I_{2 \times 2}$ denoting the $2 \times 2$ identity matrix.

In order to decouple (4.11), we proceed as follows. First, let $p_{i,\Re} + ip_{i,\Im}$ be a complex eigenvector corresponding to the complex eigenvalue $\lambda_i, \pm i\lambda_i, \overline{\lambda_i}$ of $B_i$ and let $P_i := [p_{i,\Re} \ p_{i,\Im}]$, then $P_i$ has full rank and

$$B_i = P_i \begin{bmatrix} \lambda_i, \Re & \lambda_i, \Im \\ -\lambda_i, \Im & \lambda_i, \Re \end{bmatrix} P_i^{-1}.$$ 

Hence, we obtain the following system from (4.11)

$$\begin{bmatrix} M_j + \lambda_i, \Re h_n^2 S_j & \lambda_i, \Im h_n^2 S_j \\ -\lambda_i, \Im h_n^2 S_j & M_j + \lambda_i, \Re h_n^2 S_j \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad (4.12)$$

with $(b_1^T, b_2^T)^T := [P_i^{-1} \otimes I_j] b$ and $(x_1^T, x_2^T)^T := [P_i^{-1} \otimes I_j] \xi$, i.e., the solution of (4.11) is given by

$$\xi = [P_i \otimes I_j] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$ 

Next, we employ a modified Schur complement to obtain from (4.12) the following equivalent system

$$\begin{bmatrix} M_j + \lambda_i, \Re h_n^2 S_j & \lambda_i, \Im h_n^2 S_j \\ S_j + \lambda_i, \Im h_n^{-1} [M_j + \lambda_i, \Re h_n^2 S_j][S_j^{-1} [M_j + \lambda_i, \Re h_n^2 S_j]]^T \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix}, \quad (4.13)$$

with $\tilde{b}_1$ and $\tilde{b}_2$ being certain functions of $b_1$ and $b_2$. In other words, we obtain from (4.12) the following two systems of dimension $|\Phi_j|$ \n
$$S_j x_2 = -\frac{1}{\lambda_i, \Im h_n} ([M_j + \lambda_i, \Re h_n^2 S_j] x_1 + \tilde{b}_1), \quad (4.14)$$

$$S_j x_2 = -\frac{1}{\lambda_i, \Im h_n} ([M_j + \lambda_i, \Re h_n^2 S_j] x_1 + \tilde{b}_1),$$
where
\[ [S_j + \lambda_{i}^{-2}h_n^{-2}][M_j + \lambda_{i}b_n^2S_j][S_j^{-1}][M_j + \lambda_{i}b_n^2S_j]]x_1 = \hat{b}_2. \] (4.15)

In summary, solving (4.9), and so (4.7), means solving \( q \) systems of dimension \( |\Phi_j| \) either of the form (4.10) or (4.14) or (4.15). Since (2.10) has a unique solution independently on the choice of bases, any of those \( q \) systems has also a unique solution. In particular, we obtain for (4.10):

**Lemma 4.2.** Let \( \lambda_i \in IR \) be an eigenvalue of \( [\hat{A}^{-1}A]^2 \), then \( \lambda_i > 0 \).

**Proof.** First, note that \( M_j \) and \( S_j \) are SPD. Let \( \mu \) be an eigenvalue of \( S_j^{-\frac{1}{2}}M_jS_j^{-\frac{1}{2}} \), then \( \mu > 0 \) and \( \exists \ x \neq 0 \) such that
\[ M_jx - \mu S_jx = 0. \]
Now, suppose that \( \lambda_i < 0 \). With \( h_n^2 = -\frac{\mu}{\lambda_i} > 0 \), we then have
\[ M_jx + \lambda_i h_n^2 S_jx = 0, \]
which contradicts the uniqueness of the solution of (4.10). In addition, since \( [\hat{A}^{-1}A]^2 \) is invertible, we conclude that \( \lambda_i > 0 \). \( \square \)

Since \( q \ll |\Phi_j| \) in practical applications, the temporal system matrices \( A \) and \( \hat{A} \) and their inverses, and the Schur decomposition and the eigenvalues of \( [\hat{A}^{-1}A]^2 \), and the \( 2 \times 2 \) matrices \( P_i \) and their inverses can be computed or approximated with high precision, as a once and for all preprocessing computation. Thus, there is only one problem left, which will be treated in the next subsection: how to obtain iterative solutions of (4.6), (4.10), (4.14) and (4.15) efficiently.

**Remark 4.3.** For standard bases \( \Xi^q_n \) for \( P_q(I_n) \) consisting of monomials of degree at most \( q \), since the two corresponding matrices \( A \) and \( \hat{A} \) are also invertible, one would obtain the same \( 2q \) subsystems of dimension \( |\Phi_j| \) as given in (4.6), (4.10), (4.14) and (4.15) (obviously with different right hand sides and unknowns). For orthonormal bases \( \Xi^q_n \) for \( P_q(I_n) \) consisting of Legendre polynomials of degree at most \( q \), by the invertibility of the redefined matrix \( A := (\Xi^q_{n-1}, (\xi^n_1, \cdots, \xi^n_{n-1})^T) \) and the corresponding matrix \( \hat{A} = (\Xi^q_{n-1}, ((\xi^n_1, \cdots, (\xi^n_{n-1})^T) \) as above, one would then obtain \( 2(q + 1) \) subsystems of dimension \( |\Phi_j| \) similar to the ones given in (4.6), (4.10), (4.14) and (4.15). In either case, Lemma 4.2 is still valid. Further, since the approximation spaces and the test spaces are different, we have in fact Petrov-Galerkin approximation, hence one could also choose (suitable) bases of different types for the approximation spaces and for the test spaces. \( \square \)

### 4.3. Preconditioning
In the following, we will drop the subscripts \( l^2 \) and \( l_0 \), and we will take \( \Phi_j = \Psi^q \) where \( \Psi^q = (\Psi^q_1, \cdots, \Psi^q_T)^T \) is some wavelet basis from Section 3 with parameters \( p + 1 > \gamma > 1 \) and \( \tilde{p} + 1 > \tilde{\gamma} > 0 \) such that \( \gamma \) and \( \tilde{\gamma} \) are sufficiently large. The mass matrix \( M_j \) and stiffness matrix \( S_j \) defined in Subsection 4.1 are thus
\[ M_j = \langle \Psi^j, \Psi^q \rangle \quad \text{and} \quad S_j = \langle \nabla \Psi^j, \nabla \Psi^q \rangle. \]
In terms of Subsection 3.4, \( M_j \) is the ‘stiffness’ matrix corresponding to the identity operator \( \mathcal{L}_0u = u \) and \( S_j \) is the stiffness matrix corresponding to the negative Laplacian operator \( \mathcal{L}_1u = -\Delta u \).

Note that the usual \( H^1 \)-seminorm \( | \cdot |_{H^1} = \| \nabla \cdot \| \) is equivalent with the usual \( H^1 \)-norm \( \| \cdot \|_{H^1} = (\| \nabla \cdot \|_2^2 + \| \cdot \|_2^2)^{\frac{1}{2}} \) on \( H^1_0 \) by the Poincaré-Friedrichs inequality,
so that \( \mathbf{c}^T \mathbf{S}_j \mathbf{c} = (\mathbf{c}^T \Psi^j, -\Delta \mathbf{c}^T \Psi^j) \approx \| \mathbf{c}^T \Psi^j \|^2_{\bar{H}} \) for any \( \mathbf{c} = (\mathbf{c}_0^T, \cdots, \mathbf{c}_n^T)^T \) as before. Hence, from Section 3, in particular (3.5) with \( r \in \{0, 1\} \), we conclude that

\[
\begin{align*}
\mathbf{M}_j & \simeq \mathbf{I}_j, \quad \mathbf{S}_j \simeq \mathbf{D}_j^2 \quad \text{and} \quad \mathbf{S}_j^{-1} \simeq \mathbf{D}_j^{-2} \quad (4.16)
\end{align*}
\]

independently on \( j \), with \( \mathbf{I}_j \) being the \( |\Psi^j| \times |\Psi^j| \) identity matrix as before and \( \mathbf{D}_j \) being the diagonal matrix defined in Subsection 3.4. In other words, \( \mathbf{I}_j \) and \( \mathbf{D}_j^2 \) are (SPD diagonal) preconditioners for \( \mathbf{M}_j \) and \( \mathbf{S}_j \) respectively, meaning that the condition number of \( \mathbf{M}_j \) and that of the two-sided preconditioned matrix \( \mathbf{D}_j^{-1} \mathbf{S}_j \mathbf{D}_j^{-1} \) are bounded independently on \( j \). Note that the mass- and stiffness matrix \( \mathbf{M}_j \) and \( \mathbf{S}_j \) are not truly sparse. However, as mentioned in Section 3, the applications of \( \mathbf{M}_j \) or \( \mathbf{S}_j \) to vectors of dimension \( |\Psi^j| \) can be carried out with linear complexity.

Since only matrix vector applications matter in the context of iterative methods, we conclude that iterative solutions of (4.6) and (4.14), as well as approximations to the Galerkin projections \( \mathbf{U}_0 \) and \( \mathbf{v}_0 \) of \( \mathbf{U}_0 \) and \( \mathbf{V}_0 \) from Section 2, of any desired accuracy can be obtained by, e.g., the conjugated gradient method (CG) efficiently, in the sense that the computational complexity is linear, independently on \( j \) and \( h_n \).

In the following, we propose (SPD diagonal) preconditioners for (4.10) and (4.15), via relations similar to (4.16), such that the condition numbers of the corresponding (two-sided) preconditioned matrices are bounded independently on \( j \) and \( h_n \), and so iterative solutions of (4.10) and (4.15) can also be obtained efficiently in this sense. Recall that it follows from Lemma 4.2 that the eigenvalue \( \lambda_j \) in (4.10) is positive, and note that applications of \( \mathbf{S}_j^{-1} \) in (4.15) can be iteratively approximated by, e.g., CG with the preconditioner \( \mathbf{D}_j^{-2} \).

**Proposition 4.4.** Let \( \mathbf{M}_j, \mathbf{S}_j, \mathbf{I}_j, \text{ and } \mathbf{D}_j \) be defined as above.

(i) For any \( h > 0 \), it holds that

\[
\begin{align*}
[\mathbf{M}_j + \hat{h} \mathbf{S}_j] & \simeq [\mathbf{I}_j + \hat{h} \mathbf{D}_j^2]
\end{align*}
\]

independently on \( j \) and \( \hat{h} \).

(ii) For any \( a, b \in \mathbb{R} \) and \( h > 0 \), it holds that

\[
\begin{align*}
\mathbf{S}_j + b^2 \hat{h}^{-4} [\mathbf{M}_j + a \hat{h}^2 \mathbf{S}_j] [\mathbf{S}_j^{-1}] [\mathbf{M}_j + a \hat{h}^2 \mathbf{S}_j] & \simeq \mathbf{D}_j^2 + b^2 \hat{h}^{-4} [\mathbf{I}_j + a \hat{h}^2 \mathbf{D}_j^2] [\mathbf{D}_j^{-2}] [\mathbf{I}_j + a \hat{h}^2 \mathbf{D}_j^2]
\end{align*}
\]

independently on \( j \) and \( \hat{h} \). Moreover, if \( a \geq 0 \) then it also holds independently on \( a \) and \( b \).

**Proof.**

(i) The assertion follows from (4.16).

(ii) If \( a \geq 0 \), the assertion follows from (4.16) and Lemma 3.4.

Otherwise, we write the assertion as

\[
\begin{align*}
\mathbf{S}_j + [b \hat{h}^{-2} \mathbf{M}_j - ab \mathbf{S}_j] [\mathbf{S}_j^{-1}] [b \hat{h}^{-2} \mathbf{M}_j - ab \mathbf{S}_j] & \simeq \mathbf{D}_j^2 + [b \hat{h}^{-2} \mathbf{I}_j - ab \mathbf{D}_j^2] [\mathbf{D}_j^{-2}] [b \hat{h}^{-2} \mathbf{I}_j - ab \mathbf{D}_j^2]
\end{align*}
\]

(4.17)

with \( a > 0 \) and \( b > 0 \) (for \( b = 0 \), the assertion follows from (4.16)). Now, for any \( \mathbf{c} = (\mathbf{c}_0^T, \cdots, \mathbf{c}_n^T)^T \), it follows from (4.16) that

\[
\begin{align*}
\mathbf{c}^T [\mathbf{S}_j + [b \hat{h}^{-2} \mathbf{M}_j - ab \mathbf{S}_j] [\mathbf{S}_j^{-1}] [b \hat{h}^{-2} \mathbf{M}_j - ab \mathbf{S}_j]] \mathbf{c} & \simeq \| \mathbf{D}_j^{-1} \mathbf{S}_j \mathbf{c} \|^2 + \| \mathbf{D}_j^{-1} [b \hat{h}^{-2} \mathbf{M}_j - ab \mathbf{S}_j] \mathbf{c} \|^2.
\end{align*}
\]

(4.18)
On the one hand, we have

\[
\|D_j^{-1}S_jc\|^2 + \|D_j^{-1}[b\hat{h}^{-2}M_j - abS_j]c\|^2 \\
\leq \|D_j^{-1}S_jc\|^2 + \|abD_j^{-1}S_jc\| + \|b\hat{h}^{-2}D_j^{-1}M_jc\|)^2 \\
\leq \|D_j^{-1}S_jc\|^2 + 2\|abD_j^{-1}S_jc\|^2 + 2\|b\hat{h}^{-2}D_j^{-1}M_jc\|^2 \\
\lesssim \|D_j^{-1}S_jc\|^2 + \|b\hat{h}^{-2}D_j^{-1}M_jc\|^2. \quad (4.19)
\]

where we have used Lemma 3.4 in the last line. Note that the equivalence constant in the fourth line is \(2(1 + a^2b^2)^{-1}\), independent on \(\hat{h}\). Analogously, we have on the other hand

\[
\|D_j^{-1}S_jc\|^2 + \|D_j^{-1}[b\hat{h}^{-2}M_j - abS_j]c\|^2 \\
\lesssim \|D_j^{-1}S_jc\|^2 + \|abD_j^{-1}S_jc\|^2 + \|D_j^{-1}[b\hat{h}^{-2}M_j - abS_j]c\|^2 \\
\lesssim \|D_j^{-1}S_jc\|^2 + 2^{-1}(\|abD_j^{-1}S_jc\| + \|D_j^{-1}[b\hat{h}^{-2}M_j - abS_j]c\|)^2 \\
\lesssim \|D_j^{-1}S_jc\|^2 + 2^{-1}\|b\hat{h}^{-2}D_j^{-1}M_jc\|^2 \\
\lesssim \|D_j^{-1}S_jc\|^2 + \|b\hat{h}^{-2}D_j^{-1}c\|^2,
\]

(the equivalence constant in the second line is \((1 + a^2b^2)^{-1}\), independent on \(\hat{h}\) again). It follows from (4.18), (4.19) and (4.20) that

\[
c^T[S_j + [b\hat{h}^{-2}M_j - abS_j]S_j][b\hat{h}^{-2}M_j - abS_j]c \approx c^T[D_j^2 + b^2\hat{h}^{-4}D_j^{-2}]c \quad (4.21)
\]

independently on \(\hat{h}\). With interchanged roles of \((M_j, S_j)\) and \((I_j, D_j^2)\) in (4.19) and (4.20), we directly infer that

\[
c^T[D_j^2 + [b\hat{h}^{-2}I_j - abD_j^2]D_j^{-2}[b\hat{h}^{-2}I_j - abD_j^2]c \approx c^T[D_j^2 + b^2\hat{h}^{-4}D_j^{-2}]c. \quad (4.22)
\]

The last two relations yield (4.17) independently on \(\hat{h}\).

\[\square\]

**Remark 4.5.** For \(a \geq 0, b \geq 0\) and \(\hat{h} > 0\), it holds that

\[
c^T[D_j^2 + [b\hat{h}^{-2}I_j + abD_j^2]D_j^{-2}[b\hat{h}^{-2}I_j + abD_j^2]c] \approx c^T[D_j^2 + b^2\hat{h}^{-4}D_j^{-2}]c, \quad (4.23)
\]

with equivalence constants \((1 + a^2b^2)^{-1}\) and \((1 + a^2b^2)^{-1}\). Hence, with

\[
P_+ := [D_j^2 + [b\hat{h}^{-2}I_j + abD_j^2]D_j^{-2}[b\hat{h}^{-2}I_j + abD_j^2]],
\]

\[
P_- := [D_j^2 + [b\hat{h}^{-2}I_j - abD_j^2]D_j^{-2}[b\hat{h}^{-2}I_j - abD_j^2]],
\]

\[
P_0 := [D_j^2 + b^2\hat{h}^{-4}D_j^{-2}],
\]

and similarly \(Q_\pm := [S_j + [b\hat{h}^{-2}M_j \pm abS_j]S_j^{-1}[b\hat{h}^{-2}M_j \pm abS_j]],\) it follows from (4.22), (4.23) and the proof of Proposition 4.4 that \(P_+, P_-\) and \(P_0\) are preconditioners for both \(Q_+\) and \(Q_-\). However, since \(P_+\), \(P_-\), \(Q_+\) and \(Q_-\) share the same constants \((1 + a^2b^2)^{-1}\) and \((1 + a^2b^2)^{-1}\) in their equivalences with \(P_0\), we may expect that \(P_+\) are quantitatively better than \(P_0\), being considered as preconditioners for \(Q_+\). In addition, since \(P_+ \approx Q_+\) independently on \(a\) and \(b\), we may even expect that \(P_- \approx Q_-\) independently on \(a\) and \(b\) also. Yet, we do not know how to prove it rigorously.

\[\square\]

**Remark 4.6.** Let \(D_{M_j}\) and \(D_{S_j}\) be the diagonal of \(M_j\) and \(S_j\) respectively, then \(D_{M_j} \approx I_j\) and \(D_{S_j} \approx D_j^2\). Since \(D_{M_j}\) and \(D_{S_j}\) are the diagonal of \(M_j\) and \(S_j\), we may expect that \(D_{M_j}\) and \(D_{S_j}\) are quantitatively better than \(I_j\) and \(D_j^2\), being considered as preconditioners for \(M_j\) and \(S_j\) respectively. It is well-known that this is indeed the case, and so we may expect better performances with \(D_{M_j}\) and \(D_{S_j}\) instead of \(I_j\) and \(D_j^2\) in Proposition 4.4. In our numerical experiments, we have therefore used \(D_{M_j}\) and \(D_{S_j}\) instead of \(I_j\) and \(D_j^2\) respectively.

\[\square\]
5. Numerical results. In this section, we present the numerical results of sev-
eral one and two dimensional experiments. Since the predicted rates of convergence
have been numerically observed in [11], we focus on preconditioning, by approximating
the $l_2$-condition numbers of the system matrices preconditioned accordingly to Sub-
section 4.3 and Remark 4.6. We only mention that the predicted rates of convergence
have also been observed in our one dimensional experiments.

In our experiments, $T = 5$ and $\Omega = (0, 1)^d$ for $d \in \{1, 2\}$. Further, our temporal
bases $\Xi_q^n$ and $\Xi_q^{n-1}$ are nodal bases with $q \in \{1, 2, 4\}$ and our spatial wavelet bases
$\Psi_j$ are finite element wavelet bases with parameters $p = \tilde{p} \in \{1, 2\}$ and $\gamma = \tilde{\gamma} = \frac{1}{2}$
constructed in [16]. Moreover, our spatial and temporal mesh are both uniform as
follows. Let

$$M_1 := [0, 1],$$

and

$$M_{2j} := \{(x_1, x_2) \in [0, 1]^2 : x_2 \geq x_1\} \quad \text{and} \quad M_{2j} := \{(x_1, x_2) \in [0, 1]^2 : x_2 \leq x_1\}.$$ 

For $d = 1$, our initial spatial mesh $\mathcal{M}_0$ consists of the two intervals resulted from a
dyadic refinement of $M_1$, i.e., $\mathcal{M}_0 = \{[0, \frac{1}{2}], \frac{1}{2}, 1]\}$. Analogously for $d = 2$, our initial
spatial mesh $\mathcal{M}_0$ consists of the eight triangles resulted from dyadic refinements of
$M_{21}$ and $M_{22}$. Thus, our spatial meshes $\mathcal{M}_j$ for $j \geq 0$ consist of $2^{d(j+1)}$ uniform
d-simplices for $d \in \{1, 2\}$ (and so $V_j = V_j(p)$ and $\Psi_j$ can be defined, cf. Subsection
3.2). Finally, our temporal stepsizes $h_n$ are defined by either

$$h_n = h_1(j) := \frac{T}{T \lceil q^2 \gamma^2 (j+1) \rceil} \quad \text{or} \quad h_n = h_2(j) := \frac{T}{q^{2d} \gamma^2 (j+1)},$$

with $\lceil \cdot \rceil$ being the ceiling function (thus, we focus on $L^2$ rates of convergence, cf.
Remark 2.2). Note that $h_1(j) \leq h_2(j)$. In both cases $h = h_n$ and $N h = T$.

Recall that the system matrices in (4.6) and (4.14) are $D_{M_j}$ and $S_j$ respectively,
and their preconditioners $D_{M_j}$ and $S_j$ are independent on the eigenvalues of the
matrix $[\hat{A}^{-1}A]^2$ from Subsection 4.2. The condition numbers of the corresponding
preconditioned matrices (i.e., $D_{M_j}^{-\frac{2}{q^2}} S_j D_{M_j}^{-\frac{2}{q^2}}$ and $D_{S_j}^{-\frac{2}{q^2}} S_j D_{S_j}^{-\frac{2}{q^2}}$), being dependent on $p$,
will be denoted as

$$\kappa_{p,M,j} \quad \text{and} \quad \kappa_{p,S,j}$$

respectively.

In contrary, the system matrices in (4.10) and (4.15) depend on the eigenvalues of
$[\hat{A}^{-1}A]^2$, so do their preconditioners, as followed from Proposition 4.4. In particular,
the form of the preconditioners depend on whether the corresponding eigenvalues
of $[\hat{A}^{-1}A]^2$ are real or complex (note that two system matrices corresponding to a
pair of complex conjugated eigenvalues have the same preconditioner). For $q = 1$,$[\hat{A}^{-1}A]^2$ has one (positive) real eigenvalue. For $q = 2$, it has one pair of complex
conjugated eigenvalues with positive real part. For $q = 4$ it has one pair of complex
conjugated eigenvalues with positive real part and one pair of complex conjugated
eigenvalues with negative real part. We did not numerically investigate the case
$q = 3$, because $[\hat{A}^{-1}A]^2$ has one (positive) real eigenvalue and one pair of complex
conjugated eigenvalues with positive real part in this case, and so the corresponding
preconditioners are expected to behave similarly to the preconditioners in the two cases $q \in \{1, 2\}$. For $q \in \{1, 2\}$, the condition numbers of the preconditioned matrices, being dependent also on $h_i(j)$ for $i \in \{1, 2\}$, will be denoted as

$$\kappa_{(p,q),i,j}.$$ 

For $q = 4$, the condition numbers of the preconditioned matrices will be denoted as

$$\kappa_+^{(p,q),i,j} \text{ and } \kappa_-^{(p,q),i,j},$$

corresponding to the pair of complex conjugated eigenvalues with positive and negative real part respectively.

We approximated the condition numbers with the Lanczos method. In addition, we approximated the applications of $S_j^{-1}$ for $q \in \{2, 4\}$ with (preconditioned) CG as mentioned before. The numerical results are presented in Tables 5.1 and 5.2 for $d = 1$, and in Tables 5.3 and 5.4 for $d = 2$.

### Table 5.1

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\kappa_{p,M,j}$</th>
<th>$\kappa_{p,S,j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2.2498e+00$</td>
<td>$1.5536e+01$</td>
</tr>
<tr>
<td>2</td>
<td>$2.9034e+00$</td>
<td>$9.3342e+00$</td>
</tr>
</tbody>
</table>

### Table 5.2

<table>
<thead>
<tr>
<th>$q$</th>
<th>$\kappa_{(p,q),i,j}$, and $\kappa_+^{(p,q),i,j}$ and $\kappa_-^{(p,q),i,j}$ for $d = 1$ (with $j = 10$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$\kappa_{(p,q),i,10}$</td>
</tr>
<tr>
<td>1</td>
<td>$5.1423e+00$</td>
</tr>
<tr>
<td>2</td>
<td>$5.1423e+00$</td>
</tr>
<tr>
<td>1</td>
<td>$2.9033e+00$</td>
</tr>
<tr>
<td>2</td>
<td>$2.9033e+00$</td>
</tr>
</tbody>
</table>

### Table 5.3

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\kappa_{p,M,j}$</th>
<th>$\kappa_{p,S,j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.2214e+01$</td>
<td>$4.8632e+01$</td>
</tr>
<tr>
<td>2</td>
<td>$1.2344e+01$</td>
<td>$5.9916e+01$</td>
</tr>
</tbody>
</table>

### Table 5.4

<table>
<thead>
<tr>
<th>$q$</th>
<th>$\kappa_{(p,q),i,j}$, and $\kappa_+^{(p,q),i,j}$ and $\kappa_-^{(p,q),i,j}$ for $d = 2$ (with $j = 6$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$\kappa_{(p,q),i,6}$</td>
</tr>
<tr>
<td>1</td>
<td>$2.0131e+01$</td>
</tr>
<tr>
<td>2</td>
<td>$2.0131e+01$</td>
</tr>
<tr>
<td>1</td>
<td>$1.2414e+01$</td>
</tr>
<tr>
<td>2</td>
<td>$1.2414e+01$</td>
</tr>
</tbody>
</table>
From the numerical results, we draw the following conclusions:

- As discussed in [16], the preconditioned mass- and stiffness matrices above are, in a certain sense, quite close to the optimal ones. In comparison to the preconditioned mass- and stiffness matrices, our diagonal preconditioners show thus an overall satisfactory performance.

- The preconditioners correspond to pairs of complex conjugated eigenvalues with negative real part behave similarly to those correspond to pairs of complex conjugated eigenvalues with positive real part.

- For $q = 1$, the condition numbers of the preconditioned system matrices are proportional to those of the preconditioned mass matrices (which can also be theoretically verified).

- For $q \in \{2, 4\}$, the condition numbers of the preconditioned system matrices tend to be more proportional to those of the preconditioned stiffness matrices than to those of the preconditioned mass matrices.

REFERENCES


