ESAIM: M2AN 60 (2026) 981–1019
https://doi.org/10.1051/m2an/2026024

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

CONVERGENCE ANALYSIS OF A BALANCING DOMAIN DECOMPOSITION METHOD FOR AN ELLIPTIC OPTIMAL CONTROL PROBLEM WITH HDG DISCRETIZATIONS

Sijing Liu[1] and Jinjin Zhang[2,*]

Abstract. In this work, a balancing domain decomposition by constraints (BDDC) algorithm is applied to the nonsymmetric positive definite linear system arising from the hybridizable discontinuous Galerkin (HDG) discretization of an elliptic distributed optimal control problem. Convergence analysis for the BDDC preconditioned generalized minimal residual (GMRES) solver demonstrates that, when the subdomain size is small enough, the algorithm is robust with respect to the regularization parameter, and the number of iterations is independent of the number of subdomains and depends only slightly on the subdomain problem size. Numerical experiments are performed to confirm the theoretical results.

Mathematics Subject Classification. 49J20, 49M41, 65N30, 65N55.

Received August 19, 2025. Accepted February 23, 2026.

1. Introduction

Elliptic distributed optimal control problems have been extensively studied using various discretization techniques. Standard P_1 finite element methods are applied to an elliptic optimal control problem constrained by a diffusion–reaction equation in [36] and to one constrained by a advection–diffusion–reaction equation in [8]. Discontinuous Galerkin (DG) methods have been explored in [28, 29, 32, 35] for their applicability to optimal control problems as well. Moreover, HDG methods are investigated in [11, 12, 26] and the error estimates in L^2 norm are provided. Streamline upwind/Petrov–Galerkin (SUPG) methods have also been developed and analyzed for the elliptic optimal control problems constrained by a advection-dominated state equation in [14, 24]. In particular, an error estimate in the energy norm for an HDG discretization has been established in [34], where the dependence on the regularization parameter is explicitly tracked.

In terms of existing solvers for the discretized system, multigrid methods provide an effective approach for solving elliptic optimal control problems. Typically, there are two main approaches (cf. [36]): the all-at-once approach (cf. [8,37,38]), in which the multigrid method is applied directly to the coupled optimality system, and the approach in [4,5,23], where multigrid methods are applied separately to the equations of the state variable, the dual variable, and/or the control variable as components of the overall iterative scheme. For example, in

Keywords and phrases. Elliptic distributed optimal control problems, hybridizable discontinuous Galerkin methods, BDDC algorithms.

[1] Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA, USA.

[2] Department of Mathematics, The Ohio State University, Columbus, OH, USA.

[*] Corresponding author: zhang.14647@osu.edu; jinjin14647@gmail.com

[37], an additive Schwarz-type iterative method is formulated as a preconditioned Richardson method, which enables the use of multigrid as a fully coupled all-at-once solver for the Karush–Kuhn–Tucker system, and in [4], a domain-decomposed Schur complement partial differential equation (PDE) solver with a Krylov–Schur preconditioner is proposed, which employs an approximate state/decision variable decomposition by replacing the forward PDE Jacobians with their respective preconditioners.

BDDC algorithms are popular nonoverlapping domain decomposition techniques that were first introduced in [15], they have been developed for solving symmetric positive definite problems with mixed and hybrid formulations [40, 41], DG methods [27], and HDG discretizations [46]. A variant of the BDDC method has been proposed for symmetric indefinite systems arising from finite element discretizations of the Helmholtz equation in [30]. BDDC algorithm can also be extended to solve nonsymmetric positive definite problems. In [47], BDDC algorithms are applied to the nonsymmetric positive definite discretized system resulting from using HDG discretizations on the advection–diffusion equation. In [48], BDDC algorithms are developed for the Oseen equations, where the globally coupled unknowns are reduced to the numerical trace of the velocity on element boundaries and the mean pressure on each element. The resulting discretized nonsymmetric saddle point system, derived from HDG discretizations, is solved within a benign subspace framework using GMRES. In [34], we propose a BDDC preconditioner to solve an elliptic distributed optimal control problem discretized by HDG methods. In this approach, the state and adjoint state variables on the subdomain interfaces are coupled as the primary unknowns. The original problem is reduced to a global interface system. A BDDC preconditioner is constructed by assembling the subdomain Schur complements and is used to solve the coupled interface unknowns. Once the interface variables are computed, the remaining variables in the system can be efficiently recovered.

In this article, we present a convergence analysis for the BDDC preconditioned GMRES method for the optimal control problem proposed in [34]. The analysis establishes both upper and lower bounds with respect to the regularization parameter β, as stated in Theorem 5.6, particularly for small values of β. This work is a continuation of [34]. We show that when the subdomain size H is sufficiently small, the convergence rate becomes independent of β, and the number of iterations is independent of the number of subdomains. The analysis proceeds in several steps. First, given the numerical trace $\boldsymbol{\lambda}$ defined on the union of element boundaries, we construct some extensions from the HDG discretized system of the optimal control problem. Using the results in [47], we derive estimates for the norms of these extensions by relating the optimal control equations to a pair of advection–diffusion equations. Second, due to the presence of the L^2-inner product in the nonsymmetric part of the bilinear form, we introduce a new norm in Theorem 5.6. Unlike Theorem 6.6 of [47], which only focuses on the symmetric part of the bilinear form, our norm incorporates the L^2-norm of the extension operators in addition to the symmetric component. Consequently, it is not equivalent to the triple-bar norm used in the advection–diffusion problem (cf. [47]), which behaves similarly to an H^1 semi-norm. Instead, we propose a full norm rather than a semi-norm, which behaves similarly to an energy norm as defined in equation (3.3) of [34]. This difference makes it difficult to directly estimate the h-norm, an L^2 type norm defined in Section 4, of the optimal control extension for certain key operators in the lower bound analysis. To address this issue, we estimate the h-norm of the corresponding advection–diffusion extension of these operators, then apply norm equivalence relations to obtain the desired lower bound. Third, we establish upper and lower bounds that explicitly track the dependence on the regularization parameter β. These bounds are derived by relating extensions of the optimal control problem to extensions of the advection–diffusion problem and applying known estimates for the latter.

The rest of the paper is organized as follows. In Section 2, we introduce the HDG discretizations of the optimal control problem and derive the corresponding matrix formulation. In Section 3, the subdomain local problem, a reduced subdomain interface problem, and the BDDC preconditioner are given. In Section 4, we define some extensions for the optimal control problem and introduce some useful bilinear forms and norms. In Section 5, we give some estimates of these extensions and establish the upper and lower bounds for the BDDC algorithm with the regularization parameter β explicitly tracked. In Section 6, we provide the proofs of the lemmas used in Section 5. In Section 7, we provide some numerical experiments to confirm the theoretical results. We end with some concluding remarks in Section 8. An appendix in Section A summarizes key operators and frequently used symbols.

2. Problem setting and HDG discretizations

We consider the following elliptic optimal control problem. Let $\Omega \subset \mathbb{R}^2$ be a bounded convex polygonal domain, the desire state $y_d \in L^2(\Omega)$, and the regularization parameter $\beta > 0$. Find

$$(\bar{y}, \bar{u}) = \underset{(y,u)}{\operatorname{argmin}} \left[\frac{1}{2} \|y - y_d\|^2_{L^2(\Omega)} + \frac{\beta}{2} \|u\|^2_{L^2(\Omega)} \right], \tag{2.1}$$

where (y, u) belongs to $H^1_0(\Omega) \times L^2(\Omega)$ and satisfies

$$a(y, v) = \int_\Omega uv \, \mathrm{d}x \qquad \forall v \in H^1_0(\Omega). \tag{2.2}$$

Here the bilinear form $a(\cdot, \cdot)$ is defined as

$$a(y, v) = \int_\Omega \nabla y \cdot \nabla v \, \mathrm{d}x + \int_\Omega (\boldsymbol{\zeta} \cdot \nabla y) v \, \mathrm{d}x, \tag{2.3}$$

where the vector field $\boldsymbol{\zeta} \in [W^{1,\infty}(\Omega)]^2$, and let $\nabla \cdot \boldsymbol{\zeta} = 0$ for simplicity (cf. [1, 6, 20]). Here y denotes the state variable and u is the control variable.

By the classical theory in [31, 39], the solution of (2.1) and (2.2) is characterized by the first-order optimality system (cf. [8, 25])

$$\begin{aligned}
&a(q, \bar{p}) = (\bar{y} - y_d, q)_{L^2(\Omega)} && \forall q \in H^1_0(\Omega), && (2.4a)\\
&\bar{p} + \beta \bar{u} = 0, && && (2.4b)\\
&a(\bar{y}, z) = (\bar{u}, z)_{L^2(\Omega)} && \forall z \in H^1_0(\Omega), && (2.4c)
\end{aligned}$$

where $\bar{p}$ denotes the adjoint state. By eliminating the control variable $\bar{u}$ using the relation (2.4b), we arrive at the following reduced saddle point problem:

$$\begin{aligned}
a(q, \bar{p}) - (\bar{y}, q)_{L^2(\Omega)} &= -(y_d, q)_{L^2(\Omega)} && \forall q \in H^1_0(\Omega), && (2.5a)\\
-(\bar{p}, z)_{L^2(\Omega)} - \beta a(\bar{y}, z) &= 0 && \forall z \in H^1_0(\Omega). && (2.5b)
\end{aligned}$$

In order to obtain a more balanced system, we use the change of variable $\tilde{p} = -\beta^{-\frac{1}{4}} \bar{p}$ and $\tilde{y} = \beta^{\frac{1}{4}} \bar{y}$ to arrive at

$$\begin{aligned}
\beta^{\frac{1}{2}} a(q, \tilde{p}) + (\tilde{y}, q)_{L^2(\Omega)} &= \beta^{\frac{1}{4}} (y_d, q)_{L^2(\Omega)} && \forall q \in H^1_0(\Omega), && (2.6a)\\
-(\tilde{p}, z)_{L^2(\Omega)} + \beta^{\frac{1}{2}} a(\tilde{y}, z) &= 0 && \forall z \in H^1_0(\Omega). && (2.6b)
\end{aligned}$$

Remark 2.1. By considering the reduced system (2.5), we do not explicitly discretize the control variable $\bar{u}$. This common strategy can be found in [8, 25]. The scaling argument we utilize here first appeared in [7, 8] and was used to balance the parameter β so that a suitable energy norm is straightforward to define. This scaling strategy is also utilized in [21, 22, 33, 34].

As mentioned in [34], it is more useful to consider a general problem as follows. Let f and g be sufficiently smooth, find $(y, p) \in H^1_0(\Omega) \times H^1_0(\Omega)$ such that

$$\begin{aligned}
\beta^{\frac{1}{2}} a(q, p) + (y, q)_{L^2(\Omega)} &= (g, q)_{L^2(\Omega)} && \forall q \in H^1_0(\Omega), && (2.7a)\\
-(p, z)_{L^2(\Omega)} + \beta^{\frac{1}{2}} a(y, z) &= (f, z)_{L^2(\Omega)} && \forall z \in H^1_0(\Omega). && (2.7b)
\end{aligned}$$

Therefore, the problem (2.7) is the variational form of the following equations:

$$\beta^{\frac{1}{2}}(-\Delta y+\boldsymbol{\zeta}\cdot\nabla y)-p=f \quad \text{in} \quad \Omega, \tag{2.8a}$$
$$y=0 \quad \text{on} \quad \partial\Omega, \tag{2.8b}$$
$$\beta^{\frac{1}{2}}(-\Delta p-\boldsymbol{\zeta}\cdot\nabla p)+y=g \quad \text{in} \quad \Omega, \tag{2.8c}$$
$$p=0 \quad \text{on} \quad \partial\Omega. \tag{2.8d}$$

We can also write the system (2.8) as a first-order system for the description of the HDG discretization. Let $\mathbf{q}=-\nabla y$ and $\mathbf{p}=-\nabla p$. we can write (2.8) as

$$\mathbf{q}+\nabla y=0 \quad \text{in} \quad \Omega, \tag{2.9a}$$
$$\beta^{\frac{1}{2}}(\nabla\cdot\mathbf{q}+\boldsymbol{\zeta}\cdot\nabla y)-p=f \quad \text{in} \quad \Omega, \tag{2.9b}$$
$$y=0 \quad \text{on} \quad \partial\Omega, \tag{2.9c}$$
$$\mathbf{p}+\nabla p=0 \quad \text{in} \quad \Omega, \tag{2.9d}$$
$$\beta^{\frac{1}{2}}(\nabla\cdot\mathbf{p}-\boldsymbol{\zeta}\cdot\nabla p)+y=g \quad \text{in} \quad \Omega, \tag{2.9e}$$
$$p=0 \quad \text{on} \quad \partial\Omega. \tag{2.9f}$$

2.1. HDG formulation and the matrix form

We define a shape-regular and quasi-uniform triangulation $\mathcal{T}_h$ of the domain Ω, where h represents the characteristic element size. Each element in $\mathcal{T}_h$ is denoted by K, and the set of all edges of the triangulation is denoted by $\mathcal{E}_h$. Define the inner products $(\cdot,\cdot)_{\mathcal{T}_h}=\sum_{K\in\mathcal{T}_h}(\cdot,\cdot)_K$ and $\langle\cdot,\cdot\rangle_{\partial\mathcal{T}_h}=\sum_{K\in\mathcal{T}_h}\langle\cdot,\cdot\rangle_{\partial K}$, where $(\cdot,\cdot)_K$ and $\langle\cdot,\cdot\rangle_{\partial K}$ represent the L^2-inner products for functions defined over the element K and its boundary ∂K, respectively. Define the finite element space as follows:

$$\begin{aligned}
\mathbf{V}_h &= \left\{\mathbf{v}\in(L^2(\Omega))^n : \mathbf{v}|_K\in(P^k(K))^n, \forall K\in\mathcal{T}_h\right\},\\
W_h &= \left\{w\in L^2(\Omega) : w|_K\in P^k(K), \forall K\in\mathcal{T}_h\right\},\\
X_h &= \left\{\mu\in L^2(\mathcal{E}_h) : \mu|_e\in P^k(e), \forall e\in\mathcal{E}_h\right\},\\
\Lambda_h &= \{\mu\in X_h : \mu|_e=0, \forall e\in\partial\Omega\}.
\end{aligned}$$

For notational simplicity, we will drop the subscript h and denote the corresponding spaces by $\mathbf{V}, W$ and Λ. The HDG method is to find $(\mathbf{q}_h,\mathbf{p}_h,y_h,p_h,\widehat{y}_h,\widehat{p}_h)\in\mathbf{V}\times\mathbf{V}\times W\times W\times\Lambda\times\Lambda$ such that,

$$-\beta^{\frac{1}{2}}(\mathbf{q}_h,\mathbf{r}_1)_{\mathcal{T}_h}+\beta^{\frac{1}{2}}(y_h,\nabla\cdot\mathbf{r}_1)_{\mathcal{T}_h}-\beta^{\frac{1}{2}}\langle\widehat{y}_h,\mathbf{r}_1\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h}=0, \tag{2.10a}$$
$$-\beta^{\frac{1}{2}}(\mathbf{p}_h,\mathbf{r}_2)_{\mathcal{T}_h}+\beta^{\frac{1}{2}}(p_h,\nabla\cdot\mathbf{r}_2)_{\mathcal{T}_h}-\beta^{\frac{1}{2}}\langle\widehat{p}_h,\mathbf{r}_2\cdot\mathbf{n}\rangle_{\partial\mathcal{T}_h}=0, \tag{2.10b}$$
$$\begin{aligned}\beta^{\frac{1}{2}}(\nabla\cdot\mathbf{q}_h,w_1)_{\mathcal{T}_h}-\beta^{\frac{1}{2}}(y_h,\boldsymbol{\zeta}\cdot\nabla w_1)_{\mathcal{T}_h}+\beta^{\frac{1}{2}}\langle\tau_1 y_h,w_1\rangle_{\partial\mathcal{T}_h}&\\
-(p_h,w_1)_{\mathcal{T}_h}+\beta^{\frac{1}{2}}\langle(\boldsymbol{\zeta}\cdot\mathbf{n}-\tau_1)\widehat{y}_h,w_1\rangle_{\partial\mathcal{T}_h}&=(f,w_1)_{\mathcal{T}_h},\end{aligned} \tag{2.10c}$$
$$\begin{aligned}\beta^{\frac{1}{2}}(\nabla\cdot\mathbf{p}_h,w_2)_{\mathcal{T}_h}+(y_h,w_2)_{\mathcal{T}_h}+\beta^{\frac{1}{2}}(p_h,\boldsymbol{\zeta}\cdot\nabla w_2)_{\mathcal{T}_h}&\\
+\beta^{\frac{1}{2}}\langle\tau_2 p_h,w_2\rangle_{\partial\mathcal{T}_h}-\beta^{\frac{1}{2}}\langle(\tau_2+\boldsymbol{\zeta}\cdot\mathbf{n})\widehat{p}_h,w_2\rangle_{\partial\mathcal{T}_h}&=(g,w_2)_{\mathcal{T}_h},\end{aligned} \tag{2.10d}$$
$$-\beta^{\frac{1}{2}}\langle\mathbf{q}_h\cdot\mathbf{n},\mu_1\rangle_{\partial\mathcal{T}_h}-\beta^{\frac{1}{2}}\langle\tau_1 y_h,\mu_1\rangle_{\partial\mathcal{T}_h}-\beta^{\frac{1}{2}}\langle(\boldsymbol{\zeta}\cdot\mathbf{n}-\tau_1)\widehat{y}_h,\mu_1\rangle_{\partial\mathcal{T}_h}=0, \tag{2.10e}$$
$$-\beta^{\frac{1}{2}}\langle\mathbf{p}_h\cdot\mathbf{n},\mu_2\rangle_{\partial\mathcal{T}_h}-\beta^{\frac{1}{2}}\langle\tau_2 p_h,\mu_2\rangle_{\partial\mathcal{T}_h}+\beta^{\frac{1}{2}}\langle(\boldsymbol{\zeta}\cdot\mathbf{n}+\tau_2)\widehat{p}_h,\mu_2\rangle_{\partial\mathcal{T}_h}=0, \tag{2.10f}$$

The matrix blocks in for all $(\mathbf{r}_1,\mathbf{r}_2,w_1,w_2,\mu_1,\mu_2)\in\mathbf{V}\times\mathbf{V}\times W\times W\times\Lambda\times\Lambda$, where τ_1 are τ_2 are piecewise local stabilization parameters, see [11,34] for more details. Define $\mathbf{G}=\begin{bmatrix}\mathbf{q}\\ \mathbf{p}\end{bmatrix}$, $\mathbf{u}=\begin{bmatrix}y\\ p\end{bmatrix}$, $\boldsymbol{\lambda}=\begin{bmatrix}\widehat{y}\\ \widehat{p}\end{bmatrix}$ and $\mathbf{Z}_h=\begin{bmatrix}F_h\\ G_h\end{bmatrix}$, where

$F_h = (f, w_1)_{\mathcal{T}_h}$ and $G_h = (g, w_2)_{\mathcal{T}_h}$, and $\mathbf{q}, \mathbf{p}, y, p, \widehat{y}$ and $\widehat{p}$ are the unknowns associated with $\mathbf{q}_h, \mathbf{p}_h, y_h, p_h, \widehat{y}_h$ and $\widehat{p}_h$, respectively. The matrix form of system (2.10) can be written as

$$\mathbf{A_c}\begin{bmatrix}\mathbf{G}\\ \mathbf{u}\\ \boldsymbol{\lambda}\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\ \mathbf{Z}_h\\ \mathbf{0}\end{bmatrix}, \tag{2.11}$$

where

$$\mathbf{A_c} = \begin{bmatrix}\mathbf{A_{GG}} & \mathbf{A^T_{uG}} & \mathbf{A^T_{\boldsymbol{\lambda}G}}\\ \mathbf{A_{uG}} & \mathbf{A_{uu}} & \mathbf{A_{u\boldsymbol{\lambda}}}\\ \mathbf{A_{\boldsymbol{\lambda}G}} & \mathbf{A_{\boldsymbol{\lambda}u}} & \mathbf{A_{\boldsymbol{\lambda\lambda}}}\end{bmatrix}.$$

The matrix blocks in $\mathbf{A_c}$ are defined as follows

$$\mathbf{A_{GG}} = \begin{bmatrix}A_{\mathbf{qq}} & \mathbf{0}\\ \mathbf{0} & A_{\mathbf{pp}}\end{bmatrix}, \mathbf{A_{uG}} = \begin{bmatrix}A_{y\mathbf{q}} & \mathbf{0}\\ \mathbf{0} & A_{p\mathbf{p}}\end{bmatrix}, \mathbf{A_{uu}} = \begin{bmatrix}A_{yy} & A_{yp}\\ A_{py} & A_{pp}\end{bmatrix}, \mathbf{A_{u\boldsymbol{\lambda}}} = \begin{bmatrix}A_{y\hat{y}} & \mathbf{0}\\ \mathbf{0} & A_{p\hat{p}}\end{bmatrix}$$
$$\mathbf{A_{\boldsymbol{\lambda}u}} = \begin{bmatrix}A_{\hat{y}y} & \mathbf{0}\\ \mathbf{0} & A_{\hat{p}p}\end{bmatrix}, \quad \text{and} \quad \mathbf{A_{\boldsymbol{\lambda\lambda}}} = \begin{bmatrix}A_{\hat{y}\hat{y}} & \mathbf{0}\\ \mathbf{0} & A_{\hat{p}\hat{p}}\end{bmatrix},$$

where the entries of these submatrices are obtained from the inner products defined in (2.10), evaluated on the polynomial basis functions associated with the spaces $\mathbf{V}$, W, and Λ.

For any $(\mathbf{q}_1, u_1, \lambda_1, \mathbf{p}_1, v_1, \mu_1), (\mathbf{q}_2, u_2, \lambda_2, \mathbf{p}_2, v_2, \mu_2) \in \mathbf{V}\times W\times\Lambda\times\mathbf{V}\times W\times\Lambda$, by the definition of $\mathbf{A_c}$, we can define the following bilinear form:

$$\mathcal{B}_h((\mathbf{q}_1, u_1, \lambda_1, \mathbf{p}_1, v_1, \mu_1), (\mathbf{q}_2, u_2, \lambda_2, \mathbf{p}_2, v_2, \mu_2)) = (\mathbf{q}_2, \mathbf{p}_2, u_2, v_2, \lambda_2, \mu_2)\mathbf{A_c}(\mathbf{q}_1, \mathbf{p}_1, u_1, v_1, \lambda_1, \mu_1)^T.$$

We can reorder the unknowns by regrouping the equations in (2.10) as follows:

$$\mathcal{B}_h((\mathbf{q}_1, u_1, \lambda_1, \mathbf{p}_1, v_1, \mu_1), (\mathbf{q}_2, u_2, \lambda_2, \mathbf{p}_2, v_2, \mu_2)) = (\mathbf{q}_2, u_2, \lambda_2, \mathbf{p}_2, v_2, \mu_2)\mathbf{A_a}(\mathbf{q}_1, u_1, \lambda_1, \mathbf{p}_1, v_1, \mu_1)^T, \tag{2.12}$$

where

$$\mathbf{A_a} = \begin{bmatrix}\beta^{\frac{1}{2}}A_1^{ad} & -L\\ L & \beta^{\frac{1}{2}}A_2^{ad}\end{bmatrix}. \tag{2.13}$$

Here A_1^{ad} is defined as the matrix in equation (2.23) of [47] with viscosity $\epsilon = 1$, corresponding to the advection–diffusion equation for the state variable y. The matrix A_2^{ad} is obtained by replacing $\boldsymbol{\zeta}$ with $-\boldsymbol{\zeta}$ and τ_1 with τ_2 in A_1^{ad}, corresponding to the advection–diffusion equation for the dual variable p. We see that L satisfies

$$(\mathbf{q}_2, u_2, \lambda_2)L(\mathbf{q}_1, u_1, \lambda_1)^T = (u_1, u_2)_{\mathcal{T}_h}, \tag{2.14}$$

for all $(\mathbf{q}_1, u_1, \lambda_1), (\mathbf{q}_2, u_2, \lambda_2) \in \mathbf{V}\times W\times\Lambda$. Moreover, let

$$\mathbf{L_a} = \begin{bmatrix}L & \mathbf{0}\\ \mathbf{0} & L\end{bmatrix}, \tag{2.15}$$

then

$$(\mathbf{q}_2, u_2, \lambda_2, \mathbf{p}_2, v_2, \mu_2)\mathbf{L_a}(\mathbf{q}_1, u_1, \lambda_1, \mathbf{p}_1, v_1, \mu_1)^T = (u_1, u_2)_{\mathcal{T}_h} + (v_1, v_2)_{\mathcal{T}_h},$$

for any $(\mathbf{q}_1, u_1, \lambda_1, \mathbf{p}_1, v_1, \mu_1), (\mathbf{q}_2, u_2, \lambda_2, \mathbf{p}_2, v_2, \mu_2) \in \mathbf{V}\times W\times\Lambda\times\mathbf{V}\times W\times\Lambda$.

Definition 2.2. Let $\boldsymbol{\Lambda} = \Lambda \times \Lambda$. Given $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$, we solve (2.11) locally on each element K with $\mathbf{Z}_h = \mathbf{0}$ and obtain the solution $(\mathbf{Q}\boldsymbol{\lambda}, \mathbf{U}\boldsymbol{\lambda})$ defined as the extensions of $\boldsymbol{\lambda}$ from the element boundary into the element interior for the optimal control problem. Solving (2.11) over all elements in $\mathcal{T}_h$ with $\mathbf{Z}_h = \mathbf{0}$ yields the global solution $(\mathbf{Q}\boldsymbol{\lambda}, \mathbf{U}\boldsymbol{\lambda})$, which together with $\boldsymbol{\lambda}$ satisfy

$$\begin{bmatrix} \mathbf{A_{GG}} & \mathbf{A_{uG}^T} & \mathbf{A_{\boldsymbol{\lambda} G}^T} \\ \mathbf{A_{uG}} & \mathbf{A_{uu}} & \mathbf{A_{u\boldsymbol{\lambda}}} \end{bmatrix} \begin{bmatrix} \mathbf{Q}\boldsymbol{\lambda} \\ \mathbf{U}\boldsymbol{\lambda} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{2.16}$$

Denote $\mathbf{Q}\boldsymbol{\lambda} = (Q_{A_1}\boldsymbol{\lambda}, Q_{A_2}\boldsymbol{\lambda})^T$,$\mathbf{U}\boldsymbol{\lambda} = (U_{A_1}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{\lambda})^T$, we define

$$\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}) = (Q_{A_1}\boldsymbol{\lambda}, U_{A_1}\boldsymbol{\lambda}, \lambda, Q_{A_2}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{\lambda}, \mu)^T. \tag{2.17}$$

Given $\boldsymbol{\lambda} = (\lambda, \mu)^T$ and $\boldsymbol{s} = (s, t)^T \in \boldsymbol{\Lambda}$, it follows from (2.12)–(2.14) and (2.17) that

$$\begin{aligned}
\mathcal{B}_h(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}), \mathbf{E}_{\mathcal{C}}(\boldsymbol{s})) &= \mathbf{E}_{\mathcal{C}}^T(\boldsymbol{s})\, \mathbf{A}\, \mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}) \\
&= (Q_{A_1}\boldsymbol{s}, U_{A_1}\boldsymbol{s}, s,\, Q_{A_2}\boldsymbol{s}, U_{A_2}\boldsymbol{s}, t) \begin{bmatrix} \beta^{\frac{1}{2}} A_1^{ad} & -L \\ L & \beta^{\frac{1}{2}} A_2^{ad} \end{bmatrix} (Q_{A_1}\boldsymbol{\lambda}, U_{A_1}\boldsymbol{\lambda}, \lambda,\, Q_{A_2}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{\lambda}, \mu)^T \\
&= (Q_{A_1}\boldsymbol{s}, U_{A_1}\boldsymbol{s}, s) \left(\beta^{\frac{1}{2}} A_1^{ad}\right) (Q_{A_1}\boldsymbol{\lambda}, U_{A_1}\boldsymbol{\lambda}, \lambda)^T \\
&\quad + (Q_{A_2}\boldsymbol{s}, U_{A_2}\boldsymbol{s}, t) \left(\beta^{\frac{1}{2}} A_2^{ad}\right) (Q_{A_2}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{\lambda}, \mu)^T \\
&\quad - (U_{A_1}\boldsymbol{s}, U_{A_2}\boldsymbol{\lambda})_{\mathcal{T}_h} + (U_{A_1}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{s})_{\mathcal{T}_h} \\
&= \mathbf{b}_h(\boldsymbol{\lambda}, \boldsymbol{s}) + \mathbf{z}_h(\boldsymbol{\lambda}, \boldsymbol{s}),
\end{aligned}$$

where the last equality follows from equations (2.9) and (2.19)–(2.22) of [47], together with $\langle \boldsymbol{\zeta} \cdot \mathbf{n}\, \lambda, s\rangle_{\partial\mathcal{T}_h} = 0$ and $\langle \boldsymbol{\zeta} \cdot \mathbf{n}\, \mu, t\rangle_{\partial\mathcal{T}_h} = 0$. Here $\mathbf{b}_h(\boldsymbol{\lambda}, \boldsymbol{s})$ and $\mathbf{z}_h(\boldsymbol{\lambda}, \boldsymbol{s})$ denote, respectively, the symmetric and nonsymmetric bilinear forms, defined by

$$\begin{aligned}
\mathbf{b}_h(\boldsymbol{\lambda}, \boldsymbol{s}) &= \beta^{\frac{1}{2}} (Q_{A_1}\boldsymbol{\lambda}, Q_{A_1}\boldsymbol{s})_{\mathcal{T}_h} + \beta^{\frac{1}{2}} \left\langle \left(\tau_1 - \frac{1}{2}\boldsymbol{\zeta} \cdot \mathbf{n}\right)(U_{A_1}\boldsymbol{\lambda} - \lambda), (U_{A_1}\boldsymbol{s} - s) \right\rangle_{\partial\mathcal{T}_h} \\
&\quad + \beta^{\frac{1}{2}} (Q_{A_2}\boldsymbol{\lambda}, Q_{A_2}\boldsymbol{s})_{\mathcal{T}_h} + \beta^{\frac{1}{2}} \left\langle \left(\tau_2 + \frac{1}{2}\boldsymbol{\zeta} \cdot \mathbf{n}\right)(U_{A_2}\boldsymbol{\lambda} - \mu), (U_{A_2}\boldsymbol{s} - t) \right\rangle_{\partial\mathcal{T}_h},
\end{aligned} \tag{2.18}$$

$$\begin{aligned}
\mathbf{z}_h(\boldsymbol{\lambda}, \boldsymbol{s}) &= \frac{\beta^{\frac{1}{2}}}{2} (U_{A_1}\boldsymbol{s}, \boldsymbol{\zeta} \cdot \nabla U_{A_1}\boldsymbol{\lambda})_{\mathcal{T}_h} - \frac{\beta^{\frac{1}{2}}}{2} (U_{A_1}\boldsymbol{\lambda}, \boldsymbol{\zeta} \cdot \nabla U_{A_1}\boldsymbol{s})_{\mathcal{T}_h} \\
&\quad - \frac{\beta^{\frac{1}{2}}}{2} \langle \boldsymbol{\zeta} \cdot \mathbf{n}\, s, U_{A_1}\boldsymbol{\lambda}\rangle_{\partial\mathcal{T}_h} + \frac{\beta^{\frac{1}{2}}}{2} \langle \boldsymbol{\zeta} \cdot \mathbf{n}\, \lambda, U_{A_1}\boldsymbol{s}\rangle_{\partial\mathcal{T}_h} - (U_{A_1}\boldsymbol{s}, U_{A_2}\boldsymbol{\lambda})_{\mathcal{T}_h} \\
&\quad - \frac{\beta^{\frac{1}{2}}}{2} (U_{A_2}\boldsymbol{s}, \boldsymbol{\zeta} \cdot \nabla U_{A_2}\boldsymbol{\lambda})_{\mathcal{T}_h} + \frac{\beta^{\frac{1}{2}}}{2} (U_{A_2}\boldsymbol{\lambda}, \boldsymbol{\zeta} \cdot \nabla U_{A_2}\boldsymbol{s})_{\mathcal{T}_h} \\
&\quad + \frac{\beta^{\frac{1}{2}}}{2} \langle \boldsymbol{\zeta} \cdot \mathbf{n}\, t, U_{A_2}\boldsymbol{\lambda}\rangle_{\partial\mathcal{T}_h} - \frac{\beta^{\frac{1}{2}}}{2} \langle \boldsymbol{\zeta} \cdot \mathbf{n}\, \mu, U_{A_2}\boldsymbol{s}\rangle_{\partial\mathcal{T}_h(\Omega_i)} + (U_{A_1}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{s})_{\mathcal{T}_h}.
\end{aligned} \tag{2.19}$$

The full bilinear form is defined by

$$\boldsymbol{a}_h(\boldsymbol{\lambda}, \boldsymbol{s}) = \mathbf{b}_h(\boldsymbol{\lambda}, \boldsymbol{s}) + \mathbf{z}_h(\boldsymbol{\lambda}, \boldsymbol{s}).$$

By the definitions of $\boldsymbol{a}_h$, (2.12) and (2.17), we have

$$\mathcal{B}_h(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}), \mathbf{E}_{\mathcal{C}}(\boldsymbol{s})) = \boldsymbol{a}_h(\boldsymbol{\lambda}, \boldsymbol{s}), \qquad \forall \boldsymbol{\lambda}, \boldsymbol{s} \in \boldsymbol{\Lambda}. \tag{2.20}$$

In order to establish the relations between the extensions of the optimal control problem and the extensions of the advection–diffusion problem, we introduce another definition.

Definition 2.3. Given $\boldsymbol{\lambda} = (\lambda, \mu)^T \in \boldsymbol{\Lambda}$, on each element K, we define

$$\mathbf{Q}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}) = \left(Q_{A_1^{ad}}\lambda, Q_{A_2^{ad}}\mu\right)^T, \quad \mathbf{U}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}) = \left(U_{A_1^{ad}}\lambda, U_{A_2^{ad}}\mu\right)^T,$$

where $(Q_{A_1^{ad}}\lambda, U_{A_1^{ad}}\lambda)$ denotes the extensions from the element boundary into the element interior for the advection–diffusion problem, defined similar to equation (2.14) of [47], using the extension operators associated with A_1^{ad}. Here $(Q_{A_2^{ad}}\mu, U_{A_2^{ad}}\mu)$ is defined analogously by applying the extension operators associated with A_2^{ad}. Similar as equation (2.14) of [47], we obtain that

$$A_1^{ad}\begin{bmatrix} Q_{A_1^{ad}}\lambda \\ U_{A_1^{ad}}\lambda \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad A_2^{ad}\begin{bmatrix} Q_{A_2^{ad}}\mu \\ U_{A_2^{ad}}\mu \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

We also define

$$\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}) = \left(Q_{A_1^{ad}}\lambda, U_{A_1^{ad}}\lambda, \lambda, Q_{A_2^{ad}}\mu, U_{A_2^{ad}}\mu, \mu\right)^T. \tag{2.21}$$

Let $\mathbf{A}^{\mathbf{ad}} = \begin{bmatrix} A_1^{ad} & \mathbf{0} \\ \mathbf{0} & A_2^{ad} \end{bmatrix}$. Given $\boldsymbol{\lambda}, \boldsymbol{s} \in \boldsymbol{\Lambda}$, we define the following bilinear form

$$\mathcal{B}_h^{\mathbf{ad}}(\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}), \mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{s})) = \mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}^T(\boldsymbol{s})\mathbf{A}^{\mathbf{ad}}\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}). \tag{2.22}$$

Recall in the system (2.11), we eliminate $\mathbf{G}$ and $\mathbf{u}$ in each element independently and obtain a system for $\boldsymbol{\lambda}$ as follows:

$$\mathbf{A}\boldsymbol{\lambda} = \mathbf{b}, \tag{2.23}$$

where

$$\mathbf{A} = \mathbf{A}_{\boldsymbol{\lambda}\boldsymbol{\lambda}} - \begin{bmatrix} \mathbf{A}_{\boldsymbol{\lambda}\mathbf{G}} & \mathbf{A}_{\boldsymbol{\lambda}\mathbf{u}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{\mathbf{GG}} & \mathbf{A}_{\mathbf{uG}}^{\mathbf{T}} \\ \mathbf{A}_{\mathbf{uG}} & \mathbf{A}_{\mathbf{uu}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}_{\boldsymbol{\lambda}\mathbf{G}}^{\mathbf{T}} \\ \mathbf{A}_{\mathbf{u}\boldsymbol{\lambda}} \end{bmatrix},$$

and

$$\mathbf{b} = -\begin{bmatrix} \mathbf{A}_{\boldsymbol{\lambda}\mathbf{G}} & \mathbf{A}_{\boldsymbol{\lambda}\mathbf{u}} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{\mathbf{GG}} & \mathbf{A}_{\mathbf{uG}}^{\mathbf{T}} \\ \mathbf{A}_{\mathbf{uG}} & \mathbf{A}_{\mathbf{uu}} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{Z}_h \end{bmatrix}.$$

Let $\mathbf{A} = \mathbf{B} + \mathbf{Z}$, where

$$\mathbf{B} = \frac{\mathbf{A} + \mathbf{A}^{\mathbf{T}}}{2}, \quad \mathbf{Z} = \frac{\mathbf{A} - \mathbf{A}^{\mathbf{T}}}{2} \tag{2.24}$$

denote the symmetric and nonsymmetric parts of $\mathbf{A}$, respectively. For any $\boldsymbol{\lambda}, \boldsymbol{s} \in \boldsymbol{\Lambda}$, we define the following bilinear forms

$$\langle \boldsymbol{\lambda}, \boldsymbol{s} \rangle_{\mathbf{B}} = \boldsymbol{s}^T\mathbf{B}\boldsymbol{\lambda} = \mathbf{b}_h(\boldsymbol{\lambda}, \boldsymbol{s}), \qquad \langle \boldsymbol{\lambda}, \boldsymbol{s} \rangle_{\mathbf{Z}} = \boldsymbol{s}^T\mathbf{Z}\boldsymbol{\lambda} = \mathbf{z}_h(\boldsymbol{\lambda}, \boldsymbol{s}). \tag{2.25}$$

In addition, we define the matrix $\mathbf{L}$ by

$$\langle \boldsymbol{\lambda}, \boldsymbol{s} \rangle_{\mathbf{L}} = \boldsymbol{s}^T\mathbf{L}\boldsymbol{\lambda} = (U_{A_1}\boldsymbol{\lambda}, U_{A_1}\boldsymbol{s})_{\mathcal{T}_h} + (U_{A_2}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{s})_{\mathcal{T}_h}. \tag{2.26}$$

3. The BDDC algorithm

3.1. Domain decomposition and a reduced subdomain interface problem

We decompose the domain Ω into N nonoverlapping subdomains $\Omega_i(i=1,2,\cdots,N)$ and the diameter of each subdomain is H_i. Let $H=\max_i H_i$ and $\Gamma=\left(\bigcup_{i\neq j}(\partial\Omega_i\cap\partial\Omega_j)\right)\setminus\partial\Omega$ be the subdomain interface. The spaces of finite element functions on Ω_i are denoted by $\mathbf{V}^{(i)}$, $W^{(i)}$, and $\Lambda^{(i)}$.

Given $\boldsymbol{\lambda}^{(i)}=(\lambda^{(i)},\mu^{(i)})^T, \boldsymbol{s}^{(i)}=(s^{(i)},t^{(i)})^T\in\boldsymbol{\Lambda}^{(i)}$, we define the following bilinear forms on each subdomain Ω_i:

$$
\begin{aligned}
\mathbf{b}_h\Big(\boldsymbol{\lambda}^{(i)},\boldsymbol{s}^{(i)}\Big)&=\beta^{\frac12}\Big(Q_{A_1}\boldsymbol{\lambda}^{(i)},Q_{A_1}\boldsymbol{s}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)} &(3.1)\\
&\quad+\beta^{\frac12}\Big\langle\Big(\tau_1-\frac12\boldsymbol{\zeta}\cdot\mathbf{n}\Big)\Big(U_{A_1}\boldsymbol{\lambda}^{(i)}-\lambda^{(i)}\Big),\Big(U_{A_1}\boldsymbol{s}^{(i)}-s^{(i)}\Big)\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)}\\
&\quad+\beta^{\frac12}\Big(Q_{A_2}\boldsymbol{\lambda}^{(i)},Q_{A_2}\boldsymbol{s}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)}\\
&\quad+\beta^{\frac12}\Big\langle\Big(\tau_2+\frac12\boldsymbol{\zeta}\cdot\mathbf{n}\Big)\Big(U_{A_2}\boldsymbol{\lambda}^{(i)}-\mu^{(i)}\Big),\Big(U_{A_2}\boldsymbol{s}^{(i)}-t^{(i)}\Big)\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)},\\
\mathbf{z}_h\Big(\boldsymbol{\lambda}^{(i)},\boldsymbol{s}^{(i)}\Big)&=\frac{\beta^{\frac12}}{2}\Big(U_{A_1}\boldsymbol{s}^{(i)},\boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\lambda}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)}-\frac{\beta^{\frac12}}{2}\Big(U_{A_1}\boldsymbol{\lambda}^{(i)},\boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{s}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)} &(3.2)\\
&\quad-\frac{\beta^{\frac12}}{2}\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,s^{(i)},U_{A_1}\boldsymbol{\lambda}^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)}+\frac{\beta^{\frac12}}{2}\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,\lambda^{(i)},U_{A_1}\boldsymbol{s}^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)}\\
&\quad-\Big(U_{A_1}\boldsymbol{s}^{(i)},U_{A_2}\boldsymbol{\lambda}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)}-\frac{\beta^{\frac12}}{2}\Big(U_{A_2}\boldsymbol{s}^{(i)},\boldsymbol{\zeta}\cdot\nabla U_{A_2}\boldsymbol{\lambda}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)}\\
&\quad+\frac{\beta^{\frac12}}{2}\Big(U_{A_2}\boldsymbol{\lambda}^{(i)},\boldsymbol{\zeta}\cdot\nabla U_{A_2}\boldsymbol{s}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)}+\frac{\beta^{\frac12}}{2}\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,t^{(i)},U_{A_2}\boldsymbol{\lambda}^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)}\\
&\quad-\frac{\beta^{\frac12}}{2}\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,\mu^{(i)},U_{A_2}\boldsymbol{s}^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)}+\Big(U_{A_1}\boldsymbol{\lambda}^{(i)},U_{A_2}\boldsymbol{s}^{(i)}\Big)_{\mathcal{T}_h(\Omega_i)},\\
\boldsymbol{a}_h\Big(\boldsymbol{\lambda}^{(i)},\boldsymbol{s}^{(i)}\Big)&=\mathbf{b}_h\Big(\boldsymbol{\lambda}^{(i)},\boldsymbol{s}^{(i)}\Big)+\mathbf{z}_h\Big(\boldsymbol{\lambda}^{(i)},\boldsymbol{s}^{(i)}\Big). &(3.3)
\end{aligned}
$$

The local bilinear form $\mathcal{B}_h^{(i)}$ is defined by

$$
\mathcal{B}_h^{(i)}\Big(\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{\lambda}^{(i)}\Big),\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{s}^{(i)}\Big)\Big)=\boldsymbol{a}_h\Big(\boldsymbol{\lambda}^{(i)},\boldsymbol{s}^{(i)}\Big)-\frac12\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,\lambda^{(i)},s^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)}+\frac12\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,\mu^{(i)},t^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)},
$$

which is obtained by (2.12) and restricting $\mathbf{A_a}$ to the subdomain Ω_i.

To make the subdomain local problem positive definite, we introduce the Robin boundary conditions and let

$$
\begin{aligned}
&\mathcal{B}^{(i)}\Big(\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{\lambda}^{(i)}\Big),\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{s}^{(i)}\Big)\Big)\\
&\quad=\mathcal{B}_h^{(i)}\Big(\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{\lambda}^{(i)}\Big),\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{s}^{(i)}\Big)\Big)+\frac12\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,\lambda^{(i)},s^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)}-\frac12\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\,\mu^{(i)},t^{(i)}\Big\rangle_{\partial\mathcal{T}_h(\Omega_i)} \qquad (3.4)\\
&\quad=\mathbf{E}_{\mathcal{C}}^T\Big(\boldsymbol{s}^{(i)}\Big)\mathbf{A}_{\mathbf{a}}^{(i)}\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{\lambda}^{(i)}\Big)=\mathbf{E}_{\mathcal{C}}^T\Big(\boldsymbol{s}^{(i)}\Big)\begin{bmatrix}\beta^{\frac12}A_1^{ad^{(i)}} & -L^{(i)}\\ L^{(i)} & \beta^{\frac12}A_2^{ad^{(i)}}\end{bmatrix}\mathbf{E}_{\mathcal{C}}\Big(\boldsymbol{\lambda}^{(i)}\Big)
\end{aligned}
$$

where $L^{(i)}$ is the restriction of the global matrix L in (2.14) to the subdomain Ω_i, and $\mathbf{A}_{\mathbf{a}}^{(i)}$, $A_1^{ad^{(i)}}$, $A_2^{ad^{(i)}}$ denote the subdomain matrices of $\mathbf{A_a}$, A_1^{ad}, and A_2^{ad} in (2.13), restricted to the subdomain Ω_i with Robin boundary conditions, respectively. We then have

$$\mathcal{B}^{(i)}\left(\mathbf{E}_{\mathcal{C}}\left(\boldsymbol{\lambda}^{(i)}\right), \mathbf{E}_{\mathcal{C}}\left(\boldsymbol{s}^{(i)}\right)\right) = \boldsymbol{a}_h^{(i)}\left(\boldsymbol{\lambda}^{(i)}, \boldsymbol{s}^{(i)}\right), \qquad \forall \boldsymbol{\lambda}^{(i)}, \boldsymbol{s}^{(i)} \in \boldsymbol{\Lambda}^{(i)}. \tag{3.5}$$

To reduce the global problem in (2.23) to a subdomain interface problem, we perform the following decomposition

$$\boldsymbol{\Lambda} = \left(\bigoplus_{i=1}^{N} \boldsymbol{\Lambda}_{\mathrm{I}}^{(i)}\right) \bigoplus \widehat{\boldsymbol{\Lambda}}_{\Gamma}, \tag{3.6}$$

where $\widehat{\boldsymbol{\Lambda}}_{\Gamma}$ denotes the degrees of freedom on the subdomain interface and $\boldsymbol{\Lambda}_{\mathrm{I}}^{(i)}$ denotes the degrees of freedom in the interior of subdomain Ω_i.

Given $\boldsymbol{\lambda}_{\Gamma} \in \widehat{\boldsymbol{\Lambda}}_{\Gamma}$, the original global problem (2.23) can be decomposed as

$$\begin{bmatrix} \mathbf{A}_{\mathrm{II}} & \mathbf{A}_{\mathrm{I}\Gamma} \\ \mathbf{A}_{\Gamma\mathrm{I}} & \mathbf{A}_{\Gamma\Gamma} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_{\mathrm{I}} \\ \boldsymbol{\lambda}_{\Gamma} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{\mathrm{I}} \\ \mathbf{b}_{\Gamma} \end{bmatrix}. \tag{3.7}$$

Therefore, for each subdomain Ω_i, the subdomain problem can be written as

$$\begin{bmatrix} \mathbf{A}_{\mathrm{II}}^{(i)} & \mathbf{A}_{\mathrm{I}\Gamma}^{(i)} \\ \mathbf{A}_{\Gamma\mathrm{I}}^{(i)} & \mathbf{A}_{\Gamma\Gamma}^{(i)} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda}_{\mathrm{I}}^{(i)} \\ \boldsymbol{\lambda}_{\Gamma}^{(i)} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{\mathrm{I}}^{(i)} \\ \mathbf{b}_{\Gamma}^{(i)} \end{bmatrix}. \tag{3.8}$$

The subdomain local Schur complement $\mathbf{S}_{\Gamma}^{(i)}$ can be defined as:

$$\mathbf{S}_{\Gamma}^{(i)} \boldsymbol{\lambda}_{\Gamma}^{(i)} = \mathbf{g}_{\Gamma}^{(i)}, \tag{3.9}$$

where

$$\begin{aligned} \mathbf{S}_{\Gamma}^{(i)} &= \mathbf{A}_{\Gamma\Gamma}^{(i)} - \mathbf{A}_{\Gamma\mathrm{I}}^{(i)} {\mathbf{A}_{II}^{(i)}}^{-1} \mathbf{A}_{\mathrm{I}\Gamma}^{(i)}, \\ \mathbf{g}_{\Gamma}^{(i)} &= \mathbf{b}_{\Gamma}^{(i)} - \mathbf{A}_{\Gamma I}^{(i)} {\mathbf{A}_{\mathrm{II}}^{(i)}}^{-1} \mathbf{b}_{\mathrm{I}}^{(i)}. \end{aligned}$$

Denote $\mathbf{R}_{\Gamma}^{(i)}$ as the restriction operator from $\widehat{\boldsymbol{\Lambda}}_{\Gamma}$ to the subdomain interface $\boldsymbol{\Lambda}_{\Gamma}^{(i)}$. By assembling the subdomain local Schur complement $\mathbf{S}_{\Gamma}^{(i)}$, we obtain the global interface Schur problem: find $\boldsymbol{\lambda} \in \widehat{\boldsymbol{\Lambda}}_{\Gamma}$ such that

$$\widehat{\mathbf{S}}_{\Gamma} \boldsymbol{\lambda}_{\Gamma} = \mathbf{g}_{\Gamma}, \tag{3.10}$$

where

$$\widehat{\mathbf{S}}_{\Gamma} = \sum_{i=1}^{N} {\mathbf{R}_{\Gamma}^{(i)}}^{T} \mathbf{S}_{\Gamma}^{(i)} \mathbf{R}_{\Gamma}^{(i)}, \qquad \mathbf{g}_{\Gamma} = \sum_{i=1}^{N} {\mathbf{R}_{\Gamma}^{(i)}}^{T} \mathbf{g}_{\Gamma}^{(i)}.$$

3.2. The BDDC preconditioner

We define the partially assembled interface space $\widetilde{\boldsymbol{\Lambda}}_{\Gamma}$ as

$$\widetilde{\boldsymbol{\Lambda}}_{\Gamma} = \widehat{\boldsymbol{\Lambda}}_{\Pi} \bigoplus \boldsymbol{\Lambda}_{\Delta} = \widehat{\boldsymbol{\Lambda}}_{\Pi} \bigoplus \left(\prod_{i=1}^{N} \boldsymbol{\Lambda}_{\Delta}^{(i)}\right), \tag{3.11}$$

where $\widehat{\boldsymbol{\Lambda}}_{\Pi}$ denotes the space of primal variables, which are typically continuous across subdomain interfaces. The space $\boldsymbol{\Lambda}_{\Delta}$, given by the direct sum of the local spaces $\boldsymbol{\Lambda}_{\Delta}^{(i)}$, consists of the remaining interface degrees of freedom, which are typically discontinuous across subdomain interfaces.

Let $\widetilde{\mathbf{R}}_\Gamma$ denote the injection operator from the space $\widehat{\mathbf{\Lambda}}_\Gamma$ into $\widetilde{\mathbf{\Lambda}}_\Gamma$. We also define $\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} = \mathbf{D}\widetilde{\mathbf{R}}_\Gamma$, where $\mathbf{D}$ is a diagonal scaling matrix. The diagonal entries of $\mathbf{D}$ are set to 1 for rows corresponding to primal interface variables, and to $\delta_i^\dagger(x)$ for the others. Here, $\delta_i^\dagger(x)$ is the inverse counting function defined for a subdomain interface node x as:

$$\delta_i^\dagger(x) = \frac{1}{\mathrm{card}(N_x)}, \quad \text{for } x \in \partial\Omega_i \cap \Gamma,$$

where N_x is the set of subdomain indices that contain x on their boundaries. The function $\mathrm{card}(N_x)$ gives the number of such subdomains. The scaling matrix provides a partition of unity such that

$$\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{R}}_\Gamma = \widetilde{\mathbf{R}}_\Gamma^T \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} = \mathrm{I}. \tag{3.12}$$

Various scaling strategies can be used in BDDC algorithms. Among them, deluxe scaling is particularly useful for problems with discontinuous coefficients, as it enhances the robustness of the preconditioner in the presence of parameter jumps in the model [16, 50–52].

Let $\overline{\mathbf{R}}_\Gamma$ be the direct sum of $\mathbf{R}_\Gamma^{(i)}$, which is the restriction operator from $\widetilde{\mathbf{\Lambda}}_\Gamma$ to $\mathbf{\Lambda}^{(i)}$. Let $\mathbf{S}_\Gamma$ be the direct sum of the subdomain local Schur complement $\mathbf{S}_\Gamma^{(i)}$ defined in equation (3.9). The partially assembled interface Schur complement is defined as

$$\widetilde{\mathbf{S}}_\Gamma = \overline{\mathbf{R}}_\Gamma^T \mathbf{S}_\Gamma \overline{\mathbf{R}}_\Gamma. \tag{3.13}$$

With (3.13), we can define BDDC preconditioner as

$$\mathbf{P}_{\mathbf{BDDC}}^{-1} = \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}.$$

Multiplying the BDDC preconditioner to the global interface problem (3.10) both sides, we have

$$\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \widehat{\mathbf{S}}_\Gamma \boldsymbol{\lambda}_\Gamma = \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \mathbf{g}_\Gamma. \tag{3.14}$$

Since $\widehat{\mathbf{S}}_\Gamma$ defined in (3.10) is nonsymmetric but positive definite, we employ the GMRES solver to solve (3.14).

4. Extensions and norms

4.1. Subdomain extensions

Given $\boldsymbol{\lambda}_\Gamma^{(i)} = (\lambda_\Gamma^{(i)}, \mu_\Gamma^{(i)})^T \in \mathbf{\Lambda}_\Gamma^{(i)}$, by solving the subdomain problem (3.8) with $\mathbf{b}_\mathrm{I}^{(i)} = \mathbf{0}$, we obtain the subdomain interior $\boldsymbol{\lambda}_\mathrm{I}^{(i)} \in \Omega_i$ as $\boldsymbol{\lambda}_\mathrm{I}^{(i)} = -{\mathbf{A}_{\mathrm{II}}^{(i)}}^{-1} \mathbf{A}_{\mathrm{I}\Gamma}^{(i)} \boldsymbol{\lambda}_\Gamma^{(i)}$. Let $(\lambda_\mathrm{I}^{(i)}, \mu_\mathrm{I}^{(i)})^T = \boldsymbol{\lambda}_\mathrm{I}^{(i)}$, we denote $\lambda_{A,\Gamma}^{(i)} = (\lambda_\mathrm{I}^{(i)}, \lambda_\Gamma^{(i)})^T, \mu_{A,\Gamma}^{(i)} = (\mu_\mathrm{I}^{(i)}, \mu_\Gamma^{(i)})^T$ and define $\boldsymbol{\lambda}_{\mathbf{A},\Gamma}^{(i)} = (\lambda_{A,\Gamma}^{(i)}, \mu_{A,\Gamma}^{(i)})$. Similarly, given $\boldsymbol{\lambda}_\Gamma \in \widehat{\mathbf{\Lambda}}_\Gamma$ or $\widetilde{\mathbf{\Lambda}}_\Gamma$, we obtain $\boldsymbol{\lambda}_\mathrm{I}$ as $\boldsymbol{\lambda}_\mathrm{I} = -\mathbf{A}_{\mathrm{II}}^{-1} \mathbf{A}_{\mathrm{I}\Gamma} \boldsymbol{\lambda}_\Gamma$ and denote $\lambda_{A,\Gamma} = (\lambda_\mathrm{I}, \lambda_\Gamma)^T, \mu_{A,\Gamma} = (\mu_\mathrm{I}, \mu_\Gamma)^T$, where $(\lambda_\mathrm{I}, \mu_\mathrm{I})^T = \boldsymbol{\lambda}_\mathrm{I}$. Let

$$\boldsymbol{\lambda}_{\mathbf{A},\Gamma} = (\lambda_{A,\Gamma}, \mu_{A,\Gamma})^T \tag{4.1}$$

denote the optimal control extension of $\boldsymbol{\lambda}_\Gamma$. We can also define the advection–diffusion extension of $\boldsymbol{\lambda}_\Gamma$ as

$$\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma} = \left(\lambda_{A_1^{ad},\Gamma}, \mu_{A_2^{ad},\Gamma}\right)^T, \tag{4.2}$$

where $\lambda_{A_1^{ad},\Gamma}$ is the advection–diffusion extension of λ_Γ defined similarly as equation (5.1) of [47] corresponding to A_1^{ad} in (2.13), $\mu_{A_2^{ad},\Gamma}$ can be obtained similarly as $\lambda_{A_1^{ad},\Gamma}$ corresponding to A_2^{ad} in (2.13). Let $\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}^{(i)} = (\lambda_{A_1^{ad},\Gamma}^{(i)}, \mu_{A_2^{ad},\Gamma}^{(i)})^T$, where $\lambda_{A_1^{ad},\Gamma}^{(i)}$, $\mu_{A_2^{ad},\Gamma}^{(i)}$ are the restrictions of $\lambda_{A_1^{ad},\Gamma}$ and $\mu_{A_2^{ad},\Gamma}$ in subdomain Ω_i, respectively.

4.2. A partial assembled finite element space and norms

We define a partially sub-assembled space $\widetilde{\mathbf{\Lambda}}$ as

$$\widetilde{\mathbf{\Lambda}} = \widehat{\mathbf{\Lambda}}_\Pi \bigoplus \left(\prod_{i=1}^{N} \left(\mathbf{\Lambda}_{\mathrm{I}}^{(i)} \bigoplus \mathbf{\Lambda}_{\Delta}^{(i)} \right) \right). \tag{4.3}$$

The functions in the space $\widetilde{\mathbf{\Lambda}}$ consist of a continuous primal component and a generally discontinuous dual component. Moreover, we have $\mathbf{\Lambda} \subset \widetilde{\mathbf{\Lambda}}$.

Let $\widetilde{\mathbf{B}}, \widetilde{\mathbf{Z}}, \widetilde{\mathbf{A}}$ be the partially sub-assembled matrices corresponding to the bilinear forms in (3.1), (3.2) and (3.3), respectively. Recall $\mathbf{L}$ defined in (2.26), let $\widetilde{\mathbf{L}}$ be the partially sub-assembled matrix corresponding to $\mathbf{L}$ and $\widetilde{\mathbf{R}}$ be the injection vector from $\mathbf{\Lambda}$ to $\widetilde{\mathbf{\Lambda}}$. Then, we have $\mathbf{B} = \widetilde{\mathbf{R}}^T \widetilde{\mathbf{B}} \widetilde{\mathbf{R}}, \mathbf{Z} = \widetilde{\mathbf{R}}^T \widetilde{\mathbf{Z}} \widetilde{\mathbf{R}}, \mathbf{A} = \widetilde{\mathbf{R}}^T \widetilde{\mathbf{A}} \widetilde{\mathbf{R}}, \mathbf{L} = \widetilde{\mathbf{R}}^T \widetilde{\mathbf{L}} \widetilde{\mathbf{R}}$, where $\mathbf{B}, \mathbf{Z}, \mathbf{A}$ are defined in (2.24) and (2.23). $\forall \boldsymbol{\lambda}, \boldsymbol{s} \in \widetilde{\mathbf{\Lambda}}$, we define the following bilinear forms

$$\begin{aligned}
&\langle \boldsymbol{\lambda}, \boldsymbol{s} \rangle_{\widetilde{\mathbf{B}}} = \boldsymbol{s}^T \widetilde{\mathbf{B}} \boldsymbol{\lambda} = \sum_{i=1}^{N} \mathbf{b}_h \left(\boldsymbol{\lambda}^{(i)}, \boldsymbol{s}^{(i)} \right), \quad \langle \boldsymbol{\lambda}, \boldsymbol{s} \rangle_{\widetilde{\boldsymbol{Z}}} = \boldsymbol{s}^T \widetilde{\boldsymbol{Z}} \boldsymbol{\lambda} = \sum_{i=1}^{N} \mathbf{z}_h \left(\boldsymbol{\lambda}^{(i)}, \boldsymbol{s}^{(i)} \right), \\
&\langle \boldsymbol{\lambda}, \boldsymbol{s} \rangle_{\widetilde{\mathbf{A}}} = \boldsymbol{s}^T \widetilde{\mathbf{A}} \boldsymbol{\lambda} = \sum_{i=1}^{N} \boldsymbol{a}_h \left(\boldsymbol{\lambda}^{(i)}, \boldsymbol{s}^{(i)} \right), \\
&\langle \boldsymbol{\lambda}, \boldsymbol{s} \rangle_{\widetilde{\mathbf{L}}} = \boldsymbol{s}^T \widetilde{\mathbf{L}} \boldsymbol{\lambda} = (U_{A_1} \boldsymbol{\lambda}, U_{A_1} \boldsymbol{s})_{\mathcal{T}_h} + (U_{A_2} \boldsymbol{\lambda}, U_{A_2} \boldsymbol{s})_{\mathcal{T}_h}.
\end{aligned} \tag{4.4}$$

Given the definition $\mathcal{B}^{(i)}$ in (3.4) in each subdomain, for any $\boldsymbol{\lambda}, \boldsymbol{s} \in \widetilde{\mathbf{\Lambda}}$, we can define the following bilinear form

$$\begin{aligned}
\widetilde{\mathcal{B}}(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}); \mathbf{E}_{\mathcal{C}}(\boldsymbol{s})) &= \sum_{i=1}^{N} \mathcal{B}^{(i)} \left(\mathbf{E}_{\mathcal{C}} \left(\boldsymbol{\lambda}^{(i)} \right); \mathbf{E}_{\mathcal{C}} \left(\boldsymbol{s}^{(i)} \right) \right) = \mathbf{E}_{\mathcal{C}}^T(\boldsymbol{s}) \mathbf{A}_{\mathbf{a}} \mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}) \quad (4.5) \\
&= \mathbf{E}_{\mathcal{C}}^T(\boldsymbol{s}) \begin{bmatrix} \beta^{\frac{1}{2}} A_1^{ad} & -L \\ L & \beta^{\frac{1}{2}} A_2^{ad} \end{bmatrix} \mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}), \\
\widetilde{\mathcal{B}}^{\mathbf{ad}}(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}); \mathbf{E}_{\mathcal{C}}(\boldsymbol{s})) &= \mathbf{E}_{\mathcal{C}}^T(\boldsymbol{s}) \mathbf{A}^{\mathbf{ad}} \mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}), \quad (4.6)
\end{aligned}$$

where $\mathbf{A}_{\mathbf{a}}, A_1^{ad}, A_2^{ad}$ are obtained by assembling the local matrices $\mathbf{A}_{\mathbf{a}}^{(i)}, A_1^{ad(i)}, A_2^{ad(i)}$ in (3.4) on each subdomain respectively. Here, the matrices $\mathbf{A}_{\mathbf{a}}, A_1^{ad}, A_2^{ad}, \mathbf{A}^{\mathbf{ad}}$ have been adjusted by the Robin boundary conditions.

Let

$$\mathbf{M} = \mathbf{B} + \mathbf{L}, \quad \widetilde{\mathbf{M}} = \widetilde{\mathbf{B}} + \widetilde{\mathbf{L}}, \tag{4.7}$$

by using the optimal control extension in (4.1), we can define the following bilinear forms

$$\begin{aligned}
\forall \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \in \widehat{\mathbf{\Lambda}}_\Gamma, \quad &\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\mathbf{B}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \mathbf{B} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\mathbf{Z}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \mathbf{Z} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \quad (4.8) \\
&\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\mathbf{L}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \mathbf{L} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\mathbf{M}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \mathbf{M} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}; \\
\forall \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \in \widetilde{\mathbf{\Lambda}}_\Gamma, \quad &\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\widetilde{\mathbf{B}}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \widetilde{\mathbf{B}} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\widetilde{\mathbf{Z}}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \widetilde{\mathbf{Z}} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \quad (4.9) \\
&\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \widetilde{\mathbf{L}} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\widetilde{\mathbf{M}}_\Gamma} = \boldsymbol{s}_{\mathbf{A},\Gamma}^T \widetilde{\mathbf{M}} \boldsymbol{\lambda}_{\mathbf{A},\Gamma}.
\end{aligned}$$

By (2.25), (2.26), (3.5), (4.4), (4.5), (4.7), (4.8) and (4.9), we obtain the following lemma.

Lemma 4.1. *Given $\boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \in \widetilde{\mathbf{\Lambda}}_\Gamma$, we have the following relations*

$$\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\widetilde{\mathbf{S}}_\Gamma} = \langle \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{s}_{\mathbf{A},\Gamma} \rangle_{\widetilde{\mathbf{A}}} = \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\widetilde{\mathbf{B}}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \rangle_{\widetilde{\mathbf{Z}}_\Gamma}. \tag{4.10}$$

When $\boldsymbol{\lambda}_\Gamma = \boldsymbol{s}_\Gamma$, $\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{Z}}_\Gamma} = 0, \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{S}}_\Gamma} = \langle \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{\lambda}_{\mathbf{A},\Gamma} \rangle_{\widetilde{\mathbf{A}}} = \mathbf{E}_{\mathcal{C}}^T(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}) \mathbf{A}_{\mathbf{a}} \mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}) = \langle \boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{\lambda}_{\mathbf{A},\Gamma} \rangle_{\widetilde{\mathbf{B}}} = \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{B}}_\Gamma}$, and $\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{M}}_\Gamma} = \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{B}}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma} = \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{S}}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma}$. The same results also hold for $\boldsymbol{\lambda}, \boldsymbol{s} \in \widehat{\mathbf{\Lambda}}_\Gamma$ and their corresponding bilinear forms.

4.3. Norms

Recall $\mathbf{A_a}$ defined in (2.13) and $\mathbf{L_a}$ defined in (2.15). Let

$$\mathbf{M_a} = \mathbf{A_a} + \mathbf{L_a}.$$

For any $\boldsymbol{\lambda} = (\lambda, \mu)^T \in \widetilde{\boldsymbol{\Lambda}}$, we can define the following norms

$$\begin{aligned}
\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|^2_{\mathbf{A_a}} &= \mathbf{E}^T_{\mathcal{C}}(\boldsymbol{\lambda})\mathbf{A_a}\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}) && (4.11)\\
&= \beta^{\frac{1}{2}}\|Q_{A_1}\boldsymbol{\lambda}\|^2_{\mathcal{T}_h} + \beta^{\frac{1}{2}}\left\|\left|\tau_1 - \frac{1}{2}\zeta\cdot\mathbf{n}\right|^{\frac{1}{2}}(U_{A_1}\boldsymbol{\lambda} - \lambda)\right\|^2_{\partial\mathcal{T}_h}\\
&\quad + \beta^{\frac{1}{2}}\|Q_{A_2}\boldsymbol{\lambda}\|^2_{\mathcal{T}_h} + \beta^{\frac{1}{2}}\left\|\left|\tau_2 + \frac{1}{2}\zeta\cdot\mathbf{n}\right|^{\frac{1}{2}}(U_{A_2}\boldsymbol{\lambda} - \mu)\right\|^2_{\partial\mathcal{T}_h},\\
\|\mathbf{U}\boldsymbol{\lambda}\|^2_{\mathcal{T}_h} &= \|U_{A_1}\boldsymbol{\lambda}\|^2_{\mathcal{T}_h} + \|U_{A_2}\boldsymbol{\lambda}\|^2_{\mathcal{T}_h} = \mathbf{E}^T_{\mathcal{C}}(\boldsymbol{\lambda})\mathbf{L_a}\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}), && (4.12)\\
\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|^2_{\mathbf{M_a}} &= \mathbf{E}^T_{\mathcal{C}}(\boldsymbol{\lambda})\mathbf{M_a}\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}) = \|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|^2_{\mathbf{A_a}} + \|\mathbf{U}\boldsymbol{\lambda}\|^2_{\mathcal{T}_h}. && (4.13)
\end{aligned}$$

By Lemma 4.1 and (4.13), we have

$$\forall\boldsymbol{\lambda}_\Gamma \in \widetilde{\boldsymbol{\Lambda}}_\Gamma, \|\boldsymbol{\lambda}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} = \|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|^2_{\mathbf{M_a}}, \quad \forall\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma, \|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} = \|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|^2_{\mathbf{M_a}}. \tag{4.14}$$

Additionally, by (4.6), we define the norms associated with $\mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda})$ defined in (2.21) as follows:

$$\begin{aligned}
\|\mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda})\|^2_{\mathbf{A^{ad}}} &= \mathbf{E}^T_{\mathbf{A^{ad}}}(\boldsymbol{\lambda})\mathbf{A^{ad}}\mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda}) && (4.15)\\
&= \left\|Q_{A_1^{ad}}\lambda\right\|^2_{\mathcal{T}_h} + \left\|\left|\tau_1 - \frac{1}{2}\zeta\cdot\mathbf{n}\right|^{\frac{1}{2}}\left(U_{A_1^{ad}}\lambda - \lambda\right)\right\|^2_{\partial\mathcal{T}_h}\\
&\quad + \left\|Q_{A_2^{ad}}\mu\right\|^2_{\mathcal{T}_h} + \beta^{\frac{1}{2}}\left\|\left|\tau_2 + \frac{1}{2}\zeta\cdot\mathbf{n}\right|^{\frac{1}{2}}\left(U_{A_2^{ad}}\mu - \mu\right)\right\|^2_{\partial\mathcal{T}_h},\\
\|\mathbf{U}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda})\|^2_{\mathcal{T}_h} &= \left\|U_{A_1^{ad}}\lambda\right\|^2_{\mathcal{T}_h} + \left\|U_{A_2^{ad}}\mu\right\|^2_{\mathcal{T}_h}. && (4.16)
\end{aligned}$$

Moreover, we introduce several useful norms as defined in [47], for any domain D and $\boldsymbol{\lambda} = (\lambda, \mu)^T \in \widetilde{\boldsymbol{\Lambda}}(D)$, suppose h is the diameter of D, then

$$\|\lambda\|^2_{h,D} := \sum_{K\in\mathcal{T}_h(D)} \|\lambda\|^2_{\partial K}\frac{|D|}{|\partial D|}, \tag{4.17}$$

$$|||\lambda|||^2_D := \sum_{K\in\mathcal{T}_h(D)} \|\lambda - m_K(\lambda)\|^2_{\partial K}\left(\frac{|D|}{|\partial D|}\right)^{-1}, \text{ where } m_K(\lambda) = \frac{1}{|\partial K|}\int_{\partial K}\lambda\,\mathrm{d}s. \tag{4.18}$$

For the domain Ω, we simplify the notation by omitting the subscript. Specifically, we use $\|\cdot\|_h$ and $|||\cdot|||$ to denote $\|\cdot\|_{h,\Omega}$ and $|||\cdot|||_\Omega$ respectively. Note that in this paper, the norm of $\boldsymbol{\lambda} = (\lambda, \mu)^T$, $\|\boldsymbol{\lambda}\|_{h,D}$, $|||\boldsymbol{\lambda}|||_{h,D}$ and $\|\boldsymbol{\lambda}\|_{\partial D}$, are defined as follows:

$$\|\boldsymbol{\lambda}\|_{h,D} := \left(\|\lambda\|^2_{h,D} + \|\mu\|^2_{h,D}\right)^{\frac{1}{2}}, \quad |||\boldsymbol{\lambda}|||_D := \left(|||\lambda|||^2_D + |||\mu|||^2_D\right)^{\frac{1}{2}}, \quad \|\boldsymbol{\lambda}\|_{\partial D} := \left(\|\lambda\|^2_{\partial D} + \|\mu\|^2_{\partial D}\right)^{\frac{1}{2}}.$$

4.4. Some norm estimates

Lemma 4.2 ([47], Lem. 7.8). *If $\mu \in \Lambda^{(i)}$ and μ is edge average zero on $\partial\Omega_i$, then*

$$\|\mu\|_{h,\Omega_i} \leq CH|||\mu|||_{\Omega_i}.$$

Lemma 4.3 ([47], Lem. 7.10). *If $\mu \in \widetilde{\Lambda}$, then*

$$\|\mu\|_h \leq C\,|||\mu|||.$$

From the definitions of $\|\cdot\|_h$ and $|||\cdot|||$, we see that

Lemma 4.4. *If $\mu \in \widetilde{\Lambda}$, then*

$$|||\mu||| \leq Ch^{-1}\|\mu\|_h.$$

For any $\boldsymbol{\lambda} = (\lambda, \mu)^T \in \widetilde{\boldsymbol{\Lambda}}$, by Lemma 7.2 of [47] with $\epsilon = 1$, we have

$$\left\|U_{A_1^{ad}}\lambda\right\|_K \leq C\|\lambda\|_{h,K}. \tag{4.19}$$

Replacing $\boldsymbol{\zeta}$ with $-\boldsymbol{\zeta}$, τ_1 with τ_2 and by similar proofs as Lemma 7.2 of [47], we obtain

$$\left\|U_{A_2^{ad}}\mu\right\|_K \leq C\|\mu\|_{h,K}. \tag{4.20}$$

Also, for $\boldsymbol{\lambda} \in \widetilde{\boldsymbol{\Lambda}}$, by Lemma 7.6 of [47] we have

$$c|||\boldsymbol{\lambda}||| \leq \|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda})\|_{\mathbf{A}^{\mathbf{ad}}} \leq C|||\boldsymbol{\lambda}|||, \tag{4.21}$$

where c and C are two constants such that $c \leq C$.

By Lemmas 7.3, 7.4 of [47] with $\epsilon = 1$, we obtain the following lemma.

Lemma 4.5. *For λ defined on ∂K, we have*

$$\left\|\nabla U_{A_1^{ad}}\lambda\right\|_K \leq C|||\lambda|||_K, \left\|\nabla U_{A_2^{ad}}\lambda\right\|_K \leq C|||\lambda|||_K, \tag{4.22}$$

$$\left\|U_{A_1^{ad}}\lambda - \lambda\right\|_{\partial K} \leq Ch^{\frac{1}{2}}|||\lambda|||_K, \left\|U_{A_2^{ad}}\lambda - \lambda\right\|_{\partial K} \leq Ch^{\frac{1}{2}}|||\lambda|||_K. \tag{4.23}$$

5. The Convergence rate of the BDDC preconditioned system

In this section, we examine the dependence of the GMRES convergence rate on the regularization parameter β, the subdomain size H and the mesh size h. Throughout this section, we use C to denote any constants independent of these parameters. Following A(1)–A(3) of [11] and Assumption 6.1 of [47], we make the assumption below.

Assumption 5.1. *For the stabilizers τ_1 and τ_2, we assume*

(1) *τ_1 is a piecewise positive constant on $\partial\mathcal{T}_h$ and there exists a constant C_1 such that $\tau_1 \leq C_1$;*
(2) *$\tau_1 = \tau_2 + \boldsymbol{\zeta}\cdot\mathbf{n}$;*
(3) *$\max_{\mathcal{E}\subset\partial K} \inf\limits_{x\in\mathcal{E}}(\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}) = C_K, \quad \forall K \in \mathcal{T}_h$;*
(4) *$\inf_{x\in\mathcal{E}}(\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}) \geq C^* \max_{x\in\mathcal{E}} |\boldsymbol{\zeta}(x)\cdot\mathbf{n}|, \forall\mathcal{E} \in \partial K$ and $\forall K \in \mathcal{T}_h$.*

We also make the following assumption as in Assumption 6.2 of [47].

Assumption 5.2. *For each two adjacent subdomains Ω_i and Ω_j, which share the edge $\mathcal{E}_{ij}$, for any $\boldsymbol{\lambda}^{(i)} \in \boldsymbol{\Lambda}^{(i)}$, we assume that*

$$\int_{\mathcal{E}_{ij}} \boldsymbol{\lambda}^{(i)}\, \mathrm{d}s, \quad \int_{\mathcal{E}_{ij}} \boldsymbol{\zeta} \cdot \mathbf{n}\boldsymbol{\lambda}^{(i)}\, \mathrm{d}s, \quad \int_{\mathcal{E}_{ij}} \boldsymbol{\zeta} \cdot \mathbf{n}\boldsymbol{\lambda}^{(i)} s\, \mathrm{d}s,$$

are contained in the coarse level primal subspace $\widehat{\boldsymbol{\Lambda}}_{\Pi}$, which are the same or different by a -1 factor due to the normal direction.

Remark 5.3. The first two constraints are used in Lemma 6.1 to establish estimate (6.1). Similar to Lemma 6.4 of [47], for each subdomain Ω_i $(i = 1, \ldots, N)$, assume $u \in H^2(\Omega_i)$ and let I_H denote the finite element interpolation operator into the continuous linear finite element space $\widehat{W}_H$. Applying the Bramble–Hilbert lemma yields the estimate

$$\|u - \mathrm{I}_H u\|_{H^t(\Omega_i)} \leq C H^{2-t} |u|_{H^2(\Omega_i)}, \quad t = 0, 1, 2. \tag{5.1}$$

This estimate, combined with these three constraints, is used to derive Lemma 6.3, following arguments similar to those in Lemma 7.17 of [47].

Assumption 5.4. *Each subdomain Ω_i is triangular or quadrilateral. These subdomains form a shape regular coarse mesh of Ω.*

Remark 5.5. We make this shape regularity assumption as in [30, 45, 47]. This assumption ensures the existence of an interpolation operator onto a conforming finite element space, thereby enabling the application of continuous trace theory (cf. [3, 17]). Furthermore, it guarantees that the constants C appearing in inequalities such as the trace inequality and the Bramble–Hilbert inequality (5.1) are uniformly bounded and independent of the subdomain size H and the mesh size h. These properties are essential for establishing the condition number bounds of the proposed method.

Let $\mathbf{T} = \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{S}}_{\Gamma}^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \widehat{\mathbf{S}}_{\Gamma}$ denote the BDDC preconditioned operator. The convergence behavior of the GMRES method applied to (3.14) is analyzed using the framework established in Theorem 6.5 of [47], whose proof is detailed in [18]. Based on this result, we derive the following theorem on the convergence rate of the BDDC preconditioned GMRES iterations, which constitutes the main result of this paper.

Theorem 5.6. *Given $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, let C_u and c_l be the constants established in Lemmas 5.16 and 5.17, respectively. Then, the following bounds hold:*

$$c_l \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} \leq \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma},$$
$$\langle \mathbf{T}\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} \leq C_u^2 \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma}.$$

As a result, the residual $\mathbf{r}_m$ of the GMRES iteration at iteration m satisfies

$$\frac{\|\mathbf{r}_m\|_{\mathbf{M}_\Gamma}}{\|\mathbf{r}_0\|_{\mathbf{M}_\Gamma}} \leq \left(1 - \frac{c_l^2}{C_u^2}\right)^{m/2}.$$

We outline the proof of Theorem 5.6 as follows. First, by comparing with the corresponding extensions for the advection–diffusion problem, we establish norm equivalences between the extensions associated with the optimal control problem and those for the advection–diffusion problem in Lemma 5.10. Then, we derive some estimates for the average operator in Lemma 5.12. Next, we bound some bilinear forms in Lemmas 5.13 and 5.15, and derive estimates for some key operators through the extensions of advection–diffusion problem in Lemma 5.14. These estimates are used in the proof of Theorem 5.6 to obtain the upper bound C_u and the lower bound c_l. Finally, Section 6 provides the technical details supporting the proofs in Section 5.

5.1. Average operator estimates

The relationship among the norms $|||\cdot|||$ defined in (4.18), $\|\cdot\|_{\mathbf{A_a}}$ defined in (4.11), and $\|\cdot\|_{\mathbf{M_a}}$ defined (4.13) is established in the following lemma.

Lemma 5.7. *For any $K \in \mathcal{T}_h$ and $\boldsymbol{\lambda} = (\lambda, \mu)^T \in \widetilde{\boldsymbol{\Lambda}}$ we have*

$$C|||\lambda|||_K \le \|Q_{A_1}\boldsymbol{\lambda}\|_K, \quad C|||\mu|||_K \le \|Q_{A_2}\boldsymbol{\lambda}\|_K \tag{5.2}$$

which implies that

$$C\beta^{\frac{1}{4}}|||\boldsymbol{\lambda}||| \le \|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}} \le \|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{M_a}}. \tag{5.3}$$

Proof. Equation (5.2) can be proved similarly as Lemma 3.8 of [13]. (5.3) follows from assembling (5.2) over all elements $K \in \mathcal{T}_h$, with the definitions (4.11) and (4.13). □

Estimates for the L^2-norm of $\mathbf{U}\boldsymbol{\lambda}$ is given in the lemma below, and the proof is given in Section 6.

Lemma 5.8. *For any $\boldsymbol{\lambda} = (\lambda, \mu)^T \in \widetilde{\boldsymbol{\Lambda}}$, we have the following estimates:*

$$\|\mathbf{U}\boldsymbol{\lambda}\|_K \le Cc_0\|\boldsymbol{\lambda}\|_{h,K}, \tag{5.4}$$

$$\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h} \le Cc_0\|\boldsymbol{\lambda}\|_{h}, \tag{5.5}$$

where $c_0 = h\beta^{-\frac{1}{2}} + 1$. Therefore,

$$\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h} \le Cc_0\beta^{-\frac{1}{4}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}}. \tag{5.6}$$

The following lemma provides the gradient estimates and other auxiliary bounds for the extensions of the optimal control problem, and the proof is given in Section 6.

Lemma 5.9. *For $\boldsymbol{\lambda} = (\lambda, \mu)^T \in \widetilde{\boldsymbol{\Lambda}}$, we have*

$$\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_{\mathcal{T}_h} + \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_{\mathcal{T}_h} \le Ch\beta^{-\frac{1}{2}}\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h}, \tag{5.7}$$

$$\beta^{\frac{1}{2}}\|\nabla U_{A_1}\boldsymbol{\lambda}\|_{\mathcal{T}_h} \le C\left(1+\beta^{\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{M_a}}, \quad \beta^{\frac{1}{2}}\|\nabla U_{A_2}\boldsymbol{\lambda}\|_{\mathcal{T}_h} \le C\left(1+\beta^{\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{M_a}}. \tag{5.8}$$

Moreover, we obtain that

$$\begin{aligned}\beta^{\frac{1}{4}}\left\|\left|\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac{1}{2}}(U_{A_1}\boldsymbol{\lambda} - \lambda)\right\|_{\partial\mathcal{T}_h} &+ \beta^{\frac{1}{4}}\left\|\left|\tau_2 + \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac{1}{2}}(U_{A_2}\boldsymbol{\lambda} - \mu)\right\|_{\partial\mathcal{T}_h} \\ &\le Ch^{\frac{1}{2}}\left(1+\beta^{-\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{M_a}}.\end{aligned} \tag{5.9}$$

Lemma 5.10 establishes the relation between $\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})$, as defined in (2.17) and (4.1), and $\mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma})$, as defined in (2.21) and (4.2). The proof of Lemma 5.10 can be found in Section 6.

Lemma 5.10. *Given $\boldsymbol{\lambda}_\Gamma = (\lambda_\Gamma, \mu_\Gamma)^T \in \widetilde{\Lambda}_\Gamma$, we have*

$$\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}} \le C\beta^{\frac{1}{4}}c_0\|\mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma})\|_{\mathbf{A^{ad}}} + C\left(c_0\beta^{-\frac{1}{4}}\right)\|\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}\|_h, \tag{5.10}$$

and

$$\beta^{\frac{1}{4}}\|\mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma})\|_{\mathbf{A^{ad}}} \le C\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}. \tag{5.11}$$

We define two Schur complements, $\widetilde{S}_{1,\Gamma}^{ad}$ and $\widetilde{S}_{2,\Gamma}^{ad}$, for A_1^{ad} and A_2^{ad} in (4.5). $\widetilde{S}_{1,\Gamma}^{ad}$ is defined similarly to equation (4.1) of [47] with $\epsilon = 1$ and $\tau = \tau_1$. $\widetilde{S}_{2,\Gamma}^{ad}$ is defined analogously with $\tau = \tau_2$ and $\boldsymbol{\zeta}$ replaced by $-\boldsymbol{\zeta}$ in $\widetilde{S}_{1,\Gamma}^{ad}$. Let $\widetilde{\mathbf{S}}_{\Gamma}^{\mathbf{ad}} = \begin{bmatrix} \widetilde{S}_{1,\Gamma}^{ad} & \mathbf{0} \\ \mathbf{0} & \widetilde{S}_{2,\Gamma}^{ad} \end{bmatrix}$. Then, we have

$$\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_{\Gamma}^{\mathbf{ad}}} = \|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma})\|_{\mathbf{A}^{\mathbf{ad}}}. \tag{5.12}$$

Let $\mathbf{E}_{\mathbf{D}}\boldsymbol{\lambda}_\Gamma$ be the average operator defined as $\mathbf{E}_{\mathbf{D}}\boldsymbol{\lambda}_\Gamma = \widetilde{\mathbf{R}}_\Gamma \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \boldsymbol{\lambda}_\Gamma$. An estimate for the average operator of the advection–diffusion problem is provided in Lemma 6.9 from [47] with $\epsilon = 1$. This leads to the following lemma:

Lemma 5.11. *For $\boldsymbol{\lambda}_\Gamma \in \widetilde{\boldsymbol{\Lambda}}_\Gamma$, we have*

$$\|\mathbf{E}_{\mathbf{D}}\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_{\Gamma}^{\mathbf{ad}}} \leq C\left(1 + \log\frac{H}{h}\right)\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_{\Gamma}^{\mathbf{ad}}}.$$

We are now ready to provide the estimates for the average operator.

Lemma 5.12. *Given $\boldsymbol{\lambda}_\Gamma \in \widetilde{\boldsymbol{\Lambda}}_\Gamma$, there exists a positive constant C, independent of β, H, h, such that for all $\boldsymbol{\lambda}_\Gamma \in \widetilde{\boldsymbol{\Lambda}}_\Gamma$, we have*

$$\|\mathbf{E}_{\mathbf{D}}\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} \leq CC_{ED}\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma}, \tag{5.13}$$

$$\|\mathbf{E}_{\mathbf{D}}\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} \leq C_{ED,M}\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}, \tag{5.14}$$

where $C_{ED} = c_0(1 + H\beta^{-\frac{1}{2}})\big(1 + \log\frac{H}{h}\big), C_{ED,M} = C(1 + c_0 H\beta^{-\frac{1}{4}})C_{ED}$, and c_0 is defined in Lemma 5.8.

Proof. Let $\mathbf{v}_\Gamma = \mathbf{E}_{\mathbf{D}}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma$, then we have $\|\mathbf{E}_{\mathbf{D}}\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma}^2 \leq C\|\mathbf{v}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma}^2 + C\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma}^2$. Since

$$\begin{aligned}
\|\mathbf{v}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} &= \|\mathbf{E}_{\boldsymbol{\mathcal{C}}}(\mathbf{v}_{\mathbf{A},\Gamma})\|_{\mathbf{A}_\mathbf{a}} \leq Cc_0\beta^{\frac{1}{4}}\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma})\|_{\mathbf{A}^{\mathbf{ad}}} + C\left(c_0\beta^{-\frac{1}{4}}\right)\|\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\|_h \\
&\leq Cc_0\beta^{\frac{1}{4}}\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{v}_{\mathbf{A}^{\mathbf{ad}}})\|_{\mathbf{A}^{\mathbf{ad}}} + C\left(c_0\beta^{-\frac{1}{4}}\right)H|||\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma}||| \\
&\leq C\left(c_0\beta^{\frac{1}{4}} + c_0\beta^{-\frac{1}{4}}H\right)\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma})\|_{\mathbf{A}^{\mathbf{ad}}} \\
&= C\left(c_0\beta^{\frac{1}{4}} + c_0\beta^{-\frac{1}{4}}H\right)\|\mathbf{v}_\Gamma\|_{\widetilde{\mathbf{S}}_{\Gamma}^{\mathbf{ad}}} \leq C\left(c_0\beta^{\frac{1}{4}} + c_0\beta^{-\frac{1}{4}}H\right)\left(1 + \log\frac{H}{h}\right)\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_{\Gamma}^{\mathbf{ad}}} \\
&= C\left(c_0\beta^{\frac{1}{4}} + c_0\beta^{-\frac{1}{4}}H\right)\left(1 + \log\frac{H}{h}\right)\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma})\|_{\mathbf{A}^{\mathbf{ad}}} \\
&\leq C\left(c_0 + c_0\beta^{-\frac{1}{2}}H\right)\left(1 + \log\frac{H}{h}\right)\|\mathbf{E}_{\boldsymbol{\mathcal{C}}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A}_\mathbf{a}} \\
&= C\left(c_0 + c_0\beta^{-\frac{1}{2}}H\right)\left(1 + \log\frac{H}{h}\right)\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} := CC_{ED}\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma},
\end{aligned} \tag{5.15}$$

where $C_{ED} := c_0(1+\beta^{-\frac{1}{2}}H)$. We use Lemma 4.1 for the first equality, (5.10) for the first inequality. The second inequality follows from Lemma 4.2 with the fact that

$$\|\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\|_h = \left(\sum_{i=1}^N \left\|\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma}^{(i)}\right\|_{h,\Omega_i}^2\right)^{\frac{1}{2}} \leq C\left(\sum_{i=1}^N H^2 \left|\left|\left|\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma}^{(i)}\right|\right|\right|_{\Omega_i}^2\right)^{\frac{1}{2}} \leq CH|||\mathbf{v}_{\mathbf{A}^{\mathbf{ad}},\Gamma}|||. \tag{5.16}$$

The third inequality follows from (4.21), the second and third equalities follow from (5.12), the second to last inequality follows from Lemma 5.11, and the last inequality follows from (5.11). The second to last equality follows from Lemma 4.1.

Moreover, we have

$$\begin{aligned}\|\mathbf{U}\mathbf{v}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} &\le Cc_0\|\mathbf{v}_{\mathbf{A},\Gamma}\|_h \le Cc_0H|||\mathbf{v}_{\mathbf{A},\Gamma}||| \le Cc_0H\beta^{-\frac{1}{4}}\|\mathbf{v}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma}\\ &\le Cc_0H\beta^{-\frac{1}{4}}C_{ED}\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma},\end{aligned} \tag{5.17}$$

where (5.5) is used for the first inequality. The second inequality can be obtained similarly as (5.16), the second to last inequality follows from (5.3), and the last inequality follows from (5.15).

Therefore, by Lemma 4.1, (5.15) and (5.17), we obtain that

$$\|\mathbf{v}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} \le C\Big(1 + c_0H\beta^{-\frac{1}{4}}\Big)C_{ED}\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma},$$

and

$$\|\mathbf{E}_\mathbf{D}\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} \le C\|\mathbf{v}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + C\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} \le C\Big(1 + c_0H\beta^{-\frac{1}{4}}\Big)C_{ED}\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}.$$

□

5.2. Upper bound estimate for C_u and lower bound estimate for c_l

We make use of the dual argument together with the norm equivalence between the extensions of the advection–diffusion problem and the extensions of the optimal problem stated in Lemma 5.10, to derive the following estimates for the nonsymmetric bilinear form. The proof of Lemma 5.13 is detailed in Section 6.

Lemma 5.13. *Given $\boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \in \widetilde{\boldsymbol{\Lambda}}_\Gamma$, we have*

$$\langle\boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma\rangle_{\widetilde{\mathbf{Z}}_\Gamma} \le C\gamma_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big(C\Big(H\beta^{-\frac{1}{4}}\Big)\|\boldsymbol{s}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{s}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\|_h\Big), \tag{5.18}$$

where $\gamma_1 = \big(1+\beta^{\frac{1}{4}}\big)c_0$ and c_0 is defined in Lemma 5.8. The same estimate holds for $\boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$ analogously.

Given $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, define $\mathbf{w}_\Gamma = \widetilde{\mathbf{S}}_\Gamma^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma$. Different from the extension $\mathbf{U}_{\mathbf{A}^{\mathbf{ad}}}\boldsymbol{\lambda}$ studied in [47], the extension operator $\mathbf{U}$ of the optimal control problem does not preserve constants, that is, even if $\boldsymbol{\lambda}$ is constant on the mesh element, $\mathbf{U}\boldsymbol{\lambda}$ is not a constant, as seen from (6.4). This leads to the difficulty of estimating the h-norm $\|\mathbf{w}_{\mathbf{A},\Gamma} - \boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_h$ directly using the approach in Lemma 7.20 of [47]. To overcome this difficulty, we will use a dual argument and estimate the h-norm $\|\mathbf{w}_{\mathbf{A}^{\mathbf{ad}},\Gamma} - \boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\|_h$ as in the following lemma, and its proof is presented in Section 6.

Lemma 5.14. *Given $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, let $\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}) = \boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}$, we have*

$$\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h \le C\alpha_1\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\Big), \tag{5.19}$$

$$\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h \le C\alpha_2\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\Big), \tag{5.20}$$

where $\alpha_1 = H\big(\big(1+\beta^{-\frac{1}{2}}\big)c_0^2 + \beta^{-\frac{1}{4}}\big), \alpha_2 = H\big(\big(1+\beta^{-\frac{1}{2}}\big)c_0^2 + \beta^{-\frac{1}{4}}C_{ED}\big)$, c_0 is defined in Lemma 5.8 and C_{ED} is defined in Lemma 5.12.

Lemma 5.15. *For any $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, when H is sufficiently small, let $\mathbf{z}_\Gamma = \widetilde{\mathbf{L}}_\Gamma^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\mathbf{L}_\Gamma\boldsymbol{\lambda}_\Gamma$, we have the following estimates*

$$\Big|\Big\langle\mathbf{w}_\Gamma, \mathbf{z}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\Big\rangle_{\widetilde{\mathbf{L}}_\Gamma}\Big| \le Cc_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{L}_\Gamma}\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}, \tag{5.21}$$

$$|\langle\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{Z}_\Gamma}| \le Cc_2\|\boldsymbol{\lambda}\|_{\mathbf{M}_\Gamma}\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big), \tag{5.22}$$

$$\left\|\mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\right\|_{\widetilde{\mathbf{L}}_\Gamma} \le Cc_3\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}, \tag{5.23}$$

$$\left|\left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma - \mathbf{w}_\Gamma\right\rangle_{\widetilde{\mathbf{Z}}_\Gamma}\right| \le Cc_4\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma}\Big(\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big), \tag{5.24}$$

where $c_1 = c_0 H\beta^{-\frac{1}{4}}C_{ED}, c_2 = \gamma_1\alpha_2, c_3 = c_0 H\big(1+\beta^{-\frac{1}{2}}\big), c_4 = \gamma_1\alpha_1$. α_1, α_2 are defined in Lemma 5.14. γ_1 is defined in Lemma 5.13.

The proof of Lemma 5.15 is provided in Section 6. Recall that $\mathbf{T} = \widetilde{\mathbf{R}}^T_{\mathbf{D},\Gamma}\widetilde{\mathbf{S}}^{-1}_\Gamma\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma$, we proceed to derive the upper bound C_u and lower bound c_l.

Lemma 5.16. *Given $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, when H is sufficient small, we have*

$$\langle \mathbf{T}\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} \le C_u^2\langle\boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma},$$

where $C_u = CC^2_{ED,M}$. $C_{ED,M}$ is defined in Lemma 5.12.

Proof. By (4.9) and (5.23), we have

$$\begin{aligned}\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} &= \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} + \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{L}}_\Gamma} \le \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} + \left\|\mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\right\|^2_{\widetilde{\mathbf{L}}_\Gamma} + \left\|\widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\right\|^2_{\widetilde{\mathbf{L}}_\Gamma}\\ &\le \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} + Cc_3^2\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}\|^2_{\mathbf{M}_\Gamma},\end{aligned}$$

when H is sufficiently small, we can make c_3 small enough and obtain that

$$\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} \le C\Big(\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma}\Big). \tag{5.25}$$

Also, by Lemma 4.1, (5.22) and (5.25), we have

$$\begin{aligned}\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} &= \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{S}}_\Gamma} = \mathbf{w}_\Gamma^T\widetilde{\mathbf{S}}_\Gamma\mathbf{w}_\Gamma^T = \mathbf{w}_\Gamma^T\widetilde{\mathbf{S}}_\Gamma\widetilde{\mathbf{S}}_\Gamma^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma = \left\langle\boldsymbol{\lambda}_\Gamma, \widetilde{\mathbf{R}}^T_{\mathbf{D},\Gamma}\mathbf{w}_\Gamma\right\rangle_{\widehat{\mathbf{S}}_\Gamma}\\ &= \langle\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\widehat{\mathbf{S}}_\Gamma} = \langle\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{B}_\Gamma} + \langle\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{Z}_\Gamma}\\ &\le \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{B}_\Gamma}\|\mathbf{T}\boldsymbol{\lambda}_\Gamma\|_{\mathbf{B}_\Gamma} + Cc_2\|\boldsymbol{\lambda}\|_{\mathbf{M}_\Gamma}\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big)\\ &\le \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\|\mathbf{T}\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma} + Cc_2\|\boldsymbol{\lambda}\|_{\mathbf{M}_\Gamma}\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{B}}_\Gamma} + Cc_2\|\boldsymbol{\lambda}\|^2_{\mathbf{M}_\Gamma}.\end{aligned} \tag{5.26}$$

By Young's inequality, we obtain that

$$\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} \le C\big(\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\|\mathbf{T}\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma} + \big(Cc_2^2 + Cc_2\big)\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma}\big). \tag{5.27}$$

Since

$$\begin{aligned}\langle\mathbf{T}\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} &= \left\langle\widetilde{\mathbf{R}}^T_{\mathbf{D},\Gamma}\widetilde{\mathbf{S}}^{-1}_\Gamma\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma, \widetilde{\mathbf{R}}^T_{\mathbf{D},\Gamma}\widetilde{\mathbf{S}}^{-1}_\Gamma\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma\right\rangle_{\mathbf{M}_\Gamma}\\ &= \langle\mathbf{E}_\mathbf{D}\mathbf{w}_\Gamma, \mathbf{E}_\mathbf{D}\mathbf{w}_\Gamma\rangle_{\widetilde{\mathbf{M}}_\Gamma} = \|\mathbf{E}_\mathbf{D}\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma},\end{aligned}$$

we see that

$$\begin{aligned}\|\mathbf{T}\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} &= \|\mathbf{E}_\mathbf{D}\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} \le CC^2_{ED,M}\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} \le CC^2_{ED,M}\Big(\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma}\Big),\\ &\le CC^2_{ED,M}\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\|\mathbf{T}\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma} + CC^2_{ED,M}\big(c_2^2 + c_2 + 1\big)\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma},\end{aligned}$$

where (5.14) is used for the first inequality, (5.25) for the second inequality, and (5.27) for the last inequality. Consequently, when H is sufficiently small, c_2 can be small enough, therefore applying Young's inequality yields

$$\begin{aligned}\|\mathbf{T}\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} &\le C\big(\big(C^2_{ED,M}\big)^2 + C^2_{ED,M}(c_2^2 + c_2 + 1)\big)\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma}\\ &\le C\big(C^2_{ED,M}\big)^2\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} = C_u^2\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma},\end{aligned} \tag{5.28}$$

where $C_u := CC^2_{ED,M}$. ∎

Lemma 5.17. *For any $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, when H is sufficiently small, we have*

$$c_l \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} \le \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma},$$

where $c_l = C\big(1 - Cc_2 C_{ED,M} - Cc_3 C_{ED,M}^2\big)$. c_2 and c_3 are defined in Lemma 5.15. $C_{ED,M}$ is defined in Lemma 5.16.

Proof. We have

$$\begin{aligned}
\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} &= \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{B}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{L}_\Gamma} = \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widehat{\mathbf{S}}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{L}_\Gamma} \\
&= \left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{S}}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{L}_\Gamma} \\
&= \left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{M}}_\Gamma} - \left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{L}}_\Gamma} + \left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{Z}}_\Gamma} + \left\langle \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{L}}_\Gamma} \\
&= \left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{M}}_\Gamma} - \left\langle \mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{L}}_\Gamma} + \left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma - \mathbf{w}_\Gamma \right\rangle_{\widetilde{\mathbf{Z}}_\Gamma} \\
&\le \frac{1}{2}\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} + \frac{1}{2}\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} - \left\langle \mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{L}}_\Gamma} + \left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma - \mathbf{w}_\Gamma \right\rangle_{\widetilde{\mathbf{Z}}_\Gamma},
\end{aligned}$$

where the first and the second equalities follow from Lemma 4.1. The third equality follows from (3.12) with the fact that

$$\left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{S}}_\Gamma} = \boldsymbol{\lambda}_\Gamma^T \widetilde{\mathbf{R}}_\Gamma^T \widetilde{\mathbf{S}}_\Gamma \mathbf{w}_\Gamma = \boldsymbol{\lambda}_\Gamma^T \widetilde{\mathbf{R}}_\Gamma^T \widetilde{\mathbf{S}}_\Gamma \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \widehat{\mathbf{S}}_\Gamma \boldsymbol{\lambda}_\Gamma = \langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\widehat{\mathbf{S}}_\Gamma}.$$

The last inequality follows from

$$\left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{M}}_\Gamma} \le \|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} \left\|\widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma\right\|_{\widetilde{\mathbf{M}}_\Gamma} \le \frac{1}{2}\Big(\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma}\Big).$$

Therefore, we obtain that

$$\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} \le \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} + C\left|\left\langle \mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{L}}_\Gamma}\right| + C\left|\left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma - \mathbf{w}_\Gamma \right\rangle_{\widetilde{\mathbf{Z}}_\Gamma}\right|. \tag{5.29}$$

In order to estimate $\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}$, we need to estimate $\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}$. Define $\mathbf{z}_\Gamma = \widetilde{\mathbf{L}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \mathbf{L}_\Gamma \boldsymbol{\lambda}_\Gamma$, we have

$$\begin{aligned}
\langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma \rangle_{\widetilde{\mathbf{M}}_\Gamma} &= \langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma \rangle_{\widetilde{\mathbf{B}}_\Gamma} + \langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma} = \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\widehat{\mathbf{S}}_\Gamma} + \langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma} \\
&= \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} - \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{L}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{Z}_\Gamma} + \langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma} \\
&= \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{Z}_\Gamma} - \langle \mathbf{w}_\Gamma, \mathbf{z}_\Gamma - \mathbf{w}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma} \\
&= \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{Z}_\Gamma} - \langle \mathbf{w}_\Gamma, \mathbf{z}_\Gamma - \mathbf{w}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma} \\
&= \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{M}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{Z}_\Gamma} - \left\langle \mathbf{w}_\Gamma, \mathbf{z}_\Gamma - \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{L}}_\Gamma} + \left\langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma \boldsymbol{\lambda}_\Gamma \right\rangle_{\widetilde{\mathbf{L}}_\Gamma},
\end{aligned} \tag{5.30}$$

where the first equality follows from Lemma 4.1, the second equality follows from $\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{B}}_\Gamma} = \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\widehat{\mathbf{S}}_\Gamma}$ which is proved in (5.26), and the third equality follows from the symmetry of the inner product $\langle \cdot, \cdot \rangle_{\mathbf{L}_\Gamma}$ with the fact that

$$\langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma \rangle_{\mathbf{L}_\Gamma} = \boldsymbol{\lambda}_\Gamma^T \mathbf{L}_\Gamma \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \widehat{\mathbf{S}}_\Gamma \boldsymbol{\lambda}_\Gamma = \boldsymbol{\lambda}_\Gamma^T \mathbf{L}_\Gamma \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{L}}_\Gamma^{-1} \widetilde{\mathbf{L}}_\Gamma \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \widehat{\mathbf{S}}_\Gamma \boldsymbol{\lambda}_\Gamma = \langle \mathbf{w}_\Gamma, \mathbf{z}_\Gamma \rangle_{\widetilde{\mathbf{L}}_\Gamma}.$$

Also, when H is sufficiently small, by (5.25), (5.27), (5.28), we obtain

$$\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} \le C C_{ED,M} \|\boldsymbol{\lambda}\|_{\mathbf{M}_\Gamma}, \tag{5.31}$$

where $C_{ED,M}$ is defined in Lemma 5.16. Then, we have

$$\begin{aligned}
\langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma\rangle_{\widetilde{\mathbf{M}}_\Gamma} &\le \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} + \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{Z}_\Gamma} + \left|\left\langle \mathbf{w}_\Gamma, \mathbf{z}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\right\rangle_{\widetilde{\mathbf{L}}_\Gamma}\right| \\
&\quad + \left|\left\langle \mathbf{w}_\Gamma, \mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\right\rangle_{\widetilde{\mathbf{L}}_\Gamma}\right| \\
&\le \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} + Cc_2\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}) + Cc_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} \\
&\quad + Cc_3\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} \\
&\le \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} + Cc_2C_{ED,M}\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} + Cc_3C^2_{ED,M}\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma},
\end{aligned} \tag{5.32}$$

where the first inequality follows from (5.30), the second inequality follows from (5.22), (5.21) and (5.23), the last inequality follows from (5.31), and $c_1 \le c_2$. Therefore,

$$\begin{aligned}
\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} &\le \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} + C\left\|\mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\right\|_{\widetilde{\mathbf{L}}_\Gamma}\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{L}_\Gamma} + C\left|\left\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma - \mathbf{w}_\Gamma\right\rangle_{\widetilde{\mathbf{Z}}_\Gamma}\right| \\
&\le \|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} + Cc_3\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma} + Cc_4\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma}\Big(\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big) \\
&\le (C + Cc_4)\|\mathbf{w}_\Gamma\|^2_{\widetilde{\mathbf{M}}_\Gamma} + C(c_3^2 + c_4^2)\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} \\
&\le (C + Cc_4)\Big(\langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} + Cc_2C_{ED,M}\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} + Cc_3C^2_{ED,M}\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma}\Big) + C\big(c_3^2 + c_4^2\big)\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma} \\
&\le (C + Cc_4)\langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} + (C + Cc_4)\big(c_2C_{ED,M} + c_3C^2_{ED,M}\big)\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma},
\end{aligned}$$

where the first inequality follows from (5.29), the second inequality follows from (5.23) and (5.24), the third inequality follows from Young's inequality, the second to last inequality follows from (5.32), and the last inequality follows from the fact that $c_3 \le C_{ED,M}$, $c_4 \le c_2$, and $C_{ED,M} \ge 1$.

Therefore, when H is sufficiently small, we have c_2, c_3, c_4 small enough such that

$$\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} \le C\langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} + C\big(c_2C_{ED,M} + c_3C^2_{ED,M}\big)\|\boldsymbol{\lambda}_\Gamma\|^2_{\mathbf{M}_\Gamma},$$

and

$$c_l\langle \boldsymbol{\lambda}_\Gamma, \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma} \le \langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{M}_\Gamma},$$

where $c_l := C\big(1 - Cc_2C_{ED,M} - Cc_3C^2_{ED,M}\big)$. $\square$

Remark 5.18. Based on the estimates in Lemmas 5.16 and 5.17, we obtain the following: when $\beta = O(1)$, the bounds C_u, c_l can be simplified as

$$C_u = C\left(1 + \log\frac{H}{h}\right)^2, \quad c_l = C\left(1 - CH\left(1 + \log\frac{H}{h}\right)^2\right),$$

and these estimates are consistent with the results in [47] for the advection–diffusion problem. When β is not too small compared to h, namely $\beta = O(h^{2-\delta})$ for any fixed $\delta \in (0, 2)$, the bounds C_u, c_l can be simplified as

$$C_u = C\left(1 + \log\frac{H}{h}\right)^2, \quad c_l = C\left(1 - CHh^{-1+\frac{\delta}{2}}\left(1 + \log\frac{H}{h}\right)^2\right).$$

In this case, our results are comparable to those established in Theorem 4.2 of [36], which are based on standard finite element discretizations combined with multigrid methods. The $\mathcal{O}(h^{2-\delta})$ threshold can be reflected in Lemma 5.14, which utilizes specific properties of the HDG discretizations and plays a crucial role in establishing the lower bound. In both cases, the convergence rates are independent of β. By selecting appropriate values of H and h, c_l can be ensured to be positive. The convergence rate of the proposed BDDC algorithms is independent of the number of subdomains and depends slightly on subdomain problem size H/h. When β is very small, H is required to be tiny to maintain the positivity of c_l, which might not be practical. Nevertheless, numerical experiments still exhibit good convergence behavior in such case.

6. Proofs of lemmas

In this section, we provide the proofs of Lemmas 5.8, 5.9, 5.10, 5.13, 5.14 and 5.15, along with the supporting lemmas required for these proofs.

Lemma 6.1. *Given $\boldsymbol{\lambda} = (\lambda, \mu)^T, \boldsymbol{s} = (s,t)^T \in \widetilde{\boldsymbol{\Lambda}}$, under Assumption 5.2, we have the following estimates*

$$|\langle \boldsymbol{\zeta}\cdot\mathbf{n}\lambda, s\rangle_{\partial\mathcal{T}_h}| \le CH\beta^{-\frac{1}{2}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s})\|_{\mathbf{A_a}}, \quad |\langle \boldsymbol{\zeta}\cdot\mathbf{n}t, \mu\rangle_{\partial\mathcal{T}_h}| \le CH\beta^{-\frac{1}{2}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s})\|_{\mathbf{A_a}}.$$

Proof. Suppose that Ω_i and Ω_j are two adjacent subdomains sharing the interface edge $\mathcal{E}_{ij}$. Let $\boldsymbol{\lambda}^{(i)} = (\lambda^{(i)}, \mu^{(i)})^T$ and $\boldsymbol{s}^{(i)} = (s^{(i)}, t^{(i)})^T$ denote the restrictions of $\boldsymbol{\lambda}$ and $\boldsymbol{s}$ to subdomain Ω_i, respectively. Define $\overline{\lambda}_{\mathcal{E}_{ij}} = \frac{1}{|\mathcal{E}_{ij}|}\int_{\mathcal{E}_{ij}} \lambda^{(i)}\,\mathrm{d}s$ and $\overline{s}_{\mathcal{E}_{ij}} = \frac{1}{|\mathcal{E}_{ij}|}\int_{\mathcal{E}_{ij}} s^{(i)}\,\mathrm{d}s$, we obtain the following result

$$\begin{aligned}
|\langle \boldsymbol{\zeta}\cdot\mathbf{n}\lambda, s\rangle_{\partial\mathcal{T}_h}| &= \left|\sum_{i=1}^N \sum_{\mathcal{F}\subset\partial\Omega_i}\left\langle \boldsymbol{\zeta}\cdot\mathbf{n}\lambda^{(i)}, s^{(i)}\right\rangle_{\mathcal{F}}\right| = \left|\sum_{i=1}^N \sum_{\mathcal{F}\subset\partial\Omega_i\cap\partial\Omega_j}\left\langle \boldsymbol{\zeta}\cdot\mathbf{n}\lambda^{(i)}, s^{(i)}\right\rangle_{\mathcal{F}}\right| \\
&= \left|\sum_{i=1}^N \sum_{\mathcal{F}\subset\partial\Omega_i\cap\partial\Omega_j}\left\langle \boldsymbol{\zeta}\cdot\mathbf{n}\Big(\lambda^{(i)} - \overline{\lambda}_{\mathcal{E}_{ij}}\Big), \Big(s^{(i)} - \overline{s}_{\mathcal{E}_{ij}}\Big)\right\rangle_{\mathcal{F}}\right| \\
&\le C\sum_{i=1}^N \sum_{\mathcal{F}\subset\partial\Omega_i\cap\partial\Omega_i}\left\|\lambda^{(i)} - \overline{\lambda}_{\mathcal{E}_{ij}}\right\|_{\mathcal{F}}\left\|s^{(i)} - \overline{s}_{\mathcal{E}_{ij}}\right\|_{\mathcal{F}} \le C\sum_{i=1}^N H|\!|\!|\lambda^{(i)}|\!|\!|_{\mathcal{T}_h(\Omega_i)}|\!|\!|s^{(i)}|\!|\!|_{\mathcal{T}_h(\Omega_i)} \\
&\le CH\beta^{-\frac{1}{2}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s})\|_{\mathbf{A_a}},
\end{aligned} \tag{6.1}$$

where the last equality follows from the first two constraints in Assumption 5.2. Specifically, by the first constraint we have $\sum_{i=1}^N \sum_{\mathcal{F}\subset\partial\Omega_i\cap\partial\Omega_j}\langle \boldsymbol{\zeta}\cdot\mathbf{n}\overline{\lambda}_{\mathcal{E}_{ij}}, \overline{s}_{\mathcal{E}_{ij}}\rangle_{\mathcal{F}} = 0$, and the first and the second constraints together yield $\sum_{i=1}^N \sum_{\mathcal{F}\subset\partial\Omega_i\cap\partial\Omega_j}\langle \boldsymbol{\zeta}\cdot\mathbf{n}\overline{\lambda}_{\mathcal{E}_{ij}}, s^{(i)}\rangle_{\mathcal{F}} = \sum_{i=1}^N \sum_{\mathcal{F}\subset\partial\Omega_i\cap\partial\Omega_j}\langle \boldsymbol{\zeta}\cdot\mathbf{n}\lambda^{(i)}, \overline{s}_{\mathcal{E}_{ij}}\rangle_{\mathcal{F}} = 0$. The second inequality follows from equation (7.49) of [47], and the last inequality follows from (5.3). A similar argument holds for $(\mu, t)^T \in \widetilde{\boldsymbol{\Lambda}}$. Therefore, we obtain the desired result. $\square$

6.1. Proof of Lemma 5.8

Proof. We will use a similar method as the proof in Lemma 7.2 from [47]. By equation (2.16), equation (2.16) of [47] with $\epsilon = 1$, we have

$$\begin{aligned}
&(Q_{A_1}\boldsymbol{\lambda}, \mathbf{r})_K - (U_{A_1}\boldsymbol{\lambda}, \nabla\cdot\mathbf{r})_K + \langle\lambda, \mathbf{r}\cdot\mathbf{n}\rangle_{\partial K} = 0, \\
&\beta^{\frac{1}{2}}(\nabla\cdot Q_{A_1}\boldsymbol{\lambda}, w)_K + \beta^{\frac{1}{2}}\left\langle\left(\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\right)(U_{A_1}\boldsymbol{\lambda} - \lambda), w\right\rangle_{\partial K} \\
&\quad - \frac{\beta^{\frac{1}{2}}}{2}(\boldsymbol{\zeta}\cdot\nabla w, U_{A_1}\boldsymbol{\lambda})_K + \frac{\beta^{\frac{1}{2}}}{2}(\boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\lambda}, w)_K + \frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}\lambda, w\rangle_{\partial K} - (U_{A_2}\boldsymbol{\lambda}, w)_K = 0,
\end{aligned} \tag{6.2}$$

and

$$
\begin{aligned}
&\left(Q_{A_1^{ad}}\lambda, \mathbf{r}\right)_K - \left(U_{A_1^{ad}}\lambda, \nabla\cdot\mathbf{r}\right)_K + \langle \lambda, \mathbf{r}\cdot\mathbf{n}\rangle_{\partial K} = 0,\\
&\beta^{\frac12}\left(\nabla\cdot Q_{A_1^{ad}}\lambda, w\right)_K + \beta^{\frac12}\left\langle\left(\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right)\left(U_{A_1^{ad}}\lambda - \lambda\right), w\right\rangle_{\partial K}\\
&\quad - \frac{\beta^{\frac12}}{2}\left(\boldsymbol{\zeta}\cdot\nabla w, U_{A_1^{ad}}\lambda\right)_K + \frac{\beta^{\frac12}}{2}\left(\boldsymbol{\zeta}\cdot\nabla U_{A_1^{ad}}\lambda, w\right)_K + \frac{\beta^{\frac12}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}\lambda, w\rangle_{\partial K} = 0.
\end{aligned}
\tag{6.3}
$$

Subtracting (6.3) from (6.2), we obtain that

$$
\begin{aligned}
&\beta^{\frac12}\left(Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda, \mathbf{r}\right)_K - \beta^{\frac12}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda, \nabla\cdot\mathbf{r}\right)_K = 0\\
&\beta^{\frac12}\left(\nabla\cdot\left(Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda\right), w\right)_K + \beta^{\frac12}\left\langle\left(\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right)\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right), w\right\rangle_{\partial K}\\
&\quad - \frac{\beta^{\frac12}}{2}\left(\boldsymbol{\zeta}\cdot\nabla w, U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)_K + \frac{\beta^{\frac12}}{2}\left(\boldsymbol{\zeta}\cdot\nabla\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right), w\right)_K\\
&\quad - \left(U_{A_2}\boldsymbol{\lambda}, w\right)_K = 0.
\end{aligned}
\tag{6.4}
$$

Let $\mathbf{r} = Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda,\ w = U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda$, then

$$
\beta^{\frac12}\left\|Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda\right\|_K^2 + \beta^{\frac12}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_{\partial K}^2 = \left(U_{A_2}\boldsymbol{\lambda}, U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)_K. \tag{6.5}
$$

Similarly, by (2.16), and equation (2.16) of [47] with $\epsilon = 1$, τ_1 replaced by τ_2 and $\boldsymbol{\zeta}$ replaced by $-\boldsymbol{\zeta}$, we obtain that

$$
\beta^{\frac12}\left\|Q_{A_2}\boldsymbol{\lambda} - Q_{A_2^{ad}}\mu\right\|_K^2 + \beta^{\frac12}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_{\partial K}^2 = -\left(U_{A_1}\boldsymbol{\lambda}, U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)_K. \tag{6.6}
$$

By adding (6.5) and (6.6), we have

$$
\begin{aligned}
&\beta^{\frac12}\left\|Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\boldsymbol{\lambda}\right\|_K^2 + \beta^{\frac12}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_1}\lambda - U_{A_1^{ad}}\lambda\right)\right\|_{\partial K}^2\\
&\quad + \beta^{\frac12}\left\|Q_{A_2}\boldsymbol{\lambda} - Q_{A_2^{ad}}\mu\right\|_K^2 + \beta^{\frac12}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_{\partial K}^2\\
&\quad = \left(U_{A_2}\boldsymbol{\lambda}, -U_{A_1^{ad}}\lambda\right)_K + \left(U_{A_1}\boldsymbol{\lambda}, U_{A_2^{ad}}\mu\right)_K = \left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu, -U_{A_1^{ad}}\lambda\right)_K + \left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda, U_{A_2^{ad}}\mu\right)_K.
\end{aligned}
$$

Therefore, we obtain that

$$
\begin{aligned}
&\beta^{\frac12}\left\|Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda\right\|_K^2 + \beta^{\frac12}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_{\partial K}^2\\
&\quad + \beta^{\frac12}\left\|Q_{A_2}\boldsymbol{\lambda} - Q_{A_2^{ad}}\mu\right\|_K^2 + \beta^{\frac12}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_{\partial K}^2\\
&\le \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K\left\|U_{A_1^{ad}}\lambda\right\|_K + \left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K\left\|U_{A_2^{ad}}\mu\right\|_K\\
&\le \left(\left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K + \left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K\right)\|\boldsymbol{\lambda}\|_{h,K},
\end{aligned}
\tag{6.7}
$$

where the last inequality follows from (4.19) and (4.20).

Next, to estimate $\|\mathbf{U}\boldsymbol{\lambda}\|_K$, we invoke the result from Lemma 3.3 of [13], which states that

$$\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K \leq h\left\|\mathcal{B}^t\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_K + h^{\frac{1}{2}} \min_{\mathcal{E}\subset\partial K}\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_{\mathcal{E}}, \tag{6.8}$$

$$\left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K \leq h\left\|\mathcal{B}^t\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_K + h^{\frac{1}{2}} \min_{\mathcal{E}\subset\partial K}\left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_{\mathcal{E}}. \tag{6.9}$$

Then, by (6.7), together with the identities $\mathcal{B}^t(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\mu) = Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda$, $\mathcal{B}^t(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu) = Q_{A_2}\boldsymbol{\lambda} - Q_{A_2^{ad}}\lambda$, and (3) in Assumption 5.1, it is straightforward to see that

$$\begin{aligned}
&\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K + \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K \\
&\quad\leq Ch\left\|Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda\right\|_K + \frac{C}{C_K}h^{\frac{1}{2}}\left\|\left|\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac{1}{2}}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_{\partial K} \\
&\qquad + Ch\left\|Q_{A_2}\boldsymbol{\lambda} - Q_{A_2^{ad}}\lambda\right\|_K + \frac{C}{C_K}h^{\frac{1}{2}}\left\|\left|\tau_2 + \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac{1}{2}}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\lambda\right)\right\|_{\partial K} \\
&\quad\leq Ch^{\frac{1}{2}}\beta^{-\frac{1}{4}}\sqrt{\left(\left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K + \left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K\right)\|\boldsymbol{\lambda}\|_{h,K}},
\end{aligned}$$

which implies that

$$\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K + \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K \leq Ch\beta^{-\frac{1}{2}}\|\boldsymbol{\lambda}\|_{h,K}, \tag{6.10}$$

and

$$\begin{aligned}
\|\mathbf{U}\boldsymbol{\lambda}\|_K &\leq C(\|U_{A_1}\boldsymbol{\lambda}\|_K + \|U_{A_2}\boldsymbol{\lambda}\|_K) \\
&\leq C\left(h\beta^{-\frac{1}{2}}\|\boldsymbol{\lambda}\|_{h,K} + \left\|U_{A_1^{ad}}\lambda\right\|_K + \left\|U_{A_2^{ad}}\mu\right\|_K\right) \\
&\leq C\left(h\beta^{-\frac{1}{2}} + 1\right)\|\boldsymbol{\lambda}\|_{h,K} = Cc_0\|\boldsymbol{\lambda}\|_{h,K},
\end{aligned}$$

where $c_0 := h\beta^{-\frac{1}{2}} + 1$. The last inequality follows from (4.19) and (4.20). Therefore, we see that

$$\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h} = \sqrt{\sum_{K\in\mathcal{T}_h}\|\mathbf{U}\boldsymbol{\lambda}\|_K^2} \leq Cc_0\|\boldsymbol{\lambda}\|_h \leq Cc_0|\!|\!|\boldsymbol{\lambda}|\!|\!| \leq Cc_0\beta^{-\frac{1}{4}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}},$$

where the second inequality follows from Lemma 4.3, the last inequality follows from (5.3). ☐

6.2. Proof of Lemma 5.9

Proof. By (6.5) and (6.6) we have

$$\begin{aligned}
&\beta^{\frac{1}{2}}\left\|Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda\right\|_K^2 + \beta^{\frac{1}{2}}\left\|\left|\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac{1}{2}}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_{\partial K}^2 \\
&\quad + \beta^{\frac{1}{2}}\left\|Q_{A_2}\boldsymbol{\lambda} - Q_{A_2^{ad}}\mu\right\|_K^2 + \beta^{\frac{1}{2}}\left\|\left|\tau_2 + \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac{1}{2}}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_{\partial K}^2 \\
&\quad\leq \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K\|U_{A_1}\boldsymbol{\lambda}\|_K + \left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K\|U_{A_2}\boldsymbol{\lambda}\|_K \\
&\quad\leq \left(\left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K + \left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K\right)\|\mathbf{U}\boldsymbol{\lambda}\|_K.
\end{aligned} \tag{6.11}$$

By (6.8), (6.9) and (6.11), we have

$$
\begin{aligned}
&\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K + \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K \\
&\quad\le Ch\left\|Q_{A_1}\boldsymbol{\lambda} - Q_{A_1^{ad}}\lambda\right\|_K + \frac{C}{\widetilde{C}_K}h^{\frac12}\||\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}|^{\frac12}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\|_{\partial K} \\
&\qquad + Ch\left\|Q_{A_2}\boldsymbol{\lambda} - Q_{A_2^{ad}}\lambda\right\|_K + \frac{C}{\widetilde{C}_K}h^{\frac12}\||\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}|^{\frac12}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\lambda\right)\|_{\partial K} \\
&\quad\le Ch^{\frac12}\beta^{-\frac14}\sqrt{\left(\left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K + \left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K\right)\|\mathbf{U}\boldsymbol{\lambda}\|_K},
\end{aligned}
$$

which implies that

$$
\begin{aligned}
\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_K + \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_K &\le Ch\beta^{-\frac12}\|\mathbf{U}\boldsymbol{\lambda}\|_K, \\
\beta^{\frac14}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_{\partial K} + \beta^{\frac14}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_{\partial K} &\le Ch^{\frac12}\beta^{-\frac14}\|\mathbf{U}\boldsymbol{\lambda}\|_K.
\end{aligned}
$$

Therefore, we obtain

$$
\left\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right\|_{\mathcal{T}_h} + \left\|U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right\|_{\mathcal{T}_h} \le Ch\beta^{-\frac12}\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h}, \tag{6.12}
$$

$$
\begin{aligned}
&\beta^{\frac14}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_{\partial\mathcal{T}_h} + \beta^{\frac14}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_{\partial\mathcal{T}_h} \\
&\quad\le h^{\frac12}\beta^{-\frac14}\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h}.
\end{aligned} \tag{6.13}
$$

Moreover, we obtain that

$$
\begin{aligned}
\beta^{\frac12}\|\nabla U_{A_1}\boldsymbol{\lambda}\|_{\mathcal{T}_h} &\le \beta^{\frac12}\|\nabla U_{A_1}\boldsymbol{\lambda} - \nabla U_{A_1^{ad}}\lambda\|_{\mathcal{T}_h} + \beta^{\frac12}\|\nabla U_{A_1^{ad}}\lambda\|_{\mathcal{T}_h} \\
&\le \beta^{\frac12}h^{-1}\|U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\|_{\mathcal{T}_h} + \beta^{\frac12}\|\!|\lambda|\!\|_K \\
&\le \beta^{\frac12}h^{-1}\left(h\beta^{-\frac12}\right)\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h} + \beta^{\frac14}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}} \\
&\le C\left(1+\beta^{\frac14}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{M_a}},
\end{aligned} \tag{6.14}
$$

where the second inequality follows from an inverse inequality and (4.22), the second to last inequality follows from (6.12) and (5.3), and the last inequality follows from the definition of $\|\cdot\|_{\mathbf{M_a}}$ in (4.13). Moreover, we have

$$
\begin{aligned}
&\beta^{\frac14}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}(U_{A_1}\boldsymbol{\lambda} - \lambda)\right\|_{\partial\mathcal{T}_h} + \beta^{\frac14}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}(U_{A_2}\boldsymbol{\lambda} - \mu)\right\|_{\partial\mathcal{T}_h} \\
&\quad\le \beta^{\frac14}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_1}\boldsymbol{\lambda} - U_{A_1^{ad}}\lambda\right)\right\|_{\partial\mathcal{T}_h} + \beta^{\frac14}\left\|\left|\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_1^{ad}}\lambda - \lambda\right)\right\|_{\partial\mathcal{T}_h} \\
&\qquad + \beta^{\frac14}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_2}\boldsymbol{\lambda} - U_{A_2^{ad}}\mu\right)\right\|_{\partial\mathcal{T}_h} + \beta^{\frac14}\left\|\left|\tau_2 + \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\right|^{\frac12}\left(U_{A_2^{ad}}\mu - \mu\right)\right\|_{\partial\mathcal{T}_h} \\
&\quad\le h^{\frac12}\beta^{-\frac14}\|\mathbf{U}\boldsymbol{\lambda}\|_{\mathcal{T}_h} + \beta^{\frac14}h^{\frac12}\|\!|\boldsymbol{\lambda}|\!\| \le h^{\frac12}\left(\beta^{-\frac14}+1\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{M_a}},
\end{aligned}
$$

where the second to last inequality follows from (6.13) and (4.23), and the last inequality follows from (4.13) and (5.3). $\square$

6.3. Proof of Lemma 5.10

Proof. To prove (5.10), notice that $\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}) - \mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma})$ has $\mathbf{0}$ value on $\widetilde{\Lambda}_\Gamma$, therefore

$$\big(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}) - \mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma})\big)^T \mathbf{A_a}\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}) = 0. \tag{6.15}$$

This implies

$$\begin{aligned}
\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}^2 &= \mathbf{E}^T_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma})\mathbf{A_a}\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}) = \mathbf{E}^T_{\mathbf{A}^{\mathbf{ad}}}(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma})\begin{bmatrix}\beta^{\frac12}A_1^{ad} & -L\\ L & \beta^{\frac12}A_2^{ad}\end{bmatrix}\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\\
&= \Big(Q_{A_1}\lambda_{A_1^{ad},\Gamma}, U_{A_1}\lambda_{A_1^{ad},\Gamma}, \lambda_{A_1^{ad},\Gamma}\Big)\big[\beta^{\frac12}A_1^{ad} \quad -L\big]\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\\
&\quad + \Big(Q_{A_2}\mu_{A_2^{ad},\Gamma}, U_{A_2}\mu_{A_2^{ad},\Gamma}, \mu_{A_2^{ad},\Gamma}\Big)\big[L \quad \beta^{\frac12}A_2^{ad}\big]\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}) := \mathrm{I} + \mathrm{II}.
\end{aligned}$$

We will estimate I and II can be estimated similarly. Since

$$\begin{aligned}
\mathrm{I} &= \beta^{\frac12}\Big(Q_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, Q_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big)_{\mathcal{T}_h} + \beta^{\frac12}\Big\langle\Big(\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\Big)(U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma} - \lambda_{A,\Gamma}), U_{A_1}\lambda_{A_1^{ad},\Gamma} - \lambda_{A_1^{ad},\Gamma}\Big\rangle_{\partial\mathcal{T}_h}\\
&\quad - \frac{\beta^{\frac12}}{2}\Big(U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big)_{\mathcal{T}_h} + \frac{\beta^{\frac12}}{2}\Big(U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\Big)_{\mathcal{T}_h}\\
&\quad + \frac{\beta^{\frac12}}{2}\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\lambda_{A,\Gamma}, U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big\rangle_{\partial\mathcal{T}_h} - \frac{\beta^{\frac12}}{2}\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\Big(\lambda_{A_1^{ad},\Gamma}\Big), U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\Big\rangle_{\partial\mathcal{T}_h} - \Big(U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big)_{\mathcal{T}_h}\\
&= \beta^{\frac12}\Big(Q_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, Q_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big)_{\mathcal{T}_h} + \beta^{\frac12}\Big\langle\Big(\tau_1 - \frac12\boldsymbol{\zeta}\cdot\mathbf{n}\Big)(U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma} - \lambda_{A,\Gamma}), U_{A_1}\lambda_{A_1^{ad},\Gamma} - \lambda_{A_1^{ad},\Gamma}\Big\rangle_{\partial\mathcal{T}_h}\\
&\quad - \beta^{\frac12}\Big(U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big)_{\mathcal{T}_h} + \frac{\beta^{\frac12}}{2}\Big\langle(\boldsymbol{\zeta}\cdot\mathbf{n})\Big(U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma} - \lambda_{A_1^{ad},\Gamma}\Big), U_{A_1}\boldsymbol{\lambda}_{A,\Gamma}\Big\rangle_{\partial\mathcal{T}_h}\\
&\quad + \frac{\beta^{\frac12}}{2}\Big\langle\boldsymbol{\zeta}\cdot\mathbf{n}\lambda_{A,\Gamma}, U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma} - \lambda_{A_1^{ad},\Gamma}\Big\rangle_{\partial\mathcal{T}_h} + \frac{\beta^{\frac12}}{2}\Big\langle(\boldsymbol{\zeta}\cdot\mathbf{n})\lambda_{A,\Gamma}, \lambda_{A_1^{ad},\Gamma}\Big\rangle_{\partial\mathcal{T}_h} - \Big(U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big)_{\mathcal{T}_h},
\end{aligned}$$

where the second equality follows from the divergence theorem with $\nabla\cdot\boldsymbol{\zeta} = 0$ such that

$$\Big(U_{A_1^{\mathrm{ad}}}\lambda_{A_1^{\mathrm{ad}},\Gamma},\, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\Big)_{\mathcal{T}_h} = \Big\langle(\boldsymbol{\zeta}\cdot\mathbf{n})U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma},\, U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\Big\rangle_{\partial\mathcal{T}_h} - \Big(\boldsymbol{\zeta}\cdot\nabla U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}, U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\Big)_{\mathcal{T}_h}.$$

Then, we have

$$\begin{aligned}
|\mathrm{I}| &\le \beta^{\frac14}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\big\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}\big(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big)\big\|_{\mathbf{A}^{\mathbf{ad}}} + \beta^{\frac12}\Big\|\nabla U_{A_1^{ad}}\lambda_{A_1^{ad},\Gamma}\Big\|_{\mathcal{T}_h}\|U_{A_1}\boldsymbol{\lambda}_{A_1,\Gamma}\|_{\mathcal{T}_h}\\
&\quad + C\beta^{\frac12}\Big\||\boldsymbol{\zeta}\cdot\mathbf{n}|^{\frac12}\Big(\lambda_{A_1^{ad},\Gamma} - U_{A_1}\lambda_{A_1^{ad},\Gamma}\Big)\Big\|_{\partial\mathcal{T}_h}\Big(\|U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_{\partial\mathcal{T}_h} + \|(\boldsymbol{\zeta}\cdot\mathbf{n})\lambda_{A,\Gamma}\|_{\partial\mathcal{T}_h}\Big)\\
&\quad + \left|\frac{\beta^{\frac12}}{2}\Big\langle(\boldsymbol{\zeta}\cdot\mathbf{n})\lambda_{A,\Gamma}, \lambda_{A_1^{ad},\Gamma}\Big\rangle_{\partial\mathcal{T}_h}\right| + \|U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h}\|U_{A_1}\lambda_{A_1^{ad},\Gamma}\|_{\mathcal{T}_h}\\
&\le C\beta^{\frac14}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\big\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}\big(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big)\big\|_{\mathbf{A}^{\mathbf{ad}}} + C\beta^{\frac12}\big\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}\big(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big)\big\|_{\mathbf{A}^{\mathbf{ad}}}\|U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h}\\
&\quad + C\big\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}\big(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big)\big\|_{\mathbf{A}^{\mathbf{ad}}}\beta^{\frac12}\Big(\|U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} + h^{\frac12}\|\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_{\partial\mathcal{T}_h}\Big)\\
&\quad + \beta^{\frac14}H\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\big\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}\big(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big)\big\|_{\mathbf{A}^{\mathbf{ad}}} + \|U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h}\big\|\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big\|_h\\
&\le C\Big(c_0\beta^{\frac14}\Big)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\big\|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}\big(\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big)\big\|_{\mathbf{A}^{\mathbf{ad}}} + C\Big(c_0\beta^{-\frac14}\Big)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\big\|\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\big\|_h,
\end{aligned}$$

where the second inequality follows from (4.22) with (4.21), (4.23) with a trace inequality, (4.19), and the fact that

$$\left|\frac{\beta^{\frac{1}{2}}}{2}\left\langle(\boldsymbol{\zeta}\cdot\mathbf{n})\lambda_{A,\Gamma},\lambda_{A_1^{ad},\Gamma}\right\rangle_{\partial\mathcal{T}_h}\right|\le C\beta^{\frac{1}{4}}H\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\left\|\mathbf{E}_{\mathbf{A^{ad}}}\left(\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}\right)\right\|_{\mathbf{A^{ad}}}, \tag{6.16}$$

where the inequality can be obtained following a similar process as (6.1) with Assumption 5.2, equation (7.49) from [47], (5.3) and (4.21). The last inequality follows from (5.6), Lemma 4.3 and (5.3).

To prove (5.11), we first consider the associated elliptic problem. Given $\boldsymbol{\lambda}=(\lambda,\mu)\in\widetilde{\boldsymbol{\Lambda}}$, we define $E_{A^e}(\lambda)=(Q_{A^e}\lambda,U_{A^e}\lambda,\lambda)^T, E_{A^e}(\mu)=(Q_{A^e}\mu,U_{A^e}\mu,\mu)^T$. Here, $(Q_{A^e}\lambda,U_{A^e}\lambda)$ and $(Q_{A^e}\mu,U_{A^e}\mu)$ satisfy

$$A^e\begin{bmatrix}Q_{A^e}\lambda\\U_{A^e}\lambda\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix},\quad A^e\begin{bmatrix}Q_{A^e}\mu\\U_{A^e}\mu\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix},$$

where A^e is the matrix defined as equation (2.5) of [46] with $\tau_k=\tau_1$. We then define

$$\mathbf{E}_{\mathbf{A^e}}(\boldsymbol{\lambda})=(Q_{A^e}\lambda,U_{A^e}\lambda,\lambda,Q_{A^e}\mu,U_{A^e}\mu,\mu)^T$$

and the corresponding norm

$$\|\mathbf{E}_{\mathbf{A^e}}(\boldsymbol{\lambda})\|^2_{\mathbf{A^e}}=\|E_{A^e}(\lambda)\|^2_{A^e}+\|E_{A^e}(\mu)\|^2_{A^e}.$$

Given $\boldsymbol{\lambda}_\Gamma=(\lambda_\Gamma,s_\Gamma)^T\in\widetilde{\boldsymbol{\Lambda}}_\Gamma$, let $\lambda_{A^e,\Gamma}$ be the harmonic extension of λ_Γ. For each subdomain Ω_i, let $\lambda^{(i)}_{A^e,\Gamma}$ denote the restriction of $\lambda_{A^e,\Gamma}$ to Ω_i, which satisfies

$$\left\|E_{A^e}\left(\lambda^{(i)}_{A^e,\Gamma}\right)\right\|^2_{A^{e(i)}}=\min_{\mu^{(i)}\in\Lambda^{(i)},\mu^{(i)}=\mu^{(i)}_\Gamma\text{ on }\partial\Omega_i}\left\|E_{A^e}\left(\mu^{(i)}\right)\right\|^2_{A^{e(i)}}. \tag{6.17}$$

Similarly, let $\mu_{A^e,\Gamma}$ be the harmonic extension of μ_Γ and set $\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}=(\lambda_{A^e,\Gamma},\mu_{A^e,\Gamma})^T$. Following Lemma 5.1 of [46] and (1) in Assumption 5.1, we have

$$c|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}|\!|\!|\le\|\mathbf{E}_{\mathbf{A^e}}(\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma})\|_{\mathbf{A^e}}\le C|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}|\!|\!|, \tag{6.18}$$

where c and C are two constants such that $c\le C$.

We have

$$\begin{aligned}|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}|\!|\!|^2&\le C\left\|\mathbf{E}_{\mathbf{A^{ad}}}\left(\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}\right)\right\|^2_{\mathbf{A^{ad}}}\\&=\mathbf{E}^T_{\mathbf{A^{ad}}}\left(\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}\right)\mathbf{A^{ad}}\left(\mathbf{E}_{\mathbf{A^{ad}}}\left(\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}\right)\right)\\&=\mathbf{E}^T_{\mathbf{A^{ad}}}\left(\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}\right)\mathbf{A^{ad}}(\mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}))\\&\le C|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}|\!|\!||\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}|\!|\!|,\end{aligned}$$

where the first follows from (4.21), the second equality follows from the fact that the difference $\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}$ vanishes on the subdomain boundary, and the last inequality follows from Lemma 7.9 of [47] with $\epsilon=1$ and (4.21). Therefore, we obtain

$$|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}|\!|\!|\le C|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}|\!|\!|. \tag{6.19}$$

Then, we have

$$\begin{aligned}\beta^{\frac{1}{4}}\left\|\mathbf{E}_{\mathbf{A^{ad}}}\left(\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}\right)\right\|_{\mathbf{A^{ad}}}&\le C\beta^{\frac{1}{4}}|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma}|\!|\!|\le C\beta^{\frac{1}{4}}|\!|\!|\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma}|\!|\!|\le C\beta^{\frac{1}{4}}\|\mathbf{E}_{\mathbf{A^e}}(\boldsymbol{\lambda}_{\mathbf{A^e},\Gamma})\|_{\mathbf{A^e}}\\&\le C\beta^{\frac{1}{4}}\|\mathbf{E}_{\mathbf{A^e}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A^e}}\le C\beta^{\frac{1}{4}}|\!|\!|\boldsymbol{\lambda}_{\mathbf{A},\Gamma}|\!|\!|\le C\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}},\end{aligned}$$

where the first inequality follows from (4.21), the second inequality follows from (6.19), the third inequality follows from (6.18), the third to last inequality follows from (6.17), the second to last inequality follows from (6.18), and the last inequality follows from (5.3). $\square$

Lemma 6.2. *For any* $\boldsymbol{\lambda}_\Gamma = (\lambda_\Gamma, \mu_\Gamma)^T \in \widetilde{\boldsymbol{\Lambda}}_\Gamma$, *we have*

$$\|\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\|_h \le C\Big(H\beta^{-\frac{1}{4}}\Big)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}} + \left\|\boldsymbol{\lambda}_{\mathbf{A}^{\mathbf{ad}},\Gamma}\right\|_h.$$

Proof. We have

$$\begin{aligned}
\|\lambda_{A,\Gamma}\|_h &= \Big(h^{\frac{1}{2}}\beta^{-\frac{1}{4}}\Big)\beta^{\frac{1}{4}}\|\lambda_{A,\Gamma}\|_{\partial\mathcal{T}_h} \le \Big(h^{\frac{1}{2}}\beta^{-\frac{1}{4}}\Big)\Big(\beta^{\frac{1}{4}}\left\|\lambda_{A,\Gamma}-\lambda_{A_1^{ad},\Gamma}\right\|_{\partial\mathcal{T}_h} + \beta^{\frac{1}{4}}\left\|\lambda_{A_1^{ad},\Gamma}\right\|_{\partial\mathcal{T}_h}\Big)\\
&\le \Big(h^{\frac{1}{2}}\beta^{-\frac{1}{4}}\Big)\Big(\beta^{\frac{1}{4}}h^{-\frac{1}{2}}\left\|\lambda_{A,\Gamma}-\lambda_{A_1^{ad},\Gamma}\right\|_h + \beta^{\frac{1}{4}}\left\|\lambda_{A_1^{ad},\Gamma}\right\|_{\partial\mathcal{T}_h}\Big)\\
&\le \Big(h^{\frac{1}{2}}\beta^{-\frac{1}{4}}\Big)\Big(\beta^{\frac{1}{4}}h^{-\frac{1}{2}}CH|\!|\!|\lambda_{A,\Gamma}-\lambda_{A_1^{ad},\Gamma}|\!|\!| + \beta^{\frac{1}{4}}h^{-\frac{1}{2}}\left\|\lambda_{A_1^{ad},\Gamma}\right\|_h\Big)\\
&\le CH|\!|\!|\lambda_{A,\Gamma}-\lambda_{A_1^{ad},\Gamma}|\!|\!| + \left\|\lambda_{A_1^{ad},\Gamma}\right\|_h\\
&\le C\Big(H\beta^{-\frac{1}{4}}\Big)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}} + \left\|\lambda_{A_1^{ad},\Gamma}\right\|_h,
\end{aligned}$$

where the third inequality follows from Lemma 4.2 due to the fact that $\lambda_{A,\Gamma} - \lambda_{A_1^{ad},\Gamma}$ has $\mathbf{0}$ value on the subdomain boundary, and the last inequality follows from a triangle inequality, (5.3), (4.21), and (5.11). A similar argument yields

$$\|\mu_{A,\Gamma}\|_h \le C\Big(H\beta^{-\frac{1}{4}}\Big)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}} + \left\|\mu_{A_2^{\mathrm{ad}},\Gamma}\right\|_h.$$

This completes the proof. ∎

6.4. Proof of Lemma 5.13

Proof. By (4.4) and (4.9), we have

$$\begin{aligned}
\langle\boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma\rangle_{\widetilde{\mathbf{Z}}_\Gamma} = \Bigg(&-\frac{\beta^{\frac{1}{2}}}{2}(U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma})_{\mathcal{T}_h} + \frac{\beta^{\frac{1}{2}}}{2}(U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma})_{\mathcal{T}_h}\\
&+\frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}\lambda_{A,\Gamma}, U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\rangle_{\partial\mathcal{T}_h} - \frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}s_{A,\Gamma}, U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\rangle_{\partial\mathcal{T}_h} - (U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}, U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma})_{\mathcal{T}_h}\Bigg)\\
&+\Bigg(\frac{\beta^{\frac{1}{2}}}{2}(U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_2}\boldsymbol{s}_{\mathbf{A},\Gamma})_{\mathcal{T}_h} - \frac{\beta^{\frac{1}{2}}}{2}(U_{A_2}\boldsymbol{s}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma})_{\mathcal{T}_h}\\
&-\frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}\mu_{A,\Gamma}, U_{A_2}\boldsymbol{s}_{\mathbf{A},\Gamma}\rangle_{\partial\mathcal{T}_h} + \frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}t_{A,\Gamma}, U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\rangle_{\partial\mathcal{T}_h} + (U_{A_2}\boldsymbol{s}_{\mathbf{A},\Gamma}, U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma})_{\mathcal{T}_h}\Bigg) = \mathrm{I} + \mathrm{II}.
\end{aligned}$$

We will consider the estimate of I and II can be estimated similarly. By the divergence theorem with $\nabla\cdot\boldsymbol{\zeta} = 0$, we have

$$\langle(\boldsymbol{\zeta}\cdot\mathbf{n})U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\rangle_{\partial\mathcal{T}_h} = (\boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma})_{\mathcal{T}_h} + (U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma})_{\mathcal{T}_h},$$

then, we obtain that

$$\begin{aligned}
\mathrm{I} = &-\frac{\beta^{\frac{1}{2}}}{2}\langle(\boldsymbol{\zeta}\cdot\mathbf{n})(U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma} - \lambda_{A,\Gamma}), U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\rangle_{\partial\mathcal{T}_h} + \beta^{\frac{1}{2}}(U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma})_{\mathcal{T}_h}\\
&-\frac{\beta^{\frac{1}{2}}}{2}\langle(\boldsymbol{\zeta}\cdot\mathbf{n})s_{A,\Gamma}, U_{A_1}\boldsymbol{\lambda}_{\mathbf{A},\Gamma} - \lambda_{A,\Gamma}\rangle_{\partial\mathcal{T}_h} - (U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}, U_{A_2}\boldsymbol{\lambda}_{\mathbf{A},\Gamma})_{\mathcal{T}_h} - \frac{\beta^{\frac{1}{2}}}{2}\langle(\boldsymbol{\zeta}\cdot\mathbf{n})s_{A,\Gamma}, \lambda_{A,\Gamma}\rangle_{\partial\mathcal{T}_h}.
\end{aligned} \tag{6.20}$$

Therefore, we have

$$\begin{aligned}
|\mathrm{I}| &\le Ch^{\frac{1}{2}}\left(1+\beta^{-\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\left(\beta^{\frac{1}{4}}h^{-\frac{1}{2}}\right)\|U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} + C\left(1+\beta^{\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\|U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} \\
&\quad + Ch^{\frac{1}{2}}\left(1+\beta^{-\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\left(\beta^{\frac{1}{4}}\|\boldsymbol{s}_{\mathbf{A},\Gamma}\|_{\partial\mathcal{T}_h}\right) + \|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\|U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} \\
&\quad + CH\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}} \\
&\le C\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\left(\left(1+\beta^{\frac{1}{4}}\right)\|\boldsymbol{s}_{\mathbf{A},\Gamma}\|_h + \left(1+\beta^{\frac{1}{4}}\right)\|U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} + H\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\right) \\
&\le C\left(1+\beta^{\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\left(\|\boldsymbol{s}_{\mathbf{A},\Gamma}\|_h + \|U_{A_1}\boldsymbol{s}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} + H\beta^{-\frac{1}{4}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\right) \\
&\le C\left(1+\beta^{\frac{1}{4}}\right)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\left(Cc_0\|\boldsymbol{s}_{\mathbf{A},\Gamma}\|_h + H\beta^{-\frac{1}{4}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\right) \\
&\le C\left(1+\beta^{\frac{1}{4}}\right)c_0\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\left(CH\beta^{-\frac{1}{4}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{s}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}} + \|\boldsymbol{s}_{\mathbf{A^{ad}},\Gamma}\|_h\right) \\
&= C\left(1+\beta^{\frac{1}{4}}\right)c_0\|\boldsymbol{\lambda}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}\left(C\left(H\beta^{-\frac{1}{4}}\right)\|\boldsymbol{s}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{s}_{\mathbf{A^{ad}},\Gamma}\|_h\right),
\end{aligned} \tag{6.21}$$

where the first inequality follows from (5.9), a trace inequality, (5.8), (1) and (4) in Assumption 5.1, and Lemma 6.1, the second to last inequality follows from (5.5), the last inequality follows from Lemma 6.2, and the last equality follows from (4.14) and Lemma 4.1. The same estimate holds for $\boldsymbol{\lambda}_\Gamma, \boldsymbol{s}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$ analogously with the fact that $\langle \boldsymbol{\zeta}\cdot\mathbf{n}\lambda, s\rangle_{\partial\mathcal{T}_h} = 0$ and $\langle \boldsymbol{\zeta}\cdot\mathbf{n}\mu, t\rangle_{\partial\mathcal{T}_h} = 0$. $\square$

6.5. A dual argument

Denote $\mathbf{W} = W \times W$, given any $\mathbf{g} = (g, z_d)^T \in \mathbf{W}$, let $(y_g^*, p_{z_d}^*)$ be the solutions of the dual problem of system (2.8), then we have

$$\begin{aligned}
\mathbf{q}_g^* + \nabla y_g^* &= 0 \quad \text{in} \quad \Omega, && (6.22a)\\
\beta^{\frac{1}{2}}\left(\nabla\cdot\mathbf{q}_g^* - \boldsymbol{\zeta}\cdot\nabla y_g^*\right) + p_{z_d}^* &= g \quad \text{in} \quad \Omega, && (6.22b)\\
y_g^* &= 0 \quad \text{on} \quad \partial\Omega, && (6.22c)\\
\mathbf{p}_{z_d}^* + \nabla p_{z_d}^* &= 0 \quad \text{in} \quad \Omega, && (6.22d)\\
\beta^{\frac{1}{2}}\left(\nabla\cdot\mathbf{p}_{z_d}^* + \nabla\cdot\left(\boldsymbol{\zeta}p_{z_d}^*\right)\right) - y_g^* &= z_d \quad \text{in} \quad \Omega, && (6.22e)\\
p_{z_d}^* &= 0 \quad \text{on} \quad \partial\Omega. && (6.22f)
\end{aligned}$$

The following regularity assumption holds for the dual problem ([8], Lem. 2.2):

$$\beta^{\frac{1}{2}}\|y_g^*\|_{H^2(\Omega)} + \beta^{\frac{1}{2}}\|p_{z_d}^*\|_{H^2(\Omega)} \le C\left(\|g\|_{L^2(\Omega)} + \|z_d\|_{L^2(\Omega)}\right). \tag{6.23}$$

Given $\boldsymbol{s} = (s,t)^T \in \widetilde{\boldsymbol{\Lambda}}$, denote $\mathbf{g} = \mathbf{U}\boldsymbol{s}$. Let $\varphi_{\mathbf{g}} = (\varphi_g, \varphi_{z_d})^T$ and $\widetilde{\varphi}_{\mathbf{g}} = (\widetilde{\varphi}_g, \widetilde{\varphi}_{z_d})^T$ be the solutions of the following equations,

$$\mathcal{B}_h(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}), \mathbf{E}_{\mathcal{C}}(\varphi_{\mathbf{g}})) = (\mathbf{U}\boldsymbol{\lambda}, \mathbf{g})_{\mathcal{T}_h}, \qquad \forall\boldsymbol{\lambda} = (\lambda,\mu)^T \in \boldsymbol{\Lambda}, \tag{6.24}$$

$$\widetilde{\mathcal{B}}\left(\mathbf{E}_{\mathcal{C}}\left(\widetilde{\boldsymbol{\lambda}}\right), \mathbf{E}_{\mathcal{C}}(\widetilde{\varphi}_{\mathbf{g}})\right) = \left(\mathbf{U}\widetilde{\boldsymbol{\lambda}}, \mathbf{g}\right)_{\mathcal{T}_h}, \qquad \forall\widetilde{\boldsymbol{\lambda}} = \left(\widetilde{\lambda}, \widetilde{\mu}\right)^T \in \widetilde{\boldsymbol{\Lambda}}, \tag{6.25}$$

where $\mathcal{B}_h$ is defined as (2.12) and $\widetilde{\mathcal{B}}$ is defined as (4.5).

By adopting a similar proof strategy as in Lemma 7.17 from [47], and utilizing the regularity property (6.23), Assumptions 5.1 and 5.2, we obtain the following result:

Lemma 6.3. *Given $\boldsymbol{\lambda} = (\lambda,\mu)^T \in \widetilde{\boldsymbol{\Lambda}}$, there exists a positive constant C independent of β, H and h, such that*

$$\left|\widetilde{\mathcal{B}}(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}), \mathbf{E}_{\mathcal{C}}(\varphi_{\mathbf{g}})) - (\mathbf{U}\boldsymbol{\lambda}, \mathbf{g})_{L^2(\Omega)}\right| \le CH\beta^{-\frac{1}{4}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda})\|_{\mathbf{A_a}}\|\mathbf{g}\|_{L^2(\Omega)}.$$

Using similar proofs as Lemma 7.19 from [47], together with Lemma 6.3 and Lemma 7.11 from [45], we obtain the following two lemmas.

Lemma 6.4. *Given $\mathbf{g} \in \mathbf{W}$, under the Assumption 5.2, we have*

$$\|\mathbf{E}_{\mathcal{C}}(\widetilde{\varphi}_{\mathbf{g}} - \varphi_{\mathbf{g}})\|_{\mathbf{A_a}} \leq CH\beta^{-\frac{1}{4}}\|\mathbf{g}\|_{L^2(\Omega)}.$$

Lemma 6.5. *For any $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, recall that $\mathbf{w}_\Gamma = \widetilde{\mathbf{S}}_\Gamma^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma$, we have*

$$\widetilde{\mathcal{B}}\Big(\mathbf{E}_{\mathcal{C}}\Big(\mathbf{w}_{\mathbf{A},\Gamma} - \widetilde{\mathbf{R}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\Big), \mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi})\Big) = 0,$$

for any $\boldsymbol{\xi} \in \widetilde{\mathbf{R}}(\boldsymbol{\Lambda})$.

Lemma 6.6. *Given $\boldsymbol{\lambda}_\Gamma \in \widehat{\boldsymbol{\Lambda}}_\Gamma$, recall $\mathbf{w}_\Gamma = \widetilde{\mathbf{S}}_\Gamma^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma$. There exists a positive constant C independent of β, H and h, such that*

$$\|\mathbf{U}(\mathbf{w}_{\mathbf{A},\Gamma} - \boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathcal{T}_h} \leq CHc_0\Big(1 + \beta^{-\frac{1}{2}}\Big)\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}. \tag{6.26}$$

Proof. Given $\mathbf{g} \in \mathbf{W}$, we first consider the estimate of $(\mathbf{U}(\mathbf{w}_{\mathbf{A},\Gamma} - \boldsymbol{\lambda}_{\mathbf{A},\Gamma}), \mathbf{g})_{\mathcal{T}_h}$. Since

$$\begin{aligned}(\mathbf{U}(\mathbf{w}_{\mathbf{A},\Gamma} - \boldsymbol{\lambda}_{\mathbf{A},\Gamma}), \mathbf{g})_{\mathcal{T}_h} &= \widetilde{\mathcal{B}}(\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma}), \mathbf{E}_{\mathcal{C}}(\widetilde{\varphi}_{\mathbf{g}})) - \mathcal{B}_h(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}), \mathbf{E}_{\mathcal{C}}(\varphi_{\mathbf{g}}))\\ &= \widetilde{\mathcal{B}}(\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma}), \mathbf{E}_{\mathcal{C}}(\widetilde{\varphi}_{\mathbf{g}})) - \widetilde{\mathcal{B}}(\mathbf{E}_{\mathcal{C}}(\boldsymbol{\lambda}_{\mathbf{A},\Gamma}), \mathbf{E}_{\mathcal{C}}(\varphi_{\mathbf{g}})) = \widetilde{\mathcal{B}}(\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma}), \mathbf{E}_{\mathcal{C}}(\widetilde{\varphi}_{\mathbf{g}} - \varphi_{\mathbf{g}})),\end{aligned}$$

where the first inequality follows from (6.24) and (6.25), the last equality follows from Lemma 6.5 and the fact that $\widetilde{\mathbf{R}}$ is the injection operator from $\boldsymbol{\Lambda}$ into $\widetilde{\boldsymbol{\Lambda}}$ with $\boldsymbol{\Lambda} \subset \widetilde{\boldsymbol{\Lambda}}$. Then, we obtain that

$$\begin{aligned}(\mathbf{U}(\mathbf{w}_{\mathbf{A},\Gamma} - \boldsymbol{\lambda}_{\mathbf{A},\Gamma}), \mathbf{g})_{\mathcal{T}_h} &= \widetilde{\mathcal{B}}(\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma}), \mathbf{E}_{\mathcal{C}}(\widetilde{\varphi}_{\mathbf{g}} - \varphi_{\mathbf{g}}))\\ &= (Q_{A_1}\boldsymbol{\xi}_{\mathbf{g}}, U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}, \widetilde{\varphi}_g - \varphi_g)\big[\beta^{\frac{1}{2}}A_1^{ad} \quad -L\big]\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\\ &\quad + (Q_{A_2}\boldsymbol{\xi}_{\mathbf{g}}, U_{A_2}\boldsymbol{\xi}_{\mathbf{g}}, \widetilde{\varphi}_{z_d} - \varphi_{z_d})\big[L \quad A_2^{ad}\big]\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\\ &:= \mathrm{I} + \mathrm{II}.\end{aligned}$$

Next we will estimate the term I, and term II can be obtained similarly. Denote $\boldsymbol{\xi}_{\mathbf{g}} = \widetilde{\varphi}_{\mathbf{g}} - \varphi_{\mathbf{g}}$ and assume $\mathbf{w}_{\mathbf{A},\Gamma} = (w_{A,\Gamma}, x_{A,\Gamma})^T$, we have

$$\begin{aligned}|\mathrm{I}| &= \Bigg|\beta^{\frac{1}{2}}(Q_{A_1}\mathbf{w}_{\mathbf{A},\Gamma}, Q_{A_1}\boldsymbol{\xi}_{\mathbf{g}})_{\mathcal{T}_h} + \beta^{\frac{1}{2}}\Big\langle\Big(\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\Big)(U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma} - w_{A,\Gamma}), U_{A_1}\boldsymbol{\xi}_{\mathbf{g}} - (\widetilde{\varphi}_g - \varphi_g)\Big\rangle_{\partial\mathcal{T}_h}\\ &\quad - \frac{\beta^{\frac{1}{2}}}{2}(U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\xi}_{\mathbf{g}})_{\mathcal{T}_h} + \frac{\beta^{\frac{1}{2}}}{2}(U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma})_{\mathcal{T}_h}\\ &\quad + \frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}w_{A,\Gamma}, U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}\rangle_{\partial\mathcal{T}_h} - \frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}(\widetilde{\varphi}_g - \varphi_g), U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma}\rangle_{\partial\mathcal{T}_h} - (U_{A_2}\mathbf{w}_{\mathbf{A},\Gamma}, U_{A_1}\boldsymbol{\xi}_{\mathbf{g}})_{\mathcal{T}_h}\Bigg|\\ &= \Bigg|\beta^{\frac{1}{2}}(Q_{A_1}\mathbf{w}_{\mathbf{A},\Gamma}, Q_{A_1}\boldsymbol{\xi}_{\mathbf{g}})_{\mathcal{T}_h} + \beta^{\frac{1}{2}}\Big\langle\Big(\tau_1 - \frac{1}{2}\boldsymbol{\zeta}\cdot\mathbf{n}\Big)(U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma} - w_{A,\Gamma}), U_{A_1}\boldsymbol{\xi}_{\mathbf{g}} - (\widetilde{\varphi}_g - \varphi_g)\Big\rangle_{\partial\mathcal{T}_h}\\ &\quad - \frac{\beta^{\frac{1}{2}}}{2}\langle(\boldsymbol{\zeta}\cdot\mathbf{n})(U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma} - w_{A,\Gamma}), U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}\rangle_{\partial\mathcal{T}_h} + \beta^{\frac{1}{2}}(U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma})_{\mathcal{T}_h}\\ &\quad - \frac{\beta^{\frac{1}{2}}}{2}\langle(\boldsymbol{\zeta}\cdot\mathbf{n})(\widetilde{\varphi}_g - \varphi_g), U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma} - w_{A,\Gamma}\rangle_{\partial\mathcal{T}_h} - (U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}, U_{A_2}\mathbf{w}_{\mathbf{A},\Gamma})_{\mathcal{T}_h} - \frac{\beta^{\frac{1}{2}}}{2}\langle\boldsymbol{\zeta}\cdot\mathbf{n}w_{A,\Gamma}, \widetilde{\varphi}_g - \varphi_g\rangle_{\partial\mathcal{T}_h}\Bigg|\\ &\leq \|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}} + C\Big(1 + \beta^{\frac{1}{4}}\Big)h^{\frac{1}{2}}\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\Big(\beta^{\frac{1}{4}}h^{-\frac{1}{2}}\|U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}\|_{\mathcal{T}_h}\Big)\end{aligned}$$

$$
\begin{aligned}
&+ C\Big(1+\beta^{\frac14}\Big)\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\|U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}\|_{\mathcal{T}_h} + C\Big(1+\beta^{\frac14}\Big)h^{\frac12}\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\Big(\beta^{\frac14}\|\widetilde{\varphi}_g-\varphi_g\|_{\partial\mathcal{T}_h}\Big)\\
&+ \|U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}\|_{\mathcal{T}_h}\|U_{A_2}\mathbf{w}_{\mathbf{A},\Gamma}\|_{\mathcal{T}_h} + CH\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}}\\
\le{}& \|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}}\\
&+ C\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\bigg(\Big(1+\beta^{\frac14}\Big)^2\|U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}\|_{\mathcal{T}_h} + \Big(1+\beta^{\frac14}\Big)\beta^{\frac14}\|\widetilde{\varphi}_g-\varphi_g\|_h + H\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}}\bigg)\\
\le{}& \|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}}\\
&+ C\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\bigg(\Big(1+\beta^{\frac14}\Big)^2 c_0\beta^{-\frac14}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}} + \Big(1+\beta^{\frac14}\Big)\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}} + H\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}}\bigg)\\
\le{}& Cc_0\Big(\beta^{\frac14}+\beta^{-\frac14}\Big)\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\|\mathbf{E}_{\mathcal{C}}(\boldsymbol{\xi}_{\mathbf{g}})\|_{\mathbf{A_a}}\\
\le{}& Cc_0H\Big(1+\beta^{-\frac12}\Big)\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}\|\mathbf{g}\|_{\mathcal{T}_h},
\end{aligned}
$$

where the first equality follows from the divergence theorem with $\nabla\cdot\boldsymbol{\zeta}=0$, namely

$$
(U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma}, \boldsymbol{\zeta}\cdot\nabla U_{A_1}\boldsymbol{\xi}_{\mathbf{g}})_{\mathcal{T}_h} = \langle(\boldsymbol{\zeta}\cdot\mathbf{n})U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma}, U_{A_1}\boldsymbol{\xi}_{\mathbf{g}}\rangle_{\partial\mathcal{T}_h} - (\boldsymbol{\zeta}\cdot\nabla U_{A_1}\mathbf{w}_{\mathbf{A},\Gamma}, U_{A_1}\boldsymbol{\xi}_{\mathbf{g}})_{\mathcal{T}_h},
$$

the first inequality follows from (5.9), a trace inequality, (5.8) and Lemma 6.1, the second inequality follows from (4.13) and the fact that $\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{A_a}} \le \|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}}$, the third to last inequality follows from (5.6), Lemma 4.3 and (5.3), the last inequality follows from Lemma 6.4. Therefore, we have

$$
\begin{aligned}
\|\mathbf{U}(\mathbf{w}_{\mathbf{A},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A},\Gamma})\|_{\mathcal{T}_h} &= \sup_{\boldsymbol{s}\in\widetilde{\boldsymbol{\Lambda}}}\frac{(\mathbf{U}(\mathbf{w}_{\mathbf{A},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A},\Gamma}),\mathbf{U}\boldsymbol{s})_{L^2(\Omega)}}{\|\mathbf{U}\boldsymbol{s}\|_{L^2(\Omega)}} = \sup_{\boldsymbol{s}\in\widetilde{\boldsymbol{\Lambda}},\mathbf{g}=\mathbf{U}\boldsymbol{s}}\frac{(\mathbf{U}(\mathbf{w}_{\mathbf{A},\Gamma}-\boldsymbol{\lambda}_{\mathbf{A},\Gamma}),\mathbf{g})_{L^2(\Omega)}}{\|\mathbf{g}\|_{L^2(\Omega)}}\\
&\le Cc_0H\Big(1+\beta^{-\frac12}\Big)\|\mathbf{E}_{\mathcal{C}}(\mathbf{w}_{\mathbf{A},\Gamma})\|_{\mathbf{M_a}} = Cc_0H\Big(1+\beta^{-\frac12}\Big)\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma},
\end{aligned}
$$

where the last equality follows from (4.14). ☐

6.6. Estimates of $\|\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma-\boldsymbol{\lambda}_\Gamma)\|_h$ and $\|\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{T}\boldsymbol{\lambda}_\Gamma-\boldsymbol{\lambda}_\Gamma)\|_h$

Given any $\boldsymbol{\lambda}\in\widetilde{\boldsymbol{\Lambda}}$, let $\boldsymbol{\lambda}^K$ denote the restriction of $\boldsymbol{\lambda}$ on ∂K, where K is an element in the triangulation $\mathcal{T}_h$. Let E_i represent the edge of K that lies opposite to the i-th vertex of K. Following equations (7.33)–(7.35) from [47], we define the local lifting operator $S_i^K\boldsymbol{\lambda}^K$ in the polynomial space $P_{k+1}(K)\times P_{k+1}(K)$ such that:

$$
\begin{aligned}
\big\langle S_i^K\boldsymbol{\lambda}^K,\boldsymbol{\eta}\big\rangle_{E_i} &= \big\langle\boldsymbol{\lambda}^K,\boldsymbol{\eta}\big\rangle_{E_i}, && \forall\boldsymbol{\eta}\in P_{k+1}(E_i)\times P_{k+1}(E_i), && (6.27)\\
\big(S_i^K\boldsymbol{\lambda}^K,\mathbf{v}\big)_K &= \big(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}(\boldsymbol{\lambda}^K),\mathbf{v}\big)_K, && \forall\mathbf{v}\in P_k(K)\times P_k(K), && (6.28)
\end{aligned}
$$

for $i=1,\dots,n+1$, where $n+1$ denotes the total number of edges of K. We then define the global inner product and corresponding norm over the mesh $\mathcal{T}_h$ by:

$$
(\boldsymbol{\lambda},\boldsymbol{s})_S = \sum_{K\in\mathcal{T}_h}\frac{1}{n+1}\sum_{i=1}^{n+1}\big(S_i^K\boldsymbol{\lambda},S_i^K\boldsymbol{s}\big)_K, \quad\text{and}\quad \|\boldsymbol{\lambda}\|_S^2 = (\boldsymbol{\lambda},\boldsymbol{\lambda})_S. \tag{6.29}
$$

Denote Q_k as the L^2 projection from $P_{k+1}(K)\times P_{k+1}(K)$ into $P_k(K)\times P_k(K)$, following similar process as Lemma 5.7 from [13], we have:

Lemma 6.7.

$$
\begin{aligned}
\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}(\boldsymbol{\lambda}^K) &= Q_k\big(S_i^K\boldsymbol{\lambda}^K\big)_K, && (6.30)\\
c\|\boldsymbol{\lambda}\|_h &\le \|\boldsymbol{\lambda}\|_S, && (6.31)\\
\big|S_i^K\boldsymbol{\lambda}\big|_{H^1(K)} &\le C|\!|\!|\boldsymbol{\lambda}|\!|\!|_K. && (6.32)
\end{aligned}
$$

Recall the definitions of bilinear forms $\mathcal{B}_h^{\mathbf{ad}}$ and $\widetilde{\mathcal{B}}^{\mathbf{ad}}$ in (2.22) and (4.6). Given $\mathbf{s} \in \widetilde{\mathbf{\Lambda}}$, let $\varphi_{\mathbf{s}}$ and $\widetilde{\varphi}_{\mathbf{s}}$ be the solutions of the following equations,

$$\mathcal{B}_h^{\mathbf{ad}}(\mathbf{E}_{\mathbf{A}^{\mathrm{ad}}}(\boldsymbol{\lambda}), \mathbf{E}_{\mathbf{A}^{\mathrm{ad}}}(\varphi_{\mathbf{s}})) = (\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}\boldsymbol{\lambda}, \mathbf{g})_\Omega \qquad \forall \boldsymbol{\lambda} = (\lambda, \mu)^T \in \mathbf{\Lambda}, \tag{6.33}$$

$$\widetilde{\mathcal{B}}^{\mathbf{ad}}\Big(\mathbf{E}_{\mathbf{A}^{\mathrm{ad}}}(\widetilde{\boldsymbol{\lambda}}), \mathbf{E}_{\mathbf{A}^{\mathrm{ad}}}(\widetilde{\varphi}_{\mathbf{s}})\Big) = \Big(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}\widetilde{\boldsymbol{\lambda}}, \mathbf{g}\Big)_\Omega \qquad \forall \boldsymbol{\lambda} = \Big(\widetilde{\lambda}, \widetilde{\mu}\Big)^T \in \widetilde{\mathbf{\Lambda}}, \tag{6.34}$$

where $\mathbf{g}$ is defined as

$$\mathbf{g} = \frac{1}{n+1}\sum_{i=1}^{n+1} Q_k\big(S_i^K \boldsymbol{s}\big). \tag{6.35}$$

Now, we are ready to provide the proof of Lemma 5.14.

6.7. Proof of Lemma 5.14

Proof. Given $\boldsymbol{s} = (s,t)^T \in \widetilde{\mathbf{\Lambda}}$, and defining $\mathbf{g}$ as in (6.35), by (6.29) we observe that

$$\begin{aligned}(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma), \boldsymbol{s})_S &= \sum_{K\in\mathcal{T}_h} \frac{1}{n+1}\sum_{i=1}^{n+1}\big(S_i^K \mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma), S_i^K \boldsymbol{s}\big)_K,\\ &= (\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), \mathbf{g})_{\mathcal{T}_h} + \sum_{K\in\mathcal{T}_h(\Omega)} \frac{1}{n+1}\sum_{i=1}^{n+1}\big((\mathrm{I} - Q_k)S_i^K(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), (\mathrm{I} - Q_k)S_i^K \boldsymbol{s}\big)_K.\end{aligned}$$

Therefore, we have

$$\begin{aligned}|(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma), \boldsymbol{s})_S| \le{}& \big|(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), \mathbf{g})_{\mathcal{T}_h}\big|\\ &+ \left|\sum_{K\in\mathcal{T}_h(\Omega)} \frac{1}{n+1}\sum_{i=1}^{n+1}\big((1 - Q_k)S_i^K(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), (1 - Q_k)S_i^K \boldsymbol{s}\big)_K\right|.\end{aligned} \tag{6.36}$$

Denote $\mathcal{H}_{\mathbf{A}}(\boldsymbol{\lambda}_\Gamma) = \boldsymbol{\lambda}_{\mathbf{A},\Gamma}$. Following the similar argument as in equation (7.58) from [47], we obtain

$$\begin{aligned}&\left|\sum_{K\in\mathcal{T}_h(\Omega)} \frac{1}{n+1}\sum_{i=1}^{n+1}\big((\mathrm{I} - Q_k)S_i^K(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), (\mathrm{I} - Q_k)S_i^K \boldsymbol{s}\big)_K\right|\\ &\quad\le h|\!|\!|\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)|\!|\!|\|\boldsymbol{s}\|_S \le Ch\|\mathbf{E}_{\mathbf{A}^{\mathrm{ad}}}(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathbf{A}^{\mathrm{ad}}}\|\boldsymbol{s}\|_S\\ &\quad\le Ch\beta^{-\frac14}\|\mathbf{E}_{\mathcal{C}}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathbf{A}_{\mathbf{a}}}\|\boldsymbol{s}\|_S \le Ch\beta^{-\frac14}\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\Big)\|\boldsymbol{s}\|_S,\end{aligned} \tag{6.37}$$

where the first inequality follows from a Friedrichs inequality and (6.32) with the fact that

$$\big\|(\mathrm{I} - Q_k)S_i^K(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\big\|_{L^2(K)} \le Ch\big|S_i^K\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\big|_{H^1(K)} \le Ch|\!|\!|\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)|\!|\!|_K,$$

and $\|(\mathrm{I} - Q_k)S_i^K\boldsymbol{s}\|_K \le C\|S_i^K\boldsymbol{s}\|_K$, the second inequality follows from (4.21), the second to last inequality follows from (5.11), and the last inequality follows from Lemma 4.1.

Next, we consider the estimate of $\big|\big(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}\big(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma)\big), \mathbf{g}\big)_{\mathcal{T}_h}\big|$. We have

$$\begin{aligned}&\big|(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), \mathbf{g})_{\mathcal{T}_h}\big|\\ &\quad= \big|\big(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}\big(\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\big) - \mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}\big(\mathcal{H}_{\mathbf{A}}\big(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma\big)\big), \mathbf{g}\big)_{\mathcal{T}_h}\\ &\qquad+ \big(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)) - \mathbf{U}\big(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\big), \mathbf{g}\big)_{\mathcal{T}_h} + \big(\mathbf{U}\big(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\big), \mathbf{g}\big)_{\mathcal{T}_h}\big|\\ &\quad\le \|\mathcal{H}_{\mathbf{A}^{\mathrm{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma) - \mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h\|\mathbf{g}\|_{\mathcal{T}_h} + \|\mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathcal{T}_h}\|\mathbf{g}\|_{\mathcal{T}_h}\\ &\qquad+ \big|\big(\mathbf{U}_{\mathbf{A}^{\mathrm{ad}}}\big(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\big) - \mathbf{U}\big(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\big), \mathbf{g}\big)_{\mathcal{T}_h}\big|,\end{aligned} \tag{6.38}$$

where the inequality follows from (4.19) and (4.20).

We firstly estimate the term $\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma) - \mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h$. Since $\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma) - \mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)$ has $\mathbf{0}$ values on the subdomain boundary, we obtain that

$$\begin{aligned}
&\left\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma) - \mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\right\|_h \\
&\quad = \left(\sum_{i=1}^{N} \|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma) - \mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_{h,\Omega_i}^2\right)^{\frac{1}{2}} \\
&\quad \le \left(\sum_{i=1}^{N} H^2 \|\!|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma) - \mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)|\!\|_{\Omega_i}^2\right)^{\frac{1}{2}} \\
&\quad \le CH \|\!|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma) - \mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)|\!\| \\
&\quad \le CH \|\mathbf{E}_{\mathbf{A}^{\mathbf{ad}}}(\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathbf{A}^{\mathbf{ad}}} + CH\beta^{-\frac{1}{4}} \|\mathbf{E}_{\mathcal{C}}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathbf{A}_\mathbf{a}} \\
&\quad \le CH\beta^{-\frac{1}{4}} \|\mathbf{E}_{\mathcal{C}}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathbf{A}_\mathbf{a}} \\
&\quad \le CH\beta^{-\frac{1}{4}} \left(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\right),
\end{aligned} \tag{6.39}$$

where the first inequality follows from Lemma 4.2, the second inequality follows from (4.21) and (5.3), the second to last inequality follows from (5.11), and the last inequality follows from Lemma 4.1.

Then, for the term $\left|\left(\mathbf{U}_{\mathbf{A}^{\mathbf{ad}}}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)) - \mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), \mathbf{g}\right)_{\mathcal{T}_h}\right|$, by (5.7), we have

$$\begin{aligned}
&\left|\left(\mathbf{U}_{\mathbf{A}^{\mathbf{ad}}}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)) - \mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), \mathbf{g}\right)_{\mathcal{T}_h}\right| \\
&\quad \le \left\|\mathbf{U}_{\mathbf{A}^{\mathbf{ad}}}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)) - \mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\right\|_{\mathcal{T}_h} \|\mathbf{g}\|_{\mathcal{T}_h} \\
&\quad \le Ch\beta^{-\frac{1}{2}} \|\mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathcal{T}_h} \|\mathbf{g}\|_{\mathcal{T}_h}.
\end{aligned} \tag{6.40}$$

Combining (6.38), (6.39), (6.40), Lemma 6.6, with the fact that $\|\mathbf{g}\|_{\mathcal{T}_h} \le \|\boldsymbol{s}\|_S$ which is proved in Lemma 7.20 from [47], we have

$$\begin{aligned}
&\left|(\mathbf{U}_{\mathbf{A}^{\mathbf{ad}}}(\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)), \mathbf{g})_{\mathcal{T}_h}\right| \\
&\quad \le CH\beta^{-\frac{1}{4}} \left(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\right) \|\mathbf{g}\|_{\mathcal{T}_h} + \|\mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathcal{T}_h} \|\mathbf{g}\|_{\mathcal{T}_h} \\
&\qquad + \left(h\beta^{-\frac{1}{2}}\right) \|\mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathcal{T}_h} \|\mathbf{g}\|_{\mathcal{T}_h} \\
&\quad \le CH\beta^{-\frac{1}{4}} \left(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\right) \|\mathbf{g}\|_{\mathcal{T}_h} + C\left(1 + h\beta^{-\frac{1}{2}}\right) \|\mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma))\|_{\mathcal{T}_h} \|\mathbf{g}\|_{\mathcal{T}_h} \\
&\quad \le CH\beta^{-\frac{1}{4}} \left(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\right) \|\mathbf{g}\|_{\mathcal{T}_h} + C\left(H\left(1 + \beta^{-\frac{1}{2}}\right) c_0^2 \|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}\right) \|\mathbf{g}\|_{\mathcal{T}_h} \\
&\quad \le C\left(\left(H\beta^{-\frac{1}{2}} + H\right) c_0^2 + H\beta^{-\frac{1}{4}}\right) \left(\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\right) \|\boldsymbol{s}\|_S.
\end{aligned} \tag{6.41}$$

Thus, by (6.31), (6.36), (6.37) and (6.41), we obtain that

$$\begin{aligned}
\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h \le \|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_S &= \sup_{\boldsymbol{s} \in \widetilde{\boldsymbol{\Lambda}}} \frac{(\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_\Gamma - \boldsymbol{\lambda}_\Gamma), \boldsymbol{s})_S}{\|\boldsymbol{s}\|_S} \\
&\le C\left(\left(H\beta^{-\frac{1}{2}} + H\right) c_0^2 + H\beta^{-\frac{1}{4}}\right) \left(\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\right).
\end{aligned} \tag{6.42}$$

Moreover, let $\widehat{\mathbf{R}}_\Gamma^{(i)}, \overline{\mathbf{R}}_\Gamma^{(i)}$ be the restriction operators from $\widehat{\boldsymbol{\Lambda}}_\Gamma$ and $\widetilde{\boldsymbol{\Lambda}}_\Gamma$ to $\boldsymbol{\Lambda}_\Gamma^{(i)}$ respectively. By the similar argument as in the proof of Lemma 6.10 from [47], we have

$$\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h$$

$$\leq C\left(\sum_{i=1}^{N}\left\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}\left(\widehat{\mathbf{R}}_{\Gamma}^{(i)}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^{T}\mathbf{w}_{\Gamma}-\overline{\mathbf{R}}_{\Gamma}^{(i)}\mathbf{w}_{\Gamma}\right)\right\|_{h,\Omega_i}^{2}\right)^{\frac{1}{2}}+C\left(\sum_{i=1}^{N}\left\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}\left(\overline{\mathbf{R}}_{\Gamma}^{(i)}\mathbf{w}_{\Gamma}-\widehat{\mathbf{R}}_{\Gamma}^{(i)}\boldsymbol{\lambda}_{\Gamma}\right)\right\|_{h,\Omega_i}^{2}\right)^{\frac{1}{2}}$$
$$:= \mathrm{I}+\mathrm{II}.$$

For the estimate of I, since $\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^{T}\mathbf{w}_{\Gamma}$ and $\mathbf{w}_{\Gamma}$ have the same edge average on the subdomain boundary $\partial\Omega_i$, then

$$\begin{aligned}\mathrm{I} &\leq C\left(\sum_{i=1}^{N}H^2|||\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}\left(\widehat{\mathbf{R}}_{\Gamma}^{(i)}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^{T}\mathbf{w}_{\Gamma}-\overline{\mathbf{R}}_{\Gamma}^{(i)}\mathbf{w}_{\Gamma}\right)|||_{\mathcal{T}_h(\Omega_i)}^{2}\right)^{\frac{1}{2}}\\
&\leq CH\beta^{-\frac{1}{4}}\left(\sum_{i=1}^{N}\left\|\widehat{\mathbf{R}}_{\Gamma}^{(i)}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^{T}\mathbf{w}_{\Gamma}-\overline{\mathbf{R}}_{\Gamma}^{(i)}\mathbf{w}_{\Gamma}\right\|_{\widetilde{\mathbf{S}}_{\Gamma}}^{2}\right)^{\frac{1}{2}}\leq CH\beta^{-\frac{1}{4}}C_{ED}\|\mathbf{w}_{\Gamma}\|_{\widetilde{\mathbf{S}}_{\Gamma}},\end{aligned}$$

where the first inequality follows from Lemma 4.2, the second inequality follows from (5.3) and Lemma 4.1, and the last inequality follows from (5.15). For the estimate of II, by (6.42) we have

$$\mathrm{II}\leq\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{w}_{\Gamma}-\boldsymbol{\lambda}_{\Gamma})\|_h\leq C\Big(\Big(H\beta^{-\frac{1}{2}}+H\Big)c_0^2+H\beta^{-\frac{1}{4}}\Big)\Big(\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_{\Gamma}}+\|\boldsymbol{\lambda}_{\Gamma}\|_{\widehat{\mathbf{S}}_{\Gamma}}\Big).$$

Therefore, we obtain that

$$\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{T}\boldsymbol{\lambda}_{\Gamma}-\boldsymbol{\lambda}_{\Gamma})\|_h\leq C\Big(\Big(H\beta^{-\frac{1}{2}}+H\Big)c_0^2+H\beta^{-\frac{1}{4}}C_{ED}\Big)\Big(\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_{\Gamma}}+\|\boldsymbol{\lambda}_{\Gamma}\|_{\widehat{\mathbf{S}}_{\Gamma}}\Big).$$

□

6.8. Proof of Lemma 5.15

Proof. We first give the proof of (5.21), let

$$\mathbf{z}_{\Gamma}=\widetilde{\mathbf{L}}_{\Gamma}^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\mathbf{L}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}=\widetilde{\mathbf{L}}_{\Gamma}^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widetilde{\mathbf{R}}_{\Gamma}^{T}\widetilde{\mathbf{L}}_{\Gamma}\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}=\widetilde{\mathbf{L}}_{\Gamma}^{-1}\mathbf{E}_{\mathbf{D}}^{T}\widetilde{\mathbf{L}}_{\Gamma}\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma},$$

then

$$\begin{aligned}\langle\mathbf{w}_{\Gamma},\mathbf{z}_{\Gamma}\rangle_{\widetilde{\mathbf{L}}_{\Gamma}}-\left\langle\mathbf{w}_{\Gamma},\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}\right\rangle_{\widetilde{\mathbf{L}}_{\Gamma}}&=\boldsymbol{\lambda}_{\Gamma}^{T}\widetilde{\mathbf{R}}_{\Gamma}^{T}\widetilde{\mathbf{L}}_{\Gamma}^{T}\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\left\langle\mathbf{w}_{\Gamma},\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}\right\rangle_{\widetilde{\mathbf{L}}_{\Gamma}}=\left\langle\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma},\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}\right\rangle_{\widetilde{\mathbf{L}}_{\Gamma}}-\left\langle\mathbf{w}_{\Gamma},\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}\right\rangle_{\widetilde{\mathbf{L}}_{\Gamma}}\\
&=\left\langle\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma},\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}\right\rangle_{\widetilde{\mathbf{L}}_{\Gamma}}.\end{aligned}$$

We have

$$\begin{aligned}&\left\langle\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma},\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}\right\rangle_{\widetilde{\mathbf{L}}_{\Gamma}}\\
&\quad\leq\left\|\widetilde{\mathbf{R}}_{\Gamma}\boldsymbol{\lambda}_{\Gamma}\right\|_{\widetilde{\mathbf{L}}_{\Gamma}}\|\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma}\|_{\widetilde{\mathbf{L}}_{\Gamma}}=\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{L}_{\Gamma}}\|\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma}\|_{\widetilde{\mathbf{L}}_{\Gamma}}=\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{L}_{\Gamma}}\|\mathbf{U}(\mathcal{H}_{\mathbf{A}}(\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma}))\|_{\mathcal{T}_h}\\
&\quad\leq\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{L}_{\Gamma}}(c_0\|\mathcal{H}_{\mathbf{A}}(\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma})\|_h)\leq C\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{L}_{\Gamma}}(c_0H|||\mathcal{H}_{\mathbf{A},\Gamma}(\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma})|||)\\
&\quad\leq C\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{L}_{\Gamma}}\left(c_0H\beta^{-\frac{1}{4}}\|\mathbf{E}_{\mathbf{D}}\mathbf{w}_{\Gamma}-\mathbf{w}_{\Gamma}\|_{\widetilde{\mathbf{S}}_{\Gamma}}\right)\leq Cc_0H\beta^{-\frac{1}{4}}C_{ED}\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{L}_{\Gamma}}\|\mathbf{w}_{\Gamma}\|_{\widetilde{\mathbf{S}}_{\Gamma}}\\
&\quad=Cc_1\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{L}_{\Gamma}}\|\mathbf{w}_{\Gamma}\|_{\widetilde{\mathbf{S}}_{\Gamma}}\leq Cc_1\|\boldsymbol{\lambda}_{\Gamma}\|_{\mathbf{M}_{\Gamma}}\|\mathbf{w}_{\Gamma}\|_{\widetilde{\mathbf{M}}_{\Gamma}},\end{aligned}$$

where $c_1:=H\beta^{-\frac{1}{4}}C_{ED}$, the second inequality follows from (5.5) in Lemma 5.8, the third inequality follows from Lemma 4.3, the fourth inequality follows from (5.3), the second to last inequality follows from (5.13), and the last inequality follows from Lemma 4.1.

Next, we give the proof of (5.22). We have

$$
\begin{aligned}
&\langle \boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma\rangle_{\mathbf{Z}_\Gamma}\\
&\quad\le C\gamma_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big(C\Big(H\beta^{-\frac{1}{4}}\Big)\|\mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma} + \|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h\Big)\\
&\quad\le CH\beta^{-\frac{1}{4}}\gamma_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\|\mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma} + C\gamma_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}(\mathbf{T}\boldsymbol{\lambda}_\Gamma - \boldsymbol{\lambda}_\Gamma)\|_h\\
&\quad\le CH\beta^{-\frac{1}{4}}\gamma_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big(C_{ED}\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\Big) + C\gamma_1\|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big(\alpha_2\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma}\Big)\Big)\\
&\quad\le Cc_2\|\boldsymbol{\lambda}\|_{\mathbf{M}_\Gamma}\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big)
\end{aligned}
$$

where $c_2 := \gamma_1\alpha_2$, the first inequality follows from Lemma 5.13, the second to last inequality follows from (5.20), a triangle inequality with (5.13) and the fact that

$$
\begin{aligned}
\|\mathbf{T}\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma} = \langle \mathbf{T}\boldsymbol{\lambda}_\Gamma, \mathbf{T}\boldsymbol{\lambda}_\Gamma\rangle_{\widehat{\mathbf{S}}_\Gamma}^{\frac{1}{2}} &= \Big\langle \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T\widetilde{\mathbf{S}}_\Gamma^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma, \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T\widetilde{\mathbf{S}}_\Gamma^{-1}\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}\widehat{\mathbf{S}}_\Gamma\boldsymbol{\lambda}_\Gamma\Big\rangle_{\widehat{\mathbf{S}}_\Gamma}^{\frac{1}{2}}\\
&= \langle \mathbf{E}_\mathbf{D}\mathbf{w}_\Gamma, \mathbf{E}_\mathbf{D}\mathbf{w}_\Gamma\rangle_{\widetilde{\mathbf{S}}_\Gamma}^{\frac{1}{2}} = \|\mathbf{E}_\mathbf{D}\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma} \le CC_{ED}\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma},
\end{aligned}
$$

and the last inequality follows from the definition of α_2 and Lemma 4.1.

Then, (5.23) can be obtained by Lemma 6.6 with the fact that

$$
\Big\|\mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\Big\|_{\widetilde{\mathbf{L}}_\Gamma} = \Big\|\mathbf{U}\Big(\mathbf{w}_{\mathbf{A},\Gamma} - \widetilde{\mathbf{R}}\boldsymbol{\lambda}_{\mathbf{A},\Gamma}\Big)\Big\|_{\mathcal{T}_h} \le Cc_3\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma},
$$

where $c_3 := c_0H(1+\beta^{-\frac{1}{2}})$.

Finally, we give the proof of (5.24), since

$$
\begin{aligned}
&\Big|\Big\langle \mathbf{w}_\Gamma, \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma - \mathbf{w}_\Gamma\Big\rangle_{\widetilde{\mathbf{Z}}_\Gamma}\Big|\\
&\quad\le C\gamma_1\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}\Big(C\Big(H\beta^{-\frac{1}{4}}\Big)\Big\|\widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma - \mathbf{w}_\Gamma\Big\|_{\widetilde{\mathbf{S}}_\Gamma} + \Big\|\mathcal{H}_{\mathbf{A}^{\mathbf{ad}}}\Big(\mathbf{w}_\Gamma - \widetilde{\mathbf{R}}_\Gamma\boldsymbol{\lambda}_\Gamma\Big)\Big\|_h\Big)\\
&\quad\le CH\beta^{-\frac{1}{4}}\gamma_1\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma}\Big(\|\boldsymbol{\lambda}_\Gamma\|_{\widehat{\mathbf{S}}_\Gamma} + \|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{S}}_\Gamma}\Big) + C\gamma_1\alpha_1\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma}\Big(\|\mathbf{w}_\Gamma\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}\|_{\widehat{\mathbf{S}}_\Gamma}\Big)\\
&\quad\le Cc_4\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma}\Big(\|\mathbf{w}\|_{\widetilde{\mathbf{M}}_\Gamma} + \|\boldsymbol{\lambda}_\Gamma\|_{\mathbf{M}_\Gamma}\Big),
\end{aligned}
$$

where $c_4 := \gamma_1\alpha_1$, the first inequality follows from Lemma 5.13, the second inequality follows from (5.19), and the last inequality follows from Lemma 4.1. ☐

7. Numerical experiments

We decompose the domain $\Omega = [0,1]\times[0,1]$ into nonoverlapping subdomains and each subdomain into triangles. Using the discretization system (2.10) with HDG polynomial degrees $k = 1, 2$, we consider two velocity fields $\boldsymbol{\zeta} = (1, 0)$ and $\boldsymbol{\zeta} = (x_2, -x_1)$. For each case, the regularization parameter β is set to $1, 10^{-4}, 10^{-6}, 10^{-8}$. The right-hand side functions f and g are computed based on the exact solution (y, p) as

$$
y = \sin^3(\pi x_1)\sin^2(\pi x_2)\cos(\pi x_2), \quad p = -\sin^2(\pi x_1)\sin^2(\pi x_2)\cos(\pi x_1).
$$

The stabilization parameters are chosen as $\tau_1 = \max(\sup_{x\in\mathcal{E}}(\boldsymbol{\zeta}\cdot\mathbf{n}), 0) + 1, \tau_2 = \tau_1 - \boldsymbol{\zeta}\cdot\mathbf{n}, \forall\mathcal{E}\subset\partial K, \forall K\in\mathcal{T}_h$. These choices satisfy Assumption 5.1. We also employ the edge average constraints, the edge flux weighted average constraints, and the edge flux weighted first moment constraints from Assumption 5.2. We use the GMRES method without restart, and the iteration is stopped when the residual is reduced by 10^{-11}.

TABLE 1. Iteration counts for varying number of subdomains with fixed subdomain problem size $H/h = 6$.

# of Sub.	4^2	8^2	16^2	32^2	4^2	8^2	16^2	32^2
β (Test I)		$k=1$				$k=2$		
1	15	16	15	14	17	18	17	17
10^{-4}	17	17	15	14	19	18	17	18
10^{-6}	13	15	14	13	15	17	16	16
10^{-8}	8	10	12	12	9	12	14	14
β (Test II)		$k=1$				$k=2$		
1	6	6	6	5	8	8	9	9
10^{-4}	6	6	6	5	9	8	9	9
10^{-6}	7	6	5	5	8	8	7	7
10^{-8}	6	6	6	5	5	7	7	7

For the numerical experiments, we adopt the BDDC preconditioned algorithm described in [34], modified to impose the three constraints specified in Assumption 5.2 on the subdomain interface. In this algorithm, we eliminate the interior unknowns of the subdomain local matrices to assemble a global interface Schur complement system. The resulting system is then solved using GMRES with the BDDC preconditioner. Finally, the interface solution is mapped back to recover the local variables on each subdomain.

The proposed algorithm naturally leverages parallel architectures, since the local subdomain solvers can be carried out concurrently on separate processors. Constructing the coarse problem introduces inter-subdomain communication, and enforcing the additional constraints requires extra communication between adjacent subdomains. However, this overhead is offset by the reduction in GMRES iterations, which requires global synchronization. We therefore anticipate that the algorithm will exhibit weak scalability for a reasonably large number of subdomains. For extremely large numbers of subdomains, where the coarse problem may become a bottleneck, three-level or multilevel BDDC algorithms can be employed to maintain scalability [42–44, 52].

We report our numerical results in Tables 1 and 2. In Table 1, we fix the subdomain problem size and vary the number of subdomains. When $k = 1$, the basis functions are linear polynomials. With a fixed value of β, we observe that the number of iterations is independent of the number of subdomains. Moreover, the iteration count is robust with respect to changes in β. When $k = 2$, the basis functions are quadratic polynomials. In this case, the number of iterations is higher compared to the $k = 1$ case. Nevertheless, the iteration count continues to be independent of the number of subdomains and exhibits good scalability, even for small values of β.

In Table 2, we fix the number of subdomains and vary the subdomain problem size. We observe that the iteration count increases slightly as H/h increases. This behavior is consistent with those results in symmetric positive definite problems [41, 46, 49]. More numerical results can be found in [34].

8. Concluding remarks

In this work, we perform a thorough analysis of the BDDC preconditioner applied to the HDG discretization of an elliptic optimal control problem constrained by a advection–diffusion equation. We prove that the BDDC preconditioner is robust with respect to β, provided H is sufficiently small. The analysis in this work can be easily extended to the case where the state equation is more general, for example, a non-divergence-free advection field. The results in this work could potentially be extended to the case of a advection-dominated state equation (cf. [33, 47]), or to problems with control or state constraints (cf. [6, 9, 10, 19]), where the BDDC preconditioner can be applied to the subsystem that must be solved during the outer iterative method.

TABLE 2. Iteration counts for changing subdomain problem size with 6×6 subdomains.

H/h	4	8	16	20	4	8	16	20
β (Test I)	$k = 1$				$k = 2$			
1	15	16	18	19	17	19	20	21
10^{-4}	16	18	19	20	18	20	21	22
10^{-6}	13	16	17	18	16	18	19	19
10^{-8}	8	10	12	12	9	11	13	13
β(Test II)	$k = 1$				$k = 2$			
1	5	7	8	8	8	8	8	7
10^{-4}	5	7	8	8	9	8	9	10
10^{-6}	5	7	8	8	8	8	7	7
10^{-8}	5	7	8	8	5	7	6	6

Extending the analysis to three dimensions is conceptually straightforward but entails several technical modifications. First, the coarse problem becomes more complex, requiring the inclusion of face constraints in Assumption 5.2, analogous to Assumption 4.7 of [45]. Second, Assumption 5.4 needs to be adapted for three dimensions to define each subdomain as a tetrahedron or a hexahedron, consistent with Assumption 4.8 of [45]. Finally, following Lemma 5.1 of [45], we need to introduce the interpolation operator I_h and finite element face cut-off function $\nu_{\mathcal{F}^l}$, which evaluates to 1 on the interface face $\mathcal{F}^l$ and vanishes at all other interface nodes. This modification leads to an adjusted estimate in Lemma 6.3 and, consequently, a different constant c_l in Lemma 5.17.

Regarding polygonal and polyhedral meshes, recent advancements in non-conforming polytopal hybrid discretizations provide the necessary theoretical components. In particular, Badia *et al.* [3] develops a discrete trace theory for non-conforming hybrid methods on polytopal meshes. Building on this, Badia *et al.* [2] establishes polylogarithmic condition number bounds for the BDDC method in this setting. Additionally, Du and Sayas [17] constructs local HDG+ projections, extending projection-based error analysis to polytopal meshes. These developments suggest that extending our BDDC framework to polytopal meshes is a feasible direction for future research. A fully rigorous extension, however, would require adapting multiple technical components. This includes incorporating the mesh regularity assumptions from [2, 3, 17], integrating discrete trace theory with projection theory on polytopal meshes, and modifying the coarse constraints to accommodate our optimal control setting. Given that these steps are technically involved, we leave a detailed treatment for future work.

Acknowledgments

The authors would like to thank Prof. Xuemin Tu for the helpful discussion regarding this project.

Funding

This material is based upon work supported by the National Science Foundation under Grant No. DMS-1929284 while SL was in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the Numerical PDEs: Analysis, Algorithms, and Data Challenges semester program.

References

[1] B. Ayuso and L.D. Marini, Discontinuous Galerkin methods for advection–diffusion–reaction problems. *SIAM J. Numer. Anal.* **47** (2009) 1391–1420.

[2] S. Badia, J. Droniou, J. Manyer and J. Tushar, Analysis of BDDC preconditioners for non-conforming polytopal hybrid discretisation methods. Preprint arXiv:2506.11956 (2025).

[3] S. Badia, J. Droniou and J. Tushar, A discrete trace theory for non-conforming polytopal hybrid discretisation methods. *Found. Comput. Math.* (2025) 1–50.

[4] G. Biros and O. Ghattas, Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. Part I: the Krylov–Schur solver. *SIAM J. Sci. Comput.* **27** (2005) 687–713.

[5] G. Biros and O. Ghattas, Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. Part II: the Lagrange–Newton solver and its application to optimal control of steady viscous flows. *SIAM J. Sci. Comput.* **27** (2005) 714–739.

[6] S.B. Boyana, T. Lewis, S. Liu and Y. Zhang, Convergence analysis of a dual-wind discontinuous Galerkin method for an elliptic optimal control problem with control constraints. *Results Appl. Math.* **27** (2025) 100624.

[7] S.C. Brenner, S. Liu and L.-Y. Sung, Multigrid methods for saddle point problems: Karush–Kuhn–Tucker systems. Preprint arXiv:1811.12434 (2018).

[8] S.C. Brenner, S. Liu and L.-Y. Sung, Multigrid methods for saddle point problems: optimality systems. *J. Comput. Appl. Math.* **372** (2020) 112733.

[9] S.C. Brenner, S. Liu and L.-Y. Sung, A P_1 finite element method for a distributed elliptic optimal control problem with a general state equation and pointwise state constraints. *Comput. Methods Appl. Math.* **21** (2021) 777–790.

[10] S.C. Brenner, S. Liu and L.-Y. Sung, Multigrid methods for an elliptic optimal control problem with pointwise state constraints. *Results Appl. Math.* **17** (2023) 100356.

[11] G. Chen, W. Hu, J. Shen, J.R. Singler, Y. Zhang and X. Zheng, An HDG method for distributed control of convection diffusion PDEs. *J. Comput. Appl. Math.* **343** (2018) 643–661.

[12] G. Chen, J.R. Singler and Y. Zhang, An HDG method for Dirichlet boundary control of convection dominated diffusion PDEs. *SIAM J. Numer. Anal.* **57** (2019) 1919–1946.

[13] B. Cockburn, O. Dubois, J. Gopalakrishnan and S. Tan, Multigrid for an HDG method. *IMA J. Numer. Anal.* **34** (2014) 1386–1425.

[14] S.S. Collis and M. Heinkenschloss, Analysis of the SUPG method for the solution of optimal control problems. *CoRR*. Preprint arXiv:2411.09828 (2024).

[15] C.R. Dohrmann, Preconditioning of saddle point systems by substructuring and a penalty approach, in Domain Decomposition Methods in Science and Engineering XVI. Vol. 55 of *Lect. Notes Comput. Sci. Eng.* Springer, Berlin (2007) 53–64.

[16] C. Dohrmann and O. Widlund, Some recent tools and a BDDC algorithm for 3D problems in H(curl), in Domain Decomposition Methods in Science and Engineering XX. Vol. 91 of *Lect. Notes Comput. Sci. Eng.* Springer, Heidelberg (2013) 15–25.

[17] S. Du and F.-J. Sayas, A note on devising HDG+ projections on polyhedral elements. *Math. Comput.* **90** (2021) 65–79.

[18] S.C. Eisenstat, H.C. Elman and M.H. Schultz, Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.* **20** (1983) 345–357.

[19] M. Engel and M. Griebel, A multigrid method for constrained optimal control problems. *J. Comput. Appl. Math.* **235** (2011) 4368–4388.

[20] G. Fu, W. Qiu and W. Zhang, An analysis of HDG methods for convection-dominated diffusion problems. *ESAIM Math. Model. Numer. Anal.* **49** (2015) 225–256.

[21] F. Gaspoz, C. Kreuzer, A. Veeser and W. Wollner, Quasi-best approximation in optimization with PDE constraints. *Inv. Prob.* **36** (2019) 014004.

[22] W. Gong, Z. Tan and Z. Zhou, Optimal convergence of finite element approximation to an optimization problem with PDE constraint. *Inverse Prob.* **38** (2022) 045004.

[23] S.B. Hazra and V. Schulz, Simultaneous pseudo-timestepping for PDE-model based optimization problems. *BIT* **44** (2004) 457–472.

[24] M. Heinkenschloss and D. Leykekhman, Local error estimates for SUPG solutions of advection-dominated elliptic linear-quadratic optimal control problems. *SIAM J. Numer. Anal.* **47** (2010) 4607–4638.

[25] M. Hinze, A variational discretization concept in control constrained optimization: the linear-quadratic case. *Comput. Optim. App.* **30** (2005) 45–61.

[26] W. Hu, J. Shen, J.R. Singler, Y. Zhang and X. Zheng, A superconvergent HDG method for distributed control of convection diffusion PDEs. *J. Sci. Comput.* **76** (2018) 1436–1457.

[27] H.H. Kim, E.T. Chung and C.S. Lee, A BDDC algorithm for a class of staggered discontinuous Galerkin methods. *Comput. Math. Appl.* **67** (2014) 1373–1389.

[28] D. Leykekhman, Investigation of commutative properties of discontinuous Galerkin methods in PDE constrained optimal control problems. *J. Sci. Comput.* **53** (2012) 483–511.

[29] D. Leykekhman and M. Heinkenschloss, Local error analysis of discontinuous Galerkin methods for advection-dominated elliptic linear-quadratic optimal control problems. *SIAM J. Numer. Anal.* **50** (2012) 2012–2038.

[30] J. Li and X. Tu, Convergence analysis of a balancing domain decomposition method for solving a class of indefinite linear systems. *Numer. Linear Algebra Appl.* **16** (2009) 745–773.

[31] J.-L. Lions, Optimal Control of Systems Governed by Partial Differential Equations. Vol. 170 of *Die Grundlehren der mathematischen Wissenschaften.* Springer, Berlin, Heidelberg (1971).

[32] S. Liu, Robust multigrid methods for discontinuous Galerkin discretizations of an elliptic optimal control problem. *Comput. Methods Appl. Math.* **25** (2025) 133–151.

[33] S. Liu and V. Simoncini, Multigrid preconditioning for discontinuous Galerkin discretizations of an elliptic optimal control problem with a convection-dominated state equation. *J. Sci. Comput.* **101** (2024) 79.

[34] S. Liu and J. Zhang, A balancing domain decomposition by constraints preconditioner for a hybridizable discontinuous Galerkin discretization of an elliptic optimal control problem. Preprint arXiv:2504.02072 (2025).

[35] S. Liu, Z. Tan and Y. Zhang, Discontinuous Galerkin methods for an elliptic optimal control problem with a general state equation and pointwise state constraints. *J. Comput. Appl. Math.* **437** (2024) 115494.

[36] J. Schöberl, R. Simon and W. Zulehner, A robust multigrid method for elliptic optimal control problems. *SIAM J. Numer. Anal.* **49** (2011) 1482–1503.

[37] R. Simon and W. Zulehner, On Schwarz-type Smoothers for saddle point problems with applications to PDE-constrained optimization problems. *Numer. Math.* **111** (2009) 445–468.

[38] S. Ta'asan, "One-shot" methods for optimal control of distributed parameter systems I: finite dimensional control. Technical Report ICASE Report 91-2, NASA Langley Research Center, Hampton, VA (1991).

[39] F. Tröltzsch, Optimal Control of Partial Differential Equations: Theory, Methods, and Applications. Vol. 112 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI (2010).

[40] X. Tu, A BDDC algorithm for a mixed formulation of flows in porous media. *Electron. Trans. Numer. Anal.* **20** (2005) 164–179.

[41] X. Tu, A BDDC algorithm for flow in porous media with a hybrid finite element discretization. *Electron. Trans. Numer. Anal.* **26** (2007) 146–160.

[42] X. Tu, Three-level BDDC in three dimensions. *SIAM J. Sci. Comput.* **29** (2007) 1759–1780.

[43] X. Tu, Three-level BDDC in two dimensions. *Int. J. Numer. Methods Eng.* **69** (2007) 33–59.

[44] X. Tu, A three-level BDDC algorithm for saddle point problems. *Numer. Math.* **119** (2011) 189–217.

[45] X. Tu and J. Li, A balancing domain decomposition method by constraints for advection–diffusion problems. *Commun. Appl. Math. Comput. Sci.* **3** (2008) 25–60.

[46] X. Tu and B. Wang, A BDDC algorithm for second-order elliptic problems with hybridizable discontinuous Galerkin discretizations. *Electron. Trans. Numer. Anal.* **45** (2016) 354–370.

[47] X. Tu and J. Zhang, BDDC algorithms for advection–diffusion problems with HDG discretizations. *Comput. Math. Appl.* **101** (2021) 74–106.

[48] X. Tu and J. Zhang, BDDC algorithms for Oseen problems with HDG discretizations. *IMA J. Numer. Anal.* **43** (2023) 3478–3521.

[49] X. Tu, B. Wang and J. Zhang, Analysis of BDDC algorithms for Stokes problems with hybridizable discontinuous Galerkin discretizations. *Electron. Trans. Numer. Anal.* **52** (2020) 553–570.

[50] O.B. Widlund, BDDC domain decomposition algorithms, in 75 Years of Mathematics of Computation. Vol. 754 of *Contemp. Math.* Amer. Math. Soc., Providence, RI (2020) 261–281.

[51] O.B. Widlund and C.R. Dohrmann, BDDC deluxe domain decomposition, in Domain Decomposition Methods in Science and Engineering XXII. Vol. 104 of *Lect. Notes Comput. Sci. Eng.* Springer, Cham (2016) 93–103.

[52] S. Zampini and X. Tu, Adaptive multilevel BDDC deluxe algorithms for flow in porous media. *SIAM J. Sci. Comput.* **39** (2017) A1389–A1415.

Appendix A.

Table A.1 summarizes the key operators and frequently used symbols throughout the paper.

Table A.1. Main symbols and descriptions.

Symbol	Description
$\beta > 0$	The regularization parameter of the optimal control problem.
$\boldsymbol{\zeta}$	A given divergence-free advection field, $\nabla \cdot \boldsymbol{\zeta} = 0$.
$\mathbf{\Lambda}$	A finite element space for $\boldsymbol{\lambda}$ defined in (3.6).
$\widetilde{\mathbf{\Lambda}}$	A finite element space for $\boldsymbol{\lambda}$ defined in (4.3).
$\widehat{\mathbf{\Lambda}}_\Gamma$	The degrees of freedom on the subdomain interface of $\mathbf{\Lambda}$.
$\widetilde{\mathbf{\Lambda}}_\Gamma$	The partially assembled interface space defined in (3.11).
$\mathbf{A_a}$	A block matrix defined in (2.13).
$A_1^{\mathrm{ad}},\ A_2^{\mathrm{ad}}$	Advection–diffusion matrices for y and p.
$L, \mathbf{L}_a$	Matrices defined in (2.14) and (2.15).
$\mathbf{U_A}(\boldsymbol{\lambda}), \mathbf{E}_{\boldsymbol{\mathcal{C}}}(\boldsymbol{\lambda})$	Extensions of $\boldsymbol{\lambda}$ for the optimal control problem in Definition 2.2.
$\mathbf{U}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda}), \mathbf{E}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda})$	Extensions of $\boldsymbol{\lambda}$ for the advection–diffusion problem in Definition 2.3.
$\boldsymbol{\lambda}_{\mathbf{A},\Gamma} = \mathcal{H}_{\mathbf{A}}(\boldsymbol{\lambda}_\Gamma)$	Extension of $\boldsymbol{\lambda}_\Gamma$ for the optimal control problem defined in (4.1).
$\boldsymbol{\lambda}_{\mathbf{A^{ad}},\Gamma} = \mathcal{H}_{\mathbf{A^{ad}}}(\boldsymbol{\lambda}_\Gamma)$	Extension of $\boldsymbol{\lambda}_\Gamma$ for the advection–diffusion problem defined in (4.2).
$\widetilde{\mathbf{R}}_\Gamma$	The injection operator from the space $\widehat{\mathbf{\Lambda}}_\Gamma$ into $\widetilde{\mathbf{\Lambda}}_\Gamma$.
$\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}$	$\widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} = \mathbf{D}\widetilde{\mathbf{R}}_\Gamma$, where $\mathbf{D}$ is a diagonal scaling matrix.
$\widehat{\mathbf{S}}_\Gamma$	The global Schur complement on the interface defined in (3.10).
$\widetilde{\mathbf{S}}_\Gamma$	The partially assembled interface Schur complement defined in (3.13).
$\mathbf{E_D}$	The average operator on the subdomain interface and $\mathbf{E_D} = \widetilde{\mathbf{R}}_\Gamma \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T$.
$\mathbf{T}$	The BDDC preconditioned operator and $\mathbf{T} = \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma}^T \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \widehat{\mathbf{S}}_\Gamma$.
$\mathbf{w}_\Gamma$	$\mathbf{w}_\Gamma = \widetilde{\mathbf{S}}_\Gamma^{-1} \widetilde{\mathbf{R}}_{\mathbf{D},\Gamma} \widehat{\mathbf{S}}_\Gamma \boldsymbol{\lambda}_\Gamma$.