SIAM J. NUMER. ANAL.
Vol. 61, No. 2, pp. 675–706

A SPACE-TIME MULTISCALE MORTAR MIXED FINITE ELEMENT METHOD FOR PARABOLIC EQUATIONS*

MANU JAYADHARAN†, MICHEL KERN‡, MARTIN VOHRALÍK‡, AND IVAN YOTOV†

Abstract. We develop a space-time mortar mixed finite element method for parabolic problems. The domain is decomposed into a union of subdomains discretized with nonmatching spatial grids and asynchronous time steps. The method is based on a space-time variational formulation that couples mixed finite elements in space with discontinuous Galerkin in time. Continuity of flux (mass conservation) across space-time interfaces is imposed via a coarse-scale space-time mortar variable that approximates the primary variable. Uniqueness, existence, and stability as well as a priori error estimates for the spatial and temporal errors are established. A space-time nonoverlapping domain decomposition method is developed that reduces the global problem to a space-time coarse-scale mortar interface problem. Each interface iteration involves solving in parallel space-time subdomain problems. The spectral properties of the interface operator and the convergence of the interface iteration are analyzed. Numerical experiments are provided that illustrate the theoretical results and the flexibility of the method for modeling problems with features that are localized in space and time.

Key words. space-time domain decomposition, mortar mixed finite elements, multiscale mortar method

MSC codes. 65M12, 65M22, 65M55, 65M60, 65F08

DOI. 10.1137/21M1447945

1. Introduction.

The multiscale mortar mixed finite element method of [3, 5] allows for highly efficient and accurate discretization of elliptic problems. Let a spatial domain Ω be given, which is decomposed into subdomains Ω_i. Then, on each Ω_i, an individual mesh is set, and a standard mixed finite element scheme is considered. A stand-alone mortar variable approximating the primary variable is further introduced on an independent interface mesh, which is typically coarser but where one possibly employs polynomials of higher degree. It is used to couple the subdomain problems and to ensure (multiscale) weak continuity of the normal component of the mixed finite element flux variable over the interfaces between subdomains. Moreover, the mortar variable enables a very efficient parallelization in space via a nonoverlapping domain decomposition algorithm based on reduction to an interface problem [3, 5, 19].

We introduce here a *space-time discretization* of a model parabolic equation, extending the above philosophy to time-dependent problems. Let a time interval $(0,T)$ be given. For each subdomain Ω_i, our approach considers an individual space mesh of Ω_i along with individual time stepping on $(0,T)$. On each *space-time subdomain*

* Received by the editors September 23, 2021; accepted for publication (in revised form) September 30, 2022; published electronically April 3, 2023.
https://doi.org/10.1137/21M1447945
Funding: The work of the authors was partially supported by National Science Foundation grants DMS-1818775 and DMS-2111129 and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, grant 647134 GATIPOR.

† Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260 USA (manu.jayadharan@pitt.edu, yotov@math.pitt.edu).

‡ Inria, 75589 Paris, France, and CERMICS, Ecole des Ponts, 77455 Marne-la-Vallée, France (michel.kern@inria.fr, martin.vohralik@inria.fr).

$\Omega_i \times (0,T)$, any standard mixed finite element scheme is combined with the discontinuous Galerkin (DG) time discretization [37]. Then a stand-alone mortar variable approximating the primary variable is introduced on an independent *space-time interface mesh*, which is typically coarse and where higher polynomial degrees may be used. It is used to couple the space-time subdomain problems and to ensure (multiscale) weak continuity of the normal component of the mixed finite element flux variable (and consequently mass conservation) over the space-time interfaces. This setting allows for high flexibility with individual discretizations of each space-time subdomain $\Omega_i \times (0,T)$, and in particular for *local time stepping*. Moreover, *space-time parallelization* can be achieved via reduction to a space-time interface problem requiring the solution of discrete problems on the individual space-time subdomains $\Omega_i \times (0,T)$, exchanging space-time boundary data through transmission conditions, in the spirit of space-time domain decomposition methods [17, 18, 22].

We mention some of the related previous works on local time stepping for parabolic problems in mixed formulations. The early work [15] studies finite difference methods on grids with local refinement in space and time. A similar approach is employed in [13] for transport equations. Two space-time domain decomposition methods are considered in [23]—a space-time Steklov–Poincaré operator and optimized Schwarz waveform relaxation (OSWR) [17, 18, 22] with Robin transmission conditions. Asynchronous time stepping is allowed, but the spatial grids are assumed matching. The focus is on the analysis of the iterative convergence of the OSWR method. A posteriori error estimates for these methods are developed in [2, 4] for nested time grids, with [4] also allowing for nonmatching spatial grids through the use of mortar finite elements. The methods from [23] are extended to fracture modeling in [24]. Overlapping Schwarz domain decomposition with local grid refinement in space and time for two-phase flow in porous media is developed in [28]. Domain decomposition methods for mortar mixed finite element methods for parabolic problems with nonmatching spatial grids are studied without time discretization in [20] and with uniform time stepping in [6, 16]. Local time stepping techniques have been developed for multiphysics systems coupled through interface conditions, e.g., for the Stokes–Darcy system [25, 34]. Finally, we mention some earlier works on space-time methods for parabolic problems coupling mixed finite element discretizations in space with DG in time on a single domain. In [8], a method using continuous trial and discontinuous test functions in time is developed. A posteriori error estimation and space-time adaptivity for mixed finite element–DG methods is studied in [10, 29].

To the best of the authors' knowledge, the solvability, stability, and a priori error analysis for space-time domain decomposition methods with nonmatching spatial grids and asynchronous time stepping have not been studied in the literature, which is the main goal of this paper. A key tool in the analysis is the construction of an interpolant in a space-time weakly continuous velocity space, which is used to prove a discrete divergence inf-sup condition on this space. Another key component is establishing a discrete space-time mortar inf-sup condition under a suitable assumption on the mortar space. In addition to performing complete analysis in the general case, we also consider conforming time discretizations. In this case, we provide stability and error bounds for the velocity divergence and improved error estimates. To the best of our knowledge, such a result has not been established in the literature for space-time mixed finite element methods with a DG time discretization, even on a single domain. Finally, we develop a parallel nonoverlapping domain decomposition algorithm for the solution of the resulting algebraic problem. In particular, we utilize a time-dependent Steklov–Poincaré operator approach to reduce the global problem to a

space-time interface problem. We show that the interface operator is positive definite and analyze its spectral properties. We employ an interface GMRES algorithm, which involves solving in parallel space-time subdomain problems at each iteration. The current iteration value of the mortar variable provides Dirichlet boundary data on the space-time interfaces for the subdomain problems. We emphasize that, due to the discontinuous time discretization, the space-time subdomain problems are solved using classical time marching over the local time grid. We utilize the spectral bound of the interface operator to obtain an estimate for the number of interface GMRES iterations through field-of-values analysis.

This contribution is organized as follows. In section 2, we describe the model problem and its domain decomposition weak formulation. Our space-time multiscale mortar discretization is introduced in section 3, and we prove its existence, uniqueness, and stability with respect to data in section 4. Section 5 derives a priori error estimates. Control of the velocity divergence and improved error estimates are established in section 6. The reduction to a space-time interface problem and its analysis are presented in section 7. We finally present numerical illustrations in section 8 and close with conclusions in section 9.

2. Setting. In this section we introduce the setting for our study.

2.1. Model problem. We consider a parabolic partial differential equation in a mixed form, modeling single phase flow in porous media. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a spatial polytopal domain with Lipschitz boundary and let $T > 0$ be the final time. The governing equations are

$$\mathbf{u} = -K\nabla p, \quad \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{u} = q \quad \text{in } \Omega \times (0,T], \tag{2.1a}$$

where p is the fluid pressure, $\mathbf{u}$ is the Darcy velocity, q is a source term, and K is a tensor representing the rock permeability divided by the fluid viscosity. We assume for simplicity homogeneous Dirichlet boundary condition

$$p(x,t) = 0 \text{ on } \partial\Omega \times (0,T] \tag{2.1b}$$

and assign the initial pressure

$$p(x,0) = p_0(x) \text{ on } \Omega. \tag{2.1c}$$

We assume that $q \in L^2(0,T;L^2(\Omega))$, $p_0 \in H_0^1(\Omega)$, $\nabla \cdot K\nabla p_0 \in L^2(\Omega)$ and that K is a spatially dependent, uniformly bounded, symmetric, and positive definite tensor, i.e., for constants $0 < k_{\min} \le k_{\max} < \infty$,

$$\text{for a.e. } x \in \Omega, \quad k_{\min}\zeta^T\zeta \le \zeta^T K(x)\zeta \le k_{\max}\zeta^T\zeta \quad \forall \zeta \in \mathbb{R}^d,\ d = 2,3. \tag{2.2}$$

2.2. Space-time subdomains. Let Ω be a union of nonempty nonoverlapping open polytopal subdomains with Lipschitz boundary Ω_i, $\overline{\Omega} = \cup_{i=1}^I \overline{\Omega}_i$. Let $\Gamma_i = \partial\Omega_i \setminus \partial\Omega$ be the interior boundary of Ω_i, let $\Gamma_{ij} = \Gamma_i \cap \Gamma_j$ be the interface between two adjacent subdomains Ω_i and Ω_j, and let $\Gamma = \cup\Gamma_{ij}$ be the union of all subdomain interfaces. We also introduce the space-time counterparts $\Omega^T = \Omega \times (0,T)$, $\Omega_i^T = \Omega_i \times (0,T)$, $\Gamma_i^T = \Gamma_i \times (0,T)$, $\Gamma_{ij}^T = \Gamma_{ij} \times (0,T)$, and $\Gamma^T = \Gamma \times (0,T)$; see Figure 2.1 for an illustration. We will introduce space-time domain decomposition discretizations based on Ω_i^T.

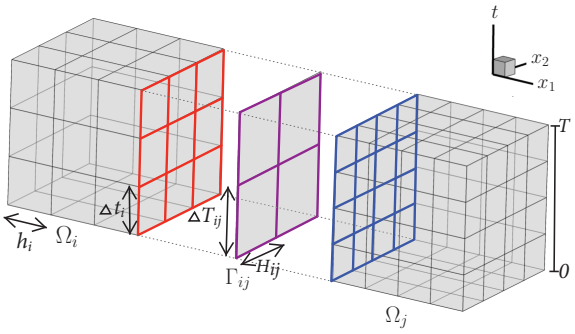


FIG. 2.1. *Nonmatching space-time subdomain and mortar grids in two spatial dimensions.*

2.3. Basic notation. We will utilize the following notation. For a domain $\mathcal{O} \subset \mathbb{R}^d$, the $L^2(\mathcal{O})$ inner product and norm for scalar and vector-valued functions are denoted by $(\cdot,\cdot)_{\mathcal{O}}$ and $\|\cdot\|_{\mathcal{O}}$, respectively. The norms and seminorms of the Sobolev spaces $W^{k,p}(\mathcal{O})$, $k \in \mathbb{R}, p \geq 1$, are denoted by $\|\cdot\|_{W^{k,p}(\mathcal{O})}$ and $|\cdot|_{W^{k,p}(\mathcal{O})}$, respectively. The norms and seminorms of the Hilbert spaces $H^k(\mathcal{O})$ are denoted by $\|\cdot\|_{H^k(\mathcal{O})}$ and $|\cdot|_{H^k(\mathcal{O})}$, respectively. For a section of a subdomain boundary $S \subset \mathbb{R}^{d-1}$ we write $\langle\cdot,\cdot\rangle_S$ and $\|\cdot\|_S$ for the $L^2(S)$ inner product (or duality pairing) and norm, respectively. By $\mathbf{M}$ we denote the vectorial counterpart of a generic scalar space M.

The above notation is extended to space-time domains as follows. For $\mathcal{O}^T = \mathcal{O} \times (0,T)$ and $S^T = S \times (0,T)$, let $(\cdot,\cdot)_{\mathcal{O}^T} = \int_0^T (\cdot,\cdot)_{\mathcal{O}}$ and $\langle\cdot,\cdot\rangle_{S^T} = \int_0^T \langle\cdot,\cdot\rangle_S$. For space-time norms we use the standard Bochner notation. For example, given a spatial norm $\|\cdot\|_V$, we denote, for $p > 0$,

$$\|\cdot\|_{L^p(0,T;V)} = \left(\int_0^T \|\cdot\|_V^p\right)^{\frac{1}{p}}, \quad \|\cdot\|_{L^\infty(0,T;V)} = \operatorname{ess\,sup} \|\cdot\|_V,$$

with the usual extension for $\|\cdot\|_{W^{k,p}(0,T;V)}$ and $\|\cdot\|_{H^k(0,T;V)}$. Let

$$\|\cdot\|_{S^T} = \|\cdot\|_{L^2(0,T;L^2(S))}.$$

We will use the space

$$\mathbf{H}(\mathrm{div};\mathcal{O}) = \left\{\mathbf{v} \in \mathbf{L}^2(\mathcal{O}) : \nabla\cdot\mathbf{v} \in L^2(\mathcal{O})\right\},$$

equipped with the norm

$$\|\mathbf{v}\|_{\mathrm{div};\mathcal{O}} = \left(\|\mathbf{v}\|_{\mathcal{O}}^2 + \|\nabla\cdot\mathbf{v}\|_{\mathcal{O}}^2\right)^{\frac{1}{2}}.$$

Finally, throughout the paper, C will denote a generic constant that is independent of the spatial and temporal discretization parameters.

2.4. Weak formulation. The weak formulation of problem (2.1) reads: Find $\mathbf{u} \in L^2(0,T;\mathbf{H}(\mathrm{div};\Omega))$ and $p \in H^1(0,T;L^2(\Omega))$ such that $p(x,0) = p_0$ and for a.e. $t \in (0,T)$,

$$(K^{-1}\mathbf{u},\mathbf{v})_\Omega - (p,\nabla\cdot\mathbf{v})_\Omega = 0 \quad \forall\,\mathbf{v} \in \mathbf{H}(\mathrm{div};\Omega), \tag{2.3a}$$

$$(\partial_t p, w)_\Omega + (\nabla\cdot\mathbf{u}, w)_\Omega = (q,w)_\Omega \quad \forall\, w \in L^2(\Omega). \tag{2.3b}$$

The following well-posedness result is rather standard and is presented in, e.g., [23, Theorem 2.1].

THEOREM 2.1 (well-posedness). *Problem* (2.3) *has a unique solution* $\mathbf{u} \in L^2(0,T;\mathbf{H}(\mathrm{div};\Omega)) \cap L^\infty(0,T;\mathbf{L}^2(\Omega))$, $p \in H^1(0,T;L^2(\Omega)) \cap L^2(0,T;H_0^1(\Omega))$.

We note that [23, Theorem 2.1] states $p \in H^1(0,T;L^2(\Omega))$. The inclusion $p \in L^2(0,T;H_0^1(\Omega))$ follows from (2.3a), which implies that for a.e. $t \in (0,T)$, $\nabla p = -K^{-1}\mathbf{u}$ in the sense of distributions.

2.5. Domain decomposition weak formulation. We now give a domain decomposition weak formulation of (2.3). Introduce the subdomain velocity and pressure spaces

$$\mathbf{V}_i = \mathbf{H}(\mathrm{div};\Omega_i),\ \mathbf{V} = \bigoplus \mathbf{V}_i, \quad W_i = L^2(\Omega_i),\ W = \bigoplus W_i = L^2(\Omega),$$

endowed with the norms

$$\|\mathbf{v}\|_{\mathbf{V}_i} = \|\mathbf{v}\|_{\mathrm{div};\Omega_i}, \quad \|\mathbf{v}\|_{\mathbf{V}} = \left(\sum_i \|\mathbf{v}\|_{\mathbf{V}_i}^2\right)^{\frac{1}{2}}, \quad \|w\|_W = \|w\|_\Omega.$$

We also introduce the following spatial bilinear forms, which will prove useful below:

$$a_i(\mathbf{u},\mathbf{v}) = (K^{-1}\mathbf{u},\mathbf{v})_{\Omega_i}, \quad a(\mathbf{u},\mathbf{v}) = \sum_i a_i(\mathbf{u},\mathbf{v}) \quad \forall\, \mathbf{u},\mathbf{v} \in \mathbf{V}, \tag{2.4a}$$

$$b_i(\mathbf{v},w) = -(\nabla\cdot\mathbf{v},w)_{\Omega_i}, \quad b(\mathbf{v},w) = \sum_i b_i(\mathbf{v},w) \quad \forall\, \mathbf{v} \in \mathbf{V}, w \in W, \tag{2.4b}$$

$$b_\Gamma(\mathbf{v},\mu) = \sum_i \langle \mathbf{v}\cdot\mathbf{n}_i, \mu\rangle_{\Gamma_i} \quad \forall\, \mathbf{v} \in \mathbf{V}, \mu \in H_0^1(\Omega)|_\Gamma. \tag{2.4c}$$

We note that $b_\Gamma(\mathbf{v},\mu)$ is well defined for $\mathbf{v} \in \mathbf{V}$ and $\mu \in H_0^1(\Omega)|_\Gamma$. Indeed, $\mu|_{\Gamma_i} \in H^{\frac{1}{2}}(\Gamma_i)$ and for $\mathbf{v} \in \mathbf{H}(\mathrm{div};\Omega_i)$, $\mathbf{v}\cdot\mathbf{n}_i \in (H^{\frac{1}{2}}(\partial\Omega_i))'$. Thus, for "interior" subdomains where $\Gamma_i = \partial\Omega_i$, $\mu|_{\partial\Omega_i} \in H^{\frac{1}{2}}(\partial\Omega_i)$ and $\langle \mathbf{v}\cdot\mathbf{n}_i,\mu\rangle_{\Gamma_i} = \langle \mathbf{v}\cdot\mathbf{n}_i,\mu\rangle_{\partial\Omega_i}$ is well defined. For "boundary" subdomains Ω_i such that $\partial\Omega_i$ has an intersection of nonzero $(d-1)$-dimensional measure with $\partial\Omega$, $\langle \mathbf{v}\cdot\mathbf{n}_i,\mu\rangle_{\Gamma_i}$ is understood as $\langle \mathbf{v}\cdot\mathbf{n}_i, E\mu\rangle_{\partial\Omega_i}$, where $E\mu \in H^{\frac{1}{2}}(\partial\Omega_i)$ is the extension by zero to $\partial\Omega_i$. We will also use $b_\Gamma(\mathbf{v},\mu)$ in the discrete setting, with $\mathbf{v}\cdot\mathbf{n}_i|_{\Gamma_i} \in L^2(\Gamma_i)$ and $\mu \in L^2(\Gamma_i)$.

For any spatial bilinear form $s(\cdot,\cdot)$, let $s^T(\cdot,\cdot) = \displaystyle\int_0^T s(\cdot,\cdot)$.

Now, since $p \in L^2(0,T;H_0^1(\Omega))$, we can consider the trace of the pressure p on the interfaces, $\lambda = p|_{\Gamma^T}$. Thus, integrating in time, it is easy to see that the solution $(\mathbf{u},p)$ of (2.3) satisfies

$$a^T(\mathbf{u},\mathbf{v}) + b^T(\mathbf{v},p) + b_\Gamma^T(\mathbf{v},\lambda) = 0 \quad \forall\, \mathbf{v} \in L^2(0,T;\mathbf{V}), \tag{2.5a}$$

$$(\partial_t p, w)_{\Omega^T} - b^T(\mathbf{u},w) = (q,w)_{\Omega^T} \quad \forall\, w \in L^2(0,T;W). \tag{2.5b}$$

3. Space-time mortar mixed finite element method. We consider a space-time discretization of (2.5), motivated by [23]. It employs a mortar finite element variable to approximate the pressure trace λ from (2.5) and uses it as a Lagrange multiplier to impose weakly the continuity of flux across space-time interfaces.

3.1. Space-time grids and spaces. Let $\mathcal{T}_{h,i}$ be a shape-regular partition of the subdomain Ω_i into parallelepipeds or simplices in the sense of [11]. We stress that this allows for grids that do not match along the interfaces Γ_{ij} between subdomains Ω_i and Ω_j. Let $h_E = \operatorname{diam} E$ for $E \in \mathcal{T}_{h,i}$, $h_i = \max_{E\in\mathcal{T}_{h,i}} h_E$, and $h = \max_i h_i$. In some parts of the analysis, we will require $\mathcal{T}_{h,i}$ to be quasi-uniform in that $h_i \leq Ch_E \ \forall E \in \mathcal{T}_{h,i}$, as well as that the mesh sizes in all subdomains be comparable in that $h \leq Ch_i \ \forall i$. Similarly, let $\mathcal{T}_i^{\Delta t} : 0 = t_i^0 < t_i^1 < \cdots < t_i^{N_i} = T$ be a partition of the time interval $(0,T)$ corresponding to subdomain Ω_i. This means that we consider different time discretizations on different subdomains. Let $\Delta t_i = \max_{1\leq k\leq N_i} |t_i^k - t_i^{k-1}|$ and $\Delta t = \max_i \Delta t_i$. Though we admit nonuniform time stepping, we will sometimes require that $\mathcal{T}_i^{\Delta t}$ be quasi-uniform in that $\Delta t_i \leq C|t_i^k - t_i^{k-1}| \ \forall (t_i^{k-1}, t_i^k) \in \mathcal{T}_i^{\Delta t}$, as well as that the time steps in all subdomains be comparable in that $\Delta t \leq C\Delta t_i \ \forall i$. Composing $\mathcal{T}_{h,i}$ and $\mathcal{T}_i^{\Delta t}$ results in a space-time tensor-product mesh

$$\mathcal{T}_{h,i}^{\Delta t} = \mathcal{T}_{h,i} \times \mathcal{T}_i^{\Delta t}$$

of the space-time subdomain Ω_i^T. An illustration is given in Figure 2.1, where yet a different mortar space-time grid is also shown in the middle.

For discretization in space, we consider any of the inf-sup stable mixed finite element spaces $\mathbf{V}_{h,i} \times W_{h,i} \subset \mathbf{V}_i \times W_i$ such as the Raviart–Thomas or Brezzi–Douglas–Marini spaces; see, e.g., [9]. For the discretization in time on the subdomains Ω_i, we will in turn utilize the DG method; cf. [37]. This is based on the space of discontinuous piecewise polynomials on the mesh $\mathcal{T}_i^{\Delta t}$ denoted by $W_i^{\Delta t}$. Composing the space and time discretization spaces by the tensor product,

$$\mathbf{V}_{h,i}^{\Delta t} = \mathbf{V}_{h,i} \bigotimes W_i^{\Delta t}, \quad W_{h,i}^{\Delta t} = W_{h,i} \bigotimes W_i^{\Delta t},$$

we obtain space-time mixed finite element spaces $\mathbf{V}_{h,i}^{\Delta t} \times W_{h,i}^{\Delta t}$ in each space-time subdomain Ω_i^T. We will also need the spatial variable only spaces

$$\mathbf{V}_h = \bigoplus \mathbf{V}_{h,i}, \quad W_h = \bigoplus W_{h,i}.$$

Let $\mathcal{T}_{H,ij}$ be a shape-regular finite element partition of Γ_{ij} (see Figure 2.1), where $H = \max_{i,j} \max_{e\in\mathcal{T}_{H,ij}} \operatorname{diam} e$. The use of index H indicates a possibly coarser interface grid compared to the subdomain grids, resulting in a multiscale approximation. Let $\mathcal{T}_{ij}^{\Delta T} : 0 = t_{ij}^0 < t_{ij}^1 < \cdots < t_{ij}^{N_{ij}} = T$ be a partition of $(0,T)$ corresponding to Γ_{ij}, which may be different from (and again possibly coarser than) the time-partitions for the neighboring subdomains. Let $\Delta T = \max_{i,j} \max_{1\leq k\leq N_{ij}} |t_{ij}^k - t_{ij}^{k-1}|$. Composing $\mathcal{T}_{H,ij}$ and $\mathcal{T}_{ij}^{\Delta T}$ gives a space-time tensor-product mesh

$$\mathcal{T}_{H,ij}^{\Delta T} = \mathcal{T}_{H,ij} \times \mathcal{T}_{ij}^{\Delta T}$$

of the space-time interface Γ_{ij}^T. Finally, let

$$\Lambda_{H,ij}^{\Delta T} = \Lambda_{H,ij} \bigotimes \Lambda_{ij}^{\Delta T}$$

be a space-time mortar finite element space on $\mathcal{T}_{H,ij}^{\Delta T}$, where $\Lambda_{H,ij}$ consists of continuous or discontinuous piecewise polynomials in space and $\Lambda_{ij}^{\Delta T}$ consists of continuous or discontinuous piecewise polynomials in time. We will also need the spatial variable space

$$\Lambda_H = \bigoplus \Lambda_{H,ij}.$$

Finally, the global space-time finite element spaces are defined as

$$\mathbf{V}_h^{\Delta t} = \bigoplus \mathbf{V}_{h,i}^{\Delta t}, \quad W_h^{\Delta t} = \bigoplus W_{h,i}^{\Delta t}, \quad \Lambda_H^{\Delta T} = \bigoplus \Lambda_{H,ij}^{\Delta T}. \tag{3.1}$$

In particular, the Lagrange multiplier will be sought in the mortar space $\Lambda_H^{\Delta T}$. For the purpose of the analysis, we also define the space of velocities with space-time weakly continuous normal components

$$\mathbf{V}_{h,0}^{\Delta t} = \left\{ \mathbf{v} \in \mathbf{V}_h^{\Delta t} : b_\Gamma^T(\mathbf{v}, \mu) = 0 \quad \forall\, \mu \in \Lambda_H^{\Delta T} \right\}. \tag{3.2}$$

The discrete velocity and pressure spaces inherit the norms $\|\cdot\|_{\mathbf{V}}$ and $\|\cdot\|_W$, respectively. The mortar space is equipped with the spatial norm $\|\mu\|_{\Lambda_H} = \|\mu\|_{L^2(\Gamma)}$.

3.2. Space-time multiscale mortar mixed finite element method. For the DG time discretization, we introduce the notation for $\varphi(x,\cdot)$, $\phi(x,\cdot) \in W_i^{\Delta t}$, $x \in \Omega_i$ (see [37]),

$$\int_0^T \tilde{\partial}_t \varphi\, \phi = \sum_{k=1}^{N_i} \int_{t_i^{k-1}}^{t_i^k} \partial_t \varphi\, \phi + \sum_{k=1}^{N_i} [\varphi]_{k-1}\, \phi_{k-1}^+, \tag{3.3}$$

where $[\varphi]_k = \varphi_k^+ - \varphi_k^-$, with $\varphi_k^+ = \lim_{t \to t_i^{k,+}} \varphi(.,t)$ and $\varphi_k^- = \lim_{t \to t_i^{k,-}} \varphi(.,t)$.

Remark 3.1 (initial value). In what follows, we will tacitly assume that a function $\varphi(x,\cdot) \in W_i^{\Delta t}$ has an associated initial value φ_0^-, which will be defined if it is explicitly used.

The space-time multiscale mortar mixed finite element method for approximating (2.5) is, Find $\mathbf{u}_h^{\Delta t} \in \mathbf{V}_h^{\Delta t}$, $p_h^{\Delta t} \in W_h^{\Delta t}$, and $\lambda_H^{\Delta T} \in \Lambda_H^{\Delta T}$ such that

$$a^T(\mathbf{u}_h^{\Delta t}, \mathbf{v}) + b^T(\mathbf{v}, p_h^{\Delta t}) + b_\Gamma^T(\mathbf{v}, \lambda_H^{\Delta T}) = 0 \quad \forall\, \mathbf{v} \in \mathbf{V}_h^{\Delta t}, \tag{3.4a}$$

$$(\tilde{\partial}_t p_h^{\Delta t}, w)_{\Omega^T} - b^T(\mathbf{u}_h^{\Delta t}, w) = (q, w)_{\Omega^T} \quad \forall\, w \in W_h^{\Delta t}, \tag{3.4b}$$

$$b_\Gamma^T(\mathbf{u}_h^{\Delta t}, \mu) = 0 \quad \forall\, \mu \in \Lambda_H^{\Delta T}, \tag{3.4c}$$

where the obvious notation $(\tilde{\partial}_t p_h^{\Delta t}, w)_{\Omega^T} = \sum_i (\tilde{\partial}_t p_h^{\Delta t}, w)_{\Omega_i^T} = \sum_i \int_0^T \int_{\Omega_i} \tilde{\partial}_t p_h^{\Delta t} w$ has been used. We note that $(\tilde{\partial}_t p_h^{\Delta t}, w)_{\Omega^T}$ involves the term $((p_h^{\Delta t})_0^+ - (p_h^{\Delta t})_0^-, w_0^+)_\Omega$; see the last term in (3.3) for $k = 1$. Here, $(p_h^{\Delta t})_0^+$ is computed by the method, while $(p_h^{\Delta t})_0^-$ is determined by the initial condition. We discuss the construction of initial data in section 4.4.

The above method provides a highly general and flexible framework, allowing for different spatial and temporal discretizations in different subdomains. We note that according to (3.4c), continuity of the flux is imposed weakly on the space-time interfaces Γ_{ij}^T, requiring that the jump in flux is orthogonal to the space-time mortar space $\Lambda_{H,ij}^{\Delta T}$. This formulation results in a correct notion of mass conservation across interfaces for time-dependent domain decomposition problems with nonmatching grids in both space and time. In the case of discontinuous mortars, (3.4c) implies that the total flux across any space-time interface cell $e \times (t_{ij}^{k-1}, t_{ij}^k)$, $e \in \mathcal{T}_{H,ij}$, is continuous.

4. Well-posedness analysis. In this section we analyze the existence, uniqueness, and stability of the solution to (3.4).

4.1. Space-time interpolants. We will make use of several space-time interpolants. Let $\mathcal{P}_{h,i}$ be the L^2-orthogonal projection onto $W_{h,i}$ and let $\mathcal{P}_i^{\Delta t}$ be the L^2-orthogonal projection onto $W_i^{\Delta t}$. We then define the L^2-orthogonal projection in space and time on the space-time subdomain Ω_i^T by

$$\mathcal{P}_{h,i}^{\Delta t} = \mathcal{P}_i^{\Delta t} \circ \mathcal{P}_{h,i} : L^2(0,T;L^2(\Omega_i)) \to W_{h,i}^{\Delta t}$$

and globally by

$$\mathcal{P}_h^{\Delta t} : L^2(0,T;L^2(\Omega)) \to W_h^{\Delta t}, \quad \mathcal{P}_h^{\Delta t}|_{\Omega_i^T} = \mathcal{P}_{h,i}^{\Delta t}.$$

Setting $\mathcal{P}_h|_{\Omega_i} = \mathcal{P}_{h,i}$ and $\mathcal{P}^{\Delta t}$ for Ω_i as $\mathcal{P}_i^{\Delta t}$, we will also write $\mathcal{P}_h^{\Delta t} = \mathcal{P}^{\Delta t} \circ \mathcal{P}_h$. Since $\nabla \cdot \mathbf{V}_{h,i} = W_{h,i}$, we have, $\forall \varphi \in L^2(0,T;L^2(\Omega_i))$,

$$(\mathcal{P}_h^{\Delta t}\varphi - \varphi, \nabla \cdot \mathbf{v})_{\Omega_i^T} = 0 \quad \forall \mathbf{v} \in \mathbf{V}_{h,i}^{\Delta t}. \tag{4.1}$$

Let $\mathbf{H}^s(\mathrm{div};\Omega_i) = \mathbf{H}^s(\Omega_i) \cap \mathbf{H}(\mathrm{div};\Omega_i)$, where $s > 0$ is a real parameter. For any real $\epsilon > 0$, let $\mathbf{\Pi}_{h,i} : \mathbf{H}^\epsilon(\mathrm{div};\Omega_i) \to \mathbf{V}_{h,i}$ be the canonical mixed interpolant [9, section III.3] and let

$$\mathbf{\Pi}_{h,i}^{\Delta t} = \mathcal{P}_i^{\Delta t} \circ \mathbf{\Pi}_{h,i} : L^2(0,T;\mathbf{H}^\epsilon(\mathrm{div};\Omega_i)) \to \mathbf{V}_{h,i}^{\Delta t}.$$

In particular, this space-time interpolant satisfies, $\forall \boldsymbol{\psi} \in L^2(0,T;\mathbf{H}^\epsilon(\mathrm{div};\Omega_i))$,

$$(\nabla \cdot (\mathbf{\Pi}_{h,i}^{\Delta t}\boldsymbol{\psi} - \boldsymbol{\psi}), w)_{\Omega_i^T} = 0 \quad \forall w \in W_{h,i}^{\Delta t}, \tag{4.2a}$$

$$\langle (\mathbf{\Pi}_{h,i}^{\Delta t}\boldsymbol{\psi} - \boldsymbol{\psi}) \cdot \mathbf{n}_i, \mathbf{v} \cdot \mathbf{n}_i \rangle_{\partial\Omega_i^T} = 0 \quad \forall \mathbf{v} \in \mathbf{V}_{h,i}^{\Delta t}, \tag{4.2b}$$

$$\|\mathbf{\Pi}_{h,i}^{\Delta t}\boldsymbol{\psi}\|_{L^2(0,T;\mathbf{V}_i)} \le C(\|\boldsymbol{\psi}\|_{L^2(0,T;\mathbf{H}^\epsilon(\Omega_i))} + \|\nabla \cdot \boldsymbol{\psi}\|_{L^2(0,T;L^2(\Omega_i))}); \tag{4.2c}$$

see [9, section III.3] for (4.2a)–(4.2b) and [32] for (4.2c).

Let $\mathbf{V}_{h,i} \cdot \mathbf{n}_i = \{\mathbf{v}_i \cdot \mathbf{n}_i : \mathbf{v}_i \in \mathbf{V}_{h,i}\}$ be the discrete normal trace space defined on the subdomain boundary $\partial\Omega_i$; it consists of discontinuous piecewise polynomials for the spaces $\mathbf{V}_{h,i}$ discussed in section 3.1. Let $\mathbf{V}_{h,i}^{\Delta t} \cdot \mathbf{n}_i$ be its tensor product with the piecewise discontinuous polynomials in time of the space $W_i^{\Delta t}$. Let $\mathcal{Q}_{h,i} : L^2(\partial\Omega_i) \to \mathbf{V}_{h,i} \cdot \mathbf{n}_i$ be the L^2-orthogonal projection and let

$$\mathcal{Q}_{h,i}^{\Delta t} = \mathcal{P}_i^{\Delta t} \circ \mathcal{Q}_{h,i} : L^2(0,T;L^2(\partial\Omega_i)) \to \mathbf{V}_{h,i}^{\Delta t} \cdot \mathbf{n}_i. \tag{4.3}$$

Finally, let $\mathcal{P}_{H,\Gamma_{ij}} : L^2(\Gamma_{ij}) \to \Lambda_{H,ij}$ and $\mathcal{P}_{ij}^{\Delta T} : L^2(0,T) \to \Lambda_{ij}^{\Delta T}$ be the L^2-orthogonal projections and let

$$\mathcal{P}_{H,\Gamma_{ij}}^{\Delta T} = \mathcal{P}_{ij}^{\Delta T} \circ \mathcal{P}_{H,\Gamma_{ij}} : L^2(0,T;L^2(\Gamma_{ij})) \to \Lambda_{H,ij}^{\Delta T}, \quad \mathcal{P}_{H,\Gamma}^{\Delta T}|_{\Gamma_{ij}} = \mathcal{P}_{H,\Gamma_{ij}}^{\Delta T} \tag{4.4}$$

be the mortar space-time L^2-orthogonal projections.

4.2. Assumptions on the mortar grids. We make the following assumptions on the mortar grids, under which we can show that the method (3.4) is well posed: there exists a positive constant C independent of the spatial mesh sizes h and H, as well as of the temporal mesh sizes Δt and ΔT, such that

$$\forall\, \mu \in \Lambda_H,\ \forall\, i,j, \quad \|\mu\|_{\Gamma_{ij}} \le C(\|\mathcal{Q}_{h,i}\,\mu\|_{\Gamma_{ij}} + \|\mathcal{Q}_{h,j}\,\mu\|_{\Gamma_{ij}}), \tag{4.5a}$$

$$\forall\, i,j, \quad \Lambda_{ij}^{\Delta T} \subset W_i^{\Delta t} \cap W_j^{\Delta t}. \tag{4.5b}$$

The spatial mortar assumption (4.5a) is the same as the assumption made in [3, 5]. Note that it is in particular satisfied with $C = \frac{1}{2}$ when $\mathcal{T}_{H,ij}$ is a coarsening of both $\mathcal{T}_{h,i}$ and $\mathcal{T}_{h,j}$ on the interface Γ_{ij} and the space $\Lambda_{H,ij}$ consists of discontinuous piecewise polynomials contained in $\mathbf{V}_{h,i}\cdot\mathbf{n}_i$ and $\mathbf{V}_{h,j}\cdot\mathbf{n}_j$ on Γ_{ij}. In general, it can be satisfied with mortar polynomial degree in space higher than the subdomain degrees and it requires that the mortar space Λ_H be sufficiently coarse, so that it is controlled by the normal traces of the neighboring subdomain velocity spaces.

The temporal mortar assumption (4.5b) similarly provides control of the mortar time discretization by the subdomain time discretizations. It requires that the mortar time discretization be a coarsening of the neighboring subdomain time discretizations and that the mortar polynomial degree in time is no higher than the subdomain degrees. We also note that (4.5a) and (4.5b) imply

$$\begin{aligned} &\forall\, \mu \in \Lambda_H^{\Delta T},\ \forall\, i,j, \\ &\|\mu\|_{L^2(0,T;L^2(\Gamma_{ij}))} \le C(\|\mathcal{Q}_{h,i}^{\Delta t}\,\mu\|_{L^2(0,T;L^2(\Gamma_{ij}))} + \|\mathcal{Q}_{h,j}^{\Delta t}\,\mu\|_{L^2(0,T;L^2(\Gamma_{ij}))}) \end{aligned} \tag{4.6}$$

for a constant C independent of h, H, Δt, and ΔT.

Remark 4.1 (assumptions (4.5) and (4.6)). We note that (4.6) is weaker than (4.5). In particular, it allows for the mortar polynomial degree in time to be of higher order compared to the subdomain time discretizations, as long as the mortar grid in time is sufficiently coarse. While we require (4.5) in the theory, the convergence studies in section 8 indicate that (4.6) may be sufficient in practice.

4.3. Discrete inf-sup conditions. Recall the form $b^T(\cdot,\cdot)$ from (2.4b). Under the above assumptions on the mortar grids, the weakly continuous velocity space $\mathbf{V}_{h,0}^{\Delta t}$ of (3.2) satisfies the following inf-sup condition.

LEMMA 4.2 (discrete divergence inf-sup condition on $\mathbf{V}_{h,0}^{\Delta t}$). *Let* (4.5) *hold. Then there exists a constant* $\beta > 0$*, independent of* h*,* H*,* Δt*, and* ΔT*, such that*

$$\forall\, w \in W_h^{\Delta t}, \qquad \sup_{0\neq \mathbf{v}\in \mathbf{V}_{h,0}^{\Delta t}} \frac{b^T(\mathbf{v},w)}{\|\mathbf{v}\|_{L^2(0,T;\mathbf{V})}} \ge \beta \|w\|_{L^2(0,T;L^2(\Omega))}. \tag{4.7}$$

Proof. Let $\mathbf{V}_{h,0} = \{\mathbf{v}\in\mathbf{V}_h : b_\Gamma(\mathbf{v},\mu) = 0 \quad \forall\, \mu\in\Lambda_H\}$. It is shown in [3, section 3] (see also [5, Lemma 3.1] for the multiscale case) that if (4.5a) holds, then there exists an interpolant $\boldsymbol{\Pi}_{h,0} : \mathbf{H}^{\frac{1}{2}+\epsilon}(\mathrm{div};\Omega) \to \mathbf{V}_{h,0}$, $\epsilon > 0$, such that, $\forall\ \boldsymbol{\psi}\in\mathbf{H}^{\frac{1}{2}+\epsilon}(\mathrm{div};\Omega)$,

$$\sum_i (\nabla\cdot(\boldsymbol{\Pi}_{h,0}\boldsymbol{\psi} - \boldsymbol{\psi}), w)_{\Omega_i} = 0 \quad \forall\, w\in W_h, \tag{4.8a}$$

$$\|\boldsymbol{\Pi}_{h,0}\boldsymbol{\psi}\|_{\mathbf{V}} \le C(\|\boldsymbol{\psi}\|_{\mathbf{H}^{\frac{1}{2}+\epsilon}(\Omega)} + \|\nabla\cdot\boldsymbol{\psi}\|_{L^2(\Omega)}) \tag{4.8b}$$

for a constant C independent of h and H. Define

$$\boldsymbol{\Pi}_{h,0}^{\Delta t} = \mathcal{P}^{\Delta t}\circ\boldsymbol{\Pi}_{h,0}.$$

We claim that $\boldsymbol{\Pi}_{h,0}^{\Delta t} : L^2(0,T;\mathbf{H}^{\frac{1}{2}+\epsilon}(\mathrm{div};\Omega)) \to \mathbf{V}_{h,0}^{\Delta t}$. To see this, note first that, for all functions $\boldsymbol{\psi}\in L^2(0,T;\mathbf{H}^{\frac{1}{2}+\epsilon}(\mathrm{div};\Omega))$, clearly $\boldsymbol{\Pi}_{h,0}^{\Delta t}\boldsymbol{\psi}\in\mathbf{V}_h^{\Delta t}$. Thus (4.5b) implies

$$\begin{aligned} b_\Gamma^T(\boldsymbol{\Pi}_{h,0}^{\Delta t}\boldsymbol{\psi},\mu) &= \sum_i \int_0^T \langle \boldsymbol{\Pi}_{h,0}^{\Delta t}\boldsymbol{\psi}\cdot\mathbf{n}_i, \mu\rangle_{\Gamma_i} = \sum_i \int_0^T \langle \boldsymbol{\Pi}_{h,0}\boldsymbol{\psi}\cdot\mathbf{n}_i, \mu\rangle_{\Gamma_i} \\ &= \int_0^T b_\Gamma(\boldsymbol{\Pi}_{h,0}\boldsymbol{\psi},\mu) = 0 \quad \forall\, \mu\in\Lambda_H^{\Delta T}, \end{aligned}$$

i.e., indeed $\boldsymbol{\Pi}_{h,0}^{\Delta t}\boldsymbol{\psi} \in \mathbf{V}_{h,0}^{\Delta t}$ by virtue of (3.2). Moreover, (4.8a) and (4.8b) imply

$$\sum_i (\nabla \cdot (\boldsymbol{\Pi}_{h,0}^{\Delta t}\boldsymbol{\psi} - \boldsymbol{\psi}), w)_{\Omega_i^T} = 0 \quad \forall\, w \in W_h^{\Delta t}, \tag{4.9a}$$

$$\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\boldsymbol{\psi}\|_{L^2(0,T;\mathbf{V})} \le C(\|\boldsymbol{\psi}\|_{L^2(0,T;\mathbf{H}^{\frac{1}{2}+\epsilon}(\Omega))} + \|\nabla\cdot\boldsymbol{\psi}\|_{L^2(0,T;L^2(\Omega))}). \tag{4.9b}$$

We complete the argument using Fortin's lemma [9, Proposition 2.8]. To that end, we first note the continuous inf-sup condition

$$\forall\, w \in L^2(0,T;L^2(\Omega)), \qquad \sup_{0\neq \mathbf{v}\in L^2(0,T;\mathbf{H}^1(\Omega))} \frac{b^T(\mathbf{v},w)}{\|\mathbf{v}\|_{L^2(0,T;\mathbf{H}^1(\Omega))}} \ge \tilde{\beta}\|w\|_{L^2(0,T;L^2(\Omega))}, \tag{4.10}$$

which can be established as follows. Given $w \in L^2(0,T;L^2(\Omega))$, let $\varphi(x,t)$ be the solution for $t \in (0,T]$ of the problem $\nabla\cdot\nabla\varphi(\cdot,t) = \tilde{w}(\cdot,t)$ in B, $\varphi = 0$ on ∂B, where B is a ball containing $\overline{\Omega}$ and $\tilde{w}$ is an extension of w by zero. Setting $\mathbf{v} = \nabla\varphi|_\Omega$ results in $\mathbf{v} \in L^2(0,T;\mathbf{H}^1(\Omega))$ such that $\nabla\cdot\mathbf{v} = w$ and, by the elliptic regularity shift (cf. [21, 30]), $\|\mathbf{v}\|_{\mathbf{H}^1(\Omega)} \le C\|w\|_{L^2(\Omega)}$, which implies (4.10). Now, using (4.9) with $0 < \epsilon \le \frac{1}{2}$ and (4.10), we have, $\forall\, w \in W_h^{\Delta t}$,

$$\begin{aligned}
&\sup_{0\neq\mathbf{v}\in\mathbf{V}_{h,0}^{\Delta t}} \frac{b^T(\mathbf{v},w)}{\|\mathbf{v}\|_{L^2(0,T;\mathbf{V})}} \\
&\quad\ge \sup_{0\neq\mathbf{v}\in L^2(0,T;\mathbf{H}^1(\Omega))} \frac{b^T(\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{v},w)}{\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{v}\|_{L^2(0,T;\mathbf{V})}} = \sup_{0\neq\mathbf{v}\in L^2(0,T;\mathbf{H}^1(\Omega))} \frac{b^T(\mathbf{v},w)}{\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{v}\|_{L^2(0,T;\mathbf{V})}} \\
&\quad\ge \frac{1}{C}\sup_{0\neq\mathbf{v}\in L^2(0,T;\mathbf{H}^1(\Omega))} \frac{b^T(\mathbf{v},w)}{\|\mathbf{v}\|_{L^2(0,T;\mathbf{H}^1(\Omega))}} \ge \frac{\tilde{\beta}}{C}\|w\|_{L^2(0,T;L^2(\Omega))},
\end{aligned}$$

which implies (4.7) with $\beta = \frac{\tilde{\beta}}{C}$. ∎

To control the mortar variable, we need the following mortar inf-sup condition.

LEMMA 4.3 (discrete mortar inf-sup condition on $\mathbf{V}_h^{\Delta t}$). *Let* (4.6) *hold. Then there exists a constant* $\beta_\Gamma > 0$*, independent of* h*,* H*,* Δt*, and* ΔT*, such that*

$$\forall\, \mu \in \Lambda_H^{\Delta T}, \qquad \sup_{0\neq\mathbf{v}\in\mathbf{V}_h^{\Delta t}} \frac{b_\Gamma^T(\mathbf{v},\mu)}{\|\mathbf{v}\|_{L^2(0,T;\mathbf{V})}} \ge \beta_\Gamma\|\mu\|_{L^2(0,T;L^2(\Gamma))}. \tag{4.11}$$

Proof. Let $\mu \in \Lambda_H^{\Delta T}$ be given. In the following we assume that μ is extended by zero on $\partial\Omega$. We consider a set of auxiliary subdomain problems. For $t \in (0,T]$, let $\varphi_i(x,t)$ be a solution of the problem (a unique weak solution in $H^1(\Omega_i)$ is obtained upon fixing the mean value)

$$\nabla\cdot\nabla\varphi_i(\cdot,t) = \frac{1}{|\Omega_i|}\langle \mathcal{Q}_{h,i}^{\Delta t}\mu, 1\rangle_{\partial\Omega_i}(t) \quad \text{in } \Omega_i, \tag{4.12a}$$

$$\nabla\varphi_i(\cdot,t)\cdot\mathbf{n}_i = (\mathcal{Q}_{h,i}^{\Delta t}\mu)(\cdot,t) \quad \text{on } \partial\Omega_i, \tag{4.12b}$$

where $|\Omega_i|$ is the measure of Ω_i. Let $\boldsymbol{\psi}_i = \nabla\varphi_i$. Elliptic regularity [21, 30] implies that $\forall\ t \in (0,T]$, there exists $s > 0$ such that

$$\|\boldsymbol{\psi}_i\|_{\mathbf{H}^s(\Omega_i)} + \|\nabla\cdot\boldsymbol{\psi}_i\|_{\Omega_i} \le C\|\mathcal{Q}_{h,i}^{\Delta t}\mu\|_{\partial\Omega_i}. \tag{4.13}$$

Let $\mathbf{v}_i = \mathbf{\Pi}_{h,i}^{\Delta t}\boldsymbol{\psi}_i \in \mathbf{V}_{h,i}^{\Delta t}$. Note that (4.2b) together with (4.3) and (4.12b) implies that $\mathbf{v}_i \cdot \mathbf{n}_i = \mathcal{Q}_{h,i}^{\Delta t}\mu$ on $\partial\Omega_i$. Thus, using definition (2.4c) of b_Γ^T, the fact that μ is extended by zero on $\partial\Omega_i \setminus \Gamma_i$, and definition (4.3) of the projection $\mathcal{Q}_{h,i}^{\Delta t}$, we have

$$
\begin{aligned}
b_\Gamma^T(\mathbf{v},\mu) &= \sum_i \langle \mathbf{\Pi}_{h,i}^{\Delta t}\boldsymbol{\psi}_i \cdot \mathbf{n}_i, \mu\rangle_{\Gamma_i^T} = \sum_i \langle \mathbf{\Pi}_{h,i}^{\Delta t}\boldsymbol{\psi}_i \cdot \mathbf{n}_i, \mu\rangle_{\partial\Omega_i^T} \\
&= \sum_i \langle \mathbf{\Pi}_{h,i}^{\Delta t}\boldsymbol{\psi}_i \cdot \mathbf{n}_i, \mathcal{Q}_{h,i}^{\Delta t}\mu\rangle_{\partial\Omega_i^T} = \sum_i \|\mathcal{Q}_{h,i}^{\Delta t}\mu\|^2_{L^2(0,T;L^2(\partial\Omega_i))} \\
&\geq C \sum_i \|\mu\|^2_{L^2(0,T;L^2(\Gamma_i))},
\end{aligned} \tag{4.14}
$$

where we used (4.6) in the inequality. On the other hand, (4.2c) with $\epsilon = s$ and (4.13), along with the stability of L^2-orthogonal projection $\mathcal{Q}_{h,i}^{\Delta t}$, imply

$$
\|\mathbf{v}_i\|_{L^2(0,T;\mathbf{V}_i)} \leq C\|\mu\|_{L^2(0,T;L^2(\Gamma_i))}. \tag{4.15}
$$

The assertion of the lemma follows from combining (4.14) and (4.15). ∎

4.4. Initial data. We next discuss the construction of discrete initial data for all variables. The data need to be compatible in the sense that they satisfy the equations without time derivatives in the method, (3.4a) and (3.4c). Recall that we are given initial pressure datum $p(0) = p_0 \in H_0^1(\Omega)$ with $\nabla \cdot K\nabla p_0 \in L^2(\Omega)$. Let us define $\mathbf{u}_0 = -K\nabla p_0$ and $\lambda_0 = p_0|_\Gamma$. Then the solution to (2.3) satisfies $\mathbf{u}(0) = \mathbf{u}_0$ and $\lambda(0) = \lambda_0$. Moreover, we have

$$
\begin{aligned}
&a(\mathbf{u}_0,\mathbf{v}) + b(\mathbf{v},p_0) + b_\Gamma(\mathbf{v},\lambda_0) = 0 \quad \forall\, \mathbf{v} \in \mathbf{V}_h, \\
&b_\Gamma(\mathbf{u}_0,\mu) = 0 \qquad \forall\, \mu \in \Lambda_H.
\end{aligned}
$$

Next, define the discrete initial data $(\mathbf{u}_{h,0}, p_{h,0}, \lambda_{H,0}) \in \mathbf{V}_h \times W_h \times \Lambda_H$ as the elliptic projection of $(\mathbf{u}_0, p_0, \lambda_0)$, i.e., the unique solution to the problem, $\forall\, \mathbf{v} \in \mathbf{V}_h$, $w \in W_h$, $\mu \in \Lambda_H$,

$$
a(\mathbf{u}_{h,0},\mathbf{v}) + b(\mathbf{v},p_{h,0}) + b_\Gamma(\mathbf{v},\lambda_{H,0}) = a(\mathbf{u}_0,\mathbf{v}) + b(\mathbf{v},p_0) + b_\Gamma(\mathbf{v},\lambda_0) = 0, \tag{4.16a}
$$

$$
b(\mathbf{u}_{h,0},w) = b(\mathbf{u}_0,w) = -(\nabla \cdot K\nabla p_0, w), \tag{4.16b}
$$

$$
b_\Gamma(\mathbf{u}_{h,0},\mu) = b_\Gamma(\mathbf{u}_0,\mu) = 0. \tag{4.16c}
$$

The well-posedness of (4.16) is shown in [3, 5] under the spatial mortar assumption (4.5a). In particular, it follows from the analysis in [3, 5] that there exist constants C_K and $\tilde{C}_K$ that depend on K, but are idependent of h and H, such that

$$
\|\mathbf{u}_{h,0}\|_{\mathbf{V}} + \|p_{h,0}\|_W + \|\lambda_{H,0}\|_{\Lambda_H} \leq C_K\|\nabla \cdot K\nabla p_0\|_\Omega, \tag{4.17}
$$

$$
\begin{aligned}
&\|\mathbf{u}_0 - \mathbf{u}_{h,0}\|_\Omega + \|p_0 - p_{h,0}\|_W + \|\lambda_0 - \lambda_{H,0}\|_{\Lambda_H} \\
&\qquad \leq \tilde{C}_K(\|\mathbf{u}_0 - \mathbf{\Pi}_{h,0}\mathbf{u}_0\|_\Omega + \|p_0 - \mathcal{P}_h p_0\|_W + \|\lambda_0 - \mathcal{P}_{H,\Gamma}\lambda_0\|_{\Lambda_H}).
\end{aligned} \tag{4.18}
$$

We now set

$$
(p_h^{\Delta t})_0^- = p_{h,0}, \quad (\mathbf{u}_h^{\Delta t})_0^- = \mathbf{u}_{h,0}, \quad (\lambda_H^{\Delta T})_0^- = \lambda_{h,0}. \tag{4.19}
$$

As we noted earlier, $(p_h^{\Delta t})_0^-$ provides an initial condition for the method (3.4); cf. (3.4b). The data $(\mathbf{u}_h^{\Delta t})_0^-$ and $(\lambda_H^{\Delta T})_0^-$ are not needed in the method, but they will be utilized in the analysis of $\nabla \cdot \mathbf{u}_h^{\Delta t}$ in section 6.

4.5. Existence, uniqueness, and stability with respect to data. In the analysis we will utilize the following auxiliary result.

LEMMA 4.4 (summation in time). *For all Ω_i and for any $\varphi(x,\cdot) \in W_i^{\Delta t}$, $x \in \Omega_i$, there holds*

$$\int_0^T \tilde{\partial}_t \varphi\, \varphi = \frac{1}{2}\left((\varphi_{N_i}^-)^2 - (\varphi_0^-)^2\right) + \frac{1}{2}\sum_{k=1}^{N_i}([\varphi]_{k-1})^2. \tag{4.20}$$

Proof. Using the definition (3.3) of $\tilde{\partial}_t \varphi$, we have

$$\begin{aligned}\int_0^T \tilde{\partial}_t \varphi\, \varphi &= \sum_{k=1}^{N_i} \int_{t_i^{k-1}}^{t_i^k} \frac{1}{2}\frac{\partial}{\partial t}\varphi^2 + \sum_{k=1}^{N_i}[\varphi]_{k-1}\varphi_{k-1}^+ \\ &= \frac{1}{2}\sum_{k=1}^{N_i}\left((\varphi_k^-)^2 - (\varphi_{k-1}^+)^2 + (\varphi_{k-1}^+)^2 - (\varphi_{k-1}^-)^2 + (\varphi_{k-1}^+ - \varphi_{k-1}^-)^2\right) \\ &= \frac{1}{2}\left((\varphi_{N_i}^-)^2 - (\varphi_0^-)^2\right) + \frac{1}{2}\sum_{k=1}^{N_i}(\varphi_{k-1}^+ - \varphi_{k-1}^-)^2. \end{aligned}$$

∎

To simplify the presentation, we introduce the notation

$$\|\varphi\|_{\mathrm{DG}}^2 = \sum_i \left(\|\varphi_{N_i}^-\|_{\Omega_i}^2 + \sum_{k=1}^{N_i} \|[\varphi]_{k-1}\|_{\Omega_i}^2 \right). \tag{4.21}$$

THEOREM 4.5 (existence and uniqueness of the discrete solution, stability with respect to data). *Assume that conditions* (4.5) *hold. Then the space-time mortar method* (3.4) *has a unique solution. Moreover, for some constant $C > 0$ independent of h, H, Δt, ΔT, $k_{\min}$, $k_{\max}$, and C_K, it holds that*

$$\begin{aligned}&\|p_h^{\Delta t}\|_{\mathrm{DG}} + k_{\max}^{-\frac{1}{2}}\|\mathbf{u}_h^{\Delta t}\|_{\Omega^T} + k_{\min}^{\frac{1}{2}}\|p_h^{\Delta t}\|_{\Omega^T} + k_{\min}^{\frac{1}{2}}\|\lambda_H^{\Delta T}\|_{\Gamma^T} \\ &\qquad \le C\left(k_{\min}^{-\frac{1}{2}}\|q\|_{\Omega^T} + C_K\|\nabla\cdot K\nabla p_0\|_{\Omega}\right).\end{aligned} \tag{4.22}$$

Proof. We begin with establishing the stability bound (4.22). Taking $\mathbf{v} = \mathbf{u}_h^{\Delta t}$, $w = p_h^{\Delta t}$, and $\mu = \lambda_H^{\Delta T}$ in (3.4) and combining the equations, we obtain, using (4.20) and Young's inequality,

$$\begin{aligned}&\frac{1}{2}\sum_i\left(\|(p_h^{\Delta t})_{N_i}^-\|_{\Omega_i}^2 + \sum_{k=1}^{N_i}\|[p_h^{\Delta t}]_{k-1}\|_{\Omega_i}^2\right) + \|K^{-\frac{1}{2}}\mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2 \\ &\qquad \le \frac{\epsilon}{2}\|p_h^{\Delta t}\|_{\Omega^T}^2 + \frac{1}{2\epsilon}\|q\|_{\Omega^T}^2 + \frac{1}{2}\|p_{h,0}\|_{\Omega}^2.\end{aligned}$$

The inf-sup condition for the weakly continuous velocity (4.7), (3.4a), and (2.2) imply

$$\|p_h^{\Delta t}\|_{\Omega^T} \le \beta^{-1}k_{\min}^{-\frac{1}{2}}\|K^{-\frac{1}{2}}\mathbf{u}_h^{\Delta t}\|_{\Omega^T}. \tag{4.23}$$

Then, taking $\epsilon = \beta^2 k_{\min}$ and using (4.23) results in

$$\|p_h^{\Delta t}\|_{\mathrm{DG}}^2 + \|K^{-\frac{1}{2}}\mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2 \le \beta^{-2}k_{\min}^{-1}\|q\|_{\Omega^T}^2 + \|p_{h,0}\|_{\Omega}^2. \tag{4.24}$$

Furthermore, the mortar inf-sup condition (4.11) and (3.4a) imply

$$\|\lambda_H^{\Delta T}\|_{\Gamma^T} \le \beta_\Gamma^{-1}(k_{\min}^{-\frac{1}{2}}\|K^{-\frac{1}{2}}\mathbf{u}_h^{\Delta t}\|_{\Omega^T} + \|p_h^{\Delta t}\|_{\Omega^T}). \tag{4.25}$$

Combining (4.23)–(4.25) and using (2.2) and (4.17), we obtain (4.22). The existence and uniqueness of a solution follow from (4.22) by taking $q=0$ and $p_0=0$. ∎

Remark 4.6 (dependence on $k_{\min}$ and $k_{\max}$). Bound (4.22) shows explicitly the dependence on $k_{\min}$ and $k_{\max}$. To simplify the presentation, we suppress this dependence in the forthcoming analysis.

Remark 4.7 (control of divergence). Control on $\|\nabla\cdot\mathbf{u}_h^{\Delta t}\|_{L^2(0,T;L^2(\Omega_i))}$ can be established under the assumption of matching time steps between subdomains and choosing the mortar finite element space in time to match the subdomains. We present this result later in section 6, along with improved error estimates.

5. A priori error analysis. In this section we derive a priori error estimates for the solution of the space-time mortar MFE method (3.4).

5.1. Approximation properties of the space-time interpolants. Assume that the spaces $\mathbf{V}_h^{\Delta t}$ and $W_h^{\Delta t}$ from (3.1) contain on each space-time element polynomials in P_k and P_l, respectively, in space and P_q in time, where P_r denotes the space of polynomials of degree up to r. Let $\Lambda_H^{\Delta T}$ contain on each space-time mortar element polynomials in P_m in space and P_s in time. We have the following approximation properties for the space-time interpolants $\mathcal{P}_h^{\Delta t}$ and $\mathcal{P}_{H,\Gamma}^{\Delta T}$ of section 4.1 and $\mathbf{\Pi}_{h,0}^{\Delta t}$ of the proof of Lemma 4.2:

$$\begin{aligned}\|\boldsymbol{\psi}-\mathbf{\Pi}_{h,0}^{\Delta t}\boldsymbol{\psi}\|_{\Omega^T} &\le C\sum_i \|\boldsymbol{\psi}\|_{H^{r_q}(0,T;\mathbf{H}^{r_k}(\Omega_i))}(h^{r_k}+\Delta t^{r_q})\\ &\quad + C\|\boldsymbol{\psi}\|_{H^{r_q}(0,T;\mathbf{H}^{\tilde{r}_k+\frac{1}{2}}(\Omega))}(h^{\tilde{r}_k}H^{\frac{1}{2}}+\Delta t^{r_q}),\\ &\quad 0<r_k\le k+1,\quad 0<\tilde{r}_k\le k+1,\quad 0\le r_q\le q+1,\end{aligned} \tag{5.1a}$$

$$\begin{aligned}\|\varphi-\mathcal{P}_h^{\Delta t}\varphi\|_{\Omega_i^T} &\le C\|\varphi\|_{H^{r_q}(0,T;H^{r_l}(\Omega_i))}(h^{r_l}+\Delta t^{r_q}),\\ &\quad 0\le r_l\le l+1,\quad 0\le r_q\le q+1,\end{aligned} \tag{5.1b}$$

$$\begin{aligned}\|\varphi-\mathcal{P}_{H,\Gamma}^{\Delta T}\varphi\|_{\Gamma_{ij}^T} &\le C\|\varphi\|_{H^{r_s}(0,T;H^{r_m}(\Gamma_{ij}))}(H^{r_m}+\Delta T^{r_s}),\\ &\quad 0\le r_m\le m+1,\quad 0\le r_s\le s+1.\end{aligned} \tag{5.1c}$$

Bound (5.1a) follows from the approximation properties of $\mathbf{\Pi}_{h,0}$ obtained in [3, section 3] and [5, Lemma 3.1]. Bounds (5.1b) and (5.1c) are standard approximation properties of the L^2 projection [11].

In the analysis we will also use the following approximation property, which follows from the stability of the L^2 projection in L^∞ [12]:

$$\begin{aligned}\|\varphi-\mathcal{P}_h^{\Delta t}\varphi\|_{L^\infty(0,T;L^2(\Omega_i))} &\le C\|\varphi\|_{W^{r_q,\infty}(0,T;H^{r_l}(\Omega_i))}(h^{r_l}+\Delta t^{r_q}),\\ &\quad 0\le r_l\le l+1,\quad 0\le r_q\le q+1.\end{aligned} \tag{5.2}$$

We also recall the well-known discrete trace (inverse) inequality for a quasi-uniform mesh $\mathcal{T}_{h,i}$: $\forall\ \mathbf{v}\in\mathbf{V}_{h,i}$, $\|\mathbf{v}\cdot\mathbf{n}_i\|_{\Gamma_i} \le Ch_i^{-\frac{1}{2}}\|\mathbf{v}\|_{\Omega_i}$. This implies

$$\forall\,\mathbf{v}\in\mathbf{V}_{h,i}^{\Delta t},\quad \|\mathbf{v}\cdot\mathbf{n}_i\|_{\Gamma_i^T} \le Ch_i^{-\frac{1}{2}}\|\mathbf{v}\|_{\Omega_i^T}. \tag{5.3}$$

5.2. A priori error estimate. We proceed with the error estimate for the space-time mortar MFE method (3.4).

THEOREM 5.1 (a priori error estimate). *Assume that conditions* (4.5) *hold and that the solution to* (2.3) *is sufficiently smooth. Let the space and time meshes* $\mathcal{T}_{h,i}$ *and* $\mathcal{T}_i^{\Delta t}$ *be quasi-uniform, and let* $h \le Ch_i$ *and* $\Delta t \le C\Delta t_i$ $\forall i$ *hold. Then there exists a constant* $C > 0$ *independent of the mesh sizes* h, H, Δt, *and* ΔT *such that the solution to the space-time mortar MFE method* (3.4) *satisfies*

$$
\begin{aligned}
&\|p - p_h^{\Delta t}\|_{\mathrm{DG}} + \|\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T} + \|p - p_h^{\Delta t}\|_{\Omega^T} + \|\lambda - \lambda_H^{\Delta T}\|_{\Gamma^T} \\
&\quad \le C\Big(\sum_i \|\mathbf{u}\|_{H^{r_q}(0,T;\mathbf{H}^{r_k}(\Omega_i))}(h^{r_k} + \Delta t^{r_q}) + \|\mathbf{u}\|_{H^{r_q}(0,T;\mathbf{H}^{\tilde{r}_k+\frac{1}{2}}(\Omega))}(h^{\tilde{r}_k} H^{\frac{1}{2}} + \Delta t^{r_q}) \\
&\quad + \sum_i \|p\|_{H^{r_q}(0,T;H^{r_l}(\Omega_i))}(h^{r_l} + \Delta t^{r_q}) + \sum_i \|p\|_{W^{r_q,\infty}(0,T;H^{r_l}(\Omega_i))}\Delta t^{-\frac{1}{2}}(h^{r_l} + \Delta t^{r_q}) \\
&\quad + \sum_{i,j} \|\lambda\|_{H^{r_s}(0,T;H^{r_m}(\Gamma_{ij}))} h^{-\frac{1}{2}}(H^{r_m} + \Delta T^{r_s}) + \sum_{i,j} \|\lambda_0\|_{H^{r_m}(\Gamma_{ij})} H^{r_m} \\
&\quad + \sum_i \|\mathbf{u}_0\|_{\mathbf{H}^{r_k}(\Omega_i)} h^{r_k} + \|\mathbf{u}_0\|_{\mathbf{H}^{\tilde{r}_k+\frac{1}{2}}(\Omega)} h^{\tilde{r}_k} H^{\frac{1}{2}} + \sum_i \|p_0\|_{H^{r_l}(\Omega_i)} h^{r_l} \Big), \\
&\qquad 0 < r_k \le k+1,\ 0 < \tilde{r}_k \le k+1,\ 0 \le r_q \le q+1, \\
&\qquad 0 \le r_l \le l+1,\ 0 \le r_m \le m+1,\ 0 \le r_s \le s+1.
\end{aligned} \tag{5.4}
$$

Proof. For the purpose of the analysis, we consider the following equivalent formulation of (3.4) in the space of weakly continuous velocities $\mathbf{V}_{h,0}^{\Delta t}$ given by (3.2): Find $\mathbf{u}_{h,0}^{\Delta t} \in \mathbf{V}_{h,0}^{\Delta t}$ and $p_h^{\Delta t} \in W_h^{\Delta t}$ such that $(p_{h,0}^{\Delta t})^- = p_{h,0}$ and

$$
a^T(\mathbf{u}_h^{\Delta t}, \mathbf{v}) + b^T(\mathbf{v}, p_h^{\Delta t}) = 0 \quad \forall\, \mathbf{v} \in \mathbf{V}_{h,0}^{\Delta t}, \tag{5.5a}
$$

$$
(\tilde{\partial}_t p_h^{\Delta t}, w)_{\Omega^T} - b^T(\mathbf{u}_h^{\Delta t}, w) = (q, w)_{\Omega^T} \quad \forall\, w \in W_h^{\Delta t}. \tag{5.5b}
$$

The fact that $\mathcal{P}_{H,\Gamma}^{\Delta T}$ defined in (4.4) maps to $\Lambda_H^{\Delta T}$ and definition (3.2) imply that $b_\Gamma^T(\mathbf{v}, \mathcal{P}_{H,\Gamma}^{\Delta T}\lambda) = 0 \ \ \forall\ \mathbf{v} \in \mathbf{V}_{h,0}^{\Delta t}$, where $\lambda = p|_\Gamma$ is the pressure trace from (2.5). Then, subtracting (5.5a)–(5.5b) from (2.5a)–(2.5b), we obtain the error equations

$$
a^T(\mathbf{u} - \mathbf{u}_h^{\Delta t}, \mathbf{v}) + b^T(\mathbf{v}, \mathcal{P}_h^{\Delta t} p - p_h^{\Delta t}) + b_\Gamma^T(\mathbf{v}, \lambda - \mathcal{P}_{H,\Gamma}^{\Delta T}\lambda) = 0 \quad \forall\, \mathbf{v} \in \mathbf{V}_{h,0}^{\Delta t}, \tag{5.6a}
$$

$$
\big(\partial_t p - \tilde{\partial}_t p_h^{\Delta t}, w\big)_{\Omega^T} - b^T(\mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}, w) = 0 \quad \forall\, w \in W_h^{\Delta t}, \tag{5.6b}
$$

where we have also used (4.1) and (4.9a) to incorporate the interpolants $\mathcal{P}_h^{\Delta t}$ and $\mathbf{\Pi}_{h,0}^{\Delta t}$. We now take $\mathbf{v} = \mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}$ and $w = \mathcal{P}_h^{\Delta t} p - p_h^{\Delta t}$ and sum the two equations, resulting in

$$
\begin{aligned}
&a^T(\mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}, \mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}) + \big(\partial_t p - \tilde{\partial}_t p_h^{\Delta t}, \mathcal{P}_h^{\Delta t} p - p_h^{\Delta t}\big)_{\Omega^T} \\
&\qquad = a^T(\mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}, \mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}) - b_\Gamma^T(\mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}, \lambda - \mathcal{P}_{H,\Gamma}^{\Delta T}\lambda).
\end{aligned} \tag{5.7}
$$

For the second term on the left of (5.7), restricted to a subdomain, we write

$$
\begin{aligned}
\int_0^T \big(\partial_t p - \tilde{\partial}_t p_h^{\Delta t}, \mathcal{P}_h^{\Delta t} p - p_h^{\Delta t}\big)_{\Omega_i} &= \int_0^T \big(\tilde{\partial}_t (p - p_h^{\Delta t}), \mathcal{P}_h^{\Delta t} p - p_h^{\Delta t}\big)_{\Omega_i} \\
&= \int_0^T \big(\tilde{\partial}_t (p - p_h^{\Delta t}), p - p_h^{\Delta t}\big)_{\Omega_i} + \int_0^T \big(\tilde{\partial}_t (p - p_h^{\Delta t}), \mathcal{P}_h^{\Delta t} p - p\big)_{\Omega_i} =: I_1^i + I_2^i.
\end{aligned} \tag{5.8}
$$

Using (4.20) and notation (4.21), for the first term, we have

$$\sum_i I_1^i = \frac{1}{2}\|p - p_h^{\Delta t}\|_{\mathrm{DG}}^2 - \frac{1}{2}\|p_0 - p_{h,0}\|_\Omega^2. \tag{5.9}$$

Using (3.3), the second term is

$$I_2^i = \sum_{k=1}^{N_i} \int_{t_i^{k-1}}^{t_i^k} \left(\partial_t(p - p_h^{\Delta t}), \mathcal{P}_h^{\Delta t} p - p\right)_{\Omega_i} + \sum_{k=1}^{N_i} \left([p - p_h^{\Delta t}]_{k-1}, (\mathcal{P}_h^{\Delta t} p - p)_{k-1}^+\right)_{\Omega_i}. \tag{5.10}$$

For the first term on the right above, using the orthogonality property of $\mathcal{P}_h^{\Delta t}$, we develop

$$\begin{aligned} \sum_{k=1}^{N_i} \int_{t_i^{k-1}}^{t_i^k} \left(\partial_t(p - p_h^{\Delta t}), \mathcal{P}_h^{\Delta t} p - p\right)_{\Omega_i} &= \sum_{k=1}^{N_i} \int_{t_i^{k-1}}^{t_i^k} \left(\partial_t(p - \mathcal{P}_h^{\Delta t} p), \mathcal{P}_h^{\Delta t} p - p\right)_{\Omega_i} \\ &= -\frac{1}{2} \sum_{k=1}^{N_i} \int_{t_i^{k-1}}^{t_i^k} \partial_t \|\mathcal{P}_h^{\Delta t} p - p\|_{\Omega_i}^2 = -\frac{1}{2} \sum_{k=1}^{N_i} \|\mathcal{P}_h^{\Delta t} p - p\|_{\Omega_i}^2 \Big|_{t^{k-1}}^{t^k}. \end{aligned} \tag{5.11}$$

Combining (5.7)–(5.11), and using (2.2) and the Cauchy–Schwarz and Young's inequalities, we obtain

$$\begin{aligned} &\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2 + \|p - p_h^{\Delta t}\|_{\mathrm{DG}}^2 \\ &\le C\Big(\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}\|_{\Omega^T} \|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T} + \sum_i \|(\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}) \cdot \mathbf{n}_i\|_{\Gamma_i^T} \|\lambda - \mathcal{P}_{H,\Gamma}^{\Delta T}\lambda\|_{\Gamma_i^T} \\ &\quad + \sum_i \sum_{k=1}^{N_i} \left(\|[p - p_h^{\Delta t}]_{k-1}\|_{\Omega_i} \|(\mathcal{P}_h^{\Delta t} p - p)_{k-1}^+\|_{\Omega_i} + \|\mathcal{P}_h^{\Delta t} p - p\|_{\Omega_i}^2 \Big|_{t^{k-1}}^{t^k} \right) \\ &\quad + \|p_0 - p_{h,0}\|_\Omega^2 \Big) \\ &\le \epsilon \left(\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2 + \|p - p_h^{\Delta t}\|_{\mathrm{DG}}^2 \right) \\ &\quad + C_\epsilon \Big(\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}\|_{\Omega^T}^2 + h^{-1} \|\lambda - \mathcal{P}_{H,\Gamma}^{\Delta T}\lambda\|_{\Gamma^T}^2 + \sum_i \sum_{k=1}^{N_i} \|(\mathcal{P}_h^{\Delta t} p - p)_{k-1}^+\|_{\Omega_i}^2 \Big) \\ &\quad + C\Big(\sum_i \sum_{k=1}^{N_i} \|\mathcal{P}_h^{\Delta t} p - p\|_{\Omega_i}^2 \Big|_{t^{k-1}}^{t^k} + \|p_0 - p_{h,0}\|_\Omega^2 \Big), \end{aligned}$$

where we used the trace (inverse) inequality (5.3) for a quasi-uniform mesh $\mathcal{T}_{h,i}$ and $h \le C h_i$ in the last estimate. Taking ϵ sufficiently small gives

$$\begin{aligned} \|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T} + \|p - p_h^{\Delta t}\|_{\mathrm{DG}} \le C\Big(&\|\boldsymbol{\Pi}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}\|_{\Omega^T} + h^{-\frac{1}{2}} \|\lambda - \mathcal{P}_{H,\Gamma}^{\Delta T}\lambda\|_{\Gamma^T} \\ &+ \Delta t^{-\frac{1}{2}} \|\mathcal{P}_h^{\Delta t} p - p\|_{L^\infty(0,T;L^2(\Omega))} + \|p_0 - p_{h,0}\|_\Omega \Big), \end{aligned} \tag{5.12}$$

where we used that $N_i \le \frac{CT}{\Delta t_i}$ for a quasi-uniform time mesh $\mathcal{T}_i^{\Delta t}$ and $\Delta t \le C \Delta t_i$ to obtain the factor $\Delta t^{-\frac{1}{2}}$.

Next, the inf-sup condition for the weakly continuous velocity (4.7) and (5.6a) imply, using (5.3) and $h \le C h_i$,

$$\|\mathcal{P}_h^{\Delta t} p - p_h^{\Delta t}\|_{\Omega^T} \le C \left(\|\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T} + h^{-\frac{1}{2}} \|\lambda - \mathcal{P}_{H,\Gamma}^{\Delta T}\lambda\|_{\Gamma^T} \right). \tag{5.13}$$

Finally, to obtain a bound on $\lambda_H^{\Delta T}$, we subtract (3.4a) from (2.5a), to obtain the error equation

(5.14)
$$a^T(\mathbf{u}-\mathbf{u}_h^{\Delta t},\mathbf{v})+b^T(\mathbf{v},p-p_h^{\Delta t})+b_\Gamma^T(\mathbf{v},\mathcal{P}_{H,\Gamma}^{\Delta T}\lambda-\lambda_H^{\Delta T})=b_\Gamma^T(\mathbf{v},\mathcal{P}_{H,\Gamma}^{\Delta T}\lambda-\lambda)\quad\forall\,\mathbf{v}\in\mathbf{V}_h^{\Delta t}.$$

The mortar inf-sup condition (4.11) and (5.14) imply, using (5.3) and $h\le Ch_i$,

$$\|\mathcal{P}_{H,\Gamma}^{\Delta T}\lambda-\lambda_H^{\Delta T}\|_{\Gamma^T}\le C\left(\|\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\Omega^T}+\|p-p_h^{\Delta t}\|_{\Omega^T}+h^{-\frac12}\|\lambda-\mathcal{P}_{H,\Gamma}^{\Delta T}\lambda\|_{\Gamma^T}\right).\tag{5.15}$$

The assertion of the theorem follows from combining (5.12), (5.13), and (5.15) and using the triangle inequality, (4.18), and the approximation bounds (5.1)–(5.2), where for the initial error we use the spatial-only version of (5.1). ∎

Remark 5.2 (the factors $\Delta t^{-\frac12}$ and $h^{-\frac12}$ and appropriate choice of the polynomial degrees m and s). The term $h^{-\frac12}(H^{r_m}+\Delta T^{r_s})$ in the error bound appears due to the use of the discrete trace (inverse) inequality (5.3) to control the consistency error $b_\Gamma^T(\mathbf{\Pi}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t},\lambda-\mathcal{P}_{H,\Gamma}^{\Delta T}\lambda)$. This term can be made comparable to the other error terms in (5.4) by choosing m and s sufficiently large, assuming that the solution is sufficiently smooth. Alternatively, this term can be improved if a bound on $\|\nabla\cdot(\mathbf{u}-\mathbf{u}_h^{\Delta t})\|_{\Omega_i^T}$ is available, with the use of the normal trace inequality for $\mathbf{H}(\mathrm{div};\Omega_i)$ functions. In this case the factor $\Delta t^{-\frac12}$ in the term $\Delta t^{-\frac12}(h^{r_l}+\Delta t^{r_q})$ can also be avoided by using a suitable time-interpolant for the pressure. We present this argument in the next section in the special case of matching time steps between subdomains; then, additionally, the assumptions on quasi-uniform space and time meshes $\mathcal{T}_{h,i}$ and $\mathcal{T}_i^{\Delta t}$ can be avoided.

6. Control of the velocity divergence and improved error estimates. In this section we establish stability and error estimates for the velocity divergence, along with an improved error bound for the rest of the variables, as noted in Remark 5.2 above. For this section only, we make the following assumption on the temporal discretization:

$$\forall\, i,j,\quad W_i^{\Delta t}=\Lambda_{ij}^{\Delta T}=W_j^{\Delta t}.\tag{6.1}$$

In particular, we assume that all subdomains and mortar interfaces have the same time discretization, which we denote by $W^{\Delta t}$. Let t^k, $k=0,\dots,N$, be the discrete times and let q be the polynomial degree in $W^{\Delta t}$. In this section, consequently, $\Delta T=\Delta t$ and $s=q$. Note that under assumption (6.1), $\mathbf{V}_{h,0}^{\Delta t}$ consists of functions that are in $\mathbf{V}_{h,0}$ pointwise in time on each time interval.

We will utilize the Radau reconstruction operator $\mathcal{I}$ [14, 31], which satisfies, for any $\varphi(x,\cdot)\in W^{\Delta t}$, $\mathcal{I}\varphi(x,\cdot)\in H^1(0,T)$, $\mathcal{I}\varphi(x,\cdot)|_{(t^{k-1},t^k)}\in P_{q+1}$, such that

$$\int_{t^{k-1}}^{t^k}\partial_t\mathcal{I}\varphi\,\phi=\int_{t^{k-1}}^{t^k}\partial_t\varphi\,\phi+[\varphi]_{k-1}\,\phi_{k-1}^+\quad\forall\,\phi(x,\cdot)\in W^{\Delta t}.$$

Recalling (3.3), this implies that

$$\int_0^T\partial_t\mathcal{I}\varphi\,\phi=\int_0^T\tilde\partial_t\varphi\,\phi\quad\forall\,\phi(x,\cdot)\in W^{\Delta t}.\tag{6.2}$$

Then the second equation (3.4b) of the space-time mortar mixed method can be rewritten as

$$(\partial_t\mathcal{I}p_h^{\Delta t},w)_{\Omega^T}-b^T(\mathbf{u}_h^{\Delta t},w)=(q,w)_{\Omega^T}\quad\forall\,w\in W_h^{\Delta t}.\tag{6.3}$$

For notational convenience, for $\mathbf{v} \in \mathbf{V}$, let $\nabla_h \cdot \mathbf{v} \in L^2(\Omega)$ be such that $(\nabla_h \cdot \mathbf{v})|_{\Omega_i} = \nabla \cdot (\mathbf{v}|_{\Omega_i})\ \forall\ i$.

6.1. Stability bound. We establish a stability bound for $\|\nabla_h \cdot \mathbf{u}_h^{\Delta t}\|_{L^2(0,T;L^2(\Omega))}$; under assumption (6.1), this complements the bound (4.22) of Theorem 4.5.

THEOREM 6.1 (control of divergence). *Assume that condition* (6.1) *holds. Then, for the solution of the space-time mortar method* (3.4)*, there exists a constant $C > 0$ independent of h, H, Δt, and ΔT such that*

$$\|\partial_t \mathcal{I} p_h^{\Delta t}\|_{\Omega^T} + \|\nabla_h \cdot \mathbf{u}_h^{\Delta t}\|_{\Omega^T} + \|\mathbf{u}_h^{\Delta t}\|_{\mathrm{DG}} \le C(\|q\|_{\Omega^T} + \|\nabla \cdot K \nabla p_0\|_{\Omega}).$$

Proof. Since $\partial_t \mathcal{I} p_h^{\Delta t} \in W_h^{\Delta t}$ and $\nabla_h \cdot \mathbf{u}_h^{\Delta t} \in W_h^{\Delta t}$, (6.3) implies that $\partial_t \mathcal{I} p_h^{\Delta t} + \nabla_h \cdot \mathbf{u}_h^{\Delta t} = \mathcal{P}_h^{\Delta t} q$. Therefore,

$$\|\partial_t \mathcal{I} p_h^{\Delta t}\|_{\Omega^T}^2 + \|\nabla_h \cdot \mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2 + 2\big(\partial_t \mathcal{I} p_h^{\Delta t}, \nabla_h \cdot \mathbf{u}_h^{\Delta t}\big)_{\Omega^T} = \|\mathcal{P}_h^{\Delta t} q\|_{\Omega^T}^2. \tag{6.4}$$

To control the third term on the left, we note that, due to (6.1), (3.4a) implies that for each $k = 1, \dots, N$ and every $t \in [t^{k-1}, t^k]$, it holds that

$$a(\mathbf{u}_h^{\Delta t}, \mathbf{v}) + b(\mathbf{v}, p_h^{\Delta t}) + b_\Gamma(\mathbf{v}, \lambda_H^{\Delta T}) = 0 \quad \forall\, \mathbf{v} \in \mathbf{V}_h.$$

Therefore, using that the initial data satisfy the above equation (see (4.19) and (4.16a)), we have

$$a^T(\partial_t \mathcal{I} \mathbf{u}_h^{\Delta t}, \mathbf{v}) + b^T(\mathbf{v}, \partial_t \mathcal{I} p_h^{\Delta t}) + b_\Gamma^T(\mathbf{v}, \partial_t \mathcal{I} \lambda_H^{\Delta T}) = 0 \quad \forall\, \mathbf{v} \in \mathbf{V}_h^{\Delta t}. \tag{6.5}$$

Taking $\mathbf{v} = \mathbf{u}_h^{\Delta t}$ in (6.5) and $\mu = \partial_t \mathcal{I} \lambda_H^{\Delta T}$ in (3.4c) and combining the equations, we obtain

$$-b^T(\mathbf{u}_h^{\Delta t}, \partial_t \mathcal{I} p_h^{\Delta t}) = a^T(\partial_t \mathcal{I} \mathbf{u}_h^{\Delta t}, \mathbf{u}_h^{\Delta t}) = \frac{1}{2}\|K^{-\frac{1}{2}} \mathbf{u}_h^{\Delta t}\|_{\mathrm{DG}}^2 - \frac{1}{2}\|K^{-\frac{1}{2}} (\mathbf{u}_h^{\Delta t})_0^-\|_{\Omega}^2 \tag{6.6}$$

using (6.2), (4.20), and (4.21) for the second equality. The assertion of the lemma follows by combining (6.4) and (6.6) and using (2.2), (4.19), and (4.17). ∎

6.2. Improved a priori error estimate. In this section we utilize the control on $\|\nabla_h \cdot \mathbf{u}_h^{\Delta t}\|_{\Omega^T}$ to obtain error estimates that avoid the factors $h^{-\frac{1}{2}}$ and $\Delta t^{-\frac{1}{2}}$ that appear in the error estimate (5.4), as well as the quasi-uniformity assumption on the space and time meshes $\mathcal{T}_{h,i}$ and $\mathcal{T}_i^{\Delta t}$. To this end, we will use an alternative time interpolant. Let $\tilde{\mathcal{P}}^{\Delta t} : H^1(0,T) \to W^{\Delta t}$ be such that, for any $\varphi \in H^1(0,T)$,

$$\forall\, k = 1, \dots, N, \quad \int_{t^{k-1}}^{t^k} (\tilde{\mathcal{P}}^{\Delta t} \varphi - \varphi) w = 0 \quad \forall\, w \in P_{q-1}, \quad (\tilde{\mathcal{P}}^{\Delta t} \varphi)_k^- = \varphi(t^k). \tag{6.7}$$

Let us further set $(\tilde{\mathcal{P}}^{\Delta t} \varphi)_0^- = \varphi(0)$ and define the space-time interpolant

$$\tilde{\mathcal{P}}_h^{\Delta t} = \tilde{\mathcal{P}}^{\Delta t} \circ \mathcal{P}_h. \tag{6.8}$$

The following properties of $\tilde{\mathcal{P}}^{\Delta t}$ and $\tilde{\mathcal{P}}_h^{\Delta t}$ will be useful in the analysis.

LEMMA 6.2 (time derivative orthogonality). *For all* $\varphi \in H^1(0,T)$ *and* $w \in W^{\Delta t}$,

$$\int_0^T \partial_t \varphi\, w = \int_0^T \tilde{\partial}_t \tilde{\mathcal{P}}^{\Delta t} \varphi\, w. \tag{6.9}$$

Furthermore, $\forall\ \varphi \in H^1(0,T;L^2(\Omega))$ *and* $w \in W_h^{\Delta t}$,

$$\int_0^T (\partial_t \varphi, w)_\Omega = \int_0^T (\tilde{\partial}_t \tilde{\mathcal{P}}_h^{\Delta t} \varphi, w)_\Omega. \tag{6.10}$$

Proof. Using (6.7), we write

$$\begin{aligned}
\int_0^T \partial_t \varphi\, w &= \sum_{k=1}^N \int_{t^{k-1}}^{t^k} \partial_t \varphi\, w = \sum_{k=1}^N \Big(- \int_{t^{k-1}}^{t^k} \varphi\, \partial_t w + \varphi(t^k) w_k^- - \varphi(t^{k-1}) w_{k-1}^+ \Big) \\
&= \sum_{k=1}^N \Big(- \int_{t^{k-1}}^{t^k} \tilde{\mathcal{P}}^{\Delta t} \varphi\, \partial_t w + (\tilde{\mathcal{P}}^{\Delta t} \varphi)_k^- w_k^- - (\tilde{\mathcal{P}}^{\Delta t} \varphi)_{k-1}^- w_{k-1}^+ \Big) \\
&= \sum_{k=1}^N \Big(\int_{t^{k-1}}^{t^k} \partial_t \tilde{\mathcal{P}}^{\Delta t} \varphi\, w + [\tilde{\mathcal{P}}^{\Delta t} \varphi]_{k-1} w_{k-1}^+ \Big) = \int_0^T \tilde{\partial}_t \tilde{\mathcal{P}}^{\Delta t} \varphi\, w,
\end{aligned}$$

where we used the definition (3.3) in the last equality. This establishes (6.9). The identity (6.10) follows from (6.9), using that $\int_0^T (\partial_t \varphi, w)_\Omega = \int_0^T (\partial_t \mathcal{P}_h \varphi, w)_\Omega$. ∎

Consider the space-time interpolants $\tilde{\boldsymbol{\Pi}}_{h,0}^{\Delta t} = \tilde{\mathcal{P}}^{\Delta t} \circ \boldsymbol{\Pi}_{h,0}$ and $\tilde{\mathcal{P}}_h^{\Delta t}$ from (6.8) in $W_h^{\Delta t}$. It is easy to see that, due to (6.1), $\tilde{\boldsymbol{\Pi}}_{h,0}^{\Delta t} \boldsymbol{\psi} \in \mathbf{V}_{h,0}^{\Delta t}$. Indeed, let $\{\mathbf{v}_j(\mathbf{x})\}$ be a basis of $\mathbf{V}_{h,0}$ and $\boldsymbol{\Pi}_{h,0} \boldsymbol{\psi}(\mathbf{x},t) = \sum_j \alpha_j(t) \mathbf{v}_j(\mathbf{x})$. Then $\tilde{\boldsymbol{\Pi}}_{h,0}^{\Delta t} \boldsymbol{\psi} = \sum_j \tilde{\mathcal{P}}^{\Delta t} \alpha_j(t) \mathbf{v}_j(\mathbf{x}) \in \mathbf{V}_{h,0}$ for any fixed t, implying that $\tilde{\boldsymbol{\Pi}}_{h,0}^{\Delta t} \boldsymbol{\psi} \in \mathbf{V}_{h,0}^{\Delta t}$. Similarly to (5.1a) and (5.1b), these interpolants satisfy

$$\begin{aligned}
&\|\boldsymbol{\psi} - \tilde{\boldsymbol{\Pi}}_{h,0}^{\Delta t} \boldsymbol{\psi}\|_{\Omega^T} \le C \sum_i \|\boldsymbol{\psi}\|_{H^{r_q}(0,T;\mathbf{H}^{r_k}(\Omega_i))} (h^{r_k} + \Delta t^{r_q}) \\
&\qquad + C \|\boldsymbol{\psi}\|_{H^{r_q}(0,T;\mathbf{H}^{\tilde{r}_k + \frac{1}{2}}(\Omega))} (h^{\tilde{r}_k} H^{\frac{1}{2}} + \Delta t^{r_q}), \\
&\qquad 0 < r_k \le k+1, \quad 0 < \tilde{r}_k \le k+1, \quad 1 \le r_q \le q+1,
\end{aligned} \tag{6.11}$$

$$\begin{aligned}
&\|\nabla \cdot (\boldsymbol{\psi} - \tilde{\boldsymbol{\Pi}}_{h,0}^{\Delta t} \boldsymbol{\psi})\|_{\Omega_i^T} \le C \|\nabla \cdot \boldsymbol{\psi}\|_{H^{r_q}(0,T;H^{r_l}(\Omega_i))} (h^{r_l} + \Delta t^{r_q}), \\
&\qquad 0 \le r_l \le l+1, \quad 1 \le r_q \le q+1,
\end{aligned} \tag{6.12}$$

$$\begin{aligned}
&\|\varphi - \tilde{\mathcal{P}}_h^{\Delta t} \varphi\|_{\Omega_i^T} \le C \|\varphi\|_{H^{r_q}(0,T;H^{r_l}(\Omega_i))} (h^{r_l} + \Delta t^{r_q}), \\
&\qquad 0 \le r_l \le l+1, \quad 1 \le r_q \le q+1.
\end{aligned} \tag{6.13}$$

Bound (6.12) follows from $\nabla \cdot \tilde{\boldsymbol{\Pi}}_{h,0} \boldsymbol{\psi} = \mathcal{P}_h \nabla \cdot \boldsymbol{\psi}$; cf. (4.8a). The use of $\tilde{\mathcal{P}}_h^{\Delta t}$ will allow us to avoid the term $\Delta t^{-\frac{1}{2}}$ in the error estimate.

We will also utilize the Scott–Zhang interpolant [35] $\mathcal{S}_{H,\Gamma} : H^1(\Gamma) \to \Lambda_{H,\Gamma} \cap C(\Gamma)$, which can be defined to preserve the trace on $\partial\Gamma$ for a function $\varphi \in H^1(\Gamma)$ such that $\varphi = 0$ on $\partial\Gamma$. Thus $\varphi - \mathcal{S}_{H,\Gamma} \varphi$ can be extended continuously by zero to $\partial\Omega \setminus \Gamma$ in the $H^{\frac{1}{2}}$-norm. Let us denote this extension by $E(\varphi - \mathcal{S}_{H,\Gamma} \varphi)$. Using the normal trace inequality, $\forall \mathbf{v} \in \mathbf{H}(\mathrm{div};\Omega_i)$, $\|\mathbf{v} \cdot \mathbf{n}_i\|_{H^{-\frac{1}{2}}(\partial\Omega_i)} \le C \|\mathbf{v}\|_{\mathbf{H}(\mathrm{div};\Omega_i)}$, we have

$$\begin{aligned}
\langle \mathbf{v} \cdot \mathbf{n}_i, \varphi - \mathcal{S}_{H,\Gamma} \varphi \rangle_{\Gamma_i} &= \langle \mathbf{v} \cdot \mathbf{n}_i, E(\varphi - \mathcal{S}_{H,\Gamma} \varphi) \rangle_{\partial\Omega_i} \\
&\le \|\mathbf{v} \cdot \mathbf{n}_i\|_{H^{-\frac{1}{2}}(\partial\Omega_i)} \|E(\varphi - \mathcal{S}_{H,\Gamma} \varphi)\|_{H^{\frac{1}{2}}(\partial\Omega_i)} \\
&\le C \|\mathbf{v}\|_{\mathbf{H}(\mathrm{div};\Omega_i)} \|\varphi - \mathcal{S}_{H,\Gamma} \varphi\|_{H^{\frac{1}{2}}(\Gamma_i)}.
\end{aligned} \tag{6.14}$$

Let $\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T} = \tilde{\mathcal{P}}^{\Delta t} \circ \mathcal{S}_{H,\Gamma}$ be the space-time mortar interpolant. It has the approximation property [35]

$$\begin{aligned} &\|\varphi - \tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\varphi\|_{L^2(0,T;H^\alpha(\Gamma_{ij}))} \le C\|\varphi\|_{H^{r_s}(0,T;H^{r_m}(\Gamma_{ij}))}(H^{r_m-\alpha} + \Delta T^{r_s}), \\ &\qquad 1 \le r_m \le m+1, \quad 1 \le r_s \le s+1, \quad 0 \le \alpha \le 1. \end{aligned} \tag{6.15}$$

Similarly to (6.14), we have $\forall\ \varphi \in L^2(0,T;H^1(\Gamma))$ such that $\varphi = 0$ on $\partial\Gamma$ for a.e. $t \in (0,T)$ and $\forall\ \mathbf{v} \in L^2(0,T;\mathbf{H}(\mathrm{div};\Omega_i))$,

$$\langle \mathbf{v}\cdot\mathbf{n}_i, \varphi - \tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\varphi\rangle_{\Gamma_i^T} \le C\|\mathbf{v}\|_{L^2(0,T;\mathbf{H}(\mathrm{div};\Omega_i))}\|\varphi - \tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\varphi\|_{L^2(0,T;H^{\frac{1}{2}}(\Gamma_i))}. \tag{6.16}$$

The use of (6.14) and (6.16) will allow us to avoid the $h^{-\frac{1}{2}}$ term in the error estimate.

We are ready to prove the following error estimate that improves the bound (5.4) of Theorem 5.1 under assumption (6.1).

THEOREM 6.3 (improved a priori error estimate). *Assume that conditions* (4.5a) *and* (6.1) *hold and that the solution to* (2.3) *is sufficiently smooth. Then there exists a constant $C > 0$ independent of the mesh sizes h, H, Δt, and ΔT, such that the solution to the space-time mortar MFE method* (3.4) *satisfies*

$$\begin{aligned} &\|\mathbf{u}(t^N) - (\mathbf{u}_h^{\Delta t})_N^-\|_\Omega + \|\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T} + \|\nabla_h\cdot(\mathbf{u} - \mathbf{u}_h^{\Delta t})\|_{\Omega^T} \\ &\quad + \|p(t^N) - (p_h^{\Delta t})_N^-\|_\Omega + \|p - p_h^{\Delta t}\|_{\Omega^T} + \|\lambda - \lambda_H^{\Delta T}\|_{\Gamma^T} \\ &\le C\Big(\sum_i \|\mathbf{u}\|_{H^{r_q}(0,T;\mathbf{H}^{r_k}(\Omega_i))}(h^{r_k} + \Delta t^{r_q}) + \|\mathbf{u}\|_{H^{r_q}(0,T;\mathbf{H}^{\tilde{r}_k+\frac{1}{2}}(\Omega))}(h^{\tilde{r}_k}H^{\frac{1}{2}} + \Delta t^{r_q}) \\ &\quad + \sum_i \|\nabla\cdot\mathbf{u}\|_{H^{r_q}(0,T;H^{r_l}(\Omega_i))}(h^{r_l} + \Delta t^{r_q}) + \sum_i \|p\|_{H^{r_q}(0,T;H^{r_l}(\Omega_i))}(h^{r_l} + \Delta t^{r_q}) \\ &\quad + \sum_i \|\mathbf{u}\|_{H^1(0,T;\mathbf{H}^{r_k}(\Omega_i))}h^{r_k} + \|\mathbf{u}\|_{H^1(0,T;\mathbf{H}^{\tilde{r}_k+\frac{1}{2}}(\Omega))}h^{\tilde{r}_k}H^{\frac{1}{2}} \\ &\quad + \sum_{i,j} \|\lambda\|_{H^{r_s}(0,T;H^{r_m}(\Gamma_{ij}))}(H^{r_m-\frac{1}{2}} + \Delta T^{r_s}) + \sum_{i,j}\|\lambda\|_{H^1(0,T;H^{r_m}(\Gamma_{ij}))}H^{r_m-\frac{1}{2}} \\ &\quad + \sum_i \|\mathbf{u}_0\|_{\mathbf{H}^{r_k}(\Omega_i)}h^{r_k} + \|\mathbf{u}_0\|_{\mathbf{H}^{\tilde{r}_k+\frac{1}{2}}(\Omega)}h^{\tilde{r}_k}H^{\frac{1}{2}} \\ &\quad + \sum_i \|p_0\|_{H^{r_l}(\Omega_i)}h^{r_l} + \sum_{i,j}\|\lambda_0\|_{H^{r_m}(\Gamma_{ij})}H^{r_m}\Big), \\ &\qquad 0 < r_k \le k+1,\ 0 < \tilde{r}_k \le k+1,\ 1 \le r_q \le q+1, \\ &\qquad 0 \le r_l \le l+1,\ \frac{1}{2} \le r_m \le m+1,\ 1 \le r_s \le s+1. \end{aligned} \tag{6.17}$$

Proof. Subtracting (5.5a)–(5.5b) from (2.5a)–(2.5b) and reasoning as below (5.5), we obtain the error equations, $\forall\,\mathbf{v} \in \mathbf{V}_{h,0}^{\Delta t}$, $w \in W_h^{\Delta t}$,

$$a^T(\mathbf{u} - \mathbf{u}_h^{\Delta t}, \mathbf{v}) + b^T(\mathbf{v}, \tilde{\mathcal{P}}_h^{\Delta t}p - p_h^{\Delta t}) + b^T(\mathbf{v}, p - \tilde{\mathcal{P}}_h^{\Delta t}p) + b_\Gamma^T(\mathbf{v}, \lambda - \tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda) = 0, \tag{6.18a}$$

$$\big(\partial_t p - \tilde{\partial}_t p_h^{\Delta t}, w\big)_{\Omega^T} - b^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}, w) - b^T(\mathbf{u} - \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}, w) = 0. \tag{6.18b}$$

We note the extra third terms in (6.18a) and (6.18b) compared to (5.6a) and (5.6b), which appear since $\tilde{\mathcal{P}}^{\Delta t}$ has orthogonality only for P_{q-1}. We take $\mathbf{v} = \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}$ and $w = \tilde{\mathcal{P}}_h^{\Delta t}p - p_h^{\Delta t}$ and sum the two equations, obtaining

$$
\begin{aligned}
&a^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t},\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t})+\big(\partial_t p-\tilde{\partial}_t p_h^{\Delta t},\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\big)_{\Omega^T}\\
&\quad=a^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u},\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t})-b^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t},p-\tilde{\mathcal{P}}_h^{\Delta t}p)\\
&\qquad-b_\Gamma^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t},\lambda-\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda)+b^T(\mathbf{u}-\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u},\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}).
\end{aligned}
\tag{6.19}
$$

For the second term on the left, using (6.10), (4.20), and (4.21), we write

$$
\begin{aligned}
\big(\partial_t p-\tilde{\partial}_t p_h^{\Delta t},\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\big)_{\Omega^T}&=\big(\tilde{\partial}_t(\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}),\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\big)_{\Omega^T}\\
&=\frac{1}{2}\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\mathrm{DG}}^2-\frac{1}{2}\|\mathcal{P}_h p_0-p_{h,0}\|_\Omega^2.
\end{aligned}
\tag{6.20}
$$

Combining (6.19) and (6.20), and using the Cauchy–Schwarz and Young's inequalities, we obtain

$$
\begin{aligned}
&\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2+\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\mathrm{DG}}^2\\
&\quad\le C\Big(\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}\|_{\Omega^T}\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\Omega^T}+\|\nabla_h\cdot(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t})\|_{\Omega^T}\|p-\tilde{\mathcal{P}}_h^{\Delta t}p\|_{\Omega^T}\\
&\qquad+\|\mathcal{P}_h p_0-p_{h,0}\|_\Omega^2+\|\nabla_h\cdot(\mathbf{u}-\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u})\|_{\Omega^T}\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\Omega^T}\\
&\qquad+\sum_i\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{L^2(0,T;\mathbf{H}(\mathrm{div};\Omega_i))}\|\lambda-\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda\|_{L^2(0,T;H^{\frac{1}{2}}(\Gamma_i))}\Big)\\
&\quad\le\epsilon_1\Big(\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2+\|\nabla_h\cdot(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t})\|_{\Omega^T}^2+\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\Omega^T}^2\Big)\\
&\qquad+C\|\mathcal{P}_h p_0-p_{h,0}\|_\Omega^2+C_{\epsilon_1}\Big(\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}\|_{\Omega^T}^2+\|\nabla_h\cdot(\mathbf{u}-\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u})\|_{\Omega^T}^2\\
&\qquad+\|p-\tilde{\mathcal{P}}_h^{\Delta t}p\|_{\Omega^T}^2+\|\lambda-\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda\|_{L^2(0,T;H^{\frac{1}{2}}(\Gamma))}^2\Big),
\end{aligned}
\tag{6.21}
$$

where we used (6.16) in the first inequality. To complete the estimate, we bound the pressure, mortar, and divergence errors.

The inf-sup condition for the weakly continuous velocity (4.7) and (6.18a) imply, using (6.16) and the triangle inequality,

$$
\begin{aligned}
\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\Omega^T}\le C\Big(&\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\Omega^T}+\|\mathbf{u}-\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}\|_{\Omega^T}\\
&+\|p-\tilde{\mathcal{P}}_h^{\Delta t}p\|_{\Omega^T}+\|\lambda-\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda\|_{L^2(0,T;H^{\frac{1}{2}}(\Gamma))}\Big).
\end{aligned}
\tag{6.22}
$$

To bound the mortar error using the mortar inf-sup condition (4.11), we subtract (3.4a) from (2.5a), to obtain the error equation

$$
a^T(\mathbf{u}-\mathbf{u}_h^{\Delta t},\mathbf{v})+b^T(\mathbf{v},p-p_h^{\Delta t})+b_\Gamma^T(\mathbf{v},\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda-\lambda_H^{\Delta T})+b_\Gamma^T(\mathbf{v},\lambda-\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda)=0\quad\forall\,\mathbf{v}\in\mathbf{V}_h^{\Delta t}.
\tag{6.23}
$$

The mortar inf-sup condition (4.11) and (6.23) imply, using (6.16) and the triangle inequality,

$$
\begin{aligned}
\|\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda-\lambda_H^{\Delta T}\|_{\Gamma^T}\le C\Big(&\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\Omega^T}+\|\mathbf{u}-\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}\|_{\Omega^T}\\
&+\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\Omega^T}+\|p-\tilde{\mathcal{P}}_h^{\Delta t}p\|_{\Omega^T}+\|\lambda-\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda\|_{L^2(0,T;H^{\frac{1}{2}}(\Gamma))}\Big).
\end{aligned}
\tag{6.24}
$$

Next, to bound the divergence error, using (6.2) and (6.10), we rewrite (6.18b) as

$$
\big(\partial_t(\mathcal{I}\tilde{\mathcal{P}}_h^{\Delta t}p-\mathcal{I}p_h^{\Delta t}),w\big)_{\Omega^T}-b^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t},w)=b^T(\mathbf{u}-\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u},w)\quad\forall\,w\in W_h^{\Delta t},
$$

concluding that $\partial_t(\mathcal{I}\tilde{\mathcal{P}}_h^{\Delta t}p - \mathcal{I}p_h^{\Delta t}) + \nabla_h \cdot (\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}) = -\mathcal{P}_h^{\Delta t}(\nabla_h \cdot (\mathbf{u} - \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}))$. Therefore,

$$\begin{aligned}\|\partial_t(\mathcal{I}\tilde{\mathcal{P}}_h^{\Delta t}p - \mathcal{I}p_h^{\Delta t})\|_{\Omega^T}^2 + \|\nabla_h \cdot (\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t})\|_{\Omega^T}^2 \\ + 2\big(\partial_t(\mathcal{I}\tilde{\mathcal{P}}_h^{\Delta t}p - \mathcal{I}p_h^{\Delta t}), \nabla_h \cdot (\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t})\big)_{\Omega^T} \\ = \|\mathcal{P}_h^{\Delta t}\big(\nabla_h \cdot (\mathbf{u} - \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u})\big)\|_{\Omega^T}^2.\end{aligned} \tag{6.25}$$

To control the third term on the left, we note that, due to (6.1), (6.5) holds, and combined with (2.5a) it implies that

$$a^T(\partial_t(\mathbf{u}-\mathcal{I}\mathbf{u}_h^{\Delta t}), \mathbf{v}) + b^T(\mathbf{v}, \partial_t(p-\mathcal{I}p_h^{\Delta t})) + b_\Gamma^T(\mathbf{v}, \partial_t(\lambda - \mathcal{I}\lambda_H^{\Delta T})) = 0 \quad \forall\, \mathbf{v} \in \mathbf{V}_h^{\Delta t}. \tag{6.26}$$

The first term is manipulated as

$$\begin{aligned}a^T(\partial_t(\mathbf{u} - \mathcal{I}\mathbf{u}_h^{\Delta t}), \mathbf{v}) &= a^T(\partial_t(\mathbf{u} - \mathbf{\Pi}_{h,0}\mathbf{u}), \mathbf{v}) + a^T(\partial_t(\mathbf{\Pi}_{h,0}\mathbf{u} - \mathcal{I}\mathbf{u}_h^{\Delta t}), \mathbf{v}) \\ &= a^T(\partial_t(\mathbf{u} - \mathbf{\Pi}_{h,0}\mathbf{u}), \mathbf{v}) + a^T(\tilde{\partial}_t(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}), \mathbf{v}),\end{aligned} \tag{6.27}$$

using (6.9) and (6.2) in the last equality. For the second term in (6.26) we write

$$b^T(\mathbf{v}, \partial_t(p - \mathcal{I}p_h^{\Delta t})) = b^T(\mathbf{v}, \partial_t(\mathcal{I}\tilde{\mathcal{P}}_h^{\Delta t}p - \mathcal{I}p_h^{\Delta t})) \tag{6.28}$$

using (6.10), (6.2), and the fact that $\nabla_h \cdot \mathbf{v} \in W_h^{\Delta t}$. The third term in (6.26) is manipulated as

$$\begin{aligned}b_\Gamma^T(\mathbf{v}, \partial_t(\lambda - \mathcal{I}\lambda_H^{\Delta T})) &= b_\Gamma^T(\mathbf{v}, \partial_t(\lambda - \mathcal{S}_{H,\Gamma}\lambda)) + b_\Gamma^T(\mathbf{v}, \partial_t(\mathcal{S}_{H,\Gamma}\lambda - \mathcal{I}\lambda_H^{\Delta T})) \\ &= b_\Gamma^T(\mathbf{v}, \partial_t(\lambda - \mathcal{S}_{H,\Gamma}\lambda)) + b_\Gamma^T(\mathbf{v}, \partial_t(\mathcal{I}\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta t}\lambda - \mathcal{I}\lambda_H^{\Delta T}))\end{aligned} \tag{6.29}$$

using (6.9) and (6.2) in the last equality. Now, combining (6.26)–(6.29), taking $\mathbf{v} = \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t} \in \mathbf{V}_{h,0}^{\Delta t}$, and using that $\partial_t(\mathcal{I}\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta t}\lambda - \mathcal{I}\lambda_H^{\Delta T}) \in \Lambda_H^{\Delta T}$ to deduce that the last term in (6.29) is zero, we obtain

$$\begin{aligned}\big(\partial_t(\mathcal{I}\tilde{\mathcal{P}}_h^{\Delta t}p - \mathcal{I}p_h^{\Delta t}), \nabla_h \cdot (\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t})\big)_{\Omega^T} &= a^T(\tilde{\partial}_t(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}), \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}) \\ &\quad + a^T(\partial_t(\mathbf{u} - \mathbf{\Pi}_{h,0}\mathbf{u}), \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}) + b_\Gamma^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}, \partial_t(\lambda - \mathcal{S}_{H,\Gamma}\lambda)).\end{aligned} \tag{6.30}$$

Combining (6.25) and (6.30) and using (4.20), we arrive at

$$\begin{aligned}&\|\partial_t(\mathcal{I}\tilde{\mathcal{P}}_h^{\Delta t}p - \mathcal{I}p_h^{\Delta t})\|_{\Omega^T}^2 + \|\nabla_h \cdot (\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t})\|_{\Omega^T}^2 + \|K^{-1/2}(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t})\|_{\mathrm{DG}}^2 \\ &\quad = -2a^T(\partial_t(\mathbf{u} - \mathbf{\Pi}_{h,0}\mathbf{u}), \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}) - 2b_\Gamma^T(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}, \partial_t(\lambda - \mathcal{S}_{H,\Gamma}\lambda)) \\ &\qquad + \|\mathbf{\Pi}_{h,0}\mathbf{u}_0 - \mathbf{u}_{h,0}\|_\Omega^2 + \|\mathcal{P}_h^{\Delta t}\big(\nabla_h \cdot (\mathbf{u} - \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u})\big)\|_{\Omega^T}^2 \\ &\quad \le \epsilon_2\Big(\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2 + \|\nabla_h \cdot (\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u} - \mathbf{u}_h^{\Delta t})\|_{\Omega^T}^2\Big) + \|\nabla_h \cdot (\mathbf{u} - \tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u})\|_{\Omega^T}^2 \\ &\qquad + C\Big(\|\partial_t(\mathbf{u} - \mathbf{\Pi}_{h,0}\mathbf{u})\|_{\Omega^T}^2 + \|\partial_t(\lambda - \mathcal{S}_{H,\Gamma}\lambda)\|^2_{L^2(0,T;H^{\frac{1}{2}}(\Gamma))}\Big) + \|\mathbf{\Pi}_{h,0}\mathbf{u}_0 - \mathbf{u}_{h,0}\|_\Omega^2,\end{aligned} \tag{6.31}$$

where we used the Cauchy–Schwarz and Young's inequalities, as well as (6.14). Combining (6.21), (6.22), (6.24), and (6.31), taking ϵ_1 and ϵ_2 small enough, and using (2.2), we obtain

$$
\begin{aligned}
&\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\Omega^T}^2+\|\nabla_h\cdot(\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t})\|_{\Omega^T}^2+\|\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{\mathrm{DG}}^2\\
&\quad+\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\Omega^T}^2+\|\tilde{\mathcal{P}}_h^{\Delta t}p-p_h^{\Delta t}\|_{\mathrm{DG}}^2+\|\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda-\lambda_H^{\Delta T}\|_{\Gamma^T}^2\\
&\le C\Big(\|\mathbf{u}-\tilde{\mathbf{\Pi}}_{h,0}^{\Delta t}\mathbf{u}\|_{L^2(0,T;\mathbf{V})}^2+\|p-\tilde{\mathcal{P}}_h^{\Delta t}p\|_{\Omega^T}^2+\|\lambda-\tilde{\mathcal{S}}_{H,\Gamma}^{\Delta T}\lambda\|_{L^2(0,T;H^{\frac12}(\Gamma))}^2\\
&\quad+\|\partial_t(\mathbf{u}-\mathbf{\Pi}_{h,0}\mathbf{u})\|_{\Omega^T}^2+\|\partial_t(\lambda-\mathcal{S}_{H,\Gamma}\lambda)\|_{L^2(0,T;H^{\frac12}(\Gamma))}^2\\
&\quad+\|\mathbf{\Pi}_{h,0}\mathbf{u}_0-\mathbf{u}_{h,0}\|_{\Omega}^2+\|\mathcal{P}_hp_0-p_{h,0}\|_{\Omega}^2\Big).
\end{aligned}
$$

Finally, using the triangle inequality, the approximation properties (6.11)–(6.15), and the initial error bound (4.18), we arrive at (6.17). We remark that in the final bound, we kept only the term at time t^N from the norm $\|\cdot\|_{\mathrm{DG}}$, since the approximation error in the full $\|\cdot\|_{\mathrm{DG}}$ norm involves a $\Delta t^{-\frac12}$ factor, and used that $(\tilde{\mathcal{P}}^{\Delta t}\varphi)_N^-=\varphi(t^N)$; cf. (6.7). ∎

Remark 6.4 (significance of the improved error estimate). The error estimate (6.17) avoids the factors $h^{-\frac12}$ and $\Delta t^{-\frac12}$, which appeared in the earlier bound (5.4), and thus provides optimal order of convergence for all variables. Moreover, it provides a bound on the velocity divergence error. To the best of the authors' knowledge, such a result has not been established in the literature for space-time mixed finite element methods with a DG time discretization, even on a single domain.

7. Reduction to an interface problem. In this section we combine the time-dependent Steklov–Poincaré operator approach from [23] with the mortar domain decomposition algorithm from [3, 5] to reduce the global problem (3.4) to a space-time interface problem.

7.1. Decomposition of the solution. Consider a decomposition of the solution to (3.4) in the form

$$
\mathbf{u}_h^{\Delta t}=\mathbf{u}_h^{\Delta t,*}(\lambda_H^{\Delta T})+\overline{\mathbf{u}}_h^{\Delta t},\quad p_h^{\Delta t}=p_h^{\Delta t,*}(\lambda_H^{\Delta T})+\overline{p}_h^{\Delta t}. \tag{7.1}
$$

Here, $\overline{\mathbf{u}}_h^{\Delta t}\in\mathbf{V}_h^{\Delta t}$, $\overline{p}_h^{\Delta t}\in W_h^{\Delta t}$ are such that for each Ω_i^T, $(\overline{\mathbf{u}}_h^{\Delta t}|_{\Omega_i^T}\in\mathbf{V}_{h,i}^{\Delta t},\overline{p}_h^{\Delta t}|_{\Omega_i^T}\in W_{h,i}^{\Delta t})$ is the solution to the space-time subdomain problem in Ω_i^T with zero Dirichlet data on the space-time interfaces and the prescribed source term, initial data as in (4.19), and boundary data on the external boundary ($p=0$ on $\partial\Omega$ in the present setting):

$$
a_i^T(\overline{\mathbf{u}}_h^{\Delta t},\mathbf{v})+b_i^T(\mathbf{v},\overline{p}_h^{\Delta t})=0\quad\forall\,\mathbf{v}\in\mathbf{V}_{h,i}^{\Delta t}, \tag{7.2a}
$$

$$
(\tilde{\partial}_t\overline{p}_h^{\Delta t},w)_{\Omega_i^T}-b_i^T(\overline{\mathbf{u}}_h^{\Delta t},w)=(q,w)_{\Omega_i^T}\quad\forall\,w\in W_{h,i}^{\Delta t}. \tag{7.2b}
$$

Furthermore, for a given $\mu\in\Lambda_H^{\Delta T}$, $\mathbf{u}_h^{\Delta t,*}(\mu)\in\mathbf{V}_h^{\Delta t}$, $p_h^{\Delta t,*}(\mu)\in W_h^{\Delta t}$ are such that for each Ω_i^T, $(\mathbf{u}_h^{\Delta t,*}(\mu)|_{\Omega_i^T}\in\mathbf{V}_{h,i}^{\Delta t},p_h^{\Delta t,*}(\mu)|_{\Omega_i^T}\in W_{h,i}^{\Delta t})$ is the solution to the space-time subdomain problem in Ω_i^T with Dirichlet data μ on the space-time interfaces and zero source term, initial data, and boundary data on the external boundary:

$$
a_i^T(\mathbf{u}_h^{\Delta t,*}(\mu),\mathbf{v})+b_i^T(\mathbf{v},p_h^{\Delta t,*}(\mu))=-\langle\mathbf{v}\cdot\mathbf{n}_i,\mu\rangle_{\Gamma_i^T}\quad\forall\,\mathbf{v}\in\mathbf{V}_{h,i}^{\Delta t}, \tag{7.3a}
$$

$$
(\tilde{\partial}_tp_h^{\Delta t,*}(\mu),w)_{\Omega_i^T}-b_i^T(\mathbf{u}_h^{\Delta t,*}(\mu),w)=0\quad\forall\,w\in W_{h,i}^{\Delta t}. \tag{7.3b}
$$

Note that both (7.2) and (7.3) are posed in the individual space-time subdomains Ω_i^T and can thus be solved in parallel (on the entire space-time subdomains Ω_i^T, without

any synchronization on time steps). It is easy to check that (3.4) is equivalent to solving the space-time interface problem: find $\lambda_H^{\Delta T} \in \Lambda_H^{\Delta T}$ such that

$$-b_\Gamma^T(\mathbf{u}_h^{\Delta t,*}(\lambda_H^{\Delta T}),\mu) = b_\Gamma^T(\overline{\mathbf{u}}_h^{\Delta t},\mu) \quad \forall\, \mu \in \Lambda_H^{\Delta T}, \tag{7.4}$$

and obtaining $\mathbf{u}_h^{\Delta t}$ and $p_h^{\Delta t}$ from (7.1)–(7.3).

7.2. Space-time Steklov–Poincaré operator. The above problem can be written in an operator form: find $\lambda_H^{\Delta T} \in \Lambda_H^{\Delta T}$ such that

$$S\,\lambda_H^{\Delta T} = g, \tag{7.5}$$

where $S : \Lambda_H^{\Delta T} \to \Lambda_H^{\Delta T}$ is the space-time Steklov–Poincaré operator defined as

$$\langle S\lambda,\mu\rangle_{\Gamma^T} = \sum_i \langle S_i\lambda,\mu\rangle_{\Gamma_i^T}, \quad \langle S_i\lambda,\mu\rangle_{\Gamma_i^T} = -\langle \mathbf{u}_h^{\Delta t,*}(\lambda)\cdot\mathbf{n}_i,\mu\rangle_{\Gamma_i^T} \quad \forall\,\lambda,\mu\in\Lambda_H^{\Delta T}, \tag{7.6}$$

and $g \in \Lambda_H^{\Delta T}$ is defined as $\langle g,\mu\rangle_{\Gamma^T} = b_\Gamma(\overline{\mathbf{u}}_h^{\Delta t},\mu)\,\forall\,\mu\in\Lambda_H^{\Delta T}$.

LEMMA 7.1 (space-time Steklov–Poincaré operator). *Assume that conditions* (4.5) *hold. Then the operator S defined in* (7.6) *is positive definite.*

Proof. For $\lambda,\mu\in\Lambda_H^{\Delta T}$, consider (7.3a) with data μ and test function $\mathbf{v} = \mathbf{u}_h^{\Delta t,*}(\lambda)$. This implies, using (7.6),

$$\begin{aligned}\langle S\lambda,\mu\rangle_{\Gamma^T} &= a^T(\mathbf{u}_h^{\Delta t,*}(\mu),\mathbf{u}_h^{\Delta t,*}(\lambda)) + b^T(\mathbf{u}_h^{\Delta t,*}(\lambda),p_h^{\Delta t,*}(\mu))\\ &= a^T(\mathbf{u}_h^{\Delta t,*}(\mu),\mathbf{u}_h^{\Delta t,*}(\lambda)) + (\tilde\partial_t p_h^{\Delta t,*}(\lambda),p_h^{\Delta t,*}(\mu))_{\Omega^T},\end{aligned} \tag{7.7}$$

where we have used (7.3b) with data λ and test function $p_h^{\Delta t,*}(\mu)$ in the second equality. Lemma 4.4 together with $p_h^{\Delta t,*}(\mu)(x,0) = 0$ (recall that zero initial data is supposed in (7.3)) imply that

$$\langle S\mu,\mu\rangle_{\Gamma^T} \ge a^T(\mathbf{u}_h^{\Delta t,*}(\mu),\mathbf{u}_h^{\Delta t,*}(\mu)) \ge 0 \quad \forall\,\mu\in\Lambda_H^{\Delta T}, \tag{7.8}$$

hence S is positive semidefinite. Assume that $\langle S\mu,\mu\rangle_{\Gamma^T} = 0$. Then $\mathbf{u}_h^{\Delta t,*}(\mu) = 0$. The inf-sup condition for the weakly continuous velocity (4.7) and (7.3a) imply $p_h^{\Delta t,*}(\mu) = 0$. Then the mortar inf-sup condition (4.11) and (7.3a) imply $\mu = 0$, thus S is positive definite. ∎

Due to Lemma 7.1, GMRES can be employed to solve the interface problem (7.5). On each GMRES iteration, the dominant computational cost is the evaluation of the action of S, which requires solving space-time problems with prescribed Dirichlet interface data in each individual space-time subdomain $\Omega_i\times(0,T)$. The following result can be used to provide a bound on the number of GMRES iterations.

THEOREM 7.2 (spectral bound). *Assume that conditions* (4.5) *hold. Let $\mathcal{T}_{h,i}$ be quasi-uniform and $h \le Ch_i$ $\forall i$. Then there exist positive constants C_0 and C_1 independent of the mesh sizes h, H, Δt, and ΔT such that*

$$\forall\,\mu\in\Lambda_H^{\Delta T}, \quad C_0\|\mu\|_{\Gamma^T}^2 \le \langle S\mu,\mu\rangle_{\Gamma^T} \le C_1 h^{-1}\|\mu\|_{\Gamma^T}^2. \tag{7.9}$$

Proof. Using (7.6), the Cauchy–Schwarz inequality, (5.3), and $h \le Ch_i$, we obtain

$$\begin{aligned}\langle S_i\mu,\mu\rangle_{\Gamma_i^T} &\le \|\mathbf{u}_h^{\Delta t,*}(\mu)\cdot\mathbf{n}_i\|_{\Gamma_i^T}\|\mu\|_{\Gamma_i^T} \le Ch^{-\frac12}\|\mathbf{u}_h^{\Delta t,*}(\mu)\|_{\Omega_i^T}\|\mu\|_{\Gamma_i^T}\\ &\le Ch^{-\frac12}\langle S_i\mu,\mu\rangle_{\Gamma_i^T}^{\frac12}\|\mu\|_{\Gamma_i^T},\end{aligned}$$

where we used (7.8), which is also valid on each Ω_i^T, in the last inequality. This implies the upper bound in (7.9).

To prove the lower bound in (7.9), we consider the set of auxiliary subdomain problems (4.12) with data μ. Let $\mathbf{v}_i = \mathbf{\Pi}_{h,i}^{\Delta t} \boldsymbol{\psi}_i$ and recall that $\mathbf{v}_i \cdot \mathbf{n}_i = \mathcal{Q}_{h,i}^{\Delta t} \mu$. Using (4.6) and (7.3a), we have

$$\begin{aligned}
\|\mu\|_{\Gamma^T}^2 &\le C \sum_i \langle \mathcal{Q}_{h,i}^{\Delta t} \mu, \mathcal{Q}_{h,i}^{\Delta t} \mu \rangle_{\Gamma_i^T} = C \sum_i \langle \mathcal{Q}_{h,i}^{\Delta t} \mu, \mu \rangle_{\Gamma_i^T} = C \sum_i \langle \mathbf{v}_i \cdot \mathbf{n}_i, \mu \rangle_{\Gamma_i^T}, \\
&= -C \sum_i \left(a_i^T(\mathbf{u}_h^{\Delta t,*}(\mu), \mathbf{v}_i) + b_i^T(\mathbf{v}_i, p_h^{\Delta t,*}(\mu)) \right) \\
&\le C \sum_i \left(\|\mathbf{u}_h^{\Delta t,*}(\mu)\|_{\Omega_i^T}^2 + \|p_h^{\Delta t,*}(\mu)\|_{\Omega_i^T}^2 \right)^{\frac{1}{2}} \|\mathbf{v}_i\|_{L^2(0,T;\mathbf{V}_i)} \\
&\le C \left\{ \sum_i \|\mathbf{u}_h^{\Delta t,*}(\mu)\|_{\Omega_i^T}^2 \right\}^{\frac{1}{2}} \left\{ \sum_i \|\mu\|_{\Gamma_i^T}^2 \right\}^{\frac{1}{2}} \\
&\le C \langle S\mu, \mu \rangle_{\Gamma^T}^{\frac{1}{2}} \|\mu\|_{\Gamma^T}.
\end{aligned}$$

In the next to last inequality above, we used the Cauchy–Schwarz inequality together with the inf-sup condition (4.7) and (7.3a) to bound $\|p_h^{\Delta t,*}(\mu)\|_{\Omega^T}$ and the elliptic regularity (4.15) to bound $\|\mathbf{v}_i\|_{L^2(0,T;\mathbf{V}_i)}$. In the last inequality we used (7.8). This concludes the proof. ∎

Theorem 7.2, combined with field-of-values analysis [36], implies convergence of GMRES for solving the interface problem (7.5), with the kth GMRES residual bounded by

$$\|\mathbf{r}_k\| \le \left(\sqrt{1 - (C_0/C_1)^2 h^2} \right)^k \|\mathbf{r}_0\|. \tag{7.10}$$

8. Numerical results. In this section, we present several numerical results obtained with the space-time mortar method developed in section 3.2, illustrating the convergence rates and other theoretical results obtained in the previous sections, as well as its performance for problems with localized features in space and time and discontinuous permeability. The method is implemented for distributed memory parallel computers using the deal.II finite element package [7] with message passing interface (MPI).

In all examples we consider two-dimensional spatial domains and take the mixed finite element spaces $\mathbf{V}_{h,i} \times W_{h,i}$ on the spatial subdomain Ω_i to be the lowest-order Raviart–Thomas pair $RT_0 \times DGQ_0$ (i.e., $k = l = 0$) on rectangular meshes [9], where DGQ_r denotes the space of discontinuous piecewise polynomials of degree up to r in each variable. Combining this with the lowest-order DG (backward Euler, $q = 0$) for time discretization on the mesh $\mathcal{T}_i^{\Delta t}$ gives us a space-time mixed finite element space $\mathbf{V}_{h,i}^{\Delta t} \times W_{h,i}^{\Delta t}$ in Ω_i^T as detailed in section 3.1. In the convergence test in Example 1 we test two different choices for the mortar finite element space $\Lambda_{H,ij}^{\Delta T}$ on the space-time interface mesh $\mathcal{T}_{H,ij}^{\Delta T}$, with ΔT suitably chosen as a function of Δt, depending on the mortar space polynomial degree. These are discontinuous bilinear DGQ_1 ($m = s = 1$) and biquadratic DGQ_2 ($m = s = 2$) mortars, i.e., $\Lambda_{H,ij}$ and $\Lambda_{ij}^{\Delta T}$ consist respectively of discontinuous piecewise first- or second-order polynomials in space/time. In Examples 2 and 3 we use discontinuous bilinear mortars.

For solving the interface problem identified in section 7.2, we have implemented the GMRES algorithm without a preconditioner. Developing a preconditioner for

TABLE 8.1
Example 1*, mesh size and number of degrees of freedom for the subdomains.*

Ref.	Ω_1^T			Ω_2^T			Ω_3^T			Ω_4^T		
No.	n_1	N_1	#DoF	n_2	N_2	#DoF	n_3	N_3	#DoF	n_4	N_4	#DoF
0	3	3	33	2	2	16	4	4	56	3	3	33
1	6	6	120	4	4	56	8	8	208	6	6	120
2	12	12	456	8	8	208	16	16	800	12	12	456
3	24	24	1776	16	16	800	32	32	3136	24	24	1776
4	48	48	7008	32	32	3136	64	64	12416	48	48	7008
5	96	96	27840	64	64	12416	128	128	49408	96	96	27840
6	192	192	110976	128	128	49408	256	256	197120	192	192	110976

the iterative solver, which could significantly reduce the number of iterations, and its theoretical analysis is a subject of future research.

8.1. Example 1: Convergence test. In this example, we solve the parabolic problem (2.1) with a known solution to verify the accuracy of the space-time mortar method. We also discuss the correspondence of the number of interface GMRES iterations to the theoretical estimate. We further compare the accuracy and computational cost with DGQ_1 and DGQ_2 mortar spaces. We use the known pressure function $p(x,y,t) = \sin(8t)\sin(11x)\cos(11y - \frac{\pi}{4})$ along with permeability $K = I_{2\times 2}$ to determine the right-hand-side q in (2.1) and impose the Dirichlet boundary condition and an initial condition on the space-time domain $\Omega^T = (0,1)^2 \times (0,0.5)$.

We partition the space domain Ω into four identical squares Ω_i, $i = 1,2,3,4$, with $\Omega_1 = (0,0.5)\times(0,0.5)$, $\Omega_2 = (0.5,1)\times(0,0.5)$, $\Omega_3 = (0,0.5)\times(0.5,1)$, and $\Omega_4 = (0.5,1)\times(0.5,1)$. The space-time domain Ω^T is correspondingly partitioned into four space-time subdomains Ω_i^T, $i = 1,2,3,4$. We start with an initial grid for each Ω_i^T and Γ_{ij}^T and refine it successively four times to test the convergence rate of the solutions with respect to the actual known solution. The subdomains Ω_i^T maintain a checkerboard nonmatching mesh structure throughout the refinement cycles. In particular, let n_i be the number of elements in either the x- or the y-direction and let N_i be the number of elements in the t-direction in subdomain Ω_i^T. The initial grids are chosen as $n_1 = N_1 = 3$, $n_2 = N_2 = 2$, $n_3 = N_3 = 4$, and $n_4 = N_4 = 3$; see Table 8.1. Note that $h = \Delta t$. In the case of bilinear mortars, we maintain $H = 2h$ and $\Delta T = 2\Delta t$, halving the mesh sizes on each refinement cycle. For biquadratic mortars, we start with $H = 2h$ and $\Delta T = 2\Delta t$, but refine the mortar mesh only every other time to maintain $H = \sqrt{h}$ and $\Delta T = \sqrt{\Delta t}$. The coarser mesh on Γ^T in the biquadratic mortar case DGQ_2 is compensated by the higher degree of the space. In particular, the last term on the right-hand side in the bound (5.4) from Theorem 5.1 gives $\mathcal{O}(h^{-\frac{1}{2}}H^{m+1})$. With $m = 1$ and $H = \mathcal{O}(h)$, this results in $\mathcal{O}(h^{\frac{3}{2}})$, while with $m = 2$ and $H = \mathcal{O}(h^{\frac{1}{2}})$, it gives $\mathcal{O}(h)$. Similarly, the last term in the bound (6.17) from Theorem 6.3 gives $\mathcal{O}(H^{m+\frac{1}{2}})$. With $m = 1$ and $H = \mathcal{O}(h)$, this results in $\mathcal{O}(h^{\frac{3}{2}})$, while with $m = 2$ and $H = \mathcal{O}(h^{\frac{1}{2}})$, it gives $\mathcal{O}(h^{\frac{5}{4}})$. In all cases, the order is no smaller than the optimal order $\mathcal{O}(h)$ with respect to the $RT_0 \times DGQ_0$ finite element spaces. More details on the mesh refinement are given in Table 8.2. There we also report the number of spatial degrees of freedom of the spaces $RT_0 \times DGQ_0$ on Ω_i, as well as the number of space-time degrees of freedom of the mortar space DGQ_1 or DGQ_2 on Γ^T. We note that these choices of subdomain and mortar spaces satisfy the mortar

TABLE 8.2
Example 1*, mesh size and number of degrees of freedom for the interface.*

Ref.	$\Gamma^T(m=1)$			$\Gamma^T(m=2)$		
No.	n_Γ	N_Γ	#DoF	n_Γ	N_Γ	#DoF
0	1	1	16	1	1	36
1	2	2	64			
2	4	4	256	2	2	144
3	8	8	1024			
4	16	16	4096	4	4	576
5	32	32	32768			
6	64	64	131072	8	8	2304

TABLE 8.3
Example 1*, convergence with bilinear mortars.*

Ref.	# GMRES		$\|\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{L^2(\mathbf{L}^2)}$		$\|p-p_h^{\Delta t}\|_{\mathrm{DG}}$		$\|p-p_h^{\Delta t}\|_{L^2(W)}$		$\|\lambda-\lambda_H^{\Delta T}\|_{L^2(\Lambda_H)}$	
0	12	Rate	6.50e–01	Rate	1.21e+00	Rate	7.91e–01	Rate	7.98e–01	Rate
1	24	–1.00	3.63e–01	0.84	7.21e–01	0.75	4.76e–01	0.73	5.11e–01	0.64
2	40	–0.74	1.74e–01	1.06	3.19e–01	1.18	2.46e–01	0.95	2.34e–01	1.13
3	59	–0.56	8.63e–02	1.02	1.46e–01	1.13	1.25e–01	0.98	1.20e–01	0.96
4	82	–0.47	4.29e–02	1.01	6.93e–02	1.08	6.25e–02	1.00	6.11e–02	0.97
5	112	–0.45	2.14e–02	1.00	3.36e–02	1.04	3.13e–02	1.00	3.08e–02	0.99
6	149	–0.41	1.07e–02	1.00	1.65e–02	1.03	1.56e–02	1.00	1.55e–02	0.99

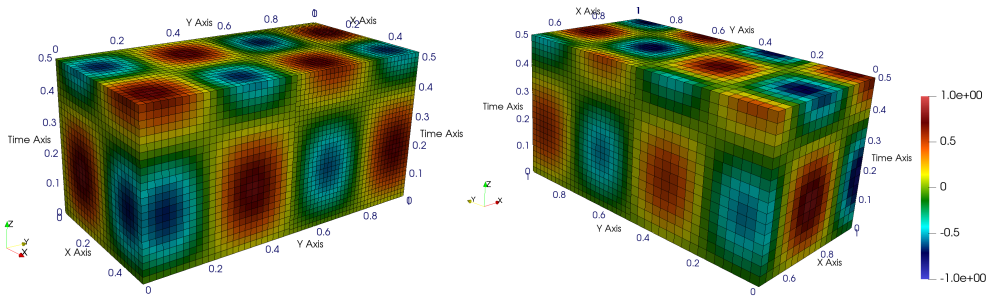
TABLE 8.4
Example 1*, convergence with biquadratic mortars.*

Ref.	# GMRES		$\|\mathbf{u}-\mathbf{u}_h^{\Delta t}\|_{L^2(\mathbf{L}^2)}$		$\|p-p_h^{\Delta t}\|_{\mathrm{DG}}$		$\|p-p_h^{\Delta t}\|_{L^2(W)}$		$\|\lambda-\lambda_H^{\Delta T}\|_{L^2(\Lambda_H)}$	
0	19	Rate	6.81e–01	Rate	1.35e+00	Rate	8.39e–01	Rate	2.13e+00	Rate
2	38	–0.50	1.70e–01	1.00	3.51e–01	0.97	2.51e–01	0.87	2.82e–01	1.46
4	60	–0.33	4.48e–02	0.96	8.59e–02	1.02	6.59e–02	0.96	9.20e–02	0.81
6	63	–0.04	1.06e–02	1.03	1.92e–02	1.08	1.55e–02	1.04	1.67e–02	1.23

space-time condition (4.6), but do not satisfy the stronger temporal discretization conditions (4.5b) and (6.1).

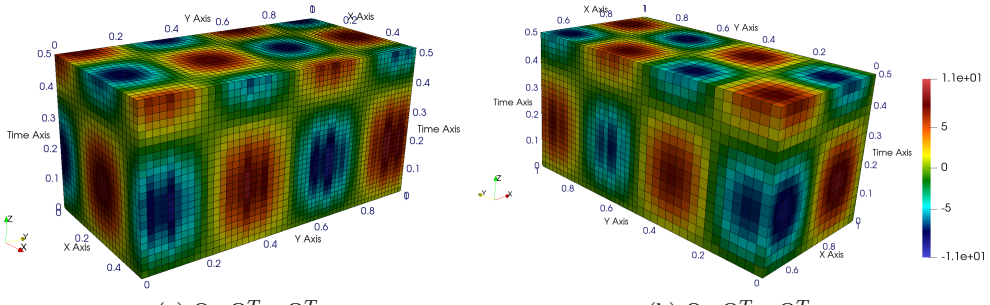
In Tables 8.3 and 8.4 we report the relative errors with respect to the norm of the true solution, as well as the convergence rates as powers of the subdomain discretization parameters h and Δt. We observe optimal first-order convergence of the method using both bilinear and biquadratic mortars. We note that the $\Delta t^{-\frac{1}{2}}$ loss in convergence rate in the bound from Theorem 5.1 is not observed in the numerical results. This indicates that the weaker mortar condition (4.6) may be sufficient in practice for the method to be well posed and optimally convergent. In Tables 8.3 and 8.4 we also report the growth rate for the number of GMRES iterations in the case of bilinear and biquadratic mortars, respectively. We recall that Theorem 7.2 bounds the spectral ratio of the interface operator S by $\mathcal{O}(h^{-1})$. Thus, up to deviation from a normal matrix, the growth rate is expected to be $\mathcal{O}(h^{-0.5})$ [26, 27]. This is close to what we observe in Tables 8.3 and 8.4.

We further compare the performance of bilinear and biquadratic mortars. As expected from the theory, the accuracy of the two cases at the same refinement level is comparable, which is evident from Tables 8.3 and 8.4. On the other hand, since



(a) On $\Omega_1^T \cup \Omega_4^T$ (b) On $\Omega_2^T \cup \Omega_3^T$

FIG. 8.1. *Example* 1, *pressure computed with bilinear mortars on the space-time grid, refinement* 3.



(a) On $\Omega_1^T \cup \Omega_4^T$ (b) On $\Omega_2^T \cup \Omega_3^T$

FIG. 8.2. *Example* 1, *x-component of the velocity computed using bilinear mortars on the space-time grid at refinement* 3.

the biquadratic mortar space DGQ_2 has far fewer degrees of freedom compared to bilinear mortar space DGQ_1, the former requires fewer GMRES iterations. Thus, higher mortar degrees m, s with a coarser mortar mesh results in a computationally more efficient method compared to smaller m, s and a finer mortar mesh.

Finally, the computed solution is presented in three-dimensional space-time plots in Figures 8.1 and 8.2. Note that the z-axis corresponds to time direction t. The plots clearly show the continuity of the pressure and the normal component of the velocity, which is imposed in a weak sense, across the space-time interfaces.

8.2. Example 2: Problem with a boundary layer. In this example we demonstrate the advantages of applying the multiscale mortar space-time domain decomposition method to a problem where the solution variables, pressure and velocity, vary on different scales across the space-time domain. For this, we use the known solution, $p(x, y, t) = 1000xyt\, e^{-10\left(x^2+y^2+\frac{1}{4}t^2\right)}$ along with permeability $K = I_{2\times 2}$ to determine the right-hand-side q in (2.1) and impose the Dirichlet boundary condition and an initial condition on the space-time domain $\Omega^T = (0,1)^2 \times (0, 0.5)$. By construction, $p(x, y, t)$ varies rapidly in both space and time along the lower-left corner of Ω^T, forming a sharp boundary layer. The pressure decays exponentially away from this corner and is close to zero over a large part of the domain. This calls for an efficient multiscale method that would take advantage of the multiscale nature of the problem and provide better resolution around the lower-left corner compared to the rest of Ω^T.

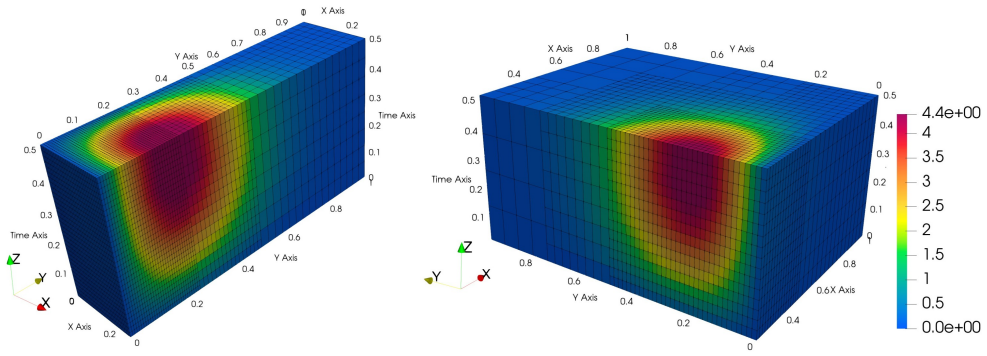


FIG. 8.3. *Example* 2*, pressure from the multiscale method, cut along the plane* $x = 0.25$.

TABLE 8.5
Example 2*, errors and GMRES iterations for the multiscale and fine-scale methods.*

Method	# GMRES	$\|\mathbf{u} - \mathbf{u}_h^{\Delta t}\|_{L^2(\mathbf{L}^2)}$	$\|p - p_h^{\Delta t}\|_{\mathrm{DG}}$	$\|p - p_h^{\Delta t}\|_{L^2(W)}$	$\|\lambda - \lambda_H^{\Delta T}\|_{L^2(\Lambda_H)}$
Multiscale	102	5.657e–02	8.425e–02	6.319e–02	5.796e–02
Fine-scale	140	1.524e–02	2.234e–02	2.154e–02	3.016e–02

We partition Ω^T into 4×4 identical space-time subdomains Ω_i^T. From knowledge about the variation of the true pressure, we design a multiscale space-time grid on Ω^T, where the refinement of the grid on each Ω_i^T is proportional to the amount of pressure variation. The finest mesh on Ω_i^T has $h = 1/128$ and $\Delta t = 1/64$, and the coarsest mesh on Ω_i^T has $h = 1/8$ and $\Delta t = 1/8$; see Figure 8.3 for the mesh refinement. The coarser meshes on the majority of the space-time subdomains reduce the computational complexity of the subdomain solves associated with them. The mortar mesh sizes in space are chosen as follows. For vertical interfaces (fixed x) between subdomains on the bottom row, the one along the boundary layer, we set $H = 1/32$. For the next row of subdomains we set $H = 1/16$, and for the other two rows, $H = 1/8$. Similarly, for the horizontal interfaces (fixed y) between subdomains on the left column we set $H = 1/32$, for the second column, $H = 1/16$, and for the other two columns, $H = 1/8$. We choose $\Delta T = 1/8$ on all interfaces. These choices guarantee that the mortar assumption (4.6) is satisfied and that the dimension of the interface problem is reduced, while at the same time providing suitable resolution to enforce weakly flux continuity across the space-time interfaces. For comparison, we solve the problem using a uniformly fine and matching subdomain mesh with $h = H = 1/128$ and $\Delta t = \Delta T = 1/64$. A comparison of the number of GMRES iterations and the relative errors from the multiscale and the fine-scale methods is given in Table 8.5. It shows that the multiscale and the fine-scale solutions attain comparable accuracy. We observe slightly smaller relative errors in the fine-scale case because of the matching grids and higher resolution throughout the space-time domain. The slightly higher errors for the multiscale method are compensated by the cheaper subdomain solves and smaller interface problem that converges in fewer iterations compared to the fine-scale method.

The computed multiscale solution is presented in Figures 8.3 and 8.4. The plots show that the multiscale method provides good resolution where it matters—in the

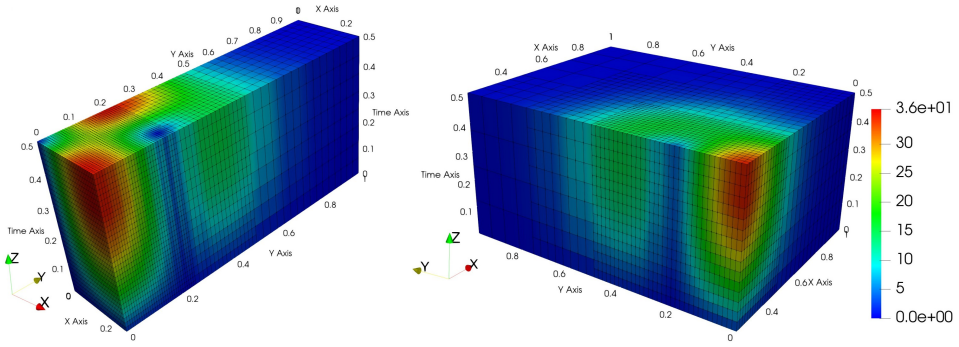


FIG. 8.4. *Example* 2, *velocity magnitude from the multiscale method, cut along the plane* $x = 0.25$.

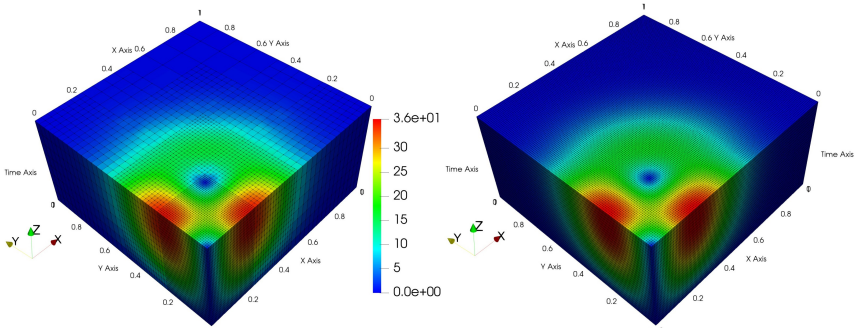


FIG. 8.5. *Example* 2. *Velocity magnitude from the multiscale (left) and fine-scale (right) methods on the whole domain.*

regions with high solution variation. Moreover, we observe very good enforcement of continuity of both pressure and velocity across various space and time interfaces. A side-by-side comparison of the multiscale and fine-scale solutions is given in Figure 8.5. It shows excellent agreement between the two solutions, illustrating that the less expensive multiscale method provides comparable accuracy to the more expensive fine-scale method.

8.3. Example 3: Heterogeneous model. In this example we show the performance of the method for a problem with a heterogeneous coefficient K, with varying permeability contrast. We still take the space-time domain $\Omega^T = (0,1)^2 \times (0,0.5)$, and we subdivide the spatial domain into three vertical layers $\Omega_1 = (0,0.4) \times (0,1)$, $\Omega_2 = (0.4,0.6) \times (0,1)$, and $\Omega_3 = (0.6,1) \times (0,1)$. Each subdomain Ω_i has a constant scalar permeability K_i. We keep the value in Ω_1 and Ω_3 fixed to 1, and we vary the value in Ω_2.

Both the source and the initial condition are taken as zero. The boundary conditions are chosen to create a diagonal flow across the domain. We impose the pressure on all sides, varying linearly in both time and space, taking the boundary trace of the function $p(x,y,t) = \min(t/0.05,1)xy$. For this example, no analytical solution is available, so no convergence results are presented.

In order to compare the respective effects of the mesh size and the permeability contrast, we use two refinement levels in both space and time. For each sudbomain Ω_i, the mesh size in both space dimensions, as well as in time, is the same and is

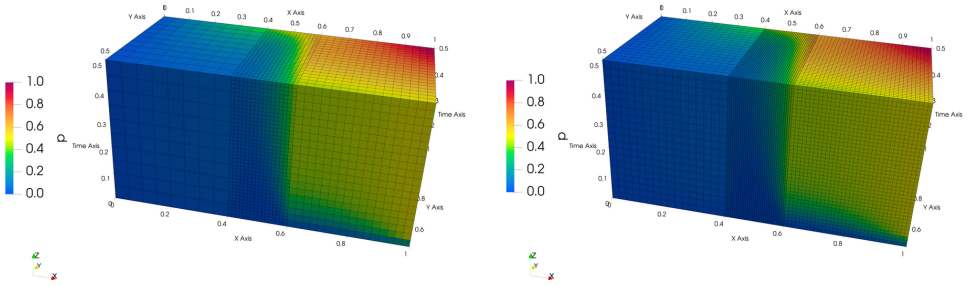


FIG. 8.6. *Example* 3, *pressure in space-time, cut along the plane* $y = 0.5$*; Left: coarse grid; right: fine grid.*

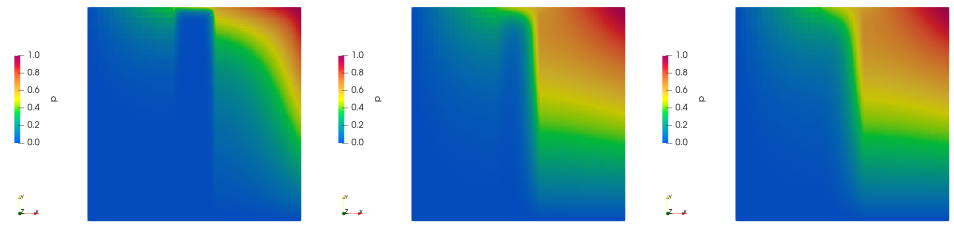


FIG. 8.7. *Example* 3, *pressure in space at different times. Left to right:* $t = 0.05$, $t = 0.25$, $t = 0.5$.

TABLE 8.6

Example 3, *number of GMRES iterations. From top to bottom: permeability in the central layer, number of iterations for the coarse grid, number of iterations for the fine grid.*

K_2	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}
Coarse	59	41	57	87	80
Fine	92	57	69	107	118

denoted by Δ_i. For both interfaces, the mesh size in space and time is the same and is denoted by Δ_Γ. At the coarse level, we take $\Delta_1 = 0.05$, $\Delta_2 = 0.0125$, $\Delta_3 = 0.025$, and $\Delta_\Gamma = 0.05$. The fine discretization is a refinement by two.

The computed solution with $K_2 = 10^{-2}$ is presented in Figures 8.6 and 8.7. Figure 8.6 shows a space-time view of the pressure for both coarse and fine grids, with the grid superimposed, so one can see the refined grid in the central domain as well as the evolution of the solution.

Figure 8.7 shows snaphosts of the pressure at several times, illustrating the sharp gradient in the central domain, which is well resolved by the finer grid in this region.

Finally, we look at the number of iterations as both the permeability contrast and the grid resolution are varied. Table 8.6 shows the number of iterations as a function of the permeability value in the central layer, for both grid resolutions. We observe only a small variation in the number of iterations over a 10^4-fold increase in the permeability contrast, which illustrates the robust behavior of the method with a discontinuous coefficient. Also, the number of iterations with the fine grid increases approximately like the square root of the number of mesh elements in each direction, similarly to Example 1.

9. Conclusions. We presented a space-time domain decomposition mixed finite element method for parabolic problems that allows for nonmatching spatial grids and local time stepping via space-time mortar finite elements. Well-posedness and a priori error estimates were established. A parallel nonoverlapping domain decomposition algorithm was developed, which reduces the algebraic problem to a coarse-scale interface problem for the mortar variable. The theoretical results and the flexibility of the method were illustrated in a series of numerical experiments. Future work may include the development and analysis of a Neumann–Neumann preconditioner for the interface problem in section 7 as in [23], using techniques from [33], as well as deriving a posteriori error estimates, possibly building upon the ideas from [1, 2, 14].

REFERENCES

[1] S. A. Hassan, C. Japhet, M. Kern, and M. Vohralík, *A posteriori stopping criteria for optimized Schwarz domain decomposition algorithms in mixed formulations*, Comput. Methods Appl. Math., 18 (2018), pp. 495–519, https://doi.org/10.1515/cmam-2018-0010.

[2] S. A. Hassan, C. Japhet, and M. Vohralík, *A posteriori stopping criteria for space-time domain decomposition for the heat equation in mixed formulations*, Electron. Trans. Numer. Anal., 49 (2018), pp. 151–181, https://doi.org/10.1553/etna.

[3] T. Arbogast, L. C. Cowsar, M. F. Wheeler, and I. Yotov, *Mixed finite element methods on nonmatching multiblock grids*, SIAM J. Numer. Anal., 37 (2000), pp. 1295–1315, https://doi.org/10.1137/S0036142996308447.

[4] T. Arbogast, D. Estep, B. Sheehan, and S. Tavener, *A posteriori error estimates for mixed finite element and finite volume methods for parabolic problems coupled through a boundary*, SIAM/ASA J. Uncertain. Quantif., 3 (2015), pp. 169–198, https://doi.org/10.1137/140964059.

[5] T. Arbogast, G. Pencheva, M. F. Wheeler, and I. Yotov, *A multiscale mortar mixed finite element method*, Multiscale Model. Simul., 6 (2007), pp. 319–346, https://doi.org/10.1137/060662587.

[6] M. Arshad, E.-J. Park, and D. Shin, *Multiscale mortar mixed domain decomposition approximations of nonlinear parabolic equations*, Comput. Math. Appl., 97 (2021), pp. 375–385, https://doi.org/10.1016/j.camwa.2021.06.009.

[7] W. Bangerth, R. Hartmann, and G, Kanschat, *deal.II—A general-purpose object-oriented finite element library*, ACM Trans. Math. Software, 33 (2007), 24, https://doi.org/10.1145/1268776.1268779.

[8] M. Bause, F. A. Radu, and U. Köcher, *Error analysis for discretizations of parabolic problems using continuous finite elements in time and mixed finite elements in space*, Numer. Math., 137 (2017), pp. 773–818.

[9] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[10] J. M. Cascón, L. Ferragut, and M. I. Asensio, *Space-time adaptive algorithm for the mixed parabolic problem*, Numer. Math., 103 (2006), pp. 367–392, https://doi.org/10.1007/s00211-006-0677-y.

[11] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Stud. Math. Appl. 4, North-Holland, Amsterdam, 1978.

[12] M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces*, Math. Comp., 48 (1987), pp. 521–532.

[13] L. Delpopolo Carciopolo, M. Cusini, L. Formaggia, and H. Hajibeygi, *Adaptive multilevel space-time-stepping scheme for transport in heterogeneous porous media (ADM-LTS)*, J. Comput. Phys. X, 6 (2020), 100052, https://doi.org/10.1016/j.jcpx.2020.100052.

[14] A. Ern, I. Smears, and M. Vohralík, *Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems*, SIAM J. Numer. Anal., 55 (2017), pp. 2811–2834, https://doi.org/10.1137/16M1097626.

[15] R. E. Ewing, R. D. Lazarov, and P. S. Vassilevski, *Finite difference schemes on grids with local refinement in time and space for parabolic problems.* I. *Derivation, stability, and error analysis*, Computing, 45 (1990), pp. 193–215, https://doi.org/10.1007/BF02250633.

[16] S. Gaiffe, R. Glowinski, and R. Masson, *Domain decomposition and splitting methods for mortar mixed finite element approximations to parabolic equations*, Numer. Math., 93 (2002), pp. 53–75, https://doi.org/10.1007/s002110100367.

[17] M. J. Gander, 50 *years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition Methods, Springer, Cham, 2013, pp. 69–113.
[18] M. J. Gander and L. Halpern, *Optimized Schwarz waveform relaxation methods for advection reaction diffusion problems*, SIAM J. Numer. Anal., 45 (2007), pp. 666–697, https://doi.org/10.1137/050642137.
[19] B. Ganis and I. Yotov, *Implementation of a mortar mixed finite element method using a multiscale flux basis*, Comput. Methods Appl. Mech. Engrg., 198 (2009), pp. 3989–3998, https://doi.org/10.1016/j.cma.2009.09.009.
[20] V. Girault, R. Glowinski, and H. López, *A domain decomposition and mixed method for a linear parabolic boundary value problem*, IMA J. Numer. Anal., 24 (2004), pp. 491–520, https://doi.org/10.1093/imanum/24.3.491.
[21] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Classics in Appl. Math. 69, SIAM, Philadelphia, 2011.
[22] L. Halpern, C. Japhet, and J. Szeftel, *Optimized Schwarz waveform relaxation and discontinuous Galerkin time stepping for heterogeneous problems*, SIAM J. Numer. Anal., 50 (2012), pp. 2588–2611, https://doi.org/10.1137/120865033.
[23] T.-T.-P. Hoang, J. Jaffré, C. Japhet, M. Kern, and J. E. Roberts, *Space-time domain decomposition methods for diffusion problems in mixed formulations*, SIAM J. Numer. Anal., 51 (2013), pp. 3532–3559, https://doi.org/10.1137/130914401.
[24] T.-T.-P. Hoang, C. Japhet, M. Kern, and J. E. Roberts, *Space-time domain decomposition for reduced fracture models in mixed formulation*, SIAM J. Numer. Anal., 54 (2016), pp. 288–316, https://doi.org/10.1137/15M1009651.
[25] T.-T.-P. Hoang and H. Lee, *A global-in-time domain decomposition method for the coupled nonlinear Stokes and Darcy flows*, J. Sci. Comput., 87 (2021), pp. 1–22.
[26] I. C. F. Ipsen, *Expressions and bounds for the GMRES residual*, BIT, 40 (2000), pp. 524–535, https://doi.org/10.1023/A:1022371814205.
[27] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, Frontiers in Appl. Math. 16, SIAM, Philadelphia, 1995.
[28] W. Kheriji, R. Masson, and A. Moncorgé, *Nearwell local space and time refinement in reservoir simulation*, Math. Comput. Simul., 118 (2015), pp. 273–292, https://doi.org/10.1016/j.matcom.2014.11.022.
[29] D. Kim, E.-J. Park, and B. Seo, *Space-time adaptive methods for the mixed formulation of a linear parabolic problem*, J. Sci. Comput., 74 (2018), pp. 1725–1756, https://doi.org/10.1007/s10915-017-0514-8.
[30] J.-L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Vol. I, Springer-Verlag, New York, 1972.
[31] C. Makridakis and R. H. Nochetto, *A posteriori error analysis for higher order dissipative methods for evolution problems*, Numer. Math., 104 (2006), pp. 489–514, https://doi.org/10.1007/s00211-006-0013-6.
[32] T. P. Mathew, *Domain Decomposition and Iterative Refinement Methods for Mixed Finite Element Discretizations of Elliptic Problems*, Ph.D. thesis, Courant Institute of Mathematical Sciences, New York University, 1989.
[33] G. Pencheva and I. Yotov, *Balancing domain decomposition for mortar mixed finite element methods*, Numer. Linear Algebra Appl., 10 (2003), pp. 159–180.
[34] I. Rybak and J. Magiera, *A multiple-time-step technique for coupled free flow and porous medium systems*, J. Comput. Phys., 272 (2014), pp. 327–342, https://doi.org/10.1016/j.jcp.2014.04.036.
[35] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[36] G. Starke, *Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems*, Numer. Math., 78 (1997), pp. 103–117, https://doi.org/10.1007/s002110050306.
[37] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Ser. Comput. Math. 25, Springer-Verlag, Berlin, 2006.