Progress toward a new implementation of the mortar finite element method in MOOSE

John W. Peterson

February 2018

Progress toward a new implementation of the mortar finite element method in MOOSE

John W. Peterson

February 2018

Progress toward a new implementation of the mortar finite element method in MOOSE

John W. Peterson

February 1, 2018

Contents

1   Introduction

The mortar finite element method has been used for many years in a variety of applications, including the enforcement of continuity conditions across decomposed domains [1], the implementation of Dirichlet boundary conditions [2–4], obtaining improved estimates of surface fluxes [5], and for solving large deformation contact mechanics problems [6–11]. There is a currently great deal of interest in developing more robust mechanical and thermal contact solution strategies in the MOOSE framework and physics modules, and schemes based on the mortar finite element approach appear to be a promising avenue of development. There are a number of challenges associated with the development of a robust Lagrange multiplier based formulation of the mortar finite element method which can be tackled using the relatively simple framework of thermal contact problems, before tackling more complicated applications such as thermomechanical contact. In this report, we describe several aspects of our mortar finite element method implementation for solving thermal contact problems. The approach and notation used are based primarily based on the work of Yang et al. [12, 13].

2   Solution continuity enforcement via Lagrange multiplier

In this section, we describe the classical Lagrange multiplier formulation for weakly enforcing solution continuity across e.g. contact interfaces or non-conforming mesh boundaries. There are many relevant research articles on this topic, here our development mainly follows that of [1, 4, 14, 15]. We consider the problem of computing the temperature distribution in two regions $\Omega^{(1)}$, $\Omega^{(2)}$ with scalar thermal conductivities $k^{(1)}$, $k^{(2)}$, which are in contact along an interface Γ_C as shown in Fig. 1.
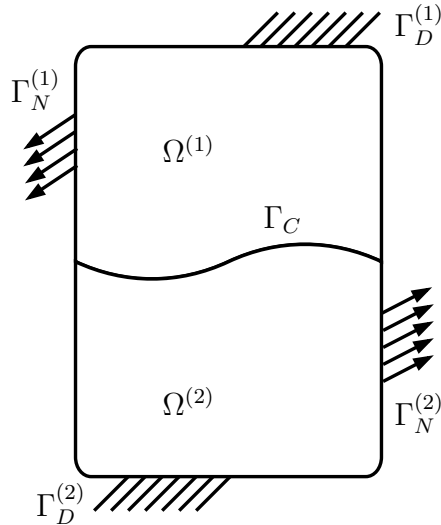

Figure 1: Boundary conditions and configuration of the bodies.

The temperature field $T^{(m)}$ in body m satisfies the heat conduction equation

$$- \nabla \cdot k^{(m)} \nabla T^{(m)} = f^{(m)} \quad \in \Omega^{(m)} \tag{1}$$

for $m = 1, 2$, where $f^{(m)}$ is a specified heat source in body m. We assume that Dirichlet and Neumann boundary conditions for body m are specified on $\Gamma_D^{(m)}$ and $\Gamma_N^{(m)}$, respectively, i.e.

$$T^{(m)} = T_D^{(m)} \in \Gamma_D^{(m)} \tag{2}$$

$$-k^{(m)} \nabla T^{(m)} \cdot \hat{\boldsymbol{n}}^{(m)} = q_N^{(m)} \in \Gamma_N^{(m)} \tag{3}$$

where $T_D^{(m)}$ and $q_N^{(m)}$ are given data, and $\hat{\boldsymbol{n}}^{(m)}$ is the outward unit normal on $\Omega^{(m)}$. For simplicity, we also assume that these regions are non-overlapping:

$$\partial \Omega^{(m)} = \Gamma_D^{(m)} \cup \Gamma_N^{(m)} \cup \Gamma_C^{(m)} \tag{4}$$

$$\emptyset = \Gamma_D^{(m)} \cap \Gamma_N^{(m)} \cap \Gamma_C^{(m)} \tag{5}$$

The thermal conductivities $k^{(m)}$ may depend on both spatial location and temperature. The temperatures in the two bodies are constrained to be equal along Γ_C, i.e.:

$$T^{(1)} = T^{(2)} \in \Gamma_C \tag{6}$$

This Dirichlet-like constraint is therefore responsible for coupling the temperature fields in the two bodies. We observe that conservation of thermal energy requires that the jump in the heat flux should be zero along the contact interface, i.e.

$$-k^{(1)} \left.\frac{\partial T^{(1)}}{\partial n^{(1)}}\right|_{\Gamma_C} = k^{(2)} \left.\frac{\partial T^{(2)}}{\partial n^{(2)}}\right|_{\Gamma_C} \equiv q_C \tag{7}$$

where the directional derivative notation $\frac{\partial T^{(m)}}{\partial n^{(m)}} \equiv \nabla T^{(m)} \cdot \hat{\boldsymbol{n}}^{(m)}$ is used. Eq. (7) implies that if the two bodies have different thermal conductivities, they must necessarily have correspondingly different temperature gradients at the contact interface in order for the conservation of energy constraint to be satisfied.

This application can be approached from the standpoint of constrained optimization theory. Using the notation of Carey and Oden [14], the abstract constrained optimization problem can be stated as:

$$\min_{v \in K} J(v) \tag{8}$$

where $K \equiv \{v \in H : Bv = g\}$ is the so-called constraint set, $J(v)$ is a functional related to the application, and $B : H \to Q$ is a given linear operator from H into the space Q of the data g. The method of Lagrange multipliers is a standard approach for solving such constrained optimization problems. To implement this method, we introduce the Lagrangian functional

$$L(v, \mu) \equiv J(v) + \langle \mu, Bv - g \rangle \tag{9}$$

where $\mu \in Q'$ is a member of the dual space of Q (the space of continuous linear operators on Q) and $\langle \cdot, \cdot \rangle$ is the duality pairing on $Q' \times Q$. Suppose that (u, λ) is a minimizer of (9). (Note: Carey and Oden [14] use (p, q) in place of (λ, μ) since their focus is on constrained optimization methods for Stokes flow, in which the Lagrange multiplier variable is associated

2

with the pressure, p.) Then, using the assumed linearity of B, the first variation of L at (u, λ) is given by

$$\delta L = (\delta J(u), v) + \langle \lambda, Bv \rangle + \langle \mu, Bu - g \rangle \tag{10}$$

Since (u, λ) is a minimizer, we must have $\delta L = 0$ for arbitrary $v \in H$ and $\mu \in Q'$, thus:

$$(\delta J(u), v) + \langle \lambda, Bv \rangle = 0 \quad \forall\, v \in H \tag{11}$$

$$\langle \mu, Bu - g \rangle = 0 \quad \forall\, \mu \in Q' \tag{12}$$

which is the variational boundary value problem associated to the constrained minimization problem (8).

In the present application, we consider the broken space

$$H \equiv H^1(\Omega^{(1)}) \oplus H^1(\Omega^{(2)}) \tag{13}$$

and seek to minimize the functional $J : H \to \mathbb{R}$ defined by:

$$J(v) \equiv J^{(1)}(v^{(1)}) + J^{(2)}(v^{(2)}) \tag{14}$$

for $v \in H$, where

$$J^{(m)}(v^{(m)}) \equiv \int_{\Omega^{(m)}} \left(k^{(m)} \frac{1}{2} |\nabla v^{(m)}|^2 - f^{(m)} v^{(m)} \right) \mathrm{d}x + \int_{\Gamma_N^{(m)}} q_N^{(m)} v^{(m)} \mathrm{d}s \tag{15}$$

subject to the constraint $B : H \to Q$ defined by

$$Bv \equiv \left(v^{(1)} - v^{(2)} \right)\big|_{\Gamma_C} = 0 \tag{16}$$

The constrained optimization problem may be stated as:

$$\min_{v \in K} J(v) \tag{17}$$

where

$$K = \{ v \in K^{(1)} \oplus K^{(2)} : v^{(1)} - v^{(2)}\big|_{\Gamma_C} = 0 \} \tag{18}$$

and

$$K^{(m)} \equiv \left\{ v^{(m)} \in H^1(\Omega^{(m)}) : v^{(m)} = T_D^{(m)} \in \Gamma_D^{(m)} \right\} \tag{19}$$

In this case, the constraint sets also enforce the Dirichlet boundary conditions (2) on the non-contact part of the boundary. The first variation of the functional (15) at $T^{(m)}$ in the direction of $v^{(m)}$ is

$$\left(\delta J^{(m)}(T^{(m)}), v^{(m)} \right) \equiv \int_{\Omega^{(m)}} \left(k^{(m)} \nabla T^{(m)} \cdot \nabla v^{(m)} - f^{(m)} v^{(m)} \right) \mathrm{d}x + \int_{\Gamma_N^{(m)}} q_N^{(m)} v^{(m)} \mathrm{d}s \tag{20}$$

Substituting the preceding terms into (11)–(12), we arrive at the variational formulation for the constrained minimization problem (17), find $(T, \lambda) \in \mathcal{S} \times Q'$ such that:

$$\sum_{m=1}^{2} \left(\delta J^{(m)}(T^{(m)}), v^{(m)} \right) + \langle \lambda, v^{(1)} - v^{(2)} \rangle = 0 \tag{21}$$

$$\langle \mu, T^{(1)} - T^{(2)} \rangle = 0 \tag{22}$$

holds for every $(v, \mu) \in \mathcal{V} \times Q'$, where the broken spaces are defined by

$$\mathcal{V} \equiv \mathcal{V}^{(1)} \oplus \mathcal{V}^{(2)} \tag{23}$$

$$\mathcal{S} \equiv \mathcal{S}^{(1)} \oplus \mathcal{S}^{(2)} \tag{24}$$

where the constituent function spaces are:

$$\mathcal{V}^{(m)} \equiv \{ v^{(m)} \in H^1 \left(\Omega^{(m)} \right) : v^{(m)} = 0 \in \Gamma_D^{(m)} \} \tag{25}$$

$$\mathcal{S}^{(m)} \equiv \{ T^{(m)} \in H^1 \left(\Omega^{(m)} \right) : T^{(m)} = T_D^{(m)} \in \Gamma_D^{(m)} \} \tag{26}$$

We observe that the space $Q' = Q = L^2(\Gamma_C)$, and the $\langle \cdot, \cdot \rangle$ inner product is over the manifold Γ_C.

Note that if we assume $T^{(m)}$ is smooth enough and integrate (21) by parts "in reverse," that is, we use the identity

$$\int_{\Omega^{(m)}} k^{(m)} \nabla T^{(m)} \cdot \nabla v^{(m)} \, \mathrm{d}x = \int_{\partial \Omega^{(m)}} v^{(m)} (k^{(m)} \nabla T^{(m)} \cdot \hat{\boldsymbol{n}}^{(m)}) \, \mathrm{d}s$$
$$- \int_{\Omega^{(m)}} v^{(m)} (\nabla \cdot k^{(m)} \nabla T^{(m)}) \, \mathrm{d}x \tag{27}$$

then we obtain:

$$\sum_{m=1}^{2} \left[\int_{\Omega^{(m)}} \left(-\nabla \cdot k^{(m)} \nabla T^{(m)} - f^{(m)} \right) v^{(m)} \, \mathrm{d}x \right]$$
$$+ \int_{\Gamma_C} \left[v^{(1)} \left(\lambda + k^{(1)} \nabla T^{(1)} \cdot \hat{\boldsymbol{n}}^{(1)} \right) + v^{(2)} \left(-\lambda + k^{(2)} \nabla T^{(2)} \cdot \hat{\boldsymbol{n}}^{(2)} \right) \right] \mathrm{d}s = 0 \tag{28}$$

where we have used the natural boundary condition (3) to eliminate the boundary integral contribution on $\Gamma_N^{(m)}$, and written the duality pairing $\langle \cdot, \cdot \rangle$ as a boundary integral over Γ_C. Recalling the definition of the flux q_C through the contact interface (7) we can write the boundary term in (28) as

$$\int_{\Gamma_C} \left[v^{(1)} \left(\lambda - q_C \right) + v^{(2)} \left(-\lambda + q_C \right) \right] \mathrm{d}s = \int_{\Gamma_C} \left(\lambda - q_C \right) \left(v^{(1)} - v^{(2)} \right) \mathrm{d}s \tag{29}$$

This boundary term vanishes when

$$\lambda = q_C \tag{30}$$

4

that is, when the Lagrange multiplier variable is simply equal to the flux through the contact interface. We are therefore left with:

$$\sum_{m=1}^{2} \left[\int_{\Omega^{(m)}} \left(-\nabla \cdot k^{(m)} \nabla T^{(m)} - f^{(m)} \right) v^{(m)} \, \mathrm{d}x \right] = 0 \tag{31}$$

which is the original weak statement of the problem. This derivation demonstrates that the constrained minimization (17) and variational boundary value formulations of the problem (21)–(22) are equivalent, and shows the relationship between the abstract Lagrange multiplier variable λ and the heat flux q_C.

A slight generalization of the solution continuity enforcement method described previously, in which the assumption of "perfect" contact is relaxed, is also possible. That is, the constraint is specified implicitly in terms of the nodal normal projection operator P as

$$Bv \equiv \left(v^{(1)} - Pv^{(2)} \right) \big|_{\Gamma_C^{(1)}} = 0 \tag{32}$$

Substituting (32) into (11)–(12) then leads to the following variational boundary value problem, find $(T, \lambda) \in \mathcal{S} \times Q'$ such that:

$$\sum_{m=1}^{2} \left(\delta J^{(m)}(T^{(m)}), v^{(m)} \right) + \int_{\Gamma_C^{(1)}} \lambda \left(v^{(1)} - Pv^{(2)} \right) \, \mathrm{d}s = 0 \tag{33}$$

$$\int_{\Gamma_C^{(1)}} \mu \left(T^{(1)} - PT^{(2)} \right) \, \mathrm{d}s = 0 \tag{34}$$

holds for every $(v, \mu) \in \mathcal{V} \times Q'$, where $Q' = Q = L^2(\Gamma_C^{(1)})$, the first variation of the functional J is defined as before in (20), and the other functional spaces are unchanged. This formulation should reduce to the formulation in §2 when the bodies are in perfect contact, but otherwise be more generally applicable. We can also show that the Lagrange multiplier variable λ is equal to the flux by an argument analogous to the one above.

3 Stability of Lagrange multiplier formulations

Once again following the description in Carey and Oden [14], we note that to guarantee the existence and uniqueness of the saddle point solution (u, λ) of (11)–(12), the following conditions are required on J, B, and the associated functional spaces (Theorem 3.2.1 in Carey and Oden [14])

(a) The functional $J : H \to \mathbb{R}$ is continuous, and "coercive" in the sense

$$J(v) \geq C_0 \|v\|_H^2 - C_1 \|v\|_H \quad \forall \, v \in H \tag{35}$$

where $C_0, C_1 > 0$, and C_1 may depend on the data $f^{(1)}$, $f^{(2)}$ of the problem.

(b) The constraint $B : H \to Q$ is a bounded linear operator, i.e.

$$\|Bv\|_Q \leq C_B \|v\|_H \quad \forall \, v \in H \tag{36}$$

where $C_B > 0$ is a constant.

5

(c) The variational operator $\delta J : H \to H'$ is bounded. That is, when $\|v\|_H \leq C_1$, we must have

$$\|\delta J(v)\|_{H'} \leq C_2 \tag{37}$$

for constants C_1, C_2.

(d) There exists a constant $\beta > 0$ such that

$$\beta \|\mu\|_{Q'} \leq \sup_{\substack{v \in H \\ (v \neq 0)}} \frac{|\langle \mu, Bv \rangle|}{\|v\|_H} \quad \forall \, \mu \in Q' \tag{38}$$

Conditions (a)–(c) are satisfied in most typical applications, but condition (d), referred to henceforth as the "inf-sup" or Ladyzhenskaya-Babuška-Brezzi (LBB) [2, 16, 17] condition, frequently imposes onerous restrictions on the function spaces (as well as their finite-dimensional approximations) that can be employed in finite element analysis.

In his early work on Lagrange multiplier formulations for enforcing Dirichlet boundary conditions in a model second-order elliptic problem, Babuška [2] showed that, in order to guarantee that the discrete form of the inf-sup condition was satisfied, the primal and Lagrange multiplier variables could not share a single discretization of the boundary. In simplistic terms, if we let h_u be the length of the edges of the finite elements used to discretize the primal variable in the interior of the domain, and h_λ be the length of the elements used in the Lagrange multiplier discretization, then

$$h_\lambda \geq C h_u \tag{39}$$

where $C > 1$ is a constant which in general depends on the domain Ω, is a sufficient (but not necessary) condition to guarantee the inf-sup stability of the formulation. The condition derived by Babuška does not explicitly mention the effects of using different polynomial degrees in the primal and Lagrange multiplier discretizations, however later publications [5, 18, 19] subsequently showed that e.g. $P^1 - P^0$ (continuous linear elements for the primal variable, discontinuous constant elements for the Lagrange multiplier variable) are also subject to the restriction (39).

Early successors to Babuška [18–20] focused primarily on designing discretizations that would allow the constant C in (39) to be as close to 1 as possible, as well as characterizing conditions under which local refinements at the boundary were possible while maintaining stability [21]. Residual-based stabilization methods were introduced by Barbosa and Hughes [3, 22] in the early 1990s, and made it possible to completely circumvent the LBB condition, rendering equal-order discretizations on shared grids convergent with optimal *a priori* error estimates for certain classes of problems. Stenberg [23] later demonstrated the connections between the residual-based stabilization methods of Barbosa with discontinuous constant Lagrange multiplier discretizations and an earlier method by Nitsche for enforcing Dirichlet boundary conditions [24]. Recent work in the area of mixed formulations for Dirichlet boundary and interface condition enforcement has focused on "global polynomial" discretizations of the Lagrange multiplier [25], stabilization methods based on jump penalty operators [26], and duality-based error representation formulas [5].

We extend the ideas of [3, 22] to the multibody continuity constraint problem by proposing the following stabilized variational statement. Find $(T^h, \lambda^h) \in \mathcal{S}^h \times Q^h$ such that:

$$\sum_{m=1}^{2} \left(\delta J^{(m)}(T^{(m,h)}), v^{(m,h)} \right) + \left\langle \lambda^h, v^{(1,h)} - v^{(2,h)} \right\rangle + F_\delta = 0 \tag{40}$$

$$\left\langle \mu^h, T^{(1,h)} - T^{(2,h)} \right\rangle + G_\delta = 0 \tag{41}$$

holds for every $(v^h, \mu^h) \in \mathcal{V}^h \times Q^h$. The stabilization operators are

$$\begin{aligned}
F_\delta \left(T^h, \lambda^h; v^h \right) &\equiv F^{(1)}(T^{(1,h)}, \lambda^h; v^{(1,h)}) + F^{(2)}(T^{(2,h)}, \lambda^h; v^{(2,h)}) \\
&= - \left\langle A^{(1)} v^{(1,h)}, \lambda^h + A^{(1)} T^{(1,h)} \right\rangle_\delta + \left\langle A^{(2)} v^{(2,h)}, \lambda^h - A^{(2)} T^{(2,h)} \right\rangle_\delta
\end{aligned} \tag{42}$$

$$\begin{aligned}
G_\delta \left(T^h, \lambda^h; \mu^h \right) &\equiv G^{(1)}(T^{(1,h)}, \lambda^h; \mu^h) + G^{(2)}(T^{(2,h)}, \lambda^h; \mu^h) \\
&= - \left\langle \mu^h, \lambda^h + A^{(1)} T^{(1,h)} \right\rangle_\delta - \left\langle \mu^h, \lambda^h - A^{(2)} T^{(2,h)} \right\rangle_\delta
\end{aligned} \tag{43}$$

the boundary operator $A^{(m)}$ is defined by

$$A^{(m)} v \equiv k^{(m)} \nabla v \cdot \hat{\boldsymbol{n}}^{(m)} \tag{44}$$

for $m = 1, 2$, and the $\langle \cdot, \cdot \rangle_\delta$ inner products are defined in terms of the discrete sums

$$\langle u, v \rangle_\delta \equiv \sum_e \delta h_e \int_{\Gamma_e} uv \, \mathrm{d}s \tag{45}$$

where the sum ranges over the boundary elements discretizing the contact interface $\Gamma_C^{(1)}$, δ is a constant which depends on the geometric element type, and h_e is the size of the current edge Γ_e.

We make the following observations about (42) and (43):

1. The stabilized formulation is "consistent" with the original problem since, for the true solution, we have $\lambda = -A^{(1)} T^{(1)} = A^{(2)} T^{(2)}$, and the additional terms vanish.

2. The formulation adds stabilizing contributions to both the primal and constraint equations. In the primal equation, the symmetric boundary contributions

$$- \left\langle A^{(m)} v^{(m,h)}, A^{(m)} T^{(m,h)} \right\rangle_\delta, \quad m = 1, 2$$

add stability. In the constraint equation, the stabilization comes from the "negative mass matrix" contribution, $-2 \left\langle \mu^h, \lambda^h \right\rangle_\delta$, and is related to the "perturbed Lagrangian" construction discussed in e.g. [14].

3. As proven in [3, 22], the additional terms ensure that the Ladyzhenskaya-Babuška-Brezzi (LBB) stability conditions for the mixed formulation are satisfied when equal-order or matching-mesh interpolations are used for both the primal and Lagrange multiplier variables.

4. The formulation maintains the symmetry of the original problem, as will be demonstrated below.

The ith residual contribution of each of the stabilizing terms is

$$F_i^{(1)} \equiv -\left\langle A^{(1)} N_i^{(1)}, \lambda^h + A^{(1)} T^{(1,h)} \right\rangle_\delta \quad i \in \mathcal{T}^{(1)} \tag{46}$$

$$F_i^{(2)} \equiv \left\langle A^{(2)} N_i^{(2)}, \lambda^h - A^{(2)} T^{(2,h)} \right\rangle_\delta \quad i \in \mathcal{T}^{(2)} \tag{47}$$

$$G_i^{(1)} \equiv -\left\langle M_i, \lambda^h + A^{(1)} T^{(1,h)} \right\rangle_\delta \quad\quad i \in \mathcal{L} \tag{48}$$

$$G_i^{(2)} \equiv -\left\langle M_i, \lambda^h - A^{(2)} T^{(2,h)} \right\rangle_\delta \quad\quad i \in \mathcal{L} \tag{49}$$

where $\mathcal{T}^{(1)}$ and $\mathcal{T}^{(2)}$, are the sets of temperature (primal) variable degrees of freedom on the slave and master sides of the contact surface, respectively, and $\mathcal{L}$ is the set of the Lagrange multiplier degrees of freedom. The Jacobian contributions corresponding to $F^{(1)}$ and $F^{(2)}$ are

$$\frac{\partial F_i^{(1)}}{\partial T_j^{(1)}} \equiv -\left\langle A^{(1)} N_i^{(1)}, A^{(1)} N_j^{(1)} \right\rangle_\delta \quad i,j \in \mathcal{T}^{(1)} \tag{50}$$

$$\frac{\partial F_i^{(1)}}{\partial \lambda_j} \equiv -\left\langle A^{(1)} N_i^{(1)}, M_j \right\rangle_\delta \quad\quad i \in \mathcal{T}^{(1)}, \, j \in \mathcal{L} \tag{51}$$

and

$$\frac{\partial F_i^{(2)}}{\partial T_j^{(2)}} \equiv -\left\langle A^{(2)} N_i^{(2)}, A^{(2)} N_j^{(2)} \right\rangle_\delta \quad i,j \in \mathcal{T}^{(2)} \tag{52}$$

$$\frac{\partial F_i^{(2)}}{\partial \lambda_j} \equiv \left\langle A^{(2)} N_i^{(2)}, M_j \right\rangle_\delta \quad\quad i \in \mathcal{T}^{(2)}, \, j \in \mathcal{L} \tag{53}$$

The Jacobian contributions corresponding to $G^{(1)}$ and $G^{(2)}$ are

$$\frac{\partial G_i^{(1)}}{\partial T_j^{(1)}} \equiv -\left\langle M_i, A^{(1)} N_j^{(1)} \right\rangle_\delta \quad i \in \mathcal{L}, \, j \in \mathcal{T}^{(1)} \tag{54}$$

$$\frac{\partial G_i^{(1)}}{\partial \lambda_j} \equiv -\left\langle M_i, M_j \right\rangle_\delta \quad\quad i,j \in \mathcal{L} \tag{55}$$

and

$$\frac{\partial G_i^{(2)}}{\partial T_j^{(2)}} \equiv \left\langle M_i, A^{(2)} N_j^{(2)} \right\rangle_\delta \quad i \in \mathcal{L}, \, j \in \mathcal{T}^{(2)} \tag{56}$$

$$\frac{\partial G_i^{(2)}}{\partial \lambda_j} \equiv -\left\langle M_i, M_j \right\rangle_\delta \quad\quad i,j \in \mathcal{L} \tag{57}$$

The "on-diagonal" contributions (50), (52), (55), and (57) are clearly symmetric. Likewise, if one exchanges the indices i and j in (51), one obtains (54), and if one exchanges i and j in (53), one obtains (56).

4 A priori error estimates

In this section, to simplify the notation, we use Γ to denote the shared contact interface $\Gamma_C = \Gamma_C^{(1)} = \Gamma_C^{(2)}$. In a series of papers in the early 1980s, Pitkäranta [18, 19, 21] expanded on the earlier work of Babuška [2], Brezzi [17], Bramble [20], and others by determining conditions under which the Lagrange multiplier formulation of the Dirichlet problem is stable, including optimal values for the constant C in the condition $h_\lambda \geq Ch_u$, and techniques for performing local refinements near the boundary in order to properly capture singularities in non-convex, piecewise-smooth domains. In [19], Pitkäranta proposed the following mesh-dependent norm for measuring the error in the combined primal/Lagrange multiplier solution

$$\||v, \mu\||_h^2 \equiv \int_\Omega |\nabla v|^2 \, \mathrm{d}x + \sum_e h_e^{-1} \int_{\Gamma_e} v^2 \, \mathrm{d}s + \sum_e h_e \int_{\Gamma_e} \mu^2 \, \mathrm{d}s \tag{58}$$

where the sums range over the elements discretizing the boundary (or, in our case, the shared contact interface) Γ, and h_e is the size of the current edge Γ_e. Note the different h_e scalings applied to each term. These ensure that the different contributions are summed in a dimensionally-consistent way, and that the resulting quantity appropriately bounds other standard norms of the primal variable. For example, in two dimensions, one has

$$\int_\Omega |\nabla v|^2 \, \mathrm{d}x \propto \frac{V^2}{L^2} L^2 = V^2 \tag{59}$$

$$h_e^{-1} \int_{\Gamma_e} v^2 \, \mathrm{d}s \propto \frac{1}{L} V^2 L = V^2 \tag{60}$$

$$h_e \int_{\Gamma_e} \mu^2 \, \mathrm{d}s \propto L \frac{V^2}{L^2} L = V^2 \tag{61}$$

where $\propto$ is used to denote the "physical units" of a given term, the units of the primal variable are V, and the flux therefore has units of $\frac{V}{L}$, where L is the unit of length. We note that in §3 of [19], p. 1116, there is a typo in the definition of the mesh-dependent norm. The middle subscript of the final term should be $+1/2$, the corrected form is given in (58).

For the Dirichlet problem

$$\nabla^2 u = f \ \in \Omega \tag{62}$$
$$u = g \ \text{on } \partial\Omega \tag{63}$$

with data $f \in L^2(\Omega)$, $g \in H^r(\partial\Omega)$, $r > 1$, Pitkäranta [19] proposed the variational formulation: find $(u, \lambda) \in H^1(\Omega) \times L^2(\partial\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x - \int_{\partial\Omega} \lambda v \, \mathrm{d}s + \int_{\partial\Omega} (u - g)\mu \, \mathrm{d}s = -\int_\Omega fv \, \mathrm{d}x \tag{64}$$

holds for every $(v, \mu) \in H^1(\Omega) \times L^2(\partial\Omega)$, where $\lambda = \frac{\partial u}{\partial n}$, the normal flux of u on $\partial\Omega$. He then showed, under certain conditions on the mesh and approximation spaces, that the error $e_u \equiv u - u^h$, $e_\lambda \equiv \lambda - \lambda^h$ is bounded in the mesh-dependent norm (58) according to

$$\||e_u, e_\lambda\||_h \leq Ch^k \left(\|f\|_{H^{k-1}(\Omega)} + \|g\|_{H^{k+1/2}(\partial\Omega)} \right) \tag{65}$$

In (65), h is an average measure of the mesh size, C is a mesh-independent constant, and $\|\cdot\|_{H^k(\Omega)}$ is the usual norm on the Sobolev space $H^k(\Omega)$. Here $k \geq 1$ can also be understood as the degree of the polynomial used to approximate u^h. In the estimate (65), it is not possible to predict how the primal or flux errors converge *individually*. In particular, it does not preclude the possibility that the flux error, e_λ, may converge *faster* than the primal error e_u in certain circumstances, or vice-versa, only that the error in the combined norm converges at $\mathcal{O}(h^k)$.

An important characteristic of Pitkäranta's approach is that, instead of working with the (conceptually much more difficult to understand) Sobolev space $H^{-1/2}(\partial\Omega)$, the analysis is carried out in the simpler space $L^2(\partial\Omega)$, albeit supplied with a specialized, mesh-dependent norm. In particular, his estimate does not require that the true flux $\lambda \in H^{k-1/2}(\partial\Omega)$, $k \geq 1$, only that the boundary data g is fairly regular (at least $g \in H^{3/2}(\partial\Omega)$) which may be an easier requirement to satisfy in applications. Furthermore, as he points out on p. 1124 [19], the regularity of $\partial\Omega$ did not play an explicit role in his analysis, and his result holds even when Ω is a Lipschitz domain with piecewise-smooth boundary.

This fact is very important for applications, which are almost always posed on polygonal domains, and therefore have a discontinuous true flux λ near the domain corners due to the discontinuity of the domain's outward normal. For example, in 1D it is possible to show, via direct calculation [27], that discontinuous functions belong to $H^{1/2-\epsilon}$, $\epsilon > 0$, and therefore $\lambda \notin H^{1/2}(\partial\Omega)$ for such problems. The *a priori* estimate we will discuss momentarily will require additional smoothness of λ, which in some cases is not available. In a related context, Hansbo [25] (p. 110) also mentions that "full elliptic regularity can only be expected if [the interface] Γ is sufficiently smooth."

Finally, we want to note some of the differences between Pitkäranta's approach and the applications of interest in this work. First, we are not using the Lagrange multiplier method to impose Dirichlet boundary conditions, but to enforce solution continuity between disjoint subdomains. Second, while Pitkäranta's result applies to the entire boundary $\partial\Omega$, we only define the Lagrange multipliers on interfaces, which may be thought of as subsets of the boundary. Finally, Pitkäranta did not explicitly consider different polynomial orders in the finite element approximations of the primal and Lagrange multiplier variables, so his error estimate involves only a single exponent, k. Nevertheless, it appears that these differences are not fundamentally in conflict, and the mesh-dependent norm and estimate are still applicable in the present work.

Barbosa and Hughes [3] also introduced a mesh-dependent norm which is closely related to (58), and additionally depends on the stabilization parameter δ. Their norm, given here in a form relevant to the present discussion is, for $(v, \mu) \in H^1(\Omega) \times L^2(\Gamma)$:

$$\|v, \mu\|_\delta^2 \equiv \frac{1}{2}\|v\|_{H^1(\Omega)}^2 + \delta \sum_e h_e \int_{\Gamma_e} \mu^2 \, \mathrm{d}s \tag{66}$$

They then prove the *a priori* estimate

$$\|e_u, e_\lambda\|_\delta \leq C \left(h^k \|u\|_{H^{k+1}(\Omega)} + h^{\ell+3/2} \|\lambda\|_{H^{\ell+1}(\Gamma)} \right) \tag{67}$$

in this mesh-dependent norm, where C is a constant independent of h, and $k \geq 1$, $\ell \geq 0$ are the polynomial orders used to approximate u^h and λ^h, respectively.

The estimate (67) effectively decouples the smoothness required for the true flux λ from the polynomial order k used to approximate the primal variable u, but may require more smoothness in λ than is actually available in typical applications. The estimate (67) also predicts the error "barriers" summarized in Table 1. In this context, "barrier" refers to the fact that increasing k or ℓ in isolation may be insufficient to improve the combined rate of convergence (even assuming sufficient regularity in the underlying true solutions).

Table 1: Error barriers for the mixed formulation according to the estimate (65). Increasing the Lagrange multiplier polynomial degree ℓ does not result in any improvement in the combined rate of convergence when linear ($k = 1$) elements are used for the primal variable. When quadratics ($k = 2$) are used, increasing the Lagrange multiplier polynomial degree from $\ell = 0$ to $\ell = 1$ improves the combined rate of convergence to $\mathcal{O}(h^2)$, but further increases in ℓ have no effect on the combined convergence rate.

k	ℓ	$\|\|e_u, e_\lambda\|\|_\delta$
1	0	$\mathcal{O}(h)$
1	1	$\mathcal{O}(h)$
2	0	$\mathcal{O}(h^{3/2})$
2	1	$\mathcal{O}(h^2)$
2	2	$\mathcal{O}(h^2)$

Finally, we note that it is possible (but not guaranteed) that the error in the flux variable alone will converge *faster* than the combined rate. For example, if a discontinuous constant interpolation ($\ell = 0$) is used for the Lagrange multiplier, we may observe better-than-linear, i.e. $\mathcal{O}(h^{3/2})$, convergence in the flux discretization, provided that the true flux is in $H^1(\Gamma)$. As discussed previously, however, for Lipschitz domains we may not even have $\lambda \in H^{1/2}(\Gamma)$, and then the higher convergence rate for the flux is not guaranteed. In fact, it may even be possible to observe suboptimal convergence rates for the primal variable, i.e. the flux error may pollute the overall estimate, in such cases. We are not currently aware of any *a priori* error estimates which cover the case of Lipschitz-continuous interfaces, however, our numerical results suggest that the Lagrange multiplier approach may still be convergent (although with lower rates) in such cases.

5 Nodal normal projection equations

In this section, we discuss the details of solving Eqs. (2.4.5) and (2.4.6) in Yang's dissertation [13] for the reference coordinate positions $\xi^{(2)}$ and $\xi^{(1)}$, respectively. This approach is also described in the related research article [12].

5.1 Projection of slave nodes onto master surface

In the notation of Yang, we have:

$$\left[N_1(\xi^{(2)})\boldsymbol{\varphi}_1^{(2)} + N_2(\xi^{(2)})\boldsymbol{\varphi}_2^{(2)} - \boldsymbol{\varphi}_s^{(1)} \right] \times \boldsymbol{n}_s^{(1)} = \boldsymbol{0} \tag{2.4.5}$$

where the superscript (1) refers to the "slave" surface, (2) refers to the "master" surface, $N_i(\xi^{(2)})$ is finite element basis function i on the master surface evaluated at $\xi^{(2)}$, $\boldsymbol{\varphi}_i^{(2)}$ is node i of the master element, $\boldsymbol{\varphi}_s^{(1)}$ is the location of the slave node which is being projected onto the master surface, and $\boldsymbol{n}_s^{(1)}$ is the outward nodal normal vector at that point. Eq. 2.4.5 is applicable when the master element into which slave node $\boldsymbol{n}_s^{(1)}$ projects has already been found via some other search algorithm. Letting

$$\boldsymbol{x}^{(2)} \equiv N_1(\xi^{(2)})\boldsymbol{\varphi}_1^{(2)} + N_2(\xi^{(2)})\boldsymbol{\varphi}_2^{(2)} \tag{68}$$

be the physical space location on the master surface of the reference coordinate $\xi^{(2)}$, we can rewrite (2.4.5) in a potentially simpler form as:

$$\left[\boldsymbol{x}^{(2)} - \boldsymbol{\varphi}_s^{(1)} \right] \times \boldsymbol{n}_s^{(1)} = \boldsymbol{0} \tag{69}$$

Eq. (69) is in general a set of three equations in the single unknown $\xi^{(2)}$. An approximate solution to this overdetermined set of equations can be found via the least squares solution method, but we will consider the special case in which the vectors $\boldsymbol{x}^{(2)} - \boldsymbol{\varphi}_s^{(1)}$ and $\boldsymbol{n}_s$ lie in the (x_1, x_2) plane, and therefore their cross product is in the x_3-direction only. We can then obtain an explicit solution to (69), which further reduces to the scalar equation:

$$u_1 v_2 - u_2 v_1 = 0 \tag{70}$$

where

$$\boldsymbol{u} \equiv \boldsymbol{x}^{(2)} - \boldsymbol{\varphi}_s^{(1)} \tag{71}$$

$$\boldsymbol{v} \equiv \boldsymbol{n}_s^{(1)} \tag{72}$$

Substituting $N_1(\xi) \equiv \frac{1-\xi}{2}$, $N_2(\xi) \equiv \frac{1+\xi}{2}$ into (70) and rearranging yields:

$$\xi^{(2)} = \frac{-\left[\varphi_{1,1}^{(2)} + \varphi_{2,1}^{(2)} - 2\varphi_{s,1}^{(1)} \right] n_{s,2}^{(1)} + \left[\varphi_{1,2}^{(2)} + \varphi_{2,2}^{(2)} - 2\varphi_{s,2}^{(1)} \right] n_{s,1}^{(1)}}{\left[-\varphi_{1,1}^{(2)} + \varphi_{2,1}^{(2)} \right] n_{s,2}^{(1)} - \left[-\varphi_{1,2}^{(2)} + \varphi_{2,2}^{(2)} \right] n_{s,1}^{(1)}} \tag{73}$$

where subscripts after commas refer to the spatial coordinate, i.e.

$$\boldsymbol{n}_s^{(1)} \equiv \left(n_{s,1}^{(1)}, n_{s,2}^{(1)} \right) \tag{74}$$

$$\boldsymbol{\varphi}_1^{(2)} \equiv \left(\varphi_{1,1}^{(2)}, \varphi_{1,2}^{(2)} \right) \tag{75}$$

etc. A representative configuration in which a slave node is mapped onto the surface of a master element is shown in Fig. 2. This formula for computing the projection also works

when the slave node $\varphi_s^{(1)}$ lies directly on the master surface, i.e. in the case where $\boldsymbol{x}^{(2)} = \varphi_s^{(1)}$, since this configuration satisfies (69). When the master element and the nodal normal are parallel, the denominator of (73) will be zero. This does not indicate a failure of the method, however, only that the normal projection of $\varphi_s^{(1)}$ cannot possibly lie on the master element in question, and the implementation should therefore prevent floating point exceptions from being raised in this case.
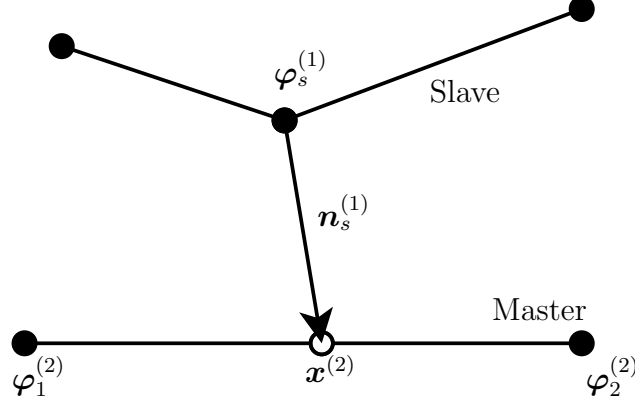
Figure 2: Diagram demonstrating projection of the slave node $\varphi_s^{(1)}$ along the nodal normal direction onto an element on the master surface defined by the nodes $(\varphi_1^{(2)}, \varphi_2^{(2)})$. The solid dots represent nodes in the finite element mesh, the open dot shows where basis functions on the master surface must be evaluated during mortar segment integral computations.

5.2   Inverse-projection of master nodes from slave surface

In the converse case, we have a similar geometric situation, but need to perform an "inverse" mapping in order to determine where a given node $\varphi_m^{(2)}$ on the master surface would have originated from, had it been projected along the slave surface nodal normal direction. The nonlinear equation to be satisfied in this case is Yang's Eq. (2.4.6) which we repeat here:

$$\left[N_1(\xi^{(1)})\varphi_1^{(1)} + N_2(\xi^{(1)})\varphi_2^{(1)} - \varphi_m^{(2)} \right] \times \left[N_1(\xi^{(1)})\boldsymbol{n}_1^{(1)} + N_2(\xi^{(1)})\boldsymbol{n}_2^{(1)} \right] = \boldsymbol{0} \qquad (2.4.6)$$

The unknown in (2.4.6) is $\xi^{(1)}$, and, as shown in Fig. 3, $\boldsymbol{n}_i^{(1)}$ is the nodal normal at slave surface node i,

$$\boldsymbol{x}^{(1)} \equiv N_1(\xi^{(1)})\varphi_1^{(1)} + N_2(\xi^{(1)})\varphi_2^{(1)} \qquad (76)$$

is the unknown physical-space coordinate of the location where master node $\varphi_m^{(2)}$ would have been projected from, and

$$\boldsymbol{n}(\boldsymbol{x}^{(1)}) \equiv N_1(\xi^{(1)})\boldsymbol{n}_1^{(1)} + N_2(\xi^{(1)})\boldsymbol{n}_2^{(1)} \qquad (77)$$

is the outward nodal normal vector on the slave surface at that location. Using (76) and (77) allows us to write (2.4.6) more compactly as:

$$\left[\boldsymbol{x}^{(1)} - \varphi_m^{(2)} \right] \times \boldsymbol{n}(\boldsymbol{x}^{(1)}) = \boldsymbol{0} \qquad (78)$$

13

It is clear that relation (78) is nonlinear (quadratic) in $\boldsymbol{x}^{(1)}$ which depends on the unknown reference coordinate $\xi^{(1)}$, and therefore we do not expect to be able to solve it directly as was the case in (73).
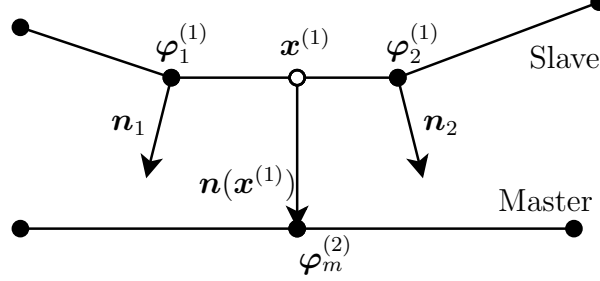
Figure 3: Diagram demonstrating "inverse" projection of the master node $\boldsymbol{\varphi}_m^{(2)}$ to the point $\boldsymbol{x}^{(1)}$ where it *would have come from* along the slave surface nodal normal direction. The solid dots represent nodes in the finite element mesh, the open dot shows the endpoint of a mortar segment used for mortar integral evaluation.

Instead, once again assuming that all vectors lie in the (x_1, x_2) plane and defining

$$\boldsymbol{u} \equiv \boldsymbol{x}^{(1)} - \boldsymbol{\varphi}_m^{(2)} \tag{79}$$

$$\boldsymbol{v} \equiv \boldsymbol{n}(\boldsymbol{x}^{(1)}) \tag{80}$$

and the residual

$$F(\xi^{(1)}) \equiv u_1 v_2 - u_2 v_1 \tag{81}$$

we will use Newton's method to compute $\xi_*^{(1)}$ such that $F(\xi_*^{(1)}) = 0$ using an initial guess of $\xi^{(1)} = 0$. Newton's method requires the first derivative of the functional, which in this case is simply

$$F'(\xi^{(1)}) \equiv (u_1 v_2' + u_1' v_2) - (u_2 v_1' + u_2' v_1) \tag{82}$$

where

$$u_i' \equiv N_1' \varphi_{1,i}^{(1)} + N_2' \varphi_{2,i}^{(1)} \tag{83}$$

$$v_i' \equiv N_1' n_{1,i} + N_2' n_{2,i} \tag{84}$$

for $i = 1, 2$, where $N_1' \equiv -\frac{1}{2}$ and $N_2' \equiv \frac{1}{2}$.

In practice, we have observed that the Newton iterations typically converge quite quickly (within 1 or 2 steps) and reliably given the initial guess of $\xi^{(1)} = 0$. When the Newton iterations converge to $|\xi^{(1)}| > 1$, this indicates that the master point in question did not come from the slave surface element we are searching, and a different candidate element should be searched instead. Finally, we observe that master nodes near the "edge" of a non-closed master surface contour may not inverse-map to *any* element on the slave surface. This does not indicate a failure of the method, however, only that the corresponding mortar segment will not have any contribution from the master side.

6 Classes of mortar segments

In this section, we describe in greater detail the various types of mortar segments which can occur in a computation. Although the types of segments do not fundamentally alter the manner in which the mortar integrals are computed, it is useful for us to be able to classify them in order for us to treat them correctly in the implementation. In what follows, we arbitrarily divide the types of segments into "interior" and "exterior" classes.

Interior mortar segments occur when all four endpoints $\xi_a^{(1)}$, $\xi_b^{(1)}$, $\xi_a^{(2)}$, $\xi_b^{(2)}$ occur in or on elements of the master and slave boundary meshes. The different interior mortar segment configurations are shown in Fig. 4. In the top two cases (I and II), the elements from the master and slave sides overlap, and one segment endpoint from each side is a mesh node from that side. For example, in Case I, $\xi_a^{(1)} = \pm 1$ and $\xi_b^{(2)} = \pm 1$, depending on the local orientation of the elements. In the bottom two cases (III and IV), a slave (resp. master) element's nodes both project within a single element on the master (resp. slave) side. In these cases, both $\xi^{(1)}$ (resp. $\xi^{(2)}$) values will be equal to either 1 or -1. A trivial case (not shown) is also possible in which the master and slave nodes exactly align and all four endpoint values are therefore ± 1.
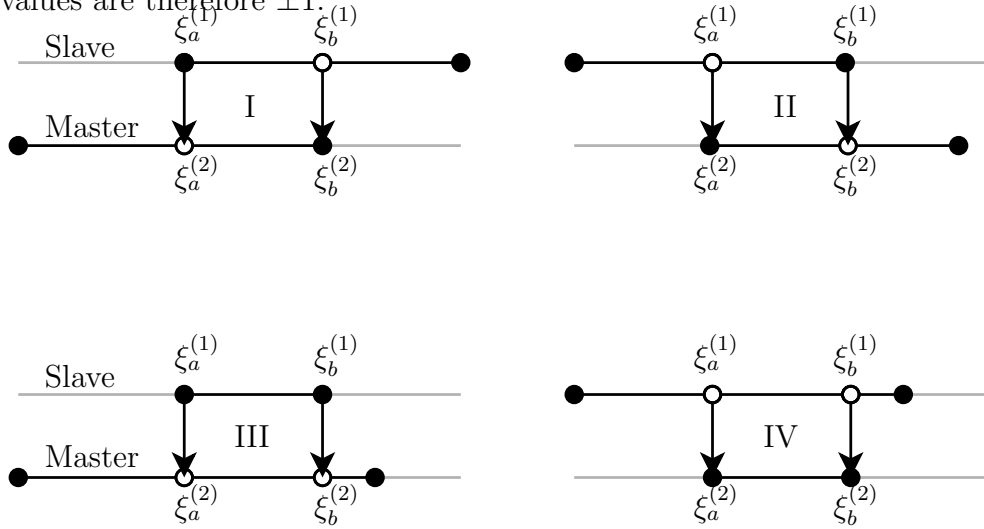
Figure 4: Different classes of "interior" mortar segments. The black lines correspond to the elements whose degrees of freedom contribute to the segment shown, the light grey lines represent other elements which do not contribute. Filled circles represent nodes from the master and slave boundary meshes, while open circles represent nodal normal projections as described in §5.

Exterior mortar segments occur when slave nodes fail to project to any elements on the master surface. Examples of exterior mortar segments are shown in Cases I and III of Fig. 5. In Case I, a slave element partially overlaps the first element on the master side, and its first node is not in contact. The first mortar segment on the slave side, denoted by the endpoints $\xi_a^{(1)}$ and $\xi_b^{(1)}$ therefore do not have corresponding $\xi_a^{(2)}$ and $\xi_b^{(2)}$ values on the master side, nor is there a corresponding master element for this segment. The slave/slave contribution for this segment is still nonzero, but the master/slave contribution will be identically zero. This situation should result in the Lagrange multiplier variable being weakly set equal to zero on the first node of the slave side.
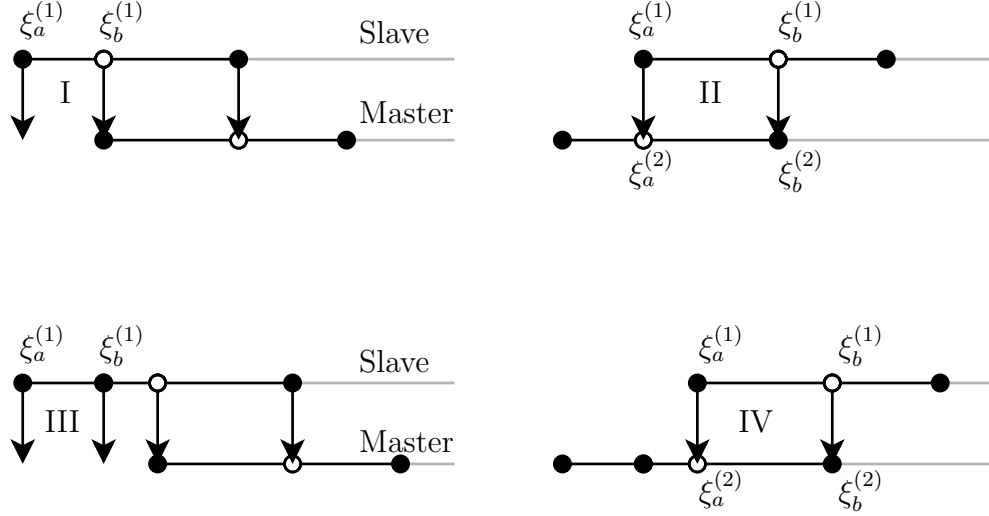
Figure 5: Different classes of "exterior" mortar segments. The notation mirrors that of Fig. 4. Nodes whose nodal projections do not intersect any elements are considered to be out of contact, and their master/slave mortar integral contributions are simply zero.

In Case III of Fig. 5, the entire first element on the slave side is not in contact with the master side. (Generally, there can be multiple consecutive slave side elements which are not in contact with the master surface; this scenario is handled analogously.) In this case, the master/slave contributions for the first two mortar segments will be identically zero, and the following segment will be treated as in Case II of Fig. 4. Cases II and IV in Fig. 5 depict situations in which nodes from the master side cannot be successfully inverse-mapped to a corresponding element on the slave side. These cases are indistinguishable from Case I of Fig. 4 as far as the mortar integral computations are concerned. The first mortar segment will have a nonzero contribution for both the slave/slave and master/slave integrals, regardless of how many master nodes are out of contact.

7 Discrete mortar integrals

In this section, we discuss the discrete form of the mortar integral terms in (33) and (34), which for brevity we will refer to as:

$$F \equiv \int_{\Gamma_C^{(1)}} \lambda \left(v^{(1)} - P v^{(2)} \right) \, \mathrm{d}s \tag{85}$$

$$G \equiv \int_{\Gamma_C^{(1)}} \mu \left(T^{(1)} - P T^{(2)} \right) \, \mathrm{d}s \tag{86}$$

The corresponding discrete forms of (85) and (86) are

$$F^h \equiv \int_{\Gamma_C^{(1,h)}} \lambda^h \left(v^{(1,h)} - P v^{(2,h)} \right) \, \mathrm{d}s \tag{87}$$

$$G^h \equiv \int_{\Gamma_C^{(1,h)}} \mu^h \left(T^{(1,h)} - P T^{(2,h)} \right) \, \mathrm{d}s \tag{88}$$

16

The discretized unknowns are written in the usual way as sums over the basis functions

$$T^{(1,h)} = \sum_{j \in \mathcal{T}^{(1)}} T_j^{(1)} N_j^{(1)} \tag{89}$$

$$T^{(2,h)} = \sum_{j \in \mathcal{T}^{(2)}} T_j^{(2)} N_j^{(2)} \tag{90}$$

$$\lambda^h = \sum_{j \in \mathcal{L}} \lambda_j M_j \tag{91}$$

where $\mathcal{T}^{(1)}$ and $\mathcal{T}^{(2)}$, are the sets of temperature (primal) variable degrees of freedom on the slave and master sides of the contact surface, respectively, and $\mathcal{L}$ is the set of the Lagrange multiplier degrees of freedom. $N_j^{(1)}$, $N_j^{(2)}$, M_j are the corresponding primal and mortar space finite element basis functions. The Lagrange multiplier is always defined on the slave (1) side, so there is no need to use a superscript for disambiguation. Correspondingly, we have

$$v^{(1,h)} = N_i^{(1)}, \quad i \in \mathcal{T}^{(1)} \tag{92}$$

$$v^{(2,h)} = N_i^{(2)}, \quad i \in \mathcal{T}^{(2)} \tag{93}$$

$$\mu^h = M_i, \quad i \in \mathcal{L} \tag{94}$$

for the test functions. Here we have assumed that the same order and family (e.g. first-order Lagrange) finite elements are used in both bodies for the primal variable, but allowed for the possibility that the Lagrange multiplier variable may be discretized with a different basis.

There will be two residual contributions (one for the master side degrees of freedom, and one for the slave side degrees of freedom) due to (87). These contributions are given by

$$F_i^{(1,h)} \equiv \int_{\Gamma_C^{(1)}} \lambda^h N_i^{(1)} \, \mathrm{d}s, \quad i \in \mathcal{T}^{(1)} \tag{95}$$

$$F_i^{(2,h)} \equiv -\int_{\Gamma_C^{(1)}} \lambda^h P N_i^{(2)} \, \mathrm{d}s, \quad i \in \mathcal{T}^{(2)} \tag{96}$$

The corresponding Jacobian contributions are

$$\frac{\partial F_i^{(1,h)}}{\partial \lambda_j} = \int_{\Gamma_C^{(1)}} M_j N_i^{(1)} \, \mathrm{d}s, \quad i \in \mathcal{T}^{(1)}, j \in \mathcal{L} \tag{97}$$

$$\frac{\partial F_i^{(2,h)}}{\partial \lambda_j} = -\int_{\Gamma_C^{(1)}} M_j P N_i^{(2)} \, \mathrm{d}s, \quad i \in \mathcal{T}^{(2)}, j \in \mathcal{L} \tag{98}$$

Likewise, the residual contribution of (88) due to Lagrange multiplier test function i is given by

$$G_i^h \equiv \int_{\Gamma_C^{(1)}} M_i \left(T^{(1,h)} - P T^{(2,h)} \right) \mathrm{d}s, \quad i \in \mathcal{L} \tag{99}$$

with corresponding Jacobian contributions

$$\frac{\partial G_i^h}{\partial T_j^{(1)}} = \int_{\Gamma_C^{(1)}} M_i N_j^{(1)} \, \mathrm{d}s, \qquad i \in \mathcal{L}, \, j \in \mathcal{T}^{(1)} \tag{100}$$

$$\frac{\partial G_i^h}{\partial T_j^{(2)}} = - \int_{\Gamma_C^{(1)}} M_i P N_j^{(2)} \, \mathrm{d}s, \quad i \in \mathcal{L}, \, j \in \mathcal{T}^{(2)} \tag{101}$$

We readily observe that the Jacobian contributions (97) and (100) are symmetric (they are equal if the indices i and j are interchanged) as are (98) and (101). Therefore, in the detailed example calculations that follow, it suffices to consider only the contributions arising from F, as the same conclusions will apply to the contributions due to G.

Terms (95) and (97) depend only on values from the slave side, and are straightforward to compute using the existing discretization of the slave boundary. On the other hand, the terms (96) and (98) depend on values from *both* the slave and master sides, and quadrature errors due to the integration of locally piecewise continuous functions will result if we attempt to compute the integral using only the slave side discretization.

8 Computing the mortar integral

In the following subsections, we discuss various details of the computation of the mortar integrals, assuming the mortar segments have already been constructed via the methods described previously. In §8.1, we compare the difference in computing element Jacobian contributions using a "naïve" approach vs. the mortar segment-based approach described in this report. Then, in §8.2 and §8.3, we walk through the process of defining a reference coordinate on the mortar segments and computing the integrals numerically (via quadrature) for a particularly simple nonconforming mesh.

8.1 Potential inaccuracies

To understand the potential accuracy issues involved with computing the mortar integral contributions, we consider a simple example in which a mesh with equal-sized elements is "perfectly unaligned" with a mesh of equal-sized elements on the other side of the contact interface as in Fig. 6. For simplicity, we also assume the mortar and primal variables are discretized with the same finite element basis, so that $M_i = N_i^{(1)} \, \forall i$ in what follows.

In this case, the nodal normal projection operation is trivial, and the master side basis functions are defined piecewise on slave side element Ω_e as:

$$N_a^{(2)} = \begin{cases} 1 + \xi^{(1)}/2, & -1 \le \xi^{(1)} \le 0 \\ 1 - \xi^{(1)}/2, & 0 \le \xi^{(1)} \le 1 \end{cases} \tag{102}$$

$$N_b^{(2)} = \begin{cases} 0 & -1 \le \xi^{(1)} \le 0 \\ \xi^{(1)}/2, & 0 \le \xi^{(1)} \le 1 \end{cases} \tag{103}$$

18
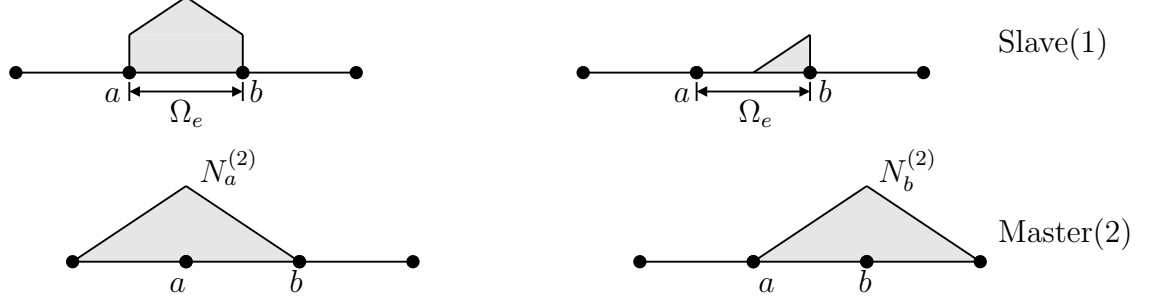
Figure 6: Master side basis functions $N_a^{(2)}$ and $N_b^{(2)}$ on slave side element Ω_e in the perfectly unaligned mesh case. Two of the three master side basis functions with support on Ω_e are shown, the last one is symmetric to $N_b^{(2)}$.

while the slave side basis functions are defined in the usual way as

$$N_a^{(1)} = \frac{1 - \xi^{(1)}}{2} \tag{104}$$

$$N_b^{(1)} = \frac{1 + \xi^{(1)}}{2} \tag{105}$$

The master/slave contributions due to (98) can be organized in matrix form as

$$K_{ms} = \begin{bmatrix} \int_{\Omega_e} N_a^{(2)} N_a^{(1)} \mathrm{d}s & \int_{\Omega_e} N_a^{(2)} N_b^{(1)} \mathrm{d}s \\ \int_{\Omega_e} N_b^{(2)} N_a^{(1)} \mathrm{d}s & \int_{\Omega_e} N_b^{(2)} N_b^{(1)} \mathrm{d}s \end{bmatrix} \tag{106}$$

The true values (using exact integration) of K_{ms} are given by

$$K_{ms} = \frac{h_e}{2} \begin{bmatrix} \frac{3}{4} & \frac{3}{4} \\ \frac{1}{24} & \frac{5}{24} \end{bmatrix} \tag{107}$$

where h_e is the length of element Ω_e. Using standard two-point quadrature, on the other hand, results in

$$\widetilde{K}_{ms} \approx \frac{h_e}{2} \begin{bmatrix} 0.711325 & 0.711325 \\ 0.061004 & 0.227671 \end{bmatrix} \tag{108}$$

This corresponds to a relative error

$$\frac{\|K_{ms} - \widetilde{K}_{ms}\|_\infty}{\|K_{ms}\|_\infty} \approx 5\% \tag{109}$$

Although relatively small in magnitude, this error will cause the usual global conservation properties of the finite element method to be lost. For instance, in contact mechanics problems, linear momentum will not be conserved if the mortar integrals are computed in this way; i.e. spurious momentum will be introduced in the gap between the two subdomains. Likewise, for thermal contact problems, the total energy will not be conserved.

8.2 Simple example 1

We next consider in detail the computation of mortar segment contributions for the simple, perfectly unaligned mesh in Fig. 7. In particular, we will compute the slave/slave and master/slave element stiffness contributions for segment S in Fig. 7. In this example, we assume the slave and master reference coordinates are aligned in opposite directions so that $\xi_a^{(1)} = 0$, $\xi_b^{(1)} = 1$, $\xi_a^{(2)} = 1$, $\xi_b^{(2)} = 0$. As discussed in Yang's dissertation [13], we parameterize the segment S with the secondary reference variable $-1 \leq \eta \leq 1$ such that

$$\xi^{(1)}(\eta) = \xi_a^{(1)} \frac{1 - \eta}{2} + \xi_b^{(1)} \frac{1 + \eta}{2} \tag{110}$$

$$\xi^{(2)}(\eta) = \xi_a^{(2)} \frac{1 - \eta}{2} + \xi_b^{(2)} \frac{1 + \eta}{2} \tag{111}$$

and again assume that the mortar and primal variables are discretized with the same finite element basis, which allows us to write the slave/slave and master/slave Jacobian contributions due to this segment as

$$K_{ss}(i,j) \equiv \int_S N_i^{(1)} N_j^{(1)} \mathrm{d}s = \frac{h_{\mathrm{seg}}}{2} \int_{-1}^{1} N_i^{(1)}(\xi^{(1)}(\eta)) N_j^{(1)}(\xi^{(1)}(\eta)) \, \mathrm{d}\eta \tag{112}$$

$$K_{ms}(i,j) \equiv \int_S N_i^{(2)} N_j^{(1)} \mathrm{d}s = \frac{h_{\mathrm{seg}}}{2} \int_{-1}^{1} N_i^{(2)}(\xi^{(2)}(\eta)) N_j^{(1)}(\xi^{(1)}(\eta)) \, \mathrm{d}\eta \tag{113}$$

where h_{seg} is the length of segment S, and the local indices $i, j = 1, 2$ correspond to the basis functions of the (single) slave and master elements which contribute to this segment (see Fig. 7). The basis functions are defined exactly as in (104) and (105), with indices a and b corresponding to 1 and 2. The method of constructing the mortar segments described in §6 guarantees that a unique element from each side contributes values to each segment, and that the integrands (112) and (113) are continuous (and therefore exactly integrable using, in this case, a two-point quadrature rule).

In this simple example, the slave/slave and master/slave Jacobian contributions can be computed explicitly as follows: the segment reference coordinate values are given by (110) and (111) as

$$\xi^{(1)}(\eta) = \frac{1 + \eta}{2} \tag{114}$$

$$\xi^{(2)}(\eta) = \frac{1 - \eta}{2} \tag{115}$$

and consequently the slave and master side basis functions in segment reference coordinates are

$$N_1^{(1)} = \frac{1 - \xi^{(1)}(\eta)}{2} = \frac{1 - \eta}{4} \tag{116}$$

$$N_2^{(1)} = \frac{1 + \xi^{(1)}(\eta)}{2} = \frac{3 + \eta}{4} \tag{117}$$

$$N_1^{(2)} = \frac{1 - \xi^{(2)}(\eta)}{2} = \frac{1 + \eta}{4} \tag{118}$$

$$N_2^{(2)} = \frac{1 + \xi^{(2)}(\eta)}{2} = \frac{3 - \eta}{4} \tag{119}$$

20
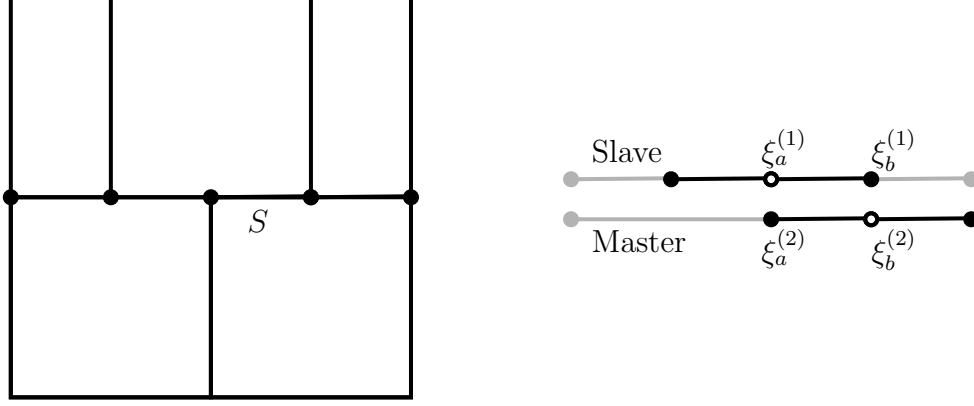
Figure 7: Perfectly unaligned mesh with five elements used for verification (left) and mortar segment detail for the segment marked S (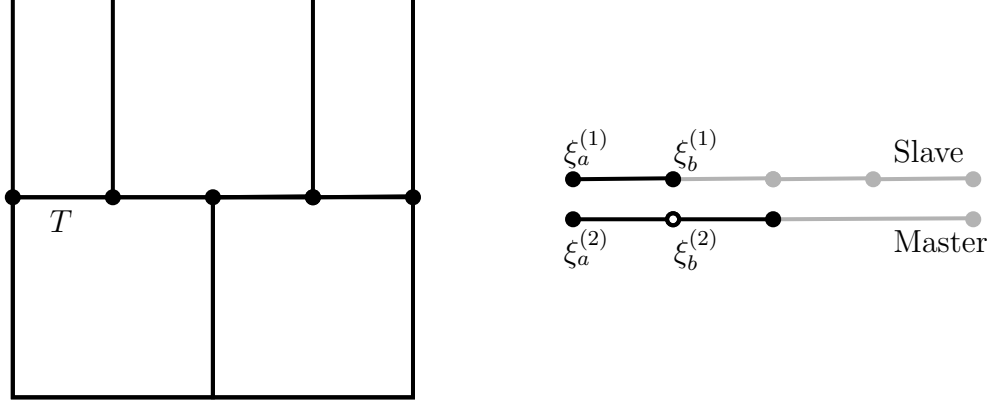right) with endpoints $\xi_a^{(1)} = 0$, $\xi_b^{(1)} = 1$, $\xi_a^{(2)} = 1$, $\xi_b^{(2)} = 0$. Dark lines indicate the faces of the elements whose degrees of freedom S contributes to, light grey lines indicate neighboring element faces which are not involved in the segment computation.

This leads to the following slave/slave

$$K_{ss}(1,1) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(1-\eta)^2}{16} \, d\eta = \frac{h_{\text{seg}}}{12} \tag{120}$$

$$K_{ss}(1,2) = K_{ss}(2,1) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(1-\eta)(3+\eta)}{16} \, d\eta = \frac{h_{\text{seg}}}{6} \tag{121}$$

$$K_{ss}(2,2) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(3+\eta)^2}{16} \, d\eta = h_{\text{seg}} \frac{7}{12} \tag{122}$$

and master/slave

$$K_{ms}(1,1) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(1+\eta)(1-\eta)}{16} \, d\eta = \frac{h_{\text{seg}}}{24} \tag{123}$$

$$K_{ms}(1,2) = K_{ms}(2,1) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(1+\eta)(3+\eta)}{16} \, d\eta = h_{\text{seg}} \frac{5}{24} \tag{124}$$

$$K_{ms}(2,2) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(3-\eta)(3+\eta)}{16} \, d\eta = h_{\text{seg}} \frac{13}{24} \tag{125}$$

Jacobian contributions. We note that the master/slave Jacobian contribution is not symmetric in general, it simply happens to be in this particular case. In summary, we have

$$K_{ss} = \frac{h_{\text{seg}}}{24} \begin{bmatrix} 2 & 4 \\ 4 & 14 \end{bmatrix} \qquad K_{ms} = \frac{h_{\text{seg}}}{24} \begin{bmatrix} 1 & 5 \\ 5 & 13 \end{bmatrix} \tag{126}$$

and, although the values above were obtained using exact integration, it is easy to verify that the same result can be computed using two-point Gaussian quadrature since the integrands are continuous.

21

Figure 8: Diagram of mortar segment T with endpoints $\xi_a^{(1)} = -1$, $\xi_b^{(1)} = 1$, $\xi_a^{(2)} = 1$, $\xi_b^{(2)} = 0$. Dark lines indicate the faces of the elements whose degrees of freedom T contributes to, light grey lines indicate neighboring element faces which are not involved in the segment computation. There is a single mesh node which corresponds to the left endpoint of this segment, so it appears as a dark circle on both the slave and master sides.

8.3 Simple example 2

As a second simple example, consider the mortar segment marked T in Fig. 8. This segment corresponds to the face of a single element on the slave side, and the left endpoint of this segment happens to be a node shared by both the slave and master surfaces. The endpoints of segment T are therefore $\xi_a^{(1)} = -1$, $\xi_b^{(1)} = 1$, $\xi_a^{(2)} = 1$, and $\xi_b^{(2)} = 0$, and the segment reference coordinates are

$$\xi^{(1)}(\eta) = \eta \tag{127}$$

$$\xi^{(2)}(\eta) = \frac{1 - \eta}{2} \tag{128}$$

which corresponds to the slave and master basis functions

$$N_1^{(1)} = \frac{1 - \eta}{2} \tag{129}$$

$$N_2^{(1)} = \frac{1 + \eta}{2} \tag{130}$$

$$N_1^{(2)} = \frac{1 + \eta}{4} \tag{131}$$

$$N_2^{(2)} = \frac{3 - \eta}{4} \tag{132}$$

Since $N_1^{(1)}$ and $N_2^{(1)}$ are the "standard" 1D finite element basis functions, the slave/slave Jacobian contribution is simply the standard one-dimensional mass matrix, i.e.

$$K_{ss} = \frac{h_{\text{seg}}}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{133}$$

The master/slave contributions are given by

$$K_{ms}(1,1) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(1+\eta)(1-\eta)}{8} \, d\eta = \frac{h_{\text{seg}}}{12} \tag{134}$$

$$K_{ms}(1,2) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(1+\eta)^2}{8} \, d\eta = \frac{h_{\text{seg}}}{6} \tag{135}$$

$$K_{ms}(2,1) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(3-\eta)(1-\eta)}{8} \, d\eta = h_{\text{seg}} \frac{5}{12} \tag{136}$$

$$K_{ms}(2,2) = \frac{h_{\text{seg}}}{2} \int_{-1}^{1} \frac{(3-\eta)(1+\eta)}{8} \, d\eta = \frac{h_{\text{seg}}}{3} \tag{137}$$

which can be summarized succinctly as

$$K_{ms} = \frac{h_{\text{seg}}}{12} \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} \tag{138}$$

As with segment S, all of the preceding values were computed using exact integration, but the same results can also be obtained with a two-point Gaussian quadrature rule.

9   Test cases

In the following subsections, we verify the correctness and accuracy of the proposed stabilized mortar finite element scheme on different meshes with different finite element schemes using the following test problem:

$$-\nabla^2 u + u = f \quad \in \Omega \tag{139}$$

$$\frac{\partial u}{\partial n} = g \quad \in \partial\Omega \tag{140}$$

where $\Omega = \left(-\frac{1}{2}, \frac{1}{2}\right)^2$ is the unit square centered at the origin, and the forcing $f = -4 + x^2 + y^2$ is chosen to give the true solution $u = x^2 + y^2$. The Neumann boundary data $g = -1$ is chosen to be consistent with the true solution on the entire boundary. For this test problem, even though the domain is only Lipschitz-continuous, the true solution $u \in C^\infty(\Omega)$, and data $g \in C^\infty(\partial\Omega)$ are both infinitely differentiable.

We compute and report the error in the following standard norms:

$$\|u - u^h\|_{L^2(\Omega)}^2 \equiv \int_\Omega |u - u^h|^2 \, dx \tag{141}$$

$$\|u - u^h\|_{H^1(\Omega)}^2 \equiv \int_\Omega |\nabla u - \nabla u^h|^2 \, dx \tag{142}$$

$$\|\lambda - \lambda^h\|_{L^2(\Gamma)}^2 \equiv \int_\Gamma |\lambda - \lambda^h|^2 \, ds \tag{143}$$

Remarks: (1) These norms are not exactly the same as the mesh-dependent norms used in the *a priori* error estimation theory, however they do allow us to conclude whether our results

23

are consistent with the *a priori* estimates. (2) The norm (142) is technically a seminorm, however since we don't consider constant exact solutions and we always look at the L^2 norm as well, we don't risk underestimating the true H^1 error in these tests.

In the following sections, we will refer to discretizations which use a piecewise continuous linear approximation for the primal variable combined with a piecewise discontinuous constant approximation for the Lagrange multiplier variable as $P^1 - P^0$ discretizations, and to discretizations that use a continuous linear Lagrange multiplier approximation as $P^1 - P^1$ discretizations. In the present work, we do not consider quadratic (P^2) approximations for the primal variable, however these discretizations are of interest in future work.

9.1   Flat interface

In this test case, a flat, vertical interface Γ is introduced along the plane $x = 0$ in Ω as shown in Fig. 9a. This interface divides the domain into left and right halves, and the left half is meshed with triangular finite elements with average diameter $h = 1/20$, while the right half is meshed with quadrilateral elements with average diameter $h = 1/21$. The coarsest grid, corresponding to $h = 1/20$, is shown in Fig. 9a. In this case, there is no discretization error introduced by the meshing of the interface since it is flat, and there is no gap between the elements on either side of the interface. Grids with $h = 1/20, 1/40, 1/60, \ldots, 1/20n$ for $1 \leq n \leq 15$ were used in the convergence study. The true flux is $\lambda = 0$ along this interface, since

$$\lambda \equiv -\nabla u \cdot \hat{\boldsymbol{n}} = -(2x, 2y) \cdot (1, 0) = -2x = 0 \tag{144}$$

The convergence results for equal-order $P^1 - P^1$ and mixed $P^1 - P^0$ discretizations of the primal and flux variables using both stabilized $\delta = 0.4$ and unstabilized $\delta = 0$ formulations are shown in Fig. 9. For this test case, all four discretizations' flux error converges at the same rate of $\mathcal{O}(h^{3/2})$, regardless of the polynomial order and presence of stabilization terms. The norms of the primal variable all converge at the optimal rates ($\mathcal{O}(h^2)$ in L^2 and $\mathcal{O}(h)$ in H^1). That said, there is still a strong argument to be made in favor of using the stabilized formulations, since the error is significantly lower (by a factor of almost 10 in the best case) as compared to the unstabilized cases.

We compare the stabilized and unstabilized $P^1 - P^0$ discretizations in Fig. 10a and the $P^1 - P^1$ discretizations in Fig. 10b for Mesh 1 ($h = 1/20$). For this mesh, both the stabilized and unstabilized solutions are fairly oscillatory throughout the domain, although the stabilized cases exhibit lower amplitude oscillations. After one uniform refinement, the $P^1 - P^0$ solutions are compared again in Fig. 10c. We see that both solutions are much more accurate than the Mesh 1 solution in the domain interior (see vertical axis) but there is still relatively high error concentrated near the boundaries.

We have also included, in Fig. 10d, a plot of the number of GMRES iterations required to solve the system matrix (recall that this is a steady, linear problem, so there is only a single matrix to solve) vs. the mesh number. In each case, the same preconditioner (ILU(1) as implemented within the PETSc numerical linear algebra library) was used, along with the default GMRES parameters, and the system iterated until a relative residual reduction of 10^{-12} was achieved. We do not currently have a firm understanding of why the solver performs so poorly for the unstabilized $P^0 - P^0$ formulation. It may be possible to achieve

improved performance by tweaking the preconditioner and GMRES parameters, but we have not invested much time in this pursuit. Until this phenomenon is better understood, the unstabilized $P^0 - P^0$ formulation should probably be avoided or only used with extreme caution.
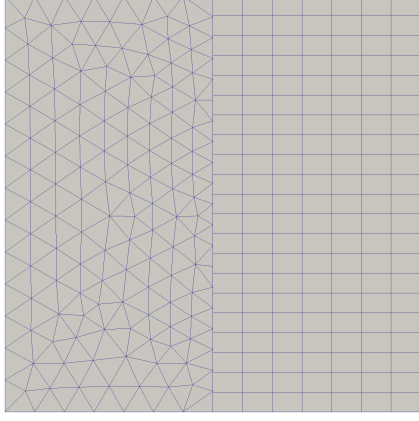
(a) Hybrid mesh, $h = 1/20$.
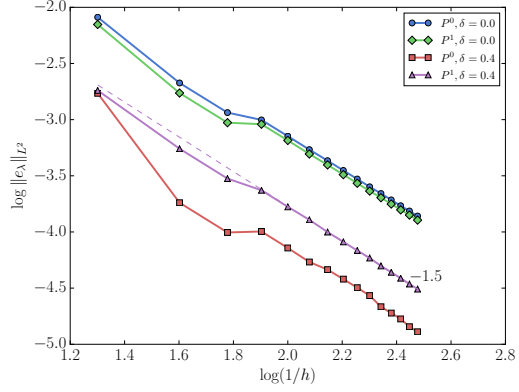
(b) $\|e_\lambda\|_{L^2}$ vs. h.
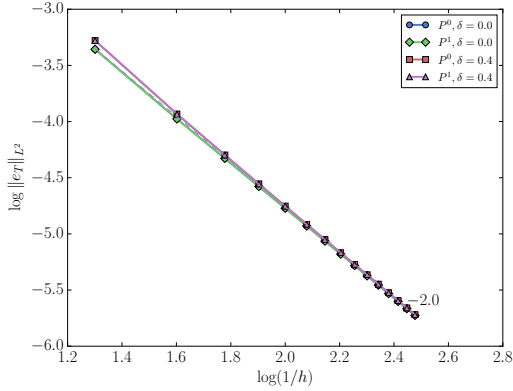
(c) $\|e_u\|_{L^2}$ vs. h.
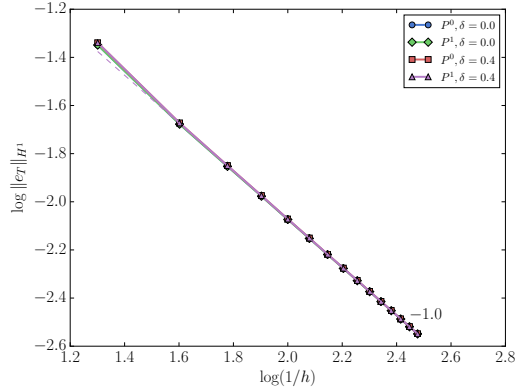
(d) $\|e_u\|_{H^1}$ vs. h.

Figure 9: Convergence plots for the flat interface test case: (a) hybrid initial mesh with $h = 1/20$, and the convergence rates of (b) the flux in $L^2(\Gamma)$ norm, (c) primal variable in $L^2(\Omega)$, and (d) primal variable in $H^1(\Omega)$ norm. Both the discontinuous constant $(P^1 - P^0)$ and continuous linear $(P^1 - P^1)$ flux discretizations converge at a rate of $\mathcal{O}(h^{3/2})$ regardless of the presence of stabilization, however the stabilized $(\delta = 0.4)$ cases have much less error overall. The most accurate method is the stabilized $P^1 - P^0$ discretization. The primal variable error norms converge at the optimal rate, and are nearly identical for each of the discretizations.
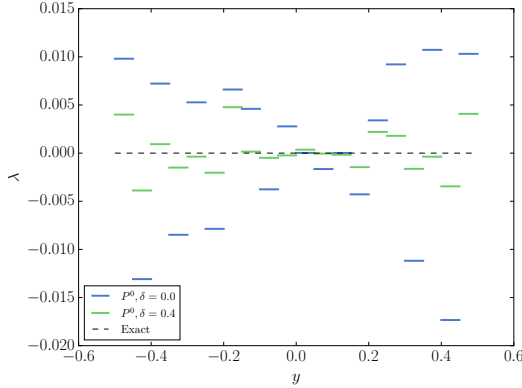
(a) $P^1 - P^0$, Mesh 1.
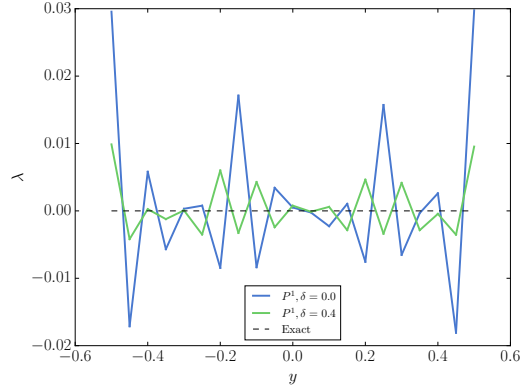
(b) $P^1 - P^1$, Mesh 1.
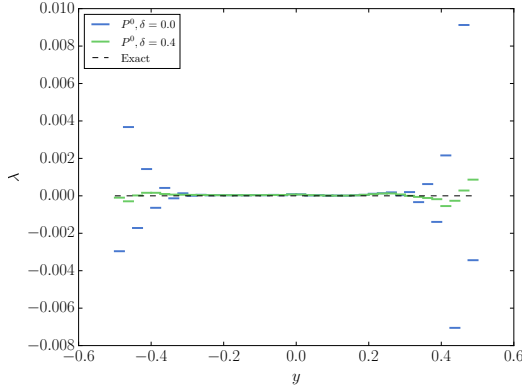
(c) $P^1 - P^0$, Mesh 2.
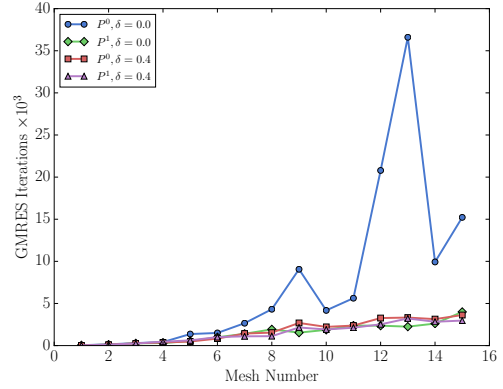
(d) Linear iterations vs. mesh number.

Figure 10: Comparison of stabilized $P^1 - P^0$ (a) and $P^1 - P^1$ (b) results on Mesh 1. Note the difference in the vertical axis between the two cases: the linear approximation has much higher error overall. (c) In Mesh 2 ($h = 1/40$), the error in the stabilized $P^1 - P^0$ formulation is significantly lower than the unstabilized $P^1 - P^0$ formulation, and the error seems to be concentrated near the domain boundaries in both cases. (d) The number of linear iterations required to solve the linear system associated with this test increases unacceptably quickly for the unstabilized $P^1 - P^0$ case. All other cases require about the same number of linear iterations.

9.2 Flat interface with 2:1 element size mismatch

In §9.1, we considered a nonconforming mesh in which the cells on either side of the interface were nevertheless of comparable size. In this section, we generalize the previous example slightly in order to test the robustness of the algorithm in the context of more strongly-mismatched elements along the interface. Using the same model problem as before, we now introduce the vertical interface Γ along the plane $x = 1/4$, and we generate the mesh so that the slave (left) side elements are roughly half the size of the elements on the master (right) side. The coarsest version of this mesh (with $h = 1/21$) can be seen in Fig. 11a. Analogously to the previous section, we considered a sequence of refined grids in this case with $h = 1/21, 1/42, 1/63, \ldots, 1/21n$ for $1 \leq n \leq 14$. The true flux is once again constant in this case, and given by:

$$\lambda \equiv -\nabla u \cdot \hat{\boldsymbol{n}} = -(2x, 2y) \cdot (1, 0) = -2 \left(\frac{1}{4} \right) = -\frac{1}{2} \tag{145}$$

In Fig. 11b, the convergence rates obtained for the L^2 norm of the flux error are seen to be $\mathcal{O}(h)$ for each of the methods, with the error in the unstabilized methods being significantly higher (over 100 times higher in some instances) than the error in the stabilized methods. This result is in contrast to the $\mathcal{O}(h^{3/2})$ result observed in §9.1, but it is still consistent with the *a priori* error estimates discussed in §4, which only require that the combined $\|e_u\|_{H^1(\Omega)}$ (see Fig. 11d) and $\|e_\lambda\|_{L^2(\Gamma)}$ errors converge at $\mathcal{O}(h)$. As in the previous case, the $\|e_u\|_{L^2(\Omega)}$ (see Fig. 11c), converges at $\mathcal{O}(h^2)$, also in accordance with the theory. It is important to note that, if one is not concerned with the flux as an independent quantity of interest, the use of an unstabilized mixed formulation does not seem to have a very large effect on the primal solution, at least in these simple test cases.

In Fig. 12, we compare the stabilized and unstabilized $P^1 - P^0$ discretizations on Meshes 1–4. This gives one a better idea of the much larger error present in the unstabilized method—it is actually not possible to distinguish between the true solution and the stabilized solution at this scale. In these plots, we can also observe a certain "periodicity" in the error which is tied to the resolution of the mesh. On inspection, we see that the largest error occurs when the nodes on the interface align, and the smallest error occurs where the mismatch in node locations is largest. We do not currently have an explanation for this phenomenon, and it deserves further investigation.

A similar sequence of plots, this time for the linear Lagrange flux discretization, is given in Fig. 13. Here, the unstabilized formulation is not nearly as inaccurate as in the discontinuous constant discretization case, but it is still clearly less accurate than the stabilized solution. The error in the linear discretization is also periodic with a period that depends on the mesh spacing. In contrast to the discontinuous constant discretization, however, the error is *smallest* when the interface nodes are aligned, and largest when they are farthest apart, so that the error is almost perfectly out of phase with the error in the $P^1 - P^0$ case. Finally, we compare the error in both stabilized methods in Fig. 14, and here, both errors are of the same relative order. This figure clearly demonstrates that the error in the two methods is out of phase, but does not suggest *why* this would be the case.
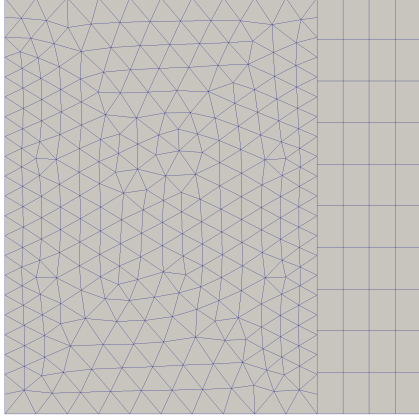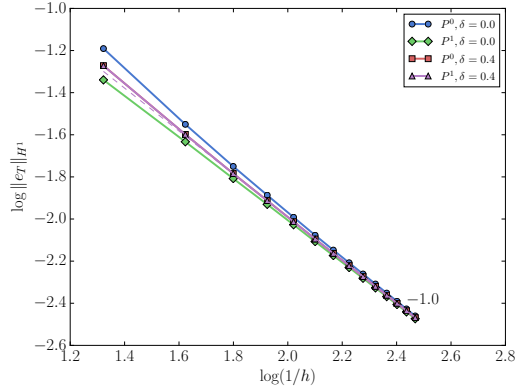
(a) Hybrid mesh, $h = 1/21$.
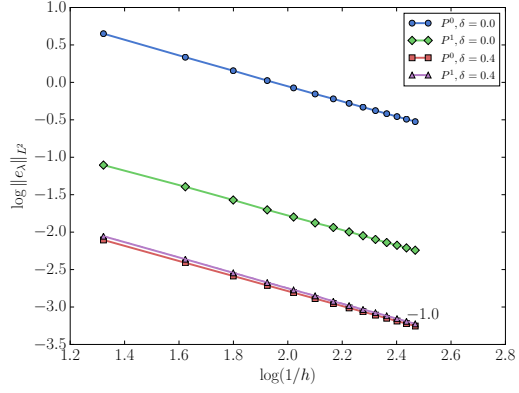
(b) $\|e_\lambda\|_{L^2}$ vs. h.
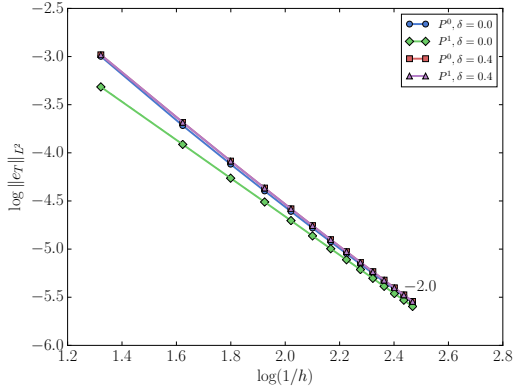
(c) $\|e_u\|_{L^2}$ vs. h.

(d) $\|e_u\|_{H^1}$ vs. h.

Figure 11: Convergence plots for the offset interface test case with 2:1 element size mismatch: (a) hybrid initial mesh with $h = 1/21$, and the convergence rates of (b) the flux in $L^2(\Gamma)$ norm, (c) primal variable in $L^2(\Omega)$, and (d) primal variable in $H^1(\Omega)$ norm. In contrast to the test case discussed in §9.1, the flux error is not superconvergent in this case, but it is still consistent with the overall $\mathcal{O}(h)$ convergence rate predicted by the *a priori* estimate (67).
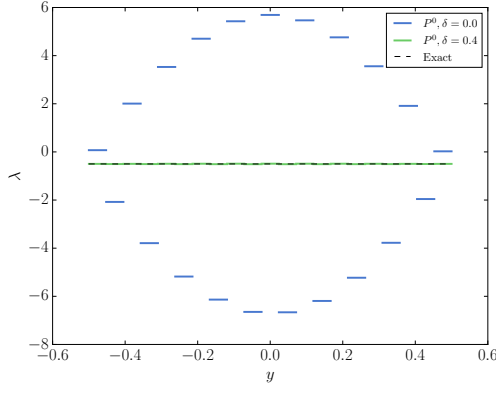
(a) $P^1 - P^0$, Mesh 1.
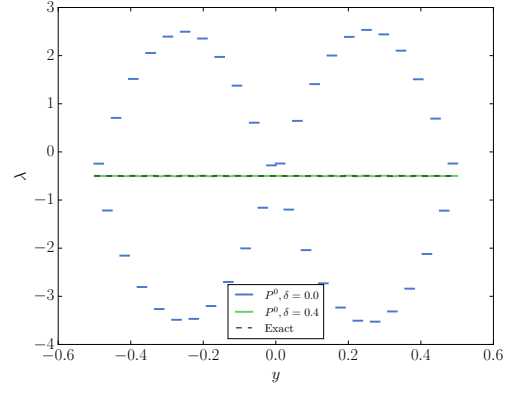
(b) $P^1 - P^0$, Mesh 2.
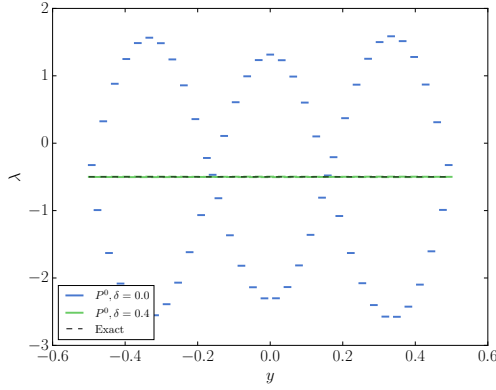
(c) $P^1 - P^0$, Mesh 3.
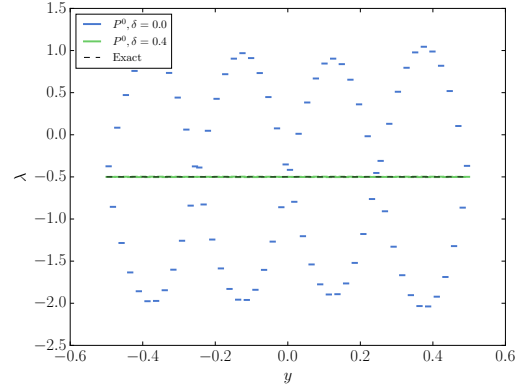
(d) $P^1 - P^0$, Mesh 4.

Figure 12: Comparison of the stabilized and unstabilized $P^1 - P^0$ discretizations on Meshes 1–4. The error in the unstabilized formulation is over 100 times larger than the error in the stabilized formulation in all cases, making it difficult to see differences between the stabilized and exact solution on this scale. Using a 2:1 element size mismatch along the interface is sufficient in these cases to "awaken" the inherent instabilities of the formulation when $\delta = 0$.
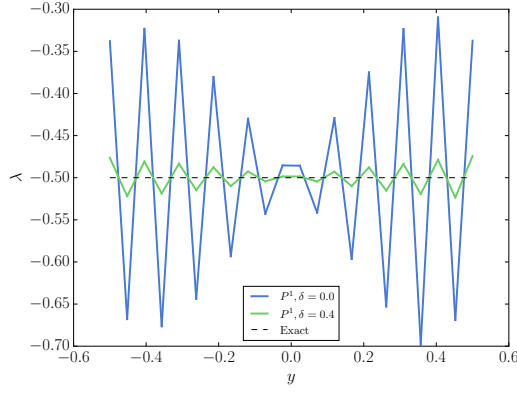
(a) $P^1 - P^1$, Mesh 1.
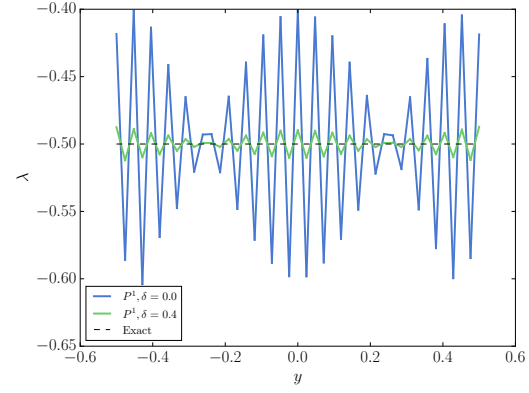
(b) $P^1 - P^1$, Mesh 2.
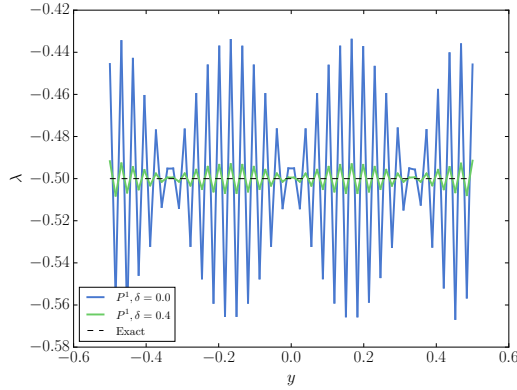
(c) $P^1 - P^1$, Mesh 3.
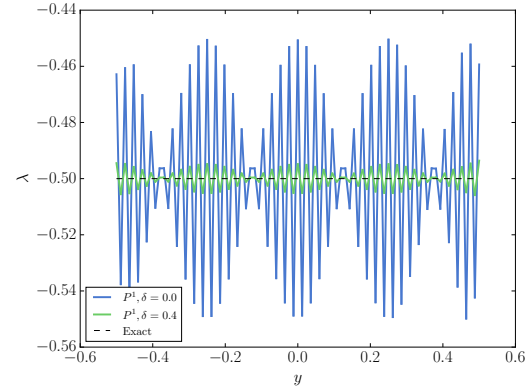
(d) $P^1 - P^1$, Mesh 4.

Figure 13: Comparison of the stabilized and unstabilized $P^1 - P^1$ discretizations on Meshes 1–4. Both methods appear to have an instability whose period depends on the mesh refinement level, however the amplitude of the instability is markedly worse for the unstabilized method. The peaks in the error coincide with points along the interface where the meshes on each side share a node in common, but it is not immediately clear what actually causes this correlation.

30
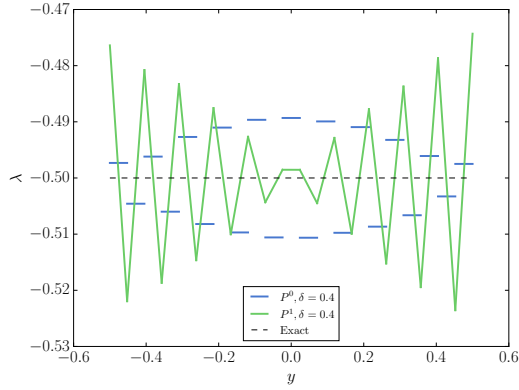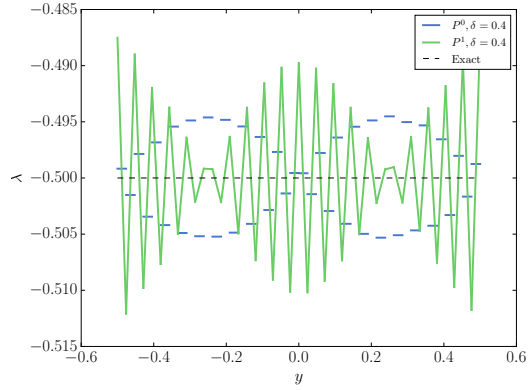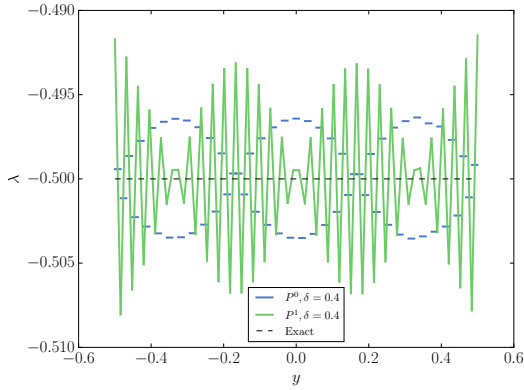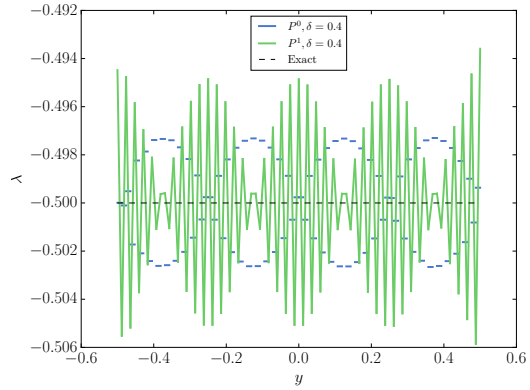
Figure 14: Comparison of the stabilized $P^1 - P^0$ and $P^1 - P^1$ discretizations on Meshes 1–4. Both discretizations appear to have a mesh-dependent instability, but the error of the two methods seems to be perfectly out of phase. Overall, the $P^1 - P^0$ discretization seems to have slightly lower error, but the two methods are almost indistinguishable on the error convergence plot.

9.3   Quarter-circle

We next investigate the accuracy of the various discretizations on a problem with a curved interface. We solve the same model problem and take the same unit square domain as in the preceding two examples, but this time allow it to be cut by a circle centered at $(x_c, y_c) = (-1/2, -1/2)$ with radius $R = \frac{\sqrt{2}}{2}$. As before, we mesh the interface with a non-conforming grid with approximately 21 intervals along the circular arc in the coarsest case. As the mesh is refined, nodes are "snapped" to the true circular geometry, so the discrete interface Γ converges to the true quarter-circular arc as $h \to 0$. Unlike the previous two test cases, there is an additional source of discretization error in this problem due to the presence of the curved interface. The coarsest grid used for the calculations is shown in Fig. 15a.

In this case, the true flux is no longer a constant along the curved interface. We can derive a formula for the true flux by defining the "shifted" polar coordinates:

$$\tilde{r} = \sqrt{(x - x_c)^2 + (y - y_c)^2} \tag{146}$$

$$\tilde{\theta} = \arctan\left(\frac{y - y_c}{x - x_c}\right) \tag{147}$$

In terms of these coordinates, we therefore have

$$x = x_c + \tilde{r}\cos\tilde{\theta} \tag{148}$$

$$y = y_c + \tilde{r}\sin\tilde{\theta} \tag{149}$$

and the outward unit normal vector to the quarter-circular arc is

$$\hat{\boldsymbol{n}} = \left(\cos\tilde{\theta}, \sin\tilde{\theta}\right) \tag{150}$$

On the circle, where $\tilde{r} = R$, we have

$$\begin{aligned}
\lambda &\equiv -\nabla u \cdot \hat{\boldsymbol{n}} \\
&= -2(x, y) \cdot \hat{\boldsymbol{n}} \\
&= -2(x_c + R\cos\tilde{\theta}, y_c + R\sin\tilde{\theta}) \cdot \left(\cos\tilde{\theta}, \sin\tilde{\theta}\right) \\
&= -2\left(x_c \cos\tilde{\theta} + y_c \sin\tilde{\theta} + R\right)
\end{aligned} \tag{151}$$

for $0 \leq \tilde{\theta} \leq \pi/2$.

The convergence results for the quarter-circular case are shown in Fig. 15, and they are essentially split between the $P^1 - P^0$ formulations which exhibits higher-order $\mathcal{O}(h^{3/2})$ convergence of the flux in the $L^2(\Gamma)$ norm, and the $P^1 - P^1$ formulations which converge at $\mathcal{O}(h)$. It is interesting to note that both the stabilized and unstabilized $P^1 - P^1$ formulations both have essentially the same error, which is in contrast to previous results where the stabilized formulation had lower error. In the $P^1 - P^0$ formulations, the stabilized formulation has significantly lower error (approximately 10 times less) than the unstabilized formulation, making it once again the obvious choice for general applications. There are once again no surprises in the L^2 or H^1 convergence rates for the primal variable; and the error magnitude for each method is nearly identical.

Some representative line plots of the flux for meshes 1 and 2 are given in Fig. 16. For the $P^1 - P^0$ discretizations, both the stabilized and unstabilized formulations are of comparable accuracy. The largest differences can be observed near the left endpoint of the interface and near the location of maximum flux, where the unstabilized formulation exhibits a slight oscillation which is not present in the stabilized solution. For the $P^1 - P^1$ discretizations, on Mesh 1 the unstabilized formulation is clearly much more oscillatory than the stabilized one, and there is a very large error near the left endpoint, where the solution appears to be abnormally flat. The stabilized $P^1 - P^1$ formulation also shows a relatively large error near the left endpoint, although it appears to be more accurate than the unstabilized method, suggesting that the inaccuracy may be related to the instability.
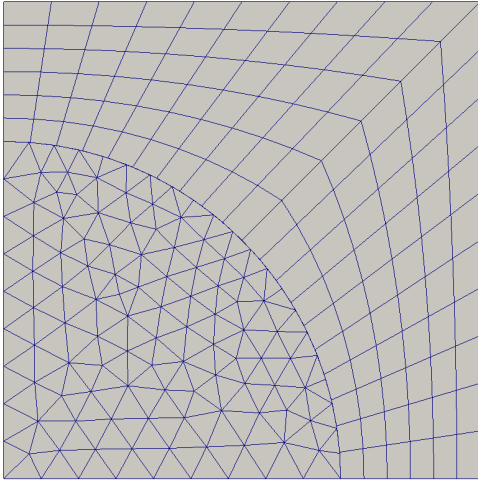
(a) Hybrid mesh, $h = \frac{\pi R}{2 \cdot 21}$, $R = \frac{\sqrt{2}}{2}$.
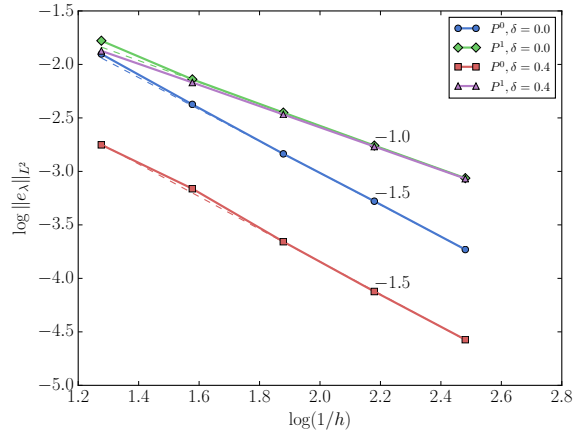
(b) $\|e_\lambda\|_{L^2}$ vs. h.
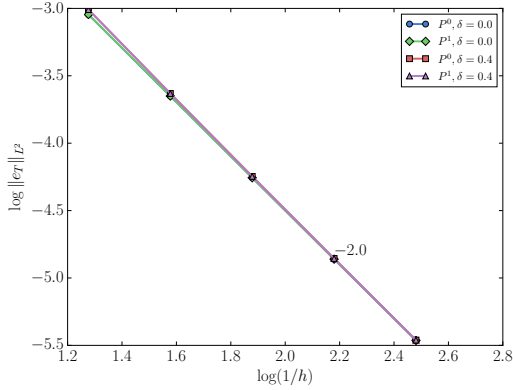
(c) $\|e_u\|_{L^2}$ vs. h.
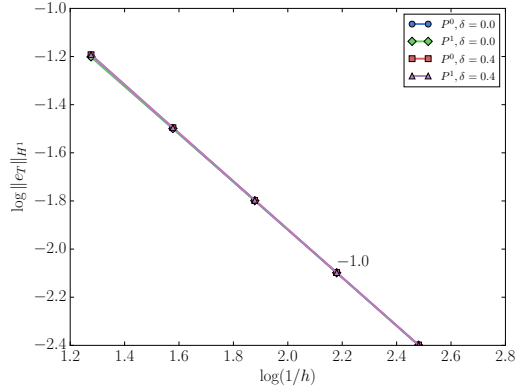
(d) $\|e_u\|_{H^1}$ vs. h.

Figure 15: Convergence plots for the quarter-circle case. (a) hybrid initial mesh, (b) convergence of the flux in the $L^2(\Gamma)$ norm, (c) convergence of the primal variable in the $L^2(\Omega)$ norm, and (d) in the $H^1(\Omega)$ norm. In this case, both the stabilized and unstabilized $P^1 - P^1$ approximations converge at $\mathcal{O}(h)$ in the flux and have approximately the same level of error, while both $P^1 - P^0$ approximations converge at the faster $\mathcal{O}(h^{3/2})$ rate, and the stabilized approximation has significantly lower error.
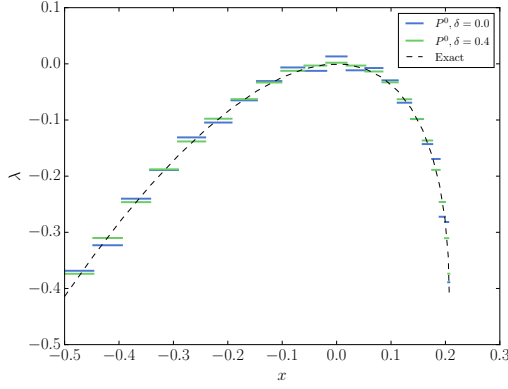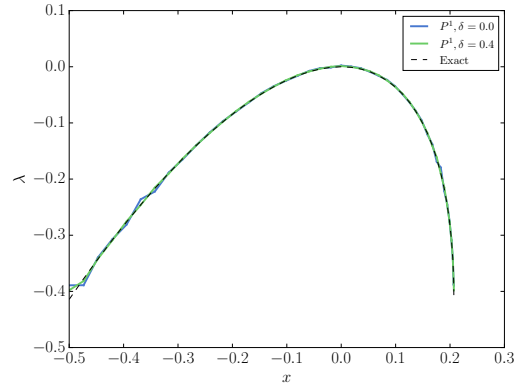
(a) $P^1 - P^0$, Mesh 1.
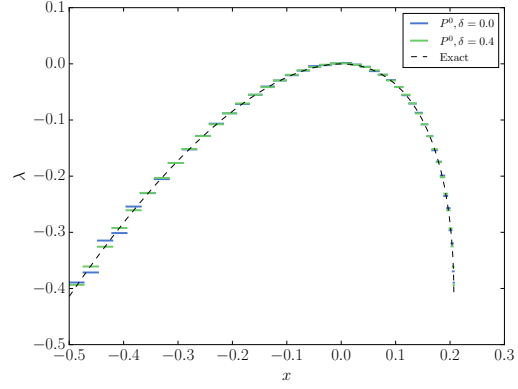
(b) $P^1 - P^0$, Mesh 2.
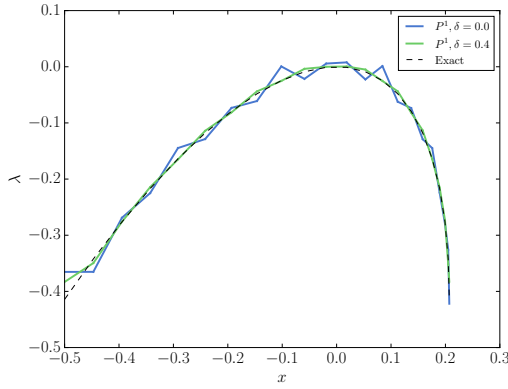
(c) $P^1 - P^1$, Mesh 1.

(d) $P^1 - P^1$, Mesh 2.

Figure 16: Comparison of the various discretizations on quarter-circle meshes 1 and 2. The top row, subfigures (a) and (b), corresponds to the $P^1 - P^0$ discretization while the bottom row, subfigures (c) and (d) corresponds to the $P^1 - P^1$ discretization. The left column corresponds to the coarsest discretization (Mesh 1) and the right column (Mesh 2) corresponds to a uniform refinement of Mesh 1. The error in all four discretizations appears to be focused near the left endpoint of the domain.

9.4 Non-smooth interface

In the preceding examples, we have considered only smooth interfaces, but in applications involving more realistic geometries, the interfaces may only be piecewise smooth (also called Lipschitz continuous). Not all of the *a priori* error estimates discussed in §4 explicitly apply to piecewise smooth interfaces, and therefore we don't necessarily expect optimal convergence rates in such cases. To test this case, we solve the same model problem, and introduce a piecewise smooth interface Γ into the computational domain Ω composed of two line segments that intersect along the centerline $x = 0$. We parameterize these line segments by the acute angle α they form with the x-axis, and y_c, the point along the line $x = 0$ where they meet. For the test case we choose $y_c = 0.2$, left angle $\alpha_1 = 15°$, and right angle $\alpha_2 = 5°$. The coarse level mesh for these parameters is shown in Fig. 17a.

It can be shown, via straightforward calculation, that the true flux for our model problem through line segment i is given by

$$\lambda_i = \frac{2y_c}{\sqrt{1 + \tan^2 \alpha_i}} \tag{152}$$

for $i = 1, 2$. We therefore note that $\lambda_1 \neq \lambda_2$ whenever $y_c \neq 0$ and $\alpha_1 \neq \alpha_2$. In other words, even though the true solution is infinitely differentiable, we have chosen an interface with a discontinuous normal vector, and therefore the flux, which is defined in terms of the interface normal, is also discontinuous.

The convergence rate of the L^2 norm of the flux error for the various discretization methods is given in Fig. 17b. Both the stabilized and unstabilized $P^1 - P^0$ formulations obtain the higher $\mathcal{O}(h^{3/2})$ rate of convergence that was observed previously in the flat interface problem, but the $P^1 - P^1$ discretizations both converge at a rate of only $\mathcal{O}(h^{1/2})$, an entire power of h slower than the lower-order flux discretization. The primal variable norms are both once again optimal for all methods, which is somewhat surprising given the poor convergence rate of the $P^1 - P^1$ flux.

We do not currently have a full explanation for the loss of accuracy observed in the $P^1 - P^1$ formulation, but the line plots in Fig. 18 suggest that over and under-shoots near both the domain boundaries and the flux discontinuity are much more pronounced for the linear flux discretization than the discontinuous constants. The stabilized formulation improves the $P^1 - P^1$ results somewhat on coarser grids, but in the limit as $h \to 0$, both the stabilized and unstabilized methods converge to the same error. It may be possible to improve the linear results further by introducing additional stabilization terms of the "artificial viscosity" variety, i.e. proportional to $\langle \nabla M_i, \nabla M_j \rangle$, but we are not currently aware of any theory covering this particular situation.
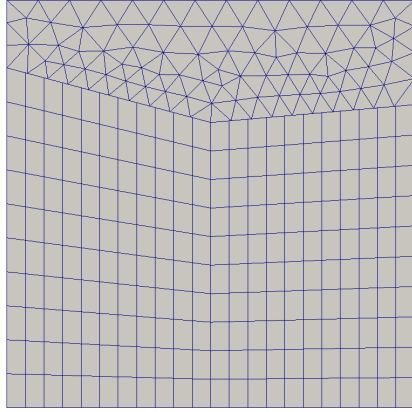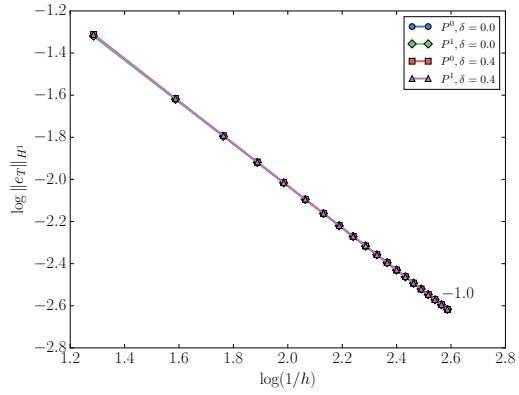
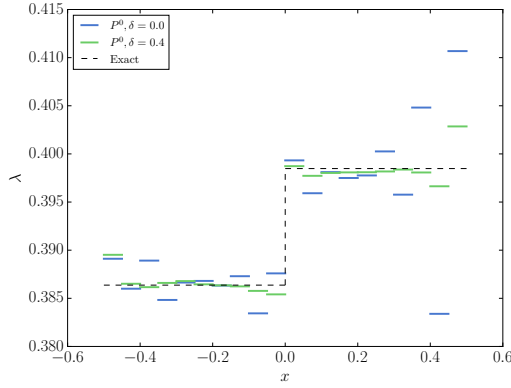(a) Coarse level hybrid mesh.
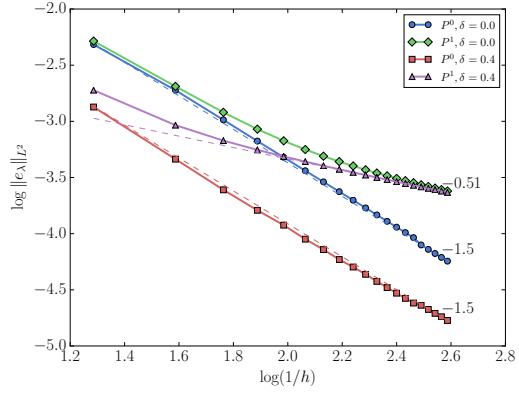
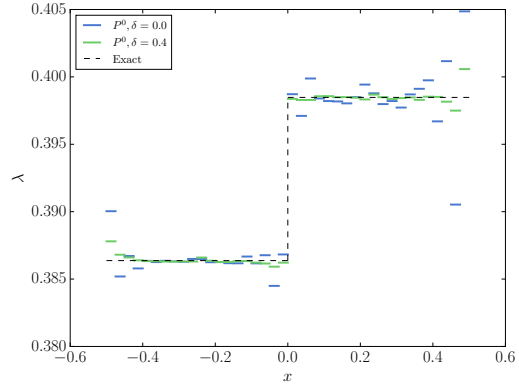(b) $\|e_\lambda\|_{L^2}$ vs. h.
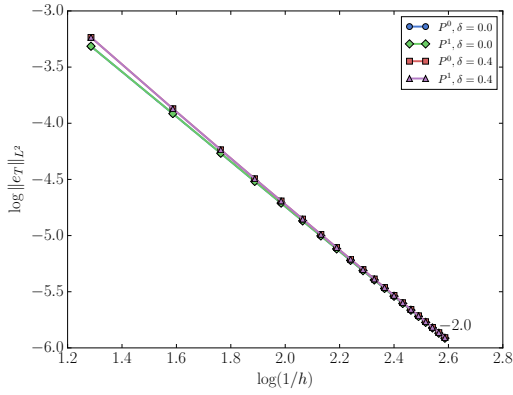
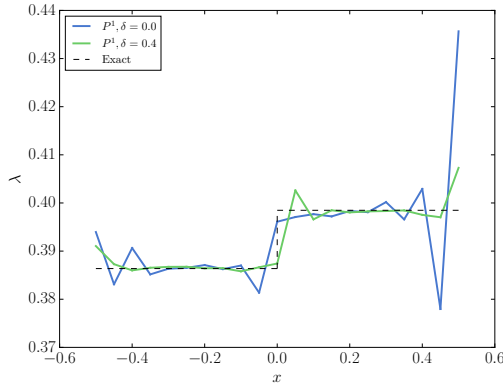(c) $\|e_u\|_{L^2}$ vs. h.

(d) $\|e_u\|_{H^1}$ vs. h.

Figure 17: Convergence plots for the non-smooth interface case. (a) hybrid initial mesh, (b) convergence of the flux in the $L^2(\Gamma)$ norm, (c) convergence of the primal variable in the $L^2(\Omega)$ norm, and (d) in the $H^1(\Omega)$ norm.
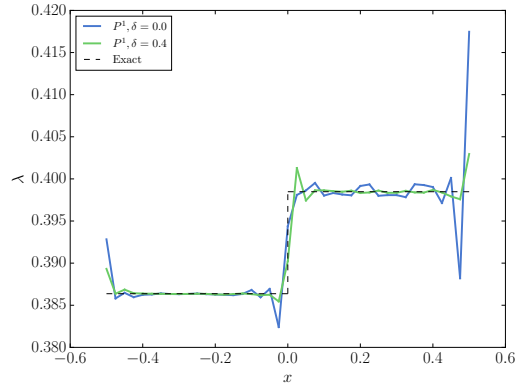
Figure 18: Comparison of the various discretizations of the non-smooth interface problem, meshes 1 and 2. The top row, subfigures (a) and (b), corresponds to the $P^1 - P^0$ discretization while the bottom row, subfigures (c) and (d) corresponds to the $P^1 - P^1$ discretization. The left column corresponds to the coarsest discretization (Mesh 1) and the right column (Mesh 2) corresponds to a uniform refinement of Mesh 1.

References

[1] C. Hesch and P. Betsch, "Energy-momentum consistent algorithms for dynamic ther-momechanical problems—Application to mortar domain decomposition methods," *International Journal for Numerical Methods in Engineering*, vol. 86, pp. 1277–1302, June 2011. `http://tinyurl.com/rm646r4`.

[2] I. Babuška, "The finite element method with Lagrange multipliers," *Numerische Mathematik*, vol. 20, pp. 179–192, June 1973. `https://doi.org/10.1007/BF01436561`.

[3] H. J. C. Barbosa and T. J. R. Hughes, "The finite element method with Lagrange multipliers on the boundary: Circumventing the Babuška-Brezzi condition," *Computer Methods in Applied Mechanics and Engineering*, vol. 85, pp. 109–128, Jan. 1991. `https://doi.org/10.1016/0045-7825(91)90125-P`.

[4] F. Ben Belgacem, "The mortar finite element method with Lagrange multipliers," *Numerische Mathematik*, vol. 84, pp. 173–197, Dec. 1999. `http://tinyurl.com/rdrfu2e`.

[5] M. G. Larson and A. Massing, "L^2-error estimates for finite element approximations of boundary fluxes." ArXiv e-print, Sept. 2014. `http://arxiv.org/abs/1401.6994`.

[6] J. C. Simo, P. Wriggers, and R. L. Taylor, "A perturbed Lagrangian formulation for the finite element solution of contact problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 50, pp. 163–180, Aug. 1985. `http://dx.doi.org/10.1016/0045-7825(85)90088-X`.

[7] P. Papadopoulos and R. L. Taylor, "A mixed formulation for the finite element solution of contact problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 94, pp. 373–389, Feb. 1992. `http://dx.doi.org/10.1016/0045-7825(92)90061-N`.

[8] T. A. Laursen and J. C. Simo, "A continuum-based finite element formulation for the implicit solution of multibody, large deformation frictional contact problems," *International Journal for Numerical Methods in Engineering*, vol. 36, pp. 3451–3485, Oct. 1993. `http://dx.doi.org/10.1002/nme.1620362005`.

[9] T. A. Laursen, *Computational Contact and Impact Mechanics*. Springer-Verlag, 2002.

[10] M. A. Puso, "A 3D mortar method for solid mechanics," *International Journal for Numerical Methods in Engineering*, vol. 59, pp. 315–336, Jan. 2004. `http://dx.doi.org/10.1002/nme.865`.

[11] M. A. Puso and T. A. Laursen, "A mortar segment-to-segment contact method for large deformation solid mechanics," *Computer Methods in Applied Mechanics and Engineering*, vol. 193, pp. 601–629, Feb. 2004. `http://dx.doi.org/10.1016/j.cma.2003.10.010`.

[12] B. Yang, T. A. Laursen, and X. Meng, "Two dimensional mortar contact methods for large deformation frictional sliding," *International Journal for Numerical Methods in Engineering*, vol. 62, pp. 1183–1225, Mar. 2005. `http://tinyurl.com/yba8fa36`.

[13] B. Yang, *Mortar finite element methods for large deformation contact mechanics*. PhD thesis, Duke University, 2006. `http://tinyurl.com/gkt5nga`.

[14] G. F. Carey and J. T. Oden, *Finite elements: A second course*. Prentice Hall, 1983.

[15] B. Flemisch and B. I. Wohlmuth, *Nonconforming discretization techniques for coupled problems*, pp. 531–560. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. `http://tinyurl.com/wqwej7u`.

[16] O. A. Ladyzhenskaya, *The mathematical theory of viscous incompressible flow*. PO Box 786, Cooper Station, New York, New York 10276: Gordon and Breach Science Publishers, 1969. 2nd edition of the English translation by R. A. Silverman, vol. 2 of the series "Mathematics and its Applications".

[17] F. Brezzi, "On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers," *ESAIM: Mathematical Modelling and Numerical Analysis – Modélisation Mathématique et Analyse Numérique*, vol. 8, no. R2, pp. 129–151, 1974. `http://tinyurl.com/y9d3v33h`.

[18] J. Pitkäranta, "Boundary subspaces for the finite element method with Lagrange multipliers," *Numerische Mathematik*, vol. 33, pp. 273–289, Sept. 1979. `https://doi.org/10.1007/BF01398644`.

[19] J. Pitkäranta, "Local stability conditions for the Babuška method of Lagrange multipliers," *Mathematics of Computation*, vol. 35, pp. 1113–1129, Oct. 1980. `http://tinyurl.com/yb3kn8dy`.

[20] J. H. Bramble, "The Lagrange multiplier method for Dirichlet's problem," *Mathematics of Computation*, vol. 37, pp. 1–11, July 1981. `http://tinyurl.com/y97ajgoq`.

[21] J. Pitkäranta, "The finite element method with Lagrange multipliers for domains with corners," *Mathematics of Computation*, vol. 37, pp. 13–30, July 1981. `http://tinyurl.com/y7cexgl9`.

[22] H. J. C. Barbosa and T. J. R. Hughes, "Boundary Lagrange multipliers in finite-element methods: Error analysis in natural norms," *Numerische Mathematik*, vol. 62, pp. 1–15, Dec. 1992. `https://doi.org/10.1007/BF01396217`.

[23] R. Stenberg, "On some techniques for approximating boundary conditions in the finite element method," *Journal of Computational and Applied Mathematics*, vol. 63, pp. 139–148, Nov. 1995. `https://doi.org/10.1016/0377-0427(95)00057-7`.

[24] M. Juntunen and R. Stenberg, "Nitsche's method for general boundary conditions," *Mathematics of Computation*, vol. 78, pp. 1353–1374, July 2009. `http://dx.doi.org/10.1090/S0025-5718-08-02183-2`.

[25] P. Hansbo, C. Lovadina, I. Perugia, and G. Sangalli, "A Lagrange multiplier method for the finite element solution of elliptic interface problems using non-matching meshes," *Numerische Mathematik*, vol. 100, pp. 91–115, Mar. 2005. `https://doi.org/10.1007/s00211-005-0587-4`.

[26] E. Burman, "Projection stabilization of Lagrange multipliers for the imposition of constraints on interfaces and boundaries," *Numerical Methods for Partial Differential Equations*, vol. 30, pp. 567–592, Mar. 2014. `https://doi.org/10.1002/num.21829`.

[27] L. Chen, "Sobolev spaces and elliptic equations." PDF on personal website, Oct. 2013. `http://tinyurl.com/yc4egah6`.