Nonoverlapping Additive Schwarz Method for hp-DGFEM with Higher-order Penalty Terms

Piotr Krzyżanowski and Marcus Sarkis

1 Introduction

Let us consider a second order elliptic equation

$$-\operatorname{div}(\varrho\nabla u) = f \text{ in } \Omega \quad \text{and} \quad u = 0 \text{ in } \partial\Omega. \tag{1}$$

The problem is discretized by an *h-p* symmetric interior higher-order [4] discontinuous Galerkin finite element method. In a K-th order multipenalty method, one penalizes the jumps of scaled normal higher-order derivatives up to order K across the interelement boundaries — so the standard interior penalty method corresponds to taking $K = 0$. The idea to penalize the discontinuity in the flux ($K = 1$) of the discrete solution was introduced by Douglas and Dupont [6]. It addresses the observation that the flux (which is an important quantity in many applications) of the accurate solution is continuous. Giving the user a possibility to control the inevitable violation of this principle makes the discretization method more robust and conservative. Recently, flux jump penalization has been used to improve stability properties of an unfitted Nitsche's method [5], the case $K > 1$ was also considered in [1] for the immersed finite element method to obtain higher-order discretizations.

A nonoverlapping additive Schwarz method [7], [3] is applied to precondition the discrete equations. For more flexibility and enhanced parallelism, we formulate our results addressing the case when the subdomains (where the local problems are solved in parallel) are potentially smaller than the coarse grid cells [8]. By allowing small subdomains of diameter $H \leq \mathcal{H}$, the local problems are cheaper to solve and the amount of concurrency of the method is substantially increased. A by-product of

Piotr Krzyżanowski
University of Warsaw, Poland, e-mail: p.krzyzanowski@mimuw.edu.pl

Marcus Sarkis
Worchester Polytechnic Institute, e-mail: msarkis@wpi.edu

this approach is more flexibility in assigning subdomain problems to processors for load balancing in coarse grain parallel processing.

The paper is organized as follows. In Section 2, the differential problem and its discontinuous Galerkin multipenalty discretization are formulated. In Section 3, a nonoverlapping two-level, three-grid additive ASM for solving the discrete problem is designed and analyzed under the assumption that the coarse mesh resolves the discontinuities of the coefficient, that the variation of the mesh size and of the polynomial degree are locally bounded, and that the original problem satisfies some regularity assumption. Section 4 presents some numerical experiments.

For nonnegative scalars x, y, we shall write $x \lesssim y$ if there exists a positive constant C, such that $x \le Cy$ with C independent of: x, y, the fine, subdomain and coarse mesh parameters $h, H, \mathcal{H}$, the orders of the finite element spaces p, q, the order of the multipenalty method (K, L), and of jumps of the diffusion coefficient ϱ as well. If both $x \lesssim y$ and $y \lesssim x$, we shall write $x \simeq y$.

The norm of a function f from the Sobolev space $H^k(S)$ will be denoted by $||f||_{k,S}$, while the seminorm of f will be denoted by $|f|_{k,S}$. For short, the L^2-norm of f will then be denoted by $|f|_{0,S}$.

2 High-order penalty *h-p* discontinuous Galerkin discretization

Let Ω be a bounded open convex polyhedral domain in R^d, $d \in \{2, 3\}$, with a Lipschitz boundary $\partial\Omega$. We consider the following variational formulation of (1): Find $U^* \in H_0^1(\Omega)$ such that for a prescribed $f \in L^2(\Omega)$ and $\varrho \in L^\infty(\Omega)$

$$a(U^*, v) = (f, v)_\Omega, \qquad \forall v \in H_0^1(\Omega), \tag{2}$$

where

$$a(u, v) = \int_\Omega \varrho \, \nabla u \cdot \nabla v \, dx, \qquad (f, v)_\Omega = \int_\Omega f v \, dx.$$

We assume that there exists a constant α such that $1 \le \varrho \le \alpha$ a.e. in Ω so that (2) is well–posed. We also assume that ϱ is piecewise constant, i.e. Ω can be partitioned into nonoverlapping polyhedral subregions with the property that ϱ restricted to any of these subregions is some positive constant, see assumption (5) later on.

Let $\mathcal{T}_h = \{\tau_1, \dots, \tau_{N_h}\}$ denote an affine nonconforming partition of Ω, where τ_i are either triangles in 2-D or tetrahedra in 3-D. For $\tau \in \mathcal{T}_h$ we set $h_\tau = \mathrm{diam}(\tau)$. By $\mathcal{E}_h^{\mathrm{in}}$ we denote the set of all common (internal) faces (edges in 2-D) of elements in $\mathcal{T}_h$, such that $e \in \mathcal{E}_h^{\mathrm{in}}$ iff $e = \partial\tau_i \cap \partial\tau_j$ is of positive measure. We will use the symbol $\mathcal{E}_h$ to denote the set of all faces (edges in 2-D) of the fine mesh $\mathcal{T}_h$, that is those either in $\mathcal{E}_h^{\mathrm{in}}$ or on the boundary $\partial\Omega$. For $e \in \mathcal{E}_h$ we set $h_e = \mathrm{diam}(e)$. We assume that $\mathcal{T}_h$ is shape- and contact–regular, that is, it admits a matching submesh $\mathcal{T}_{\hat{h}}$ which is shape–regular and such that for any $\tau \in \mathcal{T}_h$ the ratios of h_τ to diameters of simplices in $\mathcal{T}_{\hat{h}}$ covering τ are uniformly bounded by an absolute constant. As a consequence, if $e = \partial\tau_i \cap \partial\tau_j$ is of positive measure, then $h_e \simeq h_{\tau_i} \simeq h_{\tau_j}$. We shall refer to $\mathcal{T}_h$ as the "fine mesh". Throughout the paper we will assume that the fine mesh is chosen in such a way that $\varrho_{|\tau}$ is already constant for all $\tau \in \mathcal{T}_h$.

We define the finite element space V_h^p in which problem (2) is approximated,

$$V_h^p = \{v \in L^2(\Omega) : v_{|\tau} \in \mathbb{P}_{p_\tau} \text{ for } \tau \in \mathcal{T}_h\} \tag{3}$$

where $\mathbb{P}_{p_\tau}$ denotes the set of polynomials of degree not greater than p_τ. We shall assume that $1 \leq p_\tau$ and that polynomial degrees have bounded local variation, that is, if $e = \partial\tau_i \cap \partial\tau_j \in \mathcal{E}_h^{\text{in}}$, then $p_{\tau_i} \simeq p_{\tau_j}$.

On $e \in \mathcal{E}_h^{\text{in}}$ such that $e = \partial\tau^+ \cap \partial\tau^-$, we define

$$\overline{\varrho} = \frac{\varrho^+ + \varrho^-}{2}, \qquad \omega^\pm = \frac{\varrho^\pm}{\varrho^+ + \varrho^-}, \qquad \underline{\varrho} = \frac{2\varrho^+\varrho^-}{\varrho^+ + \varrho^-}$$

with the standard notation $\rho^\pm = \rho_{|\tau^\pm}$, and then define weighted averages

$$\{\varrho\nabla u\} = \omega^- \, \varrho^+\nabla u^+ + \omega^+ \, \varrho^-\nabla u^- = \frac{\underline{\varrho}}{2}(\nabla u^+ + \nabla u^-)$$

and jumps

$$[u] = u^+ n^+ + u^- n^-,$$

where $n^\pm$ denotes the outward unit normal vector to $\tau^\pm$. We note that when $\varrho^+ = \varrho^- = 1$, then $\overline{\varrho} = \underline{\varrho} = 1$ and the weighted average reduces to the usual arithmetic average. We set

$$\gamma_0 = \frac{\overline{p}^2}{\underline{h}}\,\delta_0, \qquad \gamma_k = \frac{\underline{h}^{2k-1}}{\overline{p}^{2k}}\,\delta_k, \qquad \tilde{\gamma}_k = \frac{\underline{h}^{2k-1}}{\overline{p}^{2k}}\,\tilde{\delta}_k,$$

with

$$\underline{h} = \min\{h_+, h_-\}, \qquad \overline{p} = \max\{p_+, p_-\}.$$

where for simplicity we write $h_\pm, p_\pm$ for $h_{\tau^\pm}$ (or $p_{\tau^\pm}$, respectively). The parameters $\delta_0 > 0$ and $\delta_k, \tilde{\delta}_k \geq 0$ where $k \geq 1$ are some prescribed constants. We collect all δ_k in a multi-parameter $\delta = (\delta_0, \delta_1, \tilde{\delta}_1, \ldots)$.

On e which lies on $\partial\Omega$ and belongs to the face of $\tau \in \mathcal{T}_h$, we prescribe $\underline{\varrho} = \varrho$ and

$$\{\varrho\nabla u\} = \varrho\nabla u, \quad [u] = un, \quad \gamma_0 = \frac{p_\tau^2}{h_\tau}\delta_0, \quad \gamma_k = \frac{h_\tau^{2k-1}}{p_\tau^{2k}}\delta_k, \quad \tilde{\gamma}_k = \frac{h_\tau^{2k-1}}{p_\tau^{2k}}\tilde{\delta}_k.$$

Inspired by [1], we discretize (2) by the symmetric weighted interior (K, L)-th order multipenalty discontinuous Galerkin method: Find $u^* \in V_h^p$ such that

$$\mathcal{A}_h^{p,KL}(u^*, v) = (f, v)_\Omega - \sum_{k=1}^{L} \sum_{e \in \mathcal{E}_h^{\text{in}}} \frac{\tilde{\gamma}_{k+2}}{\overline{\varrho}} \langle [\varrho \frac{\partial^k f}{\partial n^k}], [\varrho \frac{\partial^k \Delta v}{\partial n^k}] \rangle_e, , \qquad \forall v \in V_h^p, \tag{4}$$

where

$$\mathcal{A}_h^{p,KL}(u,\, v) = A_h^{p,KL}(u,\, v) - F_h^p(u,\, v) - F_h^p(v,\, u)$$

and

$$A_h^{p,KL}(u,\, v) = \sum_{\tau\in\mathcal{T}_h} (\varrho\,\nabla u, \nabla v)_\tau + \sum_{e\in\mathcal{E}_h} \gamma_0 \underline{\varrho}\langle [u], [v]\rangle_e$$
$$+ \sum_{k=1}^{K} \sum_{e\in\mathcal{E}_h^{\text{in}}} \frac{\gamma_k}{\overline{\varrho}} \langle [\varrho \frac{\partial^k u}{\partial n^k}], [\varrho \frac{\partial^k v}{\partial n^k}]\rangle_e + \sum_{k=1}^{L} \sum_{e\in\mathcal{E}_h^{\text{in}}} \frac{\tilde{\gamma}_{k+2}}{\overline{\varrho}} \langle [\varrho \frac{\partial^k \Delta u}{\partial n^k}], [\varrho \frac{\partial^k \Delta v}{\partial n^k}]\rangle_e,$$
$$F_h^p(u,\, v) = \sum_{e\in\mathcal{E}_h} \langle \{\varrho \nabla u\}\,, [v]\rangle_e.$$

Here for $\tau \in \mathcal{T}_h$ and $e \in \mathcal{E}_h$ we use the standard notation: $(u, v)_\tau = \int_\tau u\, v\, dx$ and $\langle u, v\rangle_e = \int_e u\, v\, d\sigma$. This discretization generalizes the multipenalty method, introduced by Arnold in [4] with $L = 0$, to the case of discontinuous coefficient and takes into account the explicit dependence on the polynomial degree p. In particular, for $(K, L) = (0, 0)$, a standard symmetric weighted interior penalty method is restored, with

$$A_h^{p,00}(u,\, v) = \sum_{\tau\in\mathcal{T}_h} (\varrho\,\nabla u, \nabla v)_\tau + \sum_{e\in\mathcal{E}_h} \gamma_0 \underline{\varrho}\langle [u], [v]\rangle_e.$$

Moreover, for $\varrho \equiv 1$ and $(K, L) = (1, 0)$, problem (4) corresponds to the method by Douglas and Dupont [6]. The case of $L > 0$ has been considered e.g. in [1]. It is known [4] that for sufficiently large penalty constant δ_0 problem (4) is well–defined.

3 Nonoverlapping additive Schwarz method

Let us introduce the subdomain grid $\mathcal{T}_H$ as a partition of Ω into N_H disjoint open polygons (polyhedrons in 3-D) Ω_i, $i = 1, \ldots, N_H$, such that $\bar{\Omega} = \bigcup_{i=1,\ldots,N_H} \bar{\Omega}_i$ and that each Ω_i is a union of certain elements from the fine mesh $\mathcal{T}_h$. We shall retain the common notion of "subdomains" while referring to elements of $\mathcal{T}_H$. We set $H_i = \text{diam}(\Omega_i)$ and $H = (H_1, \ldots, H_{N_H})$. We assume that there exists a reference simply-connected polygonal (polyhedral in 3-D) domain $\hat{\Omega} \subset R^d$ with Lipschitz boundary, such that every Ω_i is affinely homeomorphic to $\hat{\Omega}$ and that the aspect ratios of Ω_i are bounded independently of h and H. Moreover, we assume that the number of neighboring regions in $\mathcal{T}_H$ is uniformly bounded by an absolute constant $\mathcal{N}$.

Next, let $\mathcal{T}_\mathcal{H}$ be a shape-regular affine triangulation by triangles in 2-D or tetrahedra in 3-D, with diameter $\mathcal{H}$. We denote the elements of $\mathcal{T}_\mathcal{H}$ by D_n and we call this partition the "coarse grid" and assume that ϱ is piecewise constant on $\mathcal{T}_\mathcal{H}$:

$$\varrho_{|D_n} = \varrho_n \qquad \forall 1 \le n \le N_\mathcal{H}. \tag{5}$$

Let us define the standard decomposition of V_h^p, cf. [3], [8]:

$$V_h^p = V_0 + V_1 + \ldots + V_{N_H}, \tag{6}$$

where the coarse space consists of functions which are polynomials inside each element of the coarse grid:

$$V_0 = \{v \in V_h^p : v_{|D_n} \in \mathbb{P}_q \text{ for all } n = 1, \ldots, N_{\mathcal{H}}\} \tag{7}$$

where $1 \le q \le \min\{p_\tau : \tau \in \mathcal{T}_h\}$. Next, for $i = 1, \ldots, N_H$ we set

$$V_i = \{v \in V_h^p : v_{|\Omega_j} = 0 \text{ for all } j \ne i\}.$$

One can view V_0 as a rough approximation to V_h^p (using coarser grid and lower order polynomials), cf. condition (11). Note that V_h^p already is a direct sum of spaces $V_1, \ldots, V_{N_H}$ and when $\mathcal{T}_{\mathcal{H}} = \mathcal{T}_H$, this decomposition coincides with [3]. Next, with fixed $0 \le r \le K$ and $0 \le s \le L$, we define inexact solvers $T_i : V_h^p \to V_i$, by

$$A_h^{p,rs}(T_i u, v) = \mathcal{A}_h^{p,KL}(u, v) \qquad \forall v \in V_i, \qquad 0 \le i \le N_H, \tag{8}$$

so that for $1 \le i \le N_H$ one has to solve only a relatively small system of linear equations on subdomain Ω_i (a "local problem") for $u_i = T_i u|_{\Omega_i}$. These subdomain problems are independent of each another and can be solved in parallel. The preconditioned operator is

$$T = T_0 + T_1 + \ldots + T_{N_H}. \tag{9}$$

Obviously, T is symmetric with respect to $\mathcal{A}_h^{p,KL}(\cdot, \cdot)$. For D_n in $\mathcal{T}_{\mathcal{H}}$ let us define an auxiliary seminorm

$$|||u|||^2_{D_n,\text{in}} = \sum_{\tau \in \mathcal{T}_h(D_n)} \varrho_n |\nabla u|^2_{0,\tau} + \sum_{e \in \mathcal{E}_h^{\text{in}}(D_n)} \gamma_0 \varrho_n |[u]|^2_{0,e}, \tag{10}$$

where $\mathcal{E}_h^{\text{in}}(D_n) = \{e \in \mathcal{E}_h : e \subset \bar{D}_n \setminus \partial D_n\}$.

Theorem 1 *Let us set $r = s = 0$ in* (8) *and assume that for each $u \in V_h^p$ there exists $u^{(0)} \in V_0$ satisfying*

$$\sum_{n=1}^{N_{\mathcal{H}}} \left(\frac{\varrho_n q_n^2}{\mathcal{H}_n^2} |u - u^{(0)}|^2_{0,D_n} + |||u - u^{(0)}|||^2_{D_n,\text{in}} \right) \lesssim \mathcal{A}_h^{p,00}(u, u). \tag{11}$$

Then the operator T defined in (9) *satisfies*

$$\beta^{-1} \mathcal{A}_h^{p,KL}(u, u) \lesssim \mathcal{A}_h^{p,KL}(Tu, u) \lesssim (K+L+1) \mathcal{A}_h^{p,KL}(u, u) \qquad \forall u \in V_h^p, \tag{12}$$

where

$$\beta = \max_{n=1,\ldots,N_{\mathcal{H}}} \left\{ \frac{\mathcal{H}_n^2}{q_n} \max_{i:\Omega_i \subset D_n} \left\{ \frac{\overline{p}_i^2}{\underline{h}_i H_i} \right\} \right\}$$

with $\underline{h}_i = \min\{h_\tau : \tau \in \mathcal{T}_h(\Omega_i)\}$ *and* $\overline{p}_i = \max\{p_\tau : \tau \in \mathcal{T}_h(\Omega_i)\}$. *Therefore, the condition number of* T *is* $O(\beta \cdot (K + L + 1))$.

Proof According to the general theory of ASM [11], it suffices to check three conditions. The strengthened Cauchy–Schwarz inequality holds with a constant independent of the parameters, due to our assumption that the number of neighbouring subdomains is bounded by an absolute constant.

For the local stability condition, it suffices to prove that for any k

$$\sum_{e\in\mathcal{E}_h^{\text{in}}} \frac{\gamma_k}{\overline{\varrho}} |[\varrho \frac{\partial^k u}{\partial n^k}]|_{0,e}^2 \lesssim A_h^{p,00}(u,u) \text{ and } \sum_{e\in\mathcal{E}_h^{\text{in}}} \frac{\tilde{\gamma}_{k+2}}{\overline{\varrho}} |[\varrho \frac{\partial^k \Delta u}{\partial n^k}]|_{0,e}^2 \lesssim A_h^{p,00}(u,u). \tag{13}$$

We prove the first inequality, the other can be proved analogously. On $e = \partial\tau^+ \cap \partial\tau^-$, we have (denoting by n either n^+ or n^-)

$$\frac{1}{\overline{\varrho}} |[\varrho \frac{\partial^k u}{\partial n^k}]|_{0,e}^2 \lesssim \frac{(\varrho^+)^2}{\overline{\varrho}} |\frac{\partial^k u^+}{\partial n^k}|_{0,e}^2 + \frac{(\varrho^-)^2}{\overline{\varrho}} |\frac{\partial^k u^-}{\partial n^k}|_{0,e}^2 \lesssim \varrho^+ |\frac{\partial^k u^+}{\partial n^k}|_{0,e}^2 + \varrho^- |\frac{\partial^k u^-}{\partial n^k}|_{0,e}^2,$$

since $(\varrho^\pm)^2/\overline{\varrho} = \omega^\pm \varrho^\pm \le \varrho^\pm$. Now, by the trace inequality [4], we have $|\frac{\partial^k u}{\partial n^k}|_{0,e}^2 \lesssim \frac{1}{h_\tau}|u|_{k,\tau}^2 + h_\tau |u|_{k+1,\tau}^2$, so applying k times the inverse inequality we arrive at

$$\frac{1}{\overline{\varrho}} |[\varrho \frac{\partial^k u}{\partial n^k}]|_{0,e}^2 \lesssim \varrho^+ \frac{p_+^{2k}}{h_+^{2k-1}} |u^+|_{1,\tau^+}^2 + \varrho^- \frac{p_-^{2k}}{h_-^{2k-1}} |u^-|_{1,\tau^-}^2,$$

which yields

$$\sum_{e\in\mathcal{E}_h^{\text{in}}} \gamma_k \frac{1}{\overline{\varrho}} \langle [\varrho \frac{\partial^k u}{\partial n^k}], [\varrho \frac{\partial^k u}{\partial n^k}] \rangle_e \lesssim \sum_{\tau\in\mathcal{T}_h} \varrho\, |u|_{1,\tau}^2 \lesssim A_h^{p,00}(u,u).$$

Summing (13) over k, we complete the stability estimate

$$\mathcal{A}_h^{p,KL}(u,u) \lesssim (K + L + 1)\, A_h^{p,00}(u,u) \qquad \forall u \in V_i, \quad \forall 0 \le i \le N_H,$$

from which the right inequality in (12) already follows.

Finally, to prove the existence of a stable decomposition, from [9] we have that there exists a decomposition of $u = \sum_{i=0}^{N_H} u^{(i)}$, with $u^{(i)} \in V_i$, such that

$$\sum_{i=0}^{N_H} A_h^{p,00}(u^{(i)}, u^{(i)}) \lesssim \beta\, \mathcal{A}_h^{p,00}(u,u) \qquad \forall u \in V_h^p.$$

Since $\mathcal{A}_h^{p,00}(u,u) \le \mathcal{A}_h^{p,KL}(u,u)$, we conclude that $\sum_{i=0}^{N_H} A_h^{p,00}(u^{(i)}, u^{(i)}) \lesssim \beta \mathcal{A}_h^{p,KL}(u,u)$, which gives us the left inequality in (12). □

Remark 1 Analogous result holds if, instead of the simplified form $A_h^{p,00}(\cdot,\cdot)$, we choose $A_h^{p,KL}(\cdot,\cdot)$ while defining local and coarse solvers T_i, $i = 0, 1, \ldots, N_H$, as we do in the following section.

Remark 2 In [10], sufficient conditions are provided for (11) to hold.

4 Numerical experiments

Let us choose the unit square $[0,1]^2$ as the domain Ω and consider (2) with $\varrho = 1$ in Ω. We do not investigate the influence of the intermediate grid $\mathcal{T}_H$, referring the reader to [9] for these results. Instead, we set $\mathcal{T}_H = \mathcal{T}_{\mathcal{H}}$ and use two levels of nested grids on Ω. For a prescribed integer $\mathcal{M}$, we divide Ω into $N_{\mathcal{H}} = 2^{\mathcal{M}} \times 2^{\mathcal{M}}$ squares of equal size. This coarse grid $\mathcal{T}_{\mathcal{H}}$ is then refined into a uniform fine triangulation $\mathcal{T}_h$ based on a square $2^m \times 2^m$ grid ($m \geq \mathcal{M}$) with each square split into two triangles of identical shape. Hence, the grid parameters are $h = 2^{-m}$, $\mathcal{H} = H = 2^{-\mathcal{M}}$. We set $L = 0$ and discretize problem (2) on the fine mesh $\mathcal{T}_h$ using (4) with $\delta_0 = 8$, $\delta_1 = \ldots = \delta_K = 2$ (if not specified otherwise) and equal polynomial degree p across all elements in $\mathcal{T}_h$. For the coarse problem, we set $q = p$. We always take $(r,s) = (K,L) = (K,0)$ while defining the inexact solvers, which seems to give preferable constants in (12). Our implementation makes use of the FEniCS [2] and MATLAB software packages.

In the following tables we report the number of Preconditioned Conjugate Gradient iterations for the operator T required to reduce the initial norm of the preconditioned residual by a factor of 10^8 and (in parentheses) the condition number of T estimated from the PCG convergence history. We always choose a random vector for the solution and a zero as the initial guess.

$\mathcal{H}$	iter (cond)
1/2	120 (328)
1/4	90 (157)
1/8	64 (71)
1/16	60 (60)

Table 1: Dependence on the coarse mesh size $\mathcal{H}$. Fixed $h = 1/64$, $p = 3$, $K = 3$.

p	iter (cond)
1	26 (11)
2	34 (21)
3	42 (34)
4	50 (50)
5	59 (70)

Table 2: Dependence on the polynomial degree p. Fixed $h = 1/16$, $\mathcal{H} = 1/4$, $K = 1$.

While the results with respect to $\mathcal{H}$ and p smoothly follow the theory developed, cf. Tables 1 and 2, the dependence on K is less regular, initially with superlinear increase, as reported in Table 3. Moreover, from Table 4 we observe that higher values of the penalization parameters δ_k, $k \geq 1$, adversely influence the convergence rate which is a drawback of this, otherwise simple and efficient, domain decomposition method.

K	iter (cond)
0	61 (81)
1	59 (70)
2	81 (124)
3	142 (389)
4	214 (902)
5	222 (1016)

Table 3: Dependence on the number of penalty terms K. Fixed $h = 1/16$, $\mathcal{H} = 1/4$, $p = 5$.

δ_1	iter (cond)
$2 \cdot 10^0$	29 (15)
$2 \cdot 10^1$	41 (29)
$2 \cdot 10^2$	102 (217)
$2 \cdot 10^3$	305 (2096)

Table 4: Dependence on the flux penalty parameter γ_1. Fixed $p = 3$, $K = 1$.

Acknowledgements The authors wish to thank two anonymous referees whose comments and remarks helped to improve the paper substantially. The research of Piotr Krzyżanowski has been partially supported by the Polish National Science Centre grant 2016/21/B/ST1/00350; Marcus Sarkis was supported by NSF-MPS 1522663.

References

1. Adjerid, S., Guo, R., Lin, T.: High degree immersed finite element spaces by a least square method. Int. J. Numer. Anal. Mod. **14**(4-5) (2017)
2. Alnæs, M.S., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M.E., Wells, G.N.: The FEniCS Project Version 1.5. Archive of Numerical Software **3**(100) (2015). DOI:10.11588/ans.2015.100.20553
3. Antonietti, P.F., Houston, P.: A class of domain decomposition preconditioners for hp-discontinuous Galerkin finite element methods. J. Sci. Comput. **46**(1), 124–149 (2011). DOI:10.1007/s10915-010-9390-1. URL http://dx.doi.org/10.1007/s10915-010-9390-1
4. Arnold, D.N.: An interior penalty finite element method with discontinuous elements. SIAM J. Numer. Anal. **19**(4), 742–760 (1982)
5. Burman, E., Guzmán, J., Sánchez, M.A., Sarkis, M.: Robust flux error estimation of an unfitted Nitsche method for high-contrast interface problems. IMA Journal of Numerical Analysis **38**(2), 646–668 (2018). DOI:10.1093/imanum/drx017. URL http://dx.doi.org/10.1093/imanum/drx017
6. Douglas Jr., J., Dupont, T.: Interior penalty procedures for elliptic and parabolic Galerkin methods. In: Computing methods in applied sciences (Second Internat. Sympos., Versailles, 1975), pp. 207–216. Lecture Notes in Phys., Vol. 58. Springer, Berlin (1976)
7. Dryja, M.: On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. Comput. Methods Appl. Math. **3**(1), 76–85 (electronic) (2003)
8. Dryja, M., Krzyżanowski, P.: A massively parallel nonoverlapping additive Schwarz method for discontinuous Galerkin discretization of elliptic problems. Num. Math. **132**(2), 347–367 (2015). DOI:10.1007/s00211-015-0718-5. URL http://dx.doi.org/10.1007/s00211-015-0718-5
9. Krzyżanowski, P.: On a nonoverlapping additive Schwarz method for h-p discontinuous Galerkin discretization of elliptic problems. Num. Meth. PDEs **32**(6), 1572–1590 (2016)
10. Krzyżanowski, P.: Nonoverlapping three grid Additive Schwarz for hp-DGFEM with discontinuous coefficients. In: Domain Decomposition Methods in Science and Engineering XXIV, *Lecture Notes in Computer Science and Engineering*, vol. 125, pp. 455–463. Springer (2018)
11. Toselli, A., Widlund, O.: Domain decomposition methods—algorithms and theory, *Springer Series in Computational Mathematics*, vol. 34. Springer-Verlag, Berlin (2005)