

On the linear convergence of additive Schwarz methods for the p -Laplacian

YOUNG-JU LEE

Department of Mathematics, Texas State University, San Marcos, TX 78666, USA

AND

JONGHO PARK*

Applied Mathematics and Computational Sciences Program, Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia

*Corresponding author: jongho.park@kaust.edu.sa

[Received on 15 August 2023; revised on 9 August 2024]

We consider additive Schwarz methods for boundary value problems involving the p -Laplacian. While existing theoretical estimates suggest a sublinear convergence rate for these methods, empirical evidence from numerical experiments demonstrates a linear convergence rate. In this paper we narrow the gap between these theoretical and empirical results by presenting a novel convergence analysis. First, we present a new convergence theory for additive Schwarz methods written in terms of a quasi-norm. This quasi-norm exhibits behaviour akin to the Bregman distance of the convex energy functional associated with the problem. Secondly, we provide a quasi-norm version of the Poincaré–Friedrichs inequality, which plays a crucial role in deriving a quasi-norm stable decomposition for a two-level domain decomposition setting. By utilizing these key elements we establish the asymptotic linear convergence of additive Schwarz methods for the p -Laplacian.

Keywords: additive Schwarz method; p -Laplacian; linear convergence; quasi-norm; Poincaré–Friedrichs inequality; convergence analysis.

1. Introduction

Let Ω be a bounded polygonal domain in $\mathbb{R}^2$ with the Lipschitz boundary $\partial\Omega$. Given $p \in (1, \infty)$ we consider the following p -Laplace equation:

$$\begin{aligned} -\nabla \cdot (|\nabla u|^{p-2} \nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

where $f \in L^{p^*}(\Omega)$ with p^* being from the equation $1/p + 1/p^* = 1$.

The p -Laplacian is a standard example of nonlinear elliptic problems (Benedikt *et al.*, 2018). Furthermore, it has a number of application areas, including glaciology, non-Newtonian fluids (Shapovalov, 2017), nonlinear diffusion and nonlinear elasticity; see Díaz (1985) and references therein. Thus, there has been extensive research on (1.1), especially for numerical solutions of (1.1). Some important early results can be found in Glowinski & Marrocco (1975); Ciarlet (2002). Finite element methods for the p -Laplacian were analysed in terms of the quasi-norm in Barrett & Liu (1993, 1994). Further studies

on error estimates for the p -Laplacian in terms of the quasi-norm were conducted in [Liu & Yan \(2001, 2002\)](#); [Ebmeier & Liu \(2005\)](#); [Carstensen *et al.* \(2006\)](#). Linear convergence of adaptive finite element methods for (1.1) was shown in [Diening & Kreuzer \(2008\)](#). Numerical homogenization for multiscale p -Laplacian problems was investigated in [Liu *et al.* \(2021\)](#).

It is well known that the boundary value problem (1.1) can be formulated in the following weak form ([Glowinski & Marrocco, 1975](#); [Ciarlet, 2002](#)): find $u \in W_0^{1,p}(\Omega)$ such that

$$\int_{\Omega} |\nabla u|^{p-2} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx, \quad v \in W_0^{1,p}(\Omega),$$

where $W_0^{1,p}(\Omega)$ is a usual Sobolev space consisting of the $L^p(\Omega)$ -functions vanishing on $\partial\Omega$ with $(L^p(\Omega))^2$ -gradient. Equivalently, it is interpreted as the following convex optimization problem:

$$\min_{v \in W_0^{1,p}(\Omega)} F(v) := \frac{1}{p} \int_{\Omega} |\nabla v|^p \, dx - \int_{\Omega} f v \, dx. \quad (1.2)$$

That is, one may deal with the convex optimization problem (1.2) to obtain a solution of (1.1). Based on the convex optimization formulation (1.2) multigrid and preconditioned descent methods were proposed in [Bermejo & Infante \(2000\)](#) and [Huang *et al.* \(2007\)](#), respectively. In particular, the framework of subspace correction methods ([Xu, 1992](#)) for (1.2) were considered in [Tai & Xu \(2002\)](#); [Park \(2020\)](#).

This paper is concerned with numerical solutions of boundary value problems involving the p -Laplacian by additive Schwarz methods. Additive Schwarz methods, also known as parallel subspace correction methods, have been broadly used as efficient numerical solvers for large-scale scientific problems; see [Xu \(1992\)](#); [Toselli & Widlund \(2005\)](#) and references therein for relevant results on linear problems. In additive Schwarz methods the domain of a target problem is decomposed into a union of several subdomains, and optimal local corrections on the subdomains with respect a numerical approximation for the solution are computed in parallel. The numerical approximation for the solution is iteratively updated by collecting all the local corrections. Due to their parallel structures additive Schwarz methods are suitable for massively parallel computation using distributed memory computers. In the past decades there have been a number of results on additive Schwarz methods for large-scale convex optimization problems. The framework of additive Schwarz methods was first considered for convex optimization in [Tai & Espedal \(1998\)](#), and subsequently applied to the p -Laplacian in [Tai & Xu \(2002\)](#). These methods have since been further investigated in several studies, including [Badea \(2006, 2019\)](#); [Park \(2020, 2022\)](#).

The convergence rate of additive Schwarz methods for the p -Laplacian problem (1.1) was first analysed in [Tai & Xu \(2002\)](#); the $O(n^{\frac{p(p-1)}{(p-\underline{p})(\bar{p}-1)}})$ energy convergence of the methods was proven, where n denotes the number of iterations, $\underline{p} = \min\{p, 2\}$, and $\bar{p} = \max\{p, 2\}$. Recently, [Park \(2020\)](#) showed that the methods satisfy the improved $O(n^{\frac{p(p-1)}{p-\underline{p}}})$ convergence rate (see Proposition 2.3). The results in both [Tai & Xu \(2002\)](#) and [Park \(2020\)](#) are based on some estimates for the Bregman distance of the energy functional F in (1.2). Roughly speaking, these estimates are written as

$$\frac{\mu}{p} \|u - v\|_{W^{1,p}(\Omega)}^p \leq D_F(u, v) \leq L \frac{p}{p-\underline{p}} \|u - v\|_{W^{1,p}(\Omega)}^p, \quad u, v \in W_0^{1,p}(\Omega), \quad (1.3)$$

where $\mu_{\bar{p}}$ and L_p are positive constants independent of u and v , and $D_F(u, v)$ is the Bregman distance of F defined by

$$D_F(u, v) = F(u) - F(v) - \langle F'(v), u - v \rangle, \quad u, v \in W_0^{1,p}(\Omega). \quad (1.4)$$

Here, $F'(v)$ stands for the Frechét derivative of F at v given by

$$\langle F'(v), w \rangle = \int_{\Omega} |\nabla v|^{p-2} \nabla v \cdot \nabla w \, dx - \int_{\Omega} f w \, dx, \quad w \in W_0^{1,p}(\Omega). \quad (1.5)$$

One may refer to [Tai & Xu \(2002, Lemma 2.1\)](#) and [Park \(2020, Section 6.1\)](#) for details on the estimate (1.3).

While both [Tai & Xu \(2002\)](#) and [Park \(2020\)](#) proved the sublinear convergence of additive Schwarz methods for the p -Laplacian, it was observed numerically in several works that the methods actually converge linearly; see, e.g., [Park \(2021, Fig. 2\)](#). Indeed, as we will demonstrate in the numerical experiments presented in Section 5 of this paper, additive Schwarz methods for (1.1) exhibit linear convergence empirically under various settings on discretization and domain decomposition. More precisely, each convergence curve of the energy error with respect to the number of iterations seems linear in the x -linear y -log scale plot when the number of iterations is sufficiently large, which means that the energy error decays exponentially as the number of iterations increases. This implies that the existing convergence estimates for additive Schwarz methods for the p -Laplacian may not be optimal.

The main motivation of this paper is to discuss a linear convergence analysis for additive Schwarz methods to solve the p -Laplacian problem (1.1). As we mentioned above, while the existing theoretical estimates ([Tai & Xu, 2002](#); [Park, 2020](#)) for the convergence rate of additive Schwarz methods for the p -Laplacian are sublinear, the empirical convergence rate observed by numerical experiments is linear. This discrepancy between theoretical and empirical results motivates our work, as we aim to bridge this gap by rigorously proving the asymptotic linear convergence of additive Schwarz methods for the p -Laplacian.

In (1.3) $\underline{p}$ and $\bar{p}$ do not agree if $p \neq 2$, so that the lower and upper bounds for $D_F(u, v)$ are expressed in powers of $\|u - v\|_{W^{1,p}(\Omega)}$ with different exponents. This discrepancy indicates that a power of norm is not adequate as a tight two-sided approximation for the Bregman distance; whenever we establish a bound for $D_F(u, v)$ in terms of $\|u - v\|_{W^{1,p}(\Omega)}$ or vice versa we suffer from a kind of looseness. We claim that the sublinear convergence rates given in the existing works ([Tai & Xu, 2002](#); [Park, 2020](#)) are caused by this looseness. To overcome this issue we propose to use the quasi-norm developed in [Liu & Yan \(2001, 2002\)](#); [Ebmeyer & Liu \(2005\)](#); [Carstensen et al. \(2006\)](#), which is relevant to the problem of consideration and approximates the Bregman distance appropriately, and then to derive the convergence estimate in terms of the quasi-norm. This approach is similar to obtain the convergence measure of the iterative method using the energy-like metric relevant to the problem to be solved, as discussed in [Lee et al. \(2008, 2009\)](#). We denote the quasi-norm by $\|\cdot\|_{(\nabla v)}$ (see (3.1)) and show that

$$\mu_p \|u - v\|_{(\nabla v)}^2 \leq D_F(u, v) \leq L_p \|u - v\|_{(\nabla v)}^{1,p}, \quad u, v \in W_0^{1,p}(\Omega) \quad (1.6)$$

for some positive constants μ_p and L_p (see Lemma 3.3), i.e., $\|u - v\|_{(\nabla v)}^2$ approximates $D_F(u, v)$ well up to a multiplicative constant. Meanwhile, we note that the quasi-norm $\|\cdot\|_{(\nabla v)}$, along with several alternative versions described in [Diening & Růžička \(2007\)](#), [Diening & Kreuzer \(2008\)](#), do not induce

a norm. As a result existing convergence theories for additive Schwarz methods (Tai & Xu, 2002; Park, 2020) cannot directly utilize the estimate (1.6). A novelty in this paper is that, by extending the idea of Park (2020), a new convergence theory for additive Schwarz methods is obtained in terms of the quasi-norm, which utilizes (1.6) to obtain the asymptotic linear convergence rate of additive Schwarz methods for the p -Laplacian. In our linear convergence analysis a quasi-norm version of the Poincaré–Friedrichs inequality (see Lemmas 3.4 and 3.5) plays a critical role. We validate this asymptotic linear convergence result numerically in Section 5.

The rest of this paper is organized as follows. In Section 2 we present finite element approximations, domain decomposition settings and a two-level additive Schwarz method for the p -Laplacian problem. An asymptotic linear convergence analysis of the two-level additive Schwarz method is given in Section 3. In Section 4 we present details of the quasi-norm Poincaré–Friedrichs inequality that is used in the convergence analysis of the methods. In Section 5 we provide numerical results of the two-level additive Schwarz method for the p -Laplacian problem across various settings. Finally, we provide a concluding remark for our paper in Section 6.

2. Additive Schwarz methods

In this section we introduce finite element spaces and domain decomposition settings for the p -Laplacian problem (1.2). Based on these settings we present a two-level additive Schwarz method for (1.2) and its convergence theory, which explains the asymptotic linear convergence of the algorithm.

In what follows the notation $A \lesssim B$ means that there exists a constant $c > 0$ such that $A \leq cB$, where c is independent of the geometric parameters H , h and δ relying on discretization and domain decomposition. We also write $A \approx B$ if $A \lesssim B$ and $B \lesssim A$.

2.1 Discretization and domain decomposition

Let $\mathcal{T}_h$ be a quasi-uniform triangulation of Ω with h the characteristic element diameter. The collection of continuous and piecewise linear functions on $\mathcal{T}_h$ vanishing on $\partial\Omega$ is denoted by $V = S_h(\Omega)$. Clearly, we have $V \subset W_0^{1,\infty}(\Omega)$. For continuous functions the nodal interpolation operator I_h onto $S_h(\Omega)$ is well-defined.

In what follows we consider the following conforming finite element approximation of (1.2) defined on V :

$$\min_{u \in V} F(u). \quad (2.1)$$

A unique solution of (2.1) is denoted by $u^* \in V$. Convergence properties of (2.1) as $h \rightarrow 0$ can be found in Barrett & Liu (1993); Ciarlet (2002).

Next, we describe domain decomposition settings for the problem (2.1). We assume that Ω admits another quasi-uniform triangulation $\mathcal{T}_H$ with H the characteristic element diameter such that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$. A finite element space $S_H(\Omega)$ is defined in the same manner as $S_h(\Omega)$. In the two-level additive Schwarz method for (2.1) $\mathcal{T}_h$ and $\mathcal{T}_H$ will play roles of fine and coarse meshes, respectively. Let $\{\Omega_k\}_{k=1}^N$ be a nonoverlapping domain decomposition of Ω such that each Ω_k is the union of several coarse elements in $\mathcal{T}_H$ and the number of coarse elements consisting of Ω_k is uniformly bounded. For each subdomain Ω_k , $1 \leq k \leq N$ we consider an enlarged region Ω_k^δ consisting of the elements $T \in \mathcal{T}_h$ with $\text{dist}(T, \Omega_k) \leq \delta$. Then $\{\Omega_k^\delta\}_{k=1}^N$ forms an overlapping domain decomposition of Ω . We define

$S_h(\Omega_k) \subset W_0^{1,\infty}(\Omega_k)$ as the piecewise linear finite element space on the $T_h|_{\Omega_k}$ with the homogeneous essential boundary condition.

We set

$$V_0 = S_H(\Omega), \quad V_k = S_h(\Omega_k^t), \quad 1 \leq k \leq N.$$

A two-level domain decomposition for V is given by

$$V = \sum_{k=0}^N R_k^* V_k, \quad (2.2)$$

where $R_k^*: V_k \rightarrow V$, $1 \leq k \leq N$ is the natural extension-by-zero operator and $R_0^*: V_0 \rightarrow V$ is the natural interpolation operator. Let $\{\theta_k\}_{k=1}^N$ be the piecewise linear partition of unity for Ω subordinate to the covering $\{\Omega_k^t\}_{k=1}^N$ that was presented in [Toselli & Widlund \(2005, Eq. \(3.7\)\)](#). It is known that $\{\theta_k\}_{k=1}^N$ satisfies the following properties:

$$\theta_k = 0 \quad \text{on } \Omega \setminus \Omega_k^t, \quad (2.3a)$$

$$\sum_{k=1}^N \theta_k = 1 \quad \text{on } \overline{\Omega}, \quad (2.3b)$$

$$\|\nabla \theta_k\|_{L^\infty(\Omega_k^t)} \leq \frac{1}{\delta}, \quad 1 \leq k \leq N. \quad (2.3c)$$

The following lemma summarizes an important result on stable decomposition for the two-level domain decomposition (2.2) (see [Tai & Xu \(2002, Lemma 4.1\)](#)).

LEMMA 2.1. For $w \in V$, let $w_0 \in V_0$ be the $L^2(\Omega)$ -orthogonal projection of w onto V_0 and let $w_k \in V_k$, $1 \leq k \leq N$, such that

$$R_k^* w_k = I_h(\theta_k(w - R_0^* w_0)).$$

For $s \geq 1$ we have $w = \sum_{k=0}^N R_k^* w_k$ and

$$\sum_{k=0}^N \|R_k^* w_k\|_{W^{1,s}(\Omega)}^s \leq \left\{ 1 + \left(\frac{H}{\delta} \right)^{s-1} \right\} \|w\|_{W^{1,s}(\Omega)}^s.$$

Using the usual colouring technique one can prove that the two-level domain decomposition (2.2) enjoys the strengthened convexity condition (see [Park \(2020, Assumption 4.2\)](#)).

LEMMA 2.2. Let N_c be the minimum number of colours such that $\{\Omega_k^t\}_{k=1}^N$ is coloured in a way that the subdomains with the same colour do not intersect with each other, and let $\tau_0 = 1/(N_c + 1)$. For any

$v \in V$, $w_k \in V_k$, $0 \leq k \leq N$, and $\tau \in (0, \tau_0]$, we have

$$(1 - \tau(N + 1))F(v) + \tau \sum_{k=0}^N F(v + R_k^* w_k) \geq F\left(v + \tau \sum_{k=0}^N R_k^* w_k\right).$$

Proof. See Park (2020, Section 5.1). For suitable overlaps we have $N_c = 4$ (Tai & Xu, 2002). $\square$

2.2 Two-level additive Schwarz method

The two-level additive Schwarz method for (2.1) based on the space decomposition (2.2) is described in Algorithm 1. It is worth noting that this algorithm has been investigated in several prior works. The algorithm for smooth convex optimization was first considered in Tai & Espedal (1998), and then applied to the p -Laplacian in Tai & Xu (2002). Later, the framework was generalized to constrained and nonsmooth convex optimization problems in Badea (2006) and Park (2020, 2021), respectively. The constant τ_0 in Algorithm 1 was given in Lemma 2.2.

Algorithm 1 Two-level additive Schwarz method for (2.1)

Let $u^{(0)} \in V$ and $\tau \in (0, \tau_0]$.

for $n = 0, 1, 2, \dots$ **do**

$$w_k^{(n+1)} = \arg \min_{w_k \in V_k} F(u^{(n)} + R_k^* w_k), \quad 0 \leq k \leq N$$

$$u^{(n+1)} = u^{(n)} + \tau \sum_{k=0}^N R_k^* w_k^{(n+1)}$$

end for

The following proposition summarizes the sublinear convergence rate of Algorithm 1 analysed in Park (2020, Theorem 6.1). It was discussed in Park (2021, Remark 4.2) that the rate presented in Proposition 2.3 is the sharpest estimate among the existing ones (Tai & Xu, 2002; Badea, 2006; Badea & Krause, 2012; Park, 2020).

PROPOSITION 2.3. In Algorithm 1, we write $\zeta_n = F(u^{(n)}) - F(u^*)$ for $n \geq 0$. There exist positive constants ζ^* and c^* , depending on $u^{(0)}$, τ and H/δ , such that

$$\zeta_{n+1} \leq \begin{cases} (1 - \tau)^{\frac{p}{p-1}} \zeta_n, & \zeta_n \geq \zeta^*, \\ \zeta_n - c^* \zeta_n^{\frac{p(p-1)}{p(p-1)}}, & \zeta_n < \zeta^*, \end{cases}$$

where $\underline{p} = \min\{p, 2\}$ and $\bar{p} = \max\{p, 2\}$. Consequently, we have

$$\zeta_n \leq \frac{1}{(c^*(n+1))^{\frac{\bar{p}(\bar{p}-1)}{p-\bar{p}}}}$$

for sufficiently large $n \geq 0$.

Proof. See Park (2020, Section A.4). $\square$

While Proposition 2.3 ensures the sublinear convergence of Algorithm 1, as we will see in Section 5, the actual numerical behaviour indicates linear convergence. This observation motivates us to develop a new convergence theory for Algorithm 1 that can explain the linear convergence. We summarize our main result, the asymptotic linear convergence of Algorithm 1, in Theorem 2.4. The proof of Algorithm 1 will be provided in Section 3. We highlight that Theorem 2.4 stands as the first theoretical result that explains the linear convergence of the additive Schwarz method for the p -Laplacian.

THEOREM 2.4. If the solution $u^* \in V$ of (2.1) satisfies that $|\nabla u^*|$ does not vanish on Ω then, in Algorithm 1, we have

$$\limsup_{n \rightarrow \infty} \frac{F(u^{(n+1)}) - F(u^*)}{F(u^{(n)}) - F(u^*)} \leq 1 - \gamma^{-1},$$

where γ is a positive constant depending on p, u^*, H, δ and τ such that

$$\gamma \geq 1 + \bar{C}_{p,u^*}^{\text{PF}} \left(\frac{H}{\delta} \right)^{-p \cdot \frac{1}{\min\{p^2-1, 1\}}},$$

and the constant $\bar{C}_{p,u^*}^{\text{PF}}$ is given in either Lemma 3.4 or Lemma 3.5.

Regarding the condition in Theorem 2.4 that requires the finite element solution u^* to satisfy $|\nabla u^*| \neq 0$ on Ω , we discuss its validity for extreme values of p , particularly when p is either very large or close to 1. As we will demonstrate in Section 5, for large p , the solution may develop a singularity (see Fig. 2(e, f)). Fortunately, this singularity does not violate the condition $|\nabla u^*| \neq 0$. However, when p is close to 1 the solution may exhibit a flat region, potentially leading to a vanishing gradient (see Fig. 2(a, b)). Consequently, the applicability of Theorem 2.4 to cases near $p = 1$ may be limited.

Despite the potential limitations in applying Theorem 2.4 to such cases it remains practically relevant, as many real-world applications involving the p -Laplacian typically utilize moderate values of p . For instance, in modelling nonlinear Darcy law for fluid flows, as discussed in Benedikt *et al.* (2018), physically meaningful values for p are generally greater than 3/2.

We conclude this section by mentioning several acceleration methodologies that can be applied to Algorithm 1. In Park (2021, 2022) acceleration schemes for additive Schwarz methods for convex optimization were proposed. As the energy functional F is convex these schemes can be directly applied to Algorithm 1 to yield accelerated variants. These accelerated methods show faster convergence behaviours than the vanilla method, while they have essentially the same computational cost per iteration; see Park (2022) for relevant numerical results. We do not deal with the accelerated methods in detail because they are beyond the scope of this paper.

3. Convergence analysis

The main objective of this section is to prove Theorem 2.4, which is the asymptotic linear convergence theorem for the two-level additive Schwarz method for the p -Laplacian. We begin by presenting some useful properties of the quasi-norm $\|\cdot\|_{(\nabla_V)}$ (Liu & Yan, 2001; Ebmeyer & Liu, 2005), which is defined

as

$$\|w\|_{(\nabla v)}^2 = \int_{\Omega} (|\nabla w| + |\nabla v|)^{p-2} |\nabla w|^2 dx, \quad v, w \in W^{1,p}(\Omega). \quad (3.1)$$

Subsequently, we prove Theorem 2.4 by verifying a certain quasi-norm stable decomposition property (Tai & Xu, 2002; Park, 2020).

3.1 Properties of the quasi-norm

The quasi-norm $\|\cdot\|_{(\nabla v)}$ given in (3.1) satisfies a scaling property in the sense that the $\|tw\|_{(\nabla v)}$ is bounded by $\|w\|_{(\nabla v)}$ multiplied by t^α for some $\alpha \in \mathbb{R}$, where $v, w \in W^{1,p}(\Omega)$ and $t \in [0, 1]$. Lemma 3.1 summarizes such a property.

LEMMA 3.1. For any $v, w \in W^{1,p}(\Omega)$ and $t \in [0, 1]$, we have

$$t^{\max\{p,2\}} \|w\|_{(\nabla v)}^2 \leq \|tw\|_{(\nabla v)}^2 \leq t^{\min\{p,2\}} \|w\|_{(\nabla v)}^2.$$

Proof. Suppose that $p \in [2, \infty)$. Since the map $x \mapsto x^{p-2}$ ($x \geq 0$) is increasing we get

$$\begin{aligned} \|tw\|_{(\nabla v)}^2 &\leq \int_{\Omega} (|\nabla w| + |\nabla v|)^{p-2} |t\nabla w|^2 dx = t^2 \|w\|_{(\nabla v)}^2, \\ \|tw\|_{(\nabla v)}^2 &\geq \int_{\Omega} (t|\nabla w| + t|\nabla v|)^{p-2} |t\nabla w|^2 dx = t^p \|w\|_{(\nabla v)}^2. \end{aligned}$$

The case $p \in (1, 2)$ can be shown by a similar argument using the fact that the map $x \mapsto x^{p-2}$ ($x \geq 0$) is decreasing. $\square$

The following lemma states that $\|u - v\|_{(\nabla v)}$ is bounded by $\|u - v\|_{(\nabla u)}$ up to a multiplicative constant independent of $u, v \in W^{1,p}(\Omega)$.

LEMMA 3.2. For any $u, v \in W^{1,p}(\Omega)$, we have

$$\|u - v\|_{(\nabla v)}^2 \leq 2^{[p-2]} \|u - v\|_{(\nabla u)}^2.$$

Proof. Invoking the vector inequality

$$|\zeta + \eta| + |\zeta| \leq 2(|\zeta + \eta| + |\eta|), \quad \zeta, \eta \in \mathbb{R}^2$$

we get

$$\begin{aligned} \|u - v\|_{(\nabla v)}^2 &= \int_{\Omega} (|\nabla(u - v)| + |\nabla v|)^{p-2} |\nabla(u - v)|^2 dx \\ &\leq 2^{[p-2]} \int_{\Omega} (|\nabla(v - u)| + |\nabla u|)^{p-2} |\nabla(v - u)|^2 dx = 2^{[p-2]} \|u - v\|_{(\nabla u)}^2, \end{aligned}$$

which completes the proof. $\square$

In Barrett & Liu (1993, 1994) the following vector inequalities were established: there exist two positive constants C_1 and C_2 such that, for any $\xi, \eta \in \mathbb{R}^2$, the following hold:

$$(|\xi|^{p-2}\xi - |\eta|^{p-2}\eta) \cdot (\xi - \eta) \leq C_1 |\xi - \eta| (|\xi| + |\eta|)^{p-2}, \quad (3.2a)$$

$$(|\xi|^{p-2}\xi - |\eta|^{p-2}\eta) \cdot (\xi - \eta) \geq C_2 |\xi - \eta|^2 (|\xi| + |\eta|)^{p-2}. \quad (3.2b)$$

Using (3.2) and proceeding similarly to Barrett & Liu (1993, Theorem 2.1) we prove Lemma 3.3, which says that the estimate (1.6) actually holds. Lemma 3.3 will play an important role in proving (3.8); see also Liu *et al.* (2021, Lemma 2.3).

LEMMA 3.3. There exists positive constants μ_p and L_p depending on p such that, for any $u, v \in W^{1,p}(\Omega)$, we have

$$\mu_p \|u - v\|_{(\nabla v)}^2 \leq D_F(u, v) \leq L_p \|u - v\|_{(\nabla v)}^2.$$

Proof. By the definition of $D_F(u, v)$ given in (1.4) and the fundamental theorem of calculus, we have

$$\begin{aligned} D_F(u, v) &= \int_0^1 (F^t(v + t(u - v)), u - v) dt - (F^t(v), u - v) \\ &= \int_0^1 (F(v + t(u - v)) - F(v), t(u - v)) dt. \end{aligned}$$

With $u_t = v + t(u - v)$ we see that

$$\begin{aligned} D_F(u, v) &\stackrel{(5)}{\leq} \int_0^1 \int_{\Omega} (|\nabla u_t|^{p-2} \nabla u_t - |\nabla v|^{p-2} \nabla v) \cdot \nabla(u_t - v) dx dt \\ &\leq \int_0^1 \int_{\Omega} (|\nabla u_t|^{p-2} \nabla u_t - |\nabla v|^{p-2} \nabla v) \cdot \nabla(u_t - v) dx dt \\ &\stackrel{(3.2a)}{\leq} \int_0^1 \int_{\Omega} (|\nabla u_t| + |\nabla v|)^{p-2} |\nabla(u_t - v)|^2 dx dt \\ &= \int_0^1 \int_{\Omega} (|\nabla(v + t(u - v))| + |\nabla v|)^{p-2} |\nabla(u - v)|^2 dx dt. \end{aligned}$$

Now, we invoke the inequality

$$\frac{t}{2} (|\xi| + |\eta|) \leq |\xi + t\eta| + |\xi| \leq 2 (|\xi| + |\eta|), \quad \xi, \eta \in \mathbb{R}^2, t \in [0, 1], \quad (3.3)$$

to obtain that

$$D_F(u, v) \leq \int_0^1 \int_{\Omega} (|\nabla v| + |\nabla(u - v)|)^{p-2} |\nabla(u - v)|^2 dx dt \approx \|u - v\|_{(\nabla v)}^2.$$

Hence, we proved $D_F(u, v) \leq \|u - v\|_{(\nabla v)}^2$. The inequality $D_F(u, v) \leq 2 \|u - v\|_{(\nabla v)}^2$ can be shown in a similar manner using (3.2b) and (3.3). $\square$

3.2 Quasi-norm stable decomposition

The core step in the convergence analysis of additive Schwarz methods typically involves verifying a stable decomposition property; see, e.g., Tai & Xu (2002, Eq. (13)) and Park (2020, Assumption 4.1). In this section we derive a quasi-norm stable decomposition property associated with the space decomposition (2.2). A key distinction of the quasi-norm stable decomposition property considered in this section compared with the existing ones is that we use the quasi-norm $\|\cdot\|_{(\nabla v)}$, while the existing ones are written in terms of norms. As (1.3) implies a power of norm cannot approximate the Bregman distance of F by a multiplicative constant if $p \neq 2$. Our main insight is that if the quasi-norm can approximate the Bregman distance of F up to a multiplicative constant, i.e., if it satisfies an estimate of the form (1.6), then we can derive the asymptotic linear convergence of Algorithm 1 using this property.

We recall that two key ingredients for the stable decomposition analysis for linear elliptic problems are the Poincaré–Friedrichs inequality and interpolation error estimate; see Toselli & Widlund (2005, Chapter 3). Therefore, we need to establish these theories with respect to the quasi-norm for the stable decomposition analysis of the p -Laplacian.

In Lemmas 3.4 and 3.5 we present quasi-norm Poincaré–Friedrichs inequalities for the cases $p \in (2, \infty)$ and $p \in (1, 2)$, respectively, that are suitable for our purposes; more general results are proven in Section 4.

LEMMA 3.4. Let $p \in (2, \infty)$ and $v \in S_h(\Omega)$. Assume that every maximal polygonal region $R \subset \Omega$ with $|\nabla v| \neq 0$ satisfies that $\partial R \cap \partial\Omega$ contains an element edge. Then there exists a positive constant $C_{p,v}^{\text{PF}}$ such that

$$\int_{\Omega} (|w| + |\nabla v|)^{p-2} |w|^2 dx \leq C_{p,v}^{\text{PF}} \|w\|_{(\nabla v)}^2, \quad w \in W_0^{1,p}(\Omega).$$

Moreover, if $|\nabla v|$ does not vanish on Ω then $C_{p,v}^{\text{PF}}$ has an upper bound $\bar{C}_{p,v}^{\text{PF}}$ that is continuous at v in $S_h(\Omega)$.

LEMMA 3.5. Let $p \in (1, 2)$ and $v \in S_h(\Omega)$. Assume that every maximal polygonal region $S \subset \Omega$ with $|\nabla v| \neq 0$ satisfies that $\partial S \cap \partial\Omega$ contains an element edge. Then there exists a positive constant $C_{p,v}^{\text{PF}}$ such that

$$\int_{\Omega} (|w| + |\nabla v|)^{p-2} |w|^2 dx \leq C_{p,v}^{\text{PF}} \|w\|_{(\nabla v)}^2, \quad w \in W_0^{1,p}(\Omega).$$

Moreover, if $|\nabla v|$ does not vanish on Ω then $C_{p,v}^{\text{PF}}$ has an upper bound $\bar{C}_{p,v}^{\text{PF}}$ that is continuous at v in $S_h(\Omega)$.

As stated in Lemmas 3.4 and 3.5 the quasi-norm Poincaré–Friedrichs inequality holds for all choices of v except for certain exceptional cases, which are detailed in Examples 4.8 and 4.15. Moreover, in most

cases, the constant $C_{p,v}^{\text{PF}}$ demonstrates only a weak dependence on v . By the quasi-monotone argument (Galvis & Efendiev, 2010; Pechstein & Scheichl, 2013) presented in Section 4 we can ensure that the value of $C_{p,v}^{\text{PF}}$ is influenced by the local variation of $|\nabla v|$ only. Consequently, even if $|\nabla v|$ exhibits significant global variation $C_{p,v}^{\text{PF}}$ has a moderate value. One may refer to Scheichl *et al.* (2012) for relevant numerical evidences.

REMARK 3.6. As noted in Lemmas 3.4 and 3.5, the quasi-norm Poincaré–Friedrichs inequality may not hold in cases where ∇v vanishes in a certain pattern, which makes the convergence analysis of the algorithm challenging. In order to address this issue one may consider regularization techniques as described in Diening *et al.* (2020); Liu *et al.* (2021). However, we do not adopt such techniques since they require a delicate convergence analysis for the case when the regularization parameter tends to 0.

Next, we establish a quasi-norm error estimate for the nodal interpolation operator I_h onto the finite element space $S_h(\Omega)$, as summarized in Lemma 3.7.

LEMMA 3.7. Let $w \in W_0^{1,p}(\Omega)$ be a continuous, piecewise quadratic function defined on T_h and let $v \in S_h(\Omega)$. Then, there exists a positive constant C , independent of w , v and h , such that

$$\|I_h w\|_{(\nabla v)} \leq C \|w\|_{(\nabla v)}.$$

Proof. Take any $T \in T_h$. We first prove that I_h achieves the local $W^{1,1}$ -stability; invoking the inverse inequality (Ern & Guermond, 2021, Lemma 12.1) and the H^1 -stability (Toselli & Widlund, 2005, Lemma 3.9) yields

$$|I_h w|_{W^{1,1}(T)} \leq C h |I_h w|_{H^1(T)} \leq C h |w|_{H^1(T)} \leq C |I_h w|_{W^{1,1}(T)}. \quad (3.4)$$

Now, we proceed similarly as in the proof of Diening & Růžička (2007, Theorem 4.5). Recall that $|\nabla v|$ is constant on T , say $a = |\nabla v| \geq 0$. Since the map $x \mapsto (x + a)^{p-2} x^2$ ($x \geq 0$) is increasing and convex, we have

$$\begin{aligned} & \int_T (|\nabla(I_h w)| + a)^{p-2} |\nabla(I_h w)|^2 dx \\ & \stackrel{(i)}{\leq} \int_T h^{-2} |\nabla(I_h w)(y)| dy + a^{-p-2} \int_T h^{-2} |\nabla(I_h w)| dy^{-2} dx \\ & \stackrel{(3.4)}{\leq} \int_T h^2 h^{-2} |\nabla w| dy + a^{-p-2} \int_T h^{-2} |\nabla w| dy^{-2} \\ & \stackrel{(ii)}{\leq} \int_T (|\nabla w| + a)^{p-2} |\nabla w|^2 dy, \end{aligned} \quad (3.5)$$

where (i) is due to the inverse inequality (cf. [Diening & Růžička \(2007, Eq. \(2.4\)\)](#))

$$|u|_{W^{1,\infty}(T)} \leq S h^{-2} |u|_{W^{1,1}(T)}, \quad u \in S_h(\Omega),$$

and (ii) is due to the Jensen inequality. By summing (3.5) over all elements $T \in \mathcal{T}_h$ we arrive at the conclusion. $\square$

Using Lemmas 3.4, 3.5 and 3.7, we obtain the following quasi-norm stable decomposition result.

LEMMA 3.8. Suppose that $p \in (1, \infty)$ and $v \in V$ satisfy either the assumptions in Lemma 3.4 or those in Lemma 3.5. Then there exists a positive constant $C_{p,v}^{\text{SD}}$ depending on p and v such that the following holds: for any $w \in V$ there exist $w_k \in V_k, 0 \leq k \leq N$, such that

$$\begin{aligned} w &= \sum_{k=0}^N R_k^* w_k, \\ \sum_{k=0}^N D_F(v + R_k^* w, v) &\leq C_{p,v}^{\text{SD}} \|w\|_{(\nabla v)}^2, \\ C_{p,v}^{\text{SD}} &\leq S \left(1 + C_{p,v}^{\text{PF}} \frac{H}{\delta} \right)^{-p}, \end{aligned}$$

where the constant $C_{p,v}^{\text{PF}}$ was given in either Lemma 3.4 or Lemma 3.5.

Proof. Throughout this proof let an index k runs from 1 to N . Take any $u, v \in V$, and let $w = u - v$. We define $w_0 \in V_0$ and $w_k \in V_k$ as

$$w_0 = \tilde{I}_H w, \quad R_k^* w = I_h(\theta_k \tilde{w}),$$

where $\tilde{I}_H$ is the Scott–Zhang quasi-interpolation operator onto $S_H(\Omega)$ ([Scott & Zhang, 1990](#)), and $\tilde{w} = w - R_0^* w_0 = \sum_{k=1}^N R_k^* w_k$. Invoking Lemma 3.3 and [Diening & Růžička \(2007, Theorem 4.5\)](#), we have

$$D_F(v + R_0^* w_0, v) \leq S \|\tilde{I}_H w\|_{(\nabla v)}^2 \leq S \|w\|_{(\nabla v)}^2. \quad (3.6)$$

Similarly, Lemmas 3.3 and 3.7 imply

$$D_F(v + R_k^* w_k, v) \leq S \sum_{k=1}^N \|I_h(\theta_k \tilde{w})\|_{(\nabla v)}^2 \leq S \sum_{k=1}^N \|\theta_k \tilde{w}\|_{(\nabla v)}^2. \quad (3.7)$$

Note that the map $x \mapsto (x + a)^{p-2}x^2$ ($x \geq 0$) is increasing for any $a \geq 0$. It follows that

$$\begin{aligned}
 \|\theta_k \tilde{w}\|_{(\nabla v)}^2 &\leq \int_{\Omega} |\theta_k |\nabla \tilde{w}| + |\tilde{w}| |\nabla \theta_k| + |\nabla v|^{p-2} |\theta_k |\nabla \tilde{w}| + |\tilde{w}| |\nabla \theta_k||^2 dx \\
 &\stackrel{(i)}{\leq} \int_{\Omega} |\theta_k |\nabla \tilde{w}| + |\nabla v|^{p-2} |\theta_k |\nabla \tilde{w}||^2 dx + \int_{\Omega} |\tilde{w}| |\nabla \theta_k| + |\nabla v|^{p-2} |\tilde{w}| |\nabla \theta_k||^2 dx \\
 &\stackrel{(ii)}{\leq} \|\tilde{w}\|_{(\nabla v)}^2 + \frac{1}{\delta^p} \int_{\Omega} (|\tilde{w}| + |\nabla v|)^{p-2} |\tilde{w}|^2 dx \\
 &\stackrel{(iii)}{\leq} \left\{ 1 + \frac{C_{\mathcal{F}}}{\delta^p} \right\} \|\tilde{w}\|_{(\nabla v)}^2, \tag{3.8}
 \end{aligned}$$

where (i) is because of the triangle inequality-like result presented in [Liu & Yan \(2001, Lemma 5.4\)](#), (ii) is due to (2.3) and (iii) is due to [Lemmas 3.4 and 3.5](#). Meanwhile, we observe that [Diening & Růžička \(2007, Theorem 4.6\)](#) implies

$$\|\tilde{w}\|_{(\nabla v)}^2 = \|\tilde{w} - \tilde{I}_H \tilde{w}\|_{(\nabla v)}^2 + \| \tilde{I}_H \tilde{w} \|_{(\nabla v)}^2. \tag{3.9}$$

Combining (3.6), (3.7), (3.8) and (3.9) yields the desired result. $\square$

REMARK 3.9. The estimate presented in [Lemma 3.8](#) is not as sharp as the one in the norm-stable decomposition result given in [Lemma 2.1](#). Specifically, in [Lemma 3.8](#), the power of H/δ is p , while in [Lemma 2.1](#), it is $p - 1$. The norm-stable decomposition achieves the sharp $(H/\delta)^{p-1}$ -result using a trace theorem-type argument introduced in [Dryja & Widlund \(1994\)](#); see also [Toselli & Widlund \(2005, Lemma 3.10\)](#). Unfortunately, we were unable to make a similar argument in our quasi-norm analysis because the quasi-norm does not have a notion of trace. To obtain a sharp estimate it will be necessary to define an appropriate trace for the quasi-norm, which is remained as a topic for future research.

3.3 Proof of Theorem 2.4

The proof of [Theorem 2.4](#) presented here uses a similar argument to [Park \(2020\)](#). However, due to the nonlinearity of the quasi-norm $\|\cdot\|_{(\nabla v)}$, we have to make a careful consideration on dealing with $\|\cdot\|_{(\nabla v)}$. In [Lemma 3.10](#) we state the generalized additive Schwarz lemma (see [Park \(2020, Lemma 4.5\)](#)) applied to [Algorithm 1](#) in a form suitable for our purposes.

LEMMA 3.10 (generalized additive Schwarz lemma). Let $\{u^{(n)}\}$ be the sequence generated by [Algorithm 1](#). Then it satisfies

$$u^{(n+1)} \in \arg \min_{u \in V} \left\{ F(u^{(n)}) + (F(u^{(n)}), u - u^{(n)}) + M_{\tau}(u, u^{(n)}) \right\},$$

where the functional $M_{\tau} : V \times V \rightarrow \mathbb{R}$ is given by

$$M_{\tau}(u, v) = \tau \inf_{k=0}^N \left\| D_F(v + R_k^* w_k, v) : u - v = \tau \sum_{k=0}^N R_k^* w_k, w_k \in V_k \right\|, \quad u, v \in V.$$

Proof. Here, we provide a simple proof that does not rely on convex analysis tools. We define a functional $Q_n : V \rightarrow \mathbb{R}$ as

$$Q_n(u) = F(u^{(n)}) + (F(u^{(n)}), u - u^{(n)}) + M_\tau(u, u^{(n)}), \quad u \in V.$$

For any $u \in V$, invoking (1.4) with some direct computation yields

$$\begin{aligned} Q_n(u) &= (1 - \tau N)F(u^{(n)}) + \tau \inf_{k=0}^N F(u^{(n)} + R_k^* w_k) : u - u^{(n)} = \tau \sum_{k=0}^N R_k^* w_k, w_k \in V_k \\ &\geq (1 - \tau N)F(u^{(n)}) + \tau \min_{k=0}^N F(u^{(n)} + R_k^* w_k) \\ &= (1 - \tau N)F(u^{(n)}) + \tau \min_{k=0}^N F(u^{(n)} + R_k^* w_k^{(n+1)}) \\ &\geq Q_n(u^{(n+1)}). \end{aligned}$$

That is, $u^{(n+1)}$ minimizes Q_n , which is our desired result. $\square$

In the following text, similar to Park (2020, Lemma 4.6), we prove that M_τ defined in Lemma 3.10 is bounded below by D_F and above by $\|\cdot\|_{(\nabla_V)}^2$ up to a multiplicative constant. We note that Lemma 3.11 can be regarded as a variant of the Lipschitz-like/convexity condition discussed in Teboulle (2018).

LEMMA 3.11. Suppose that $p \in (1, \infty)$ and $v \in V$ satisfy either the assumptions in Lemma 3.4 or those in Lemma 3.5. Then we have

$$D_F(u, v) \leq M_\tau(u, v) \leq \frac{C_{p,v}^{\text{SD}}}{\tau^{\max\{p, 2\}-1}} \|u - v\|_{(\nabla_V)}^2, \quad u \in V,$$

where $C_{p,v}^{\text{SD}}$ and M_τ were defined in Lemmas 3.10 and 3.8, respectively.

Proof. We define $u^- \in V$ by the following:

$$u^- - v = \frac{1}{\tau} (u - v).$$

By Lemma 3.8 there exist $w_k \in V_k, 0 \leq k \leq N$, such that

$$\begin{aligned} u^- - v &= \sum_{k=0}^N R_k^* w_k, \\ \sum_{k=0}^N D_F(v + R_k^* w_k, v) &\leq C_{p,v}^{\text{SD}} \|u^- - v\|_{(\nabla_V)}^2. \end{aligned}$$

Note that $u - v = \tau \sum_{k=0}^N R_k^* w_k$. It follows by Lemma 3.1 that

$$\tau \sum_{k=0}^N D_F(v + R_k^* w_k, v) \leq \tau C_{p,v}^{\text{SD}} \|u - v\|_{(\nabla v)}^2 \leq \frac{C_{p,v}^{\text{SD}}}{\tau^{\max\{p, 2\}-1}} \|u - v\|_{(\nabla v)}^2.$$

Meanwhile, invoking Lemma 2.2 yields

$$\tau \sum_{k=0}^N D_F(v + R_k^* w_k, v) = \tau \sum_{k=0}^N F(v + R_k^* w_k) - \tau(N+1)F(v) - (F^t(v), u - v) \geq D_F(u, v),$$

which completes the proof. $\square$

By closely following the argument in (Park, 2020, Appendix A.4) and manipulating $\|\cdot\|_{(\nabla v)}$ -terms using the properties of $\|\cdot\|_{(\nabla v)}$ presented in Lemmas 3.1 to 3.3 we establish the following lemma, which provides an estimate for the ratio of two consecutive energy errors in Algorithm 1.

LEMMA 3.12. In Algorithm 1 suppose that $v = u^{(n)}$ satisfy either the assumptions in Lemma 3.4 or those in Lemma 3.5 for some $n \geq 0$. Then we have

$$\frac{F(u^{(n+1)}) - F(u^*)}{F(u^{(n)}) - F(u^*)} \leq 1 - \frac{1}{p} \frac{p^{\frac{1}{p-1}} \mu_p}{p^{2\bar{p}} C_{p,u^{(n)}}^{\text{SD}}},$$

where $\underline{p} = \min\{p, 2\}$, $\bar{p} = \max\{p, 2\}$, $p^* = |p - 2|$ and μ_p and $C_{p,u^{(n)}}^{\text{SD}}$ were given in Lemmas 3.3 and 3.8, respectively.

Proof. For $t \in [0, 1]$ we write

$$u_t = u^{(n)} + t(u^* - u^{(n)}). \quad (3.10)$$

Then we have

$$\begin{aligned} F(u^{(n+1)}) &= F(u^{(n)}) + (F^t(u^{(n)}), u^{(n+1)} - u^{(n)}) + D_F(u^{(n+1)}, u^{(n)}) \\ &\stackrel{(i)}{\leq} F(u^{(n)}) + (F^t(u^{(n)}), u^{(n+1)} - u^{(n)}) + M_\tau(u^{(n+1)}, u^{(n)}) \\ &\stackrel{(ii)}{\leq} \min_{u \in V} F(u^{(n)}) + (F^t(u^{(n)}), u - u^{(n)}) + M_\tau(u, u^{(n)}) \\ &\stackrel{(i)}{\leq} \min_{t \geq 0} F(u^{(n)} + t(u^* - u^{(n)})) + \frac{C_{p,u^{(n)}}^{\text{SD}}}{t^{\bar{p}-1}} \|u - u^{(n)}\|_{(\nabla u^{(n)})}^2 \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{(iii)}}{\leq} \min_{t \geq 0} \left\{ F(u^{(n)}) + t(F'(u^{(n)}), u^* - u^{(n)}) + \frac{C_{p,u^{(n)}}^{\text{SD}}}{\tau^{\bar{p}-1}} t^{\bar{p}} \|u^* - u^{(n)}\|_{(\nabla u^{(n)})}^2 \right\} \\
& \stackrel{\text{(iv)}}{\leq} \min_{t \geq 0} \left\{ F(u^{(n)}) + t(F'(u^{(n)}), u^* - u^{(n)}) + \frac{2^{\hat{p}} C_{p,u^{(n)}}^{\text{SD}}}{\tau^{\bar{p}-1} \mu_p} t^{\bar{p}} \|u^{(n)} - u^*\|^2 \right\} \\
& \stackrel{\text{(v)}}{\leq} \min_{t \geq 0} \left\{ F(u^{(n)}) + t(F'(u^{(n)}), u^* - u^{(n)}) + \frac{2^{\hat{p}} C_{p,u^{(n)}}^{\text{SD}}}{\tau^{\bar{p}-1} \mu_p} t^{\bar{p}} (F(u^{(n)}) - F(u^*)) \right\},
\end{aligned}$$

where (i)–(v) are due to Lemmas 3.11, 3.10, 3.1, 3.2 and 3.3, respectively. By the convexity of F we get

$$\begin{aligned}
F(u^{(n+1)}) - F(u^*) & \leq \min_{t \geq 0} \left\{ 1 - t + \frac{2^{\hat{p}} C_{p,u^{(n)}}^{\text{SD}}}{\tau^{\bar{p}-1} \mu_p} t^{\bar{p}} \right\} (F(u^{(n)}) - F(u^*)) \\
& = \left[1 - \frac{1}{\bar{p}} \frac{\tau^{\bar{p}-1} \mu_p}{2^{\hat{p}} C_{p,u^{(n)}}^{\text{SD}}} \right] (F(u^{(n)}) - F(u^*)),
\end{aligned}$$

which completes the proof. $\square$

Finally, we are ready to present our proof of Theorem 2.4.

Proof of Theorem 2.4. Suppose that $|\nabla u^*|$ does not vanish on Ω . Take any $E > 0$. By the continuity of $\bar{C}_{p,u^*}^{\text{PF}}$ stated in Lemmas 3.4 and 3.5 we can find a neighbourhood B_E of u^* in V such that, for any $v \in B_E$, $|\nabla v|$ does not vanish on Ω and

$$\bar{C}_{p,v}^{\text{PF}} \leq \bar{C}_{p,u^*}^{\text{PF}} + E. \quad (3.11)$$

Meanwhile, by Proposition 2.3 and (1.3) (see Park (2020, Section 6.1) for the precise statement of (1.3)) we deduce that the sequence $\{u^{(n)}\}$ converges to u^* in V . Hence, there exists n_0 such that, if $n \geq n_0$, then $u^{(n)} \in B_E$. By Lemma 3.8, for $n \geq n_0$, we have

$$C_{p,u^{(n)}}^{\text{SD}} \leq 1 + \bar{C}_{p,u^{(n)}}^{\text{PF}} \left(\frac{H}{\delta} \right)^{-p} \leq 1 + \bar{C}_{p,u^{(n)}}^{\text{PF}} \left(\frac{H}{\delta} \right)^{-p} (3.11) \leq 1 + (\bar{C}_{p,u^*}^{\text{PF}} + E) \left(\frac{H}{\delta} \right)^{-p}.$$

It follows by Lemma 3.12 that

$$\begin{aligned}
1 - \frac{F(u^{(n+1)}) - F(u^*)}{F(u^{(n)}) - F(u^*)} & \leq \frac{p}{\bar{p}-1} \left(\frac{2^{\hat{p}} C_{p,u^{(n)}}^{\text{SD}}}{\tau^{\bar{p}-1} \mu_p} \right)^{\frac{1}{\bar{p}-1}} \\
& \leq \frac{p}{\bar{p}-1} \left\{ \frac{2^{\hat{p}}}{\tau^{\bar{p}-1} \mu_p} \left(\frac{H}{\delta} \right)^{-p} + 1 + (\bar{C}_{p,u^*}^{\text{PF}} + E) \left(\frac{H}{\delta} \right)^{-p} \right\}^{\frac{1}{\bar{p}-1}}.
\end{aligned}$$

Since E is arbitrary we obtain the desired result with

$$\gamma = \frac{p}{p-1} \left\{ \frac{p2^{p^*}}{p^{p^*-1}\mu_p} - 1 + \bar{C}_{p,u^*}^{\text{PF}} \left(\frac{H}{\delta} \right)^{-p^*, \frac{1}{p^*-1}} \right\}.$$

This completes the proof. $\square$

REMARK 3.13. As stated in Lemma 3.3 the squared quasi-norm $\|u - v\|_{(\nabla v)}^2$ is equivalent to the Bregman distance $D_F(u, v)$ up to multiplicative constants for $u, v \in W^{1,p}(\Omega)$. This equivalence allows us to perform the convergence analysis presented in this section using the Bregman distance instead of the quasi-norm. Indeed, the Bregman distance is frequently utilized in the analysis of convex optimization algorithms; see, e.g., Bauschke *et al.* (2017); Teboulle (2018). Nevertheless, we opted to present the convergence analysis of Algorithm 1 in terms of the quasi-norm in this paper, because using the quasi-norm has an advantage that we can simplify our proof by borrowing some useful techniques regarding the quasi-norm introduced in from the existing literature (Liu & Yan, 2001, 2002; Ebmeyer & Liu, 2005; Carstensen *et al.*, 2006).

4. Quasi-norm Poincaré–Friedrichs inequality

This section is devoted to the proofs of Lemmas 3.4 and 3.5. Namely, we deal with quasi-norm Poincaré–Friedrichs inequalities of the form

$$\int_{\Omega} (|w| + |\nabla v|)^{p-2} |w|^2 dx \leq C \|w\|_{(\nabla v)}^2, \quad (4.1)$$

where $p \in (1, \infty)$ with $p \neq 2$. Unfortunately, the inequality (4.1) does not hold for every $v, w \in W^{1,p}(\Omega)$; see Examples 4.8 and 4.15. Based on a quasi-monotonicity argument introduced in Pechstein & Scheichl (2013) we characterize the conditions on v such that the inequality (4.1) holds and provide a precise estimate for the Poincaré–Friedrichs constant C in (4.1). Throughout this section let $W_{\Gamma}^{1,p}(\Omega)$ denote the collection of all $W^{1,p}(\Omega)$ -functions vanishing on $\Gamma \subset \partial\Omega$. In addition we use the conventions $0/0 = 1$ and $1/0 = \infty$.

We first observe that a particular case of (4.1), when $|\nabla v|$ is constant on Ω , is valid. By the same argument as in Liu & Yan (2002, Lemma 3.1) and Carstensen *et al.* (2006, Lemma. 4.1) we can prove the following lemma.

LEMMA 4.1. Let $\Gamma \subset \partial\Omega$ have nonvanishing one-dimensional measure. Then, there exists a positive constant C such that

$$\int_{\Omega} (|w| + a)^{p-2} |w|^2 dx \leq C \int_{\Omega} (|\nabla w| + a)^{p-2} |\nabla w|^2 dx, \quad w \in W_{\Gamma}^{1,p}(\Omega), a \geq 0.$$

For a non-negative function $\alpha \in L^\infty(\Omega)$ and a partition $Y = \{Y_l\}_{l=1}^m$ of Ω consisting of nonoverlapping polygonal regions we define two non-negative functions $\underline{\alpha}^Y, \bar{\alpha}^Y \in L^\infty(\Omega)$ as follows:

$$\underline{\alpha}^Y(x) = \operatorname{ess\,inf}_{Y_l} \alpha, \quad \bar{\alpha}^Y(x) = \operatorname{ess\,sup}_{Y_l} \alpha, \quad x \in Y_l, 1 \leq l \leq m.$$

In the following text we address the cases $p \in (2, \infty)$ and $p \in (1, 2)$ separately. We first focus on the case $p \in (2, \infty)$. In Definition 4.2 we introduce the concept of *quasi-monotone increase*. We note that relevant notions were explored in Galvis & Efendiev (2010); Pechstein & Scheichl (2013).

DEFINITION 4.2. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω , and let $Y = \{Y_l\}_{l=1}^m$ denote a partition of Ω into nonoverlapping polygonal regions.

1. We say that the region $P_{l_1, l_s} = (\bar{Y}_{l_1} \cup \cdots \cup \bar{Y}_{l_s})^\circ$, $1 \leq l_1, \dots, l_s \leq m$, is a quasi-monotonically increasing path from Y_{l_1} to Y_{l_s} with respect to α if the following two conditions hold:
 - a. For each $1 \leq i \leq s-1$ the regions $\bar{Y}_{l_i}$ and $\bar{Y}_{l_{i+1}}$ share a common edge.
 - b. $\underline{\alpha}^Y(Y_{l_1}) \leq \cdots \leq \underline{\alpha}^Y(Y_{l_s})$.
2. We say that α is $\partial\Omega$ -quasi-monotonically increasing on Y if, for any $1 \leq l \leq m$, there exist an index l^* and a quasi-monotonically increasing path P_{l, l^*} from Y_l to Y_{l^*} , such that $\partial Y_{l^*} \cap \partial\Omega$ has nonvanishing one-dimensional measure.

By a similar argument, as in the proof of Pechstein & Scheichl (2013, Theorem 2.9), we prove the following lemma.

LEMMA 4.3. Assume that $p \in (2, \infty)$. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω , and let $Y = \{Y_l\}_{l=1}^m$ denote a partition of Ω into nonoverlapping polygonal regions. If α is $\partial\Omega$ -quasi-monotonically increasing on Y then, for each $1 \leq l \leq m$, there exists a positive constant $c_{p, Y, l}$ independent of α , such that

$$\int_{Y_l} (|w| + \alpha)^{p-2} |w|^2 \, dx \leq c_{p, Y, l} \frac{\int_{P_{l, l^*}} (\bar{\alpha}^Y)^{p-2} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 \, dx}{\int_{Y_l} \underline{\alpha}^Y \, dx}, \quad w \in W^{1,p}(\Omega),$$

where the region P_{l, l^*} was given in Definition 4.2.

Proof. Note that the map $x \mapsto x^{p-2}$ ($x \geq 0$) is increasing. Take any l such that $1 \leq l \leq m$. Since $\alpha \leq \bar{\alpha}^Y(Y_l)$ on Y_l and $Y_l \subset P_{l, l^*}$ we get

$$\begin{aligned} \int_{Y_l} (|w| + \alpha)^{p-2} |w|^2 \, dx &\leq \int_{Y_l} (|w| + \bar{\alpha}^Y(Y_l))^{p-2} |w|^2 \, dx \\ &\leq \int_{P_{l, l^*}} (|w| + \bar{\alpha}^Y(Y_l))^{p-2} |w|^2 \, dx. \end{aligned} \quad (4.2)$$

Since $\partial P_{l,l^*} \cap \partial \Omega$ has nonvanishing one-dimensional measure Lemma 4.1 ensures that we have a positive constant c_{p,Y,l^*} independent of w and α , such that

$$\int_{P_{l,l^*}} (|w| + \alpha Y(Y_l))^{p-2} |w|^2 dx \leq c_{p,Y,l^*} \int_{P_{l,l^*}} (|\nabla w| + \bar{\alpha} Y(Y_l))^{p-2} |\nabla w|^2 dx. \quad (4.3)$$

Invoking the inequality

$$\frac{a + \bar{\alpha} Y(Y_l)}{a + \underline{\alpha} Y(Y_l)} \leq \frac{\bar{\alpha} Y(Y_l)}{\underline{\alpha} Y(Y_l)}, \quad a \geq 0,$$

we have

$$\begin{aligned} & \int_{P_{l,l^*}} (|\nabla w| + \bar{\alpha} Y(Y_l))^{p-2} |\nabla w|^2 dx \\ & \leq \frac{\{\bar{\alpha} Y(Y_l)\}^{p-2}}{\{\underline{\alpha} Y(Y_l)\}^{p-2}} \int_{P_{l,l^*}} (|\nabla w| + \alpha Y(Y_l))^{p-2} |\nabla w|^2 dx \\ & \leq \frac{\bar{\alpha} Y(Y_l)}{\{\bar{\alpha} Y(Y_l)\}^{p-2}} \int_{P_{l,l^*}} (|\nabla w| + \alpha Y(Y_l))^{p-2} |\nabla w|^2 dx \\ & \leq \frac{\{\bar{\alpha} Y(Y_l)\}^{p-2}}{\{\underline{\alpha} Y(Y_l)\}^{p-2}} \int_{P_{l,l^*}} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx, \end{aligned} \quad (4.4)$$

where the penultimate inequality is because $\underline{\alpha} Y$ increases along P_{l,l^*} . Combining Lemmas 4.2, 4.3 and 4.4 yields the desired result. $\square$

Due to Lemma 4.3 we are able to define the *quasi-monotone increase constant* $C_{p,\alpha}^{\text{QM}}$ for $p \in (2, \infty)$, as presented in Definition 4.4.

DEFINITION 4.4. Assume that $p \in (2, \infty)$. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω . The quasi-monotone increase constant $C_{p,\alpha}^{\text{QM}} \in [0, \infty]$ is defined by

$$C_{p,\alpha}^{\text{QM}} = \inf_{p,\alpha} \max_{Y \mid 1 \leq l \leq m} \left\{ \frac{\{\bar{\alpha} Y(Y_l)\}^{p-2}}{\{\underline{\alpha} Y(Y_l)\}^{p-2}} \cdot \frac{1}{c_{p,Y,l^*}} \right\},$$

where the constants c_{p,Y,l^*} 's were given in Lemma 4.3 and the infimum is taken over every nonoverlapping polygonal partition $Y = \{Y_l\}_{l=1}^m$ of Ω such that α is $\partial\Omega$ -quasi-monotonically increasing on Y .

Note that the infimum in Definition 4.2 is well-defined because α is $\partial\Omega$ -quasi-monotonically increasing on the trivial partition $\{\Omega\}$. In terms of the quasi-monotone increase constant $C_{p,\alpha}^{\text{QM}}$ we present a quasi-norm Poincaré–Friedrichs inequality for $p \in (2, \infty)$ in Theorem 4.5.

THEOREM 4.5. Assume that $p \in (2, \infty)$. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω . Then we have

$$\int_{\Omega} (|w| + \alpha)^{p-2} |w|^2 dx \leq C_{p,\alpha}^{\text{QM}} \int_{\Omega} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx, \quad w \in W_0^{1,p}(\Omega),$$

where the quasi-monotone increase constant $C_{p,\alpha}^{\text{QM}} \in [0, \infty]$ was given in Definition 4.4.

Proof. We fix any nonoverlapping polygonal partition $\mathcal{Y}$ such that α is $\partial\Omega$ -quasi-monotonically increasing on $\mathcal{Y}$. By Lemma 4.3, we have

$$\begin{aligned} \int_{Y_l} (|w| + \alpha)^{p-2} |w|^2 dx &\leq c_{p,Y,l} \frac{\alpha_{Y(Y_l)}^{p-2}}{\underline{\alpha}_{Y(Y_l)}} \int_{Y_l} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx \\ &\leq c_{p,Y,l} \max_{1 \leq l \leq m} \frac{\alpha_{Y(Y_l)}^{p-2}}{\underline{\alpha}_{Y(Y_l)}} \int_{\Omega} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx. \end{aligned} \quad (4.5)$$

Summing (4.5) over all l followed by taking the infimum over all $\mathcal{Y}$ completes the proof. $\square$

Let $W_h(\Omega)$ be the space of piecewise constant functions on the triangulation $\mathcal{T}_h$. Under an additional assumption that $\alpha \in W_h(\Omega)$ we can characterize the condition when the quasi-monotone increase constant $C_{p,\alpha}^{\text{QM}}$ is finite.

LEMMA 4.6. Assume that $p \in (2, \infty)$. Let $\alpha \in W_h(\Omega)$ be a non-negative piecewise constant function on $\mathcal{T}_h$. Then, the quasi-monotone increase constant $C_{p,\alpha}^{\text{QM}}$ is finite if and only if every maximal polygonal region $R \subset \Omega$ with $\alpha > 0$ satisfies that $\partial R \cap \partial\Omega$ has nonvanishing one-dimensional measure.

Proof. We first assume that every maximal polygonal region $R \subset \Omega$ with $\alpha > 0$ satisfies that $\partial R \cap \partial\Omega$ has nonvanishing one-dimensional measure. We consider the partition $\mathcal{Y}^*$ of Ω consisting of all maximal polygonal regions $\{R_i\}$ with $\alpha > 0$ and all maximal polygonal regions $\{S_i\}$ with $\alpha = 0$. It is obvious that each R_i forms a quasi-monotonically increasing path with respect to α from R_i to itself. For each S_i , if $\partial S_i \cap \partial\Omega$ has nonvanishing one-dimensional measure, then S_i forms a quasi-monotonically increasing path with respect to α from S_i to itself. Otherwise, the maximality of S_i implies there exists some R_j such that $\bar{S}_i$ and $\bar{R}_j$ share a common edge. Then we readily deduce that $(\bar{S}_i \cup \bar{R}_j)^\circ$ forms a quasi-monotonically increasing path with respect to α from S_i to R_j . Meanwhile, since α is piecewise constant, we have

$$\frac{\alpha_{Y^*}(R_i)}{\underline{\alpha}_{Y^*}(R_i)} < \infty, \quad \frac{\alpha_{Y^*}(S_i)}{\underline{\alpha}_{Y^*}(S_i)} = \frac{0}{0} = 1,$$

for every R and S . Hence, we conclude that $C_{p,\alpha}^{\text{QM}} < \infty$.

Next, we suppose that there exists a maximal polygonal region $R^* \subset \Omega$ with $\alpha > 0$ such that $\partial R^* \cap \partial\Omega$ is a null set. That is, every edge of R^* is shared with a region with $\alpha = 0$. Take any nonoverlapping polygonal partition $\mathcal{Y}$ of Ω . If $\mathcal{Y}$ has an element Y such that $Y \subset R^*$ then it is impossible to find any quasi-monotonically increasing path with respect to α starting from Y , since any such path would necessarily have to pass through a region where $\alpha = 0$. Otherwise, $\mathcal{Y}$ must contain an element Y such

that both $Y \cap R^*$ and $Y \setminus R^*$ are nontrivial, which implies that

$$\frac{\alpha^Y(Y)}{\underline{\alpha}^Y(Y)} = \infty.$$

Hence, we conclude that $C_{p,\alpha}^{\text{QM}} = \infty$, which completes the proof. $\square$

Combining Theorem 4.5 and Lemma 4.6 yields Corollary 4.7, in which Lemma 3.4 is a particular case $\alpha = |\nabla v|$ of this result.

COROLLARY 4.7. Assume that $p \in (2, \infty)$. Let $\alpha \in W_h(\Omega)$ be a non-negative piecewise constant function on T_h . If every maximal polygonal region $R \subset \Omega$ with $\alpha > 0$ satisfies that $\partial R \cap \partial\Omega$ has nonvanishing one-dimensional measure, then we have

$$\int_{\Omega} (|w| + \alpha)^{p-2} |w|^2 dx \leq C_{p,\alpha}^{\text{QM}} \int_{\Omega} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx, \quad w \in W_0^{1,p}(\Omega),$$

where $C_{p,\alpha}^{\text{QM}}$ is a finite constant given in Definition 4.4. Moreover, if α does not vanish on Ω , then $C_{p,\alpha}^{\text{QM}}$ has an upper bound $\epsilon_{p,\alpha}^{\text{QM}}$ that is continuous at α in $W_h(\Omega)$.

Proof. It suffices to find a continuous upper bound of $C_{p,\alpha}^{\text{QM}}$ under the given condition. We assume that α does not vanish on Ω . By Definition 4.4, we have

$$C_{p,\alpha}^{\text{QM}} \leq \frac{\max_{\Omega} \alpha}{\min_{\Omega} \alpha} c_{p,\{\Omega\},1} =: \epsilon_{p,\alpha}^{\text{QM}},$$

where the inequality is obtained by taking $Y = \{\Omega\} = \{Y_1\}$ in Definition 4.4. As $\alpha > 0$ in Ω it is clear that $C_{p,\alpha}^{\text{QM}}$ is continuous at α in $W_h(\Omega)$, which completes the proof. $\square$

We show that, under the condition presented in Lemma 4.6 for the quasi-monotone increase constant $C_{p,\alpha}$ to be infinite, the quasi-norm Poincaré–Friedrichs inequality of the form (4.1) is not valid. For simplicity we provide a counterexample in one-dimension; we note that the construction can be generalized to higher dimensions.

EXAMPLE 4.8. Let $p \in (2, \infty)$ and $\Omega = (0, 1) \subset \mathbb{R}$. We define $w \in W_0^{1,p}(\Omega)$ and $\alpha \in L^\infty(\Omega)$ as

$$w(x) = \begin{cases} 3x, & \text{if } 0 < x < \frac{1}{3}, \\ 1 & \text{if } \frac{1}{3} \leq x < \frac{2}{3}, \\ -3x + 3, & \text{if } \frac{2}{3} \leq x < 1, \end{cases} \quad \alpha(x) = \begin{cases} 1, & \text{if } \frac{1}{3} \leq x < \frac{2}{3}, \\ 0, & \text{otherwise.} \end{cases}$$

We observe that the quasi-monotone increase constant $C_{p,\alpha}$ becomes infinite because the interval $(1/3, 2/3)$ where α is nonzero does not touch $\partial\Omega$. For any $E > 0$ direct calculation yields

$$\frac{\int_0^1 (|Ew| + \alpha)^{p-2} |Ew|^2 dx}{\int_0^1 (|Ew| + \alpha)^{p-2} |Ew|^2 dx} \rightarrow 0 \quad \text{as } E \rightarrow 0^+,$$

which implies that (4.1) does not hold.

We now turn to the case $p \in (1, 2)$. In contrast to the case $p \in (2, \infty)$, which heavily relies on the quasi-monotone increase of α , the analysis of the case $p \in (1, 2)$ hinges on the *quasi-monotone decrease* of α ; see Definition 4.9.

DEFINITION 4.9. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω , and let $Y = \{Y_l\}_{l=1}^m$ denote a partition of Ω into nonoverlapping polygonal regions.

1. We say that the region $P_{l_1, l_s} = (\bar{Y}_{l_1} \cup \cdots \cup \bar{Y}_{l_s})^\circ$, $1 \leq l_1, \dots, l_s \leq m$, is a quasi-monotonically decreasing path from Y_{l_1} to Y_{l_s} with respect to α if the following two conditions hold:
 - a. For each $1 \leq i \leq s-1$ the regions $\bar{Y}_{l_i}$ and $\bar{Y}_{l_{i+1}}$ share a common edge.
 - b. $\alpha_Y(Y_{l_1}) \geq \cdots \geq \alpha_Y(Y_{l_s})$.
2. We say that α is $\partial\Omega$ -quasi-monotonically decreasing on Y if, for any $1 \leq l \leq m$, there exist an index l^* and a quasi-monotonically decreasing path P_{l, l^*} from Y_l to Y_{l^*} , such that $\partial Y_{l^*} \cap \partial\Omega$ has nonvanishing one-dimensional measure.

One can prove the following lemma using the fact that the map $x \mapsto x^{p-2}$ ($x \geq 0$) is decreasing and by following a similar argument to that used in the proof of Lemma 4.3.

LEMMA 4.10. Assume that $p \in (1, 2)$. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω , and let $Y = \{Y_l\}_{l=1}^m$ denote a partition of Ω into nonoverlapping polygonal regions. If α is $\partial\Omega$ -quasi-monotonically decreasing on Y then, for each $1 \leq l \leq m$, there exists a positive constant $c_{p, Y, l}$ independent of α , such that

$$\int_{Y_l} (|w| + \alpha)^{p-2} |w|^2 dx \leq c_{p, Y, l} \int_{P_{l, l^*}} \frac{\alpha_Y(Y_l)}{\bar{\alpha}_Y(Y_l)} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx, \quad w \in W_0^{1,p}(\Omega),$$

where the region P_{l, l^*} was given in Definition 4.9.

Similar to Definition 4.4 we present the definition of the *quasi-monotone decrease constant* $C_{p, \alpha}^{\text{QM}}$ for $p \in (1, 2)$ in the following text.

DEFINITION 4.11. Assume that $p \in (1, 2)$. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω . The quasi-monotone decrease constant $C_{p, \alpha}^{\text{QM}} \in [0, \infty]$ is defined by

$$C_{p, \alpha}^{\text{QM}} = \inf_{p, \alpha} \max_{Y} \left\{ \max_{1 \leq l \leq m} \left\{ \frac{\alpha_Y(Y_l)}{\bar{\alpha}_Y(Y_l)} \right\}^{p-2} \cdot \min_{l=1}^m c_{p, Y, l} \right\},$$

where the constants $c_{p, Y, l}$'s were given in Lemma 4.10 and the infimum is taken over every nonoverlapping polygonal partition Y of Ω such that α is $\partial\Omega$ -quasi-monotonically decreasing on Y .

In terms of the quasi-monotone decrease constant $C_{p, \alpha}^{\text{QM}}$, we present a quasi-norm Poincaré–Friedrichs inequality for $p \in (1, 2)$ in Theorem 4.12, which can be proven in a similar manner to Theorem 4.5.

THEOREM 4.12. Assume that $p \in (1, 2)$. Let $\alpha \in L^\infty(\Omega)$ be a non-negative function on Ω . Then we have

$$\int_{\Omega} (|w| + \alpha)^{p-2} |w|^2 dx \leq C_{p, \alpha}^{\text{QM}} \int_{\Omega} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx, \quad w \in W_0^{1,p}(\Omega),$$

where the quasi-monotone decrease constant $C_{p,\alpha}^{\text{QM}} \in [0, \infty]$ was given in Definition 4.11.

The following lemma characterizes the condition when the quasi-monotone decrease constant $C_{p,\alpha}$ is finite, under an additional assumption that $\alpha \in W_h(\Omega)$, i.e., α is a non-negative piecewise constant function on the triangulation T_h .

LEMMA 4.13. Assume that $p \in (1, 2)$. Let $\alpha \in W_h(\Omega)$ be a non-negative piecewise constant function on T_h . Then, the quasi-monotone decrease constant $C_{p,\alpha}^{\text{QM}}$ is finite if and only if every maximal polygonal region $S \subset \Omega$ with $\alpha = 0$ satisfies that $\partial S \cap \partial\Omega$ has nonvanishing one-dimensional measure.

Proof. The proof is analogous to that of Lemma 4.6. $\square$

We obtain Corollary 4.14, in which Lemma 3.5 is a particular case $\alpha = |\nabla u|$, as a direct consequence of Theorem 4.12 and Lemma 4.13.

COROLLARY 4.14. Assume that $p \in (1, 2)$. Let $\alpha \in W_h(\Omega)$ be a non-negative piecewise constant function on T_h . If every maximal polygonal region $S \subset \Omega$ with $\alpha = 0$ satisfies that $\partial S \cap \partial\Omega$ has nonvanishing one-dimensional measure, then we have

$$\int_{\Omega} (|w| + \alpha)^{p-2} |w|^2 dx \leq C_{p,\alpha}^{\text{QM}} \int_{\Omega} (|\nabla w| + \alpha)^{p-2} |\nabla w|^2 dx, \quad w \in W_0^{1,p}(\Omega),$$

where $C_{p,\alpha}^{\text{QM}}$ is a finite constant given in Definition 4.11. Moreover, if α does not vanish on Ω then $C_{p,\alpha}^{\text{QM}}$ has an upper bound $C_{p,\alpha}^{\text{QM}}$ that is continuous at α in $W_h(\Omega)$.

Finally, we present a counterexample of the quasi-norm Poincaré–Friedrichs inequality (4.1) under the condition presented in Lemma 4.13 for the quasi-monotone decrease constant $C_{p,\alpha}^{\text{QM}}$ to be infinite.

EXAMPLE 4.15. Let $p \in (1, 2)$ and $\Omega = (0, 1) \subset \mathbb{R}$. We define $w \in W^{1,p}(\Omega)$ and $\alpha \in L^\infty(\Omega)$ as

$$w(x) = \begin{cases} 3x, & \text{if } 0 < x < \frac{1}{3}, \\ 1 & \text{if } \frac{1}{3} \leq x < \frac{2}{3}, \\ -3x + 3, & \text{if } \frac{2}{3} \leq x < 1, \end{cases} \quad \alpha(x) = \begin{cases} 0, & \text{if } \frac{1}{3} \leq x < \frac{2}{3}, \\ 1, & \text{otherwise.} \end{cases}$$

We observe that the quasi-monotone increase constant $C_{p,\alpha}^{\text{QM}}$ becomes infinite because the interval $(1/3, 2/3)$ where α vanishes does not touch $\partial\Omega$. For $E > 0$ direct calculation yields

$$\frac{\int_0^1 (|Ew| + \alpha)^{p-2} |Ew|^2 dx}{\int_0^1 (|Ew| + \alpha)^{p-2} |Ew|^2 dx} \rightarrow 0 \quad \text{as } E \rightarrow 0^+,$$

which implies that (4.1) does not hold.

5. Numerical experiments

In this section we present numerical results of the two-level additive Schwarz method for the p -Laplacian, which support our theoretical findings. All the algorithms were implemented in MATLAB R2022b. They

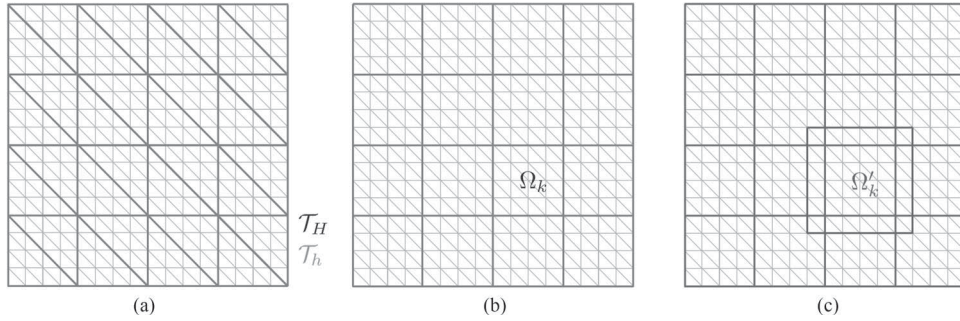


FIG. 1. Discretization and domain decomposition settings when $h = 1/2^4$, $H = 1/2^2$ and $\delta = h$. (a) Coarse triangulation T_H and fine triangulation T_h . (b) Nonoverlapping domain decomposition $\{\Omega_k\}_{k=1}^N$. (c) Overlapping domain decomposition $\{\Omega_k^t\}_{k=1}^N$

were executed on a desktop equipped with AMD Ryzen 5 5600X CPU (3.7GHz, 6C), 40GB RAM and the operating system Windows 10 Pro.

In the model p -Laplacian problem (1.1) we set $p \in \{1.05, 1.1, 1.5, 5, 10, 20\}$, $\Omega = [0, 1]^2 \subset \mathbb{R}^2$ and $f = 1$. The domain Ω is partitioned into $2 \times 1/H \times 1/H$ uniform triangles to form a coarse triangulation T_H of Ω . We further refine T_H to obtain a fine triangulation T_h , which consists of total $2 \times 1/h \times 1/h$ uniform triangles. Each subdomain Ω_k , $1 \leq k \leq N$ ($N = 1/H \times 1/H$) is defined by a rectangular region consisting of two coarse triangles sharing a diagonal edge. Then we extend Ω_k by adding its surrounding layers of fine triangles in T_h with the width δ to construct Ω_k^t , so that $\{\Omega_k^t\}_{k=1}^N$ becomes an overlapping domain decomposition for Ω . If $\delta \in (0, H/2)$ then $\{\Omega_k^t\}_{k=1}^N$ can be coloured with four colours in the way described in Lemma 2.2. The discretization and domain decomposition settings described above are illustrated in Fig. 1.

In Algorithm 1 we set $u^{(0)} = 0$ and $\tau = \tau_0 = 1/5$. Local problems defined on V_k , $1 \leq k \leq N$, and coarse problems defined on V_0 are solved by the adaptive Newton method proposed in Mishchenko (2023, Algorithm 2.1). We use the stop criterion

$$\frac{1}{1} \frac{F_k(w_k^{(n+1)}) - F_k(w_k^{(n)})}{F_k(w_k^{(n+1)})} < 10^{-12}$$

for both local and coarse problems, where F_k represents the energy functional corresponding to the local or coarse problems on V_k .

REMARK 5.1. As alternatives to the adaptive Newton method used in this paper, which is a second-order optimization algorithm, first-order optimization algorithms (Teboulle, 2018) can be adopted to solve the local and coarse problems. These algorithms are generally easier to implement as they do not require the Hessian information of the energy functional, but known to converge slower than second-order algorithms. To accelerate the convergence rate of a first-order algorithm several techniques such as the FISTA momentum (Beck & Teboulle, 2009), restart scheme (O'Donoghue & Candes, 2015) and backtracking (Scheinberg et al., 2014) can be employed.

A reference solution $u^* \in V$ for each p and h is computed by sufficiently many iterations of the adaptive Newton method applied to the full-dimension problem (2.1). The computed reference solutions

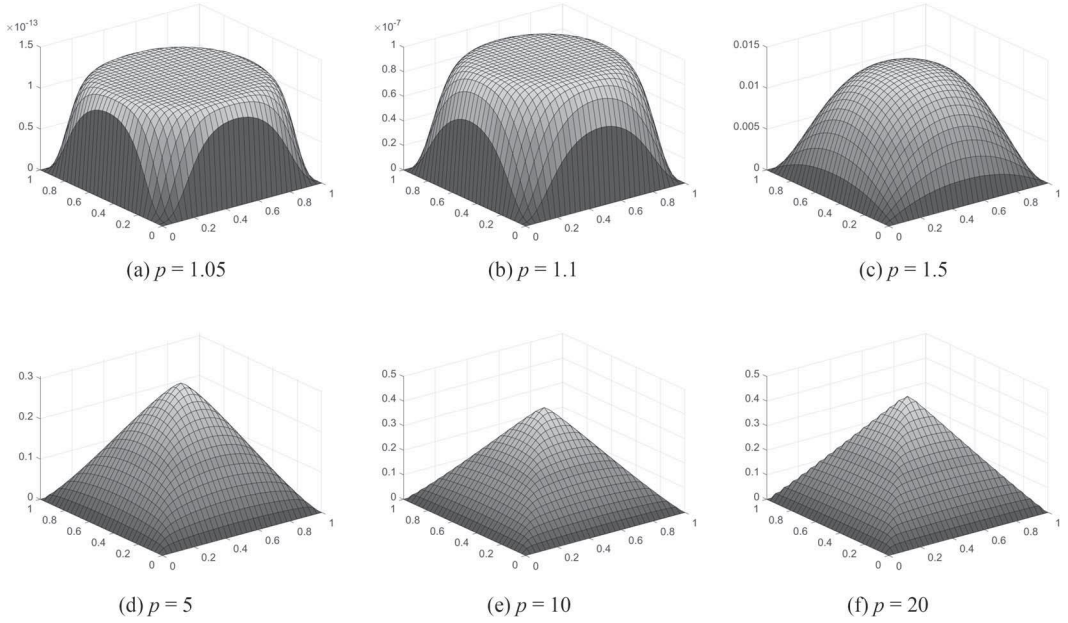


FIG. 2. Reference solutions of the p -Laplacian problem (2.1) ($p \in \{1.05, 1.1, 1.5, 5, 10, 20\}$) computed by the adaptive Newton method (Mishchenko, 2023) ($h = 2^{-5}$).

for $p \in \{1.05, 1.1, 1.5, 5, 10, 20\}$ are plotted in Fig. 2. One can observe that for cases where p is close to 1 the reference solutions exhibit flat regions where the gradient vanishes. This observation implies that when p is close to 1 the assumption in Theorem 2.4 that ∇u^* does not vanish may not hold. On the other hand, for cases where p is large, the reference solutions display peaks, leading to singular behaviour in the solution.

In Figs 3, 4, 5, 6, 7, 8, we depict the relative energy errors

$$\frac{F(u^{(n)}) - F(u^*)}{F(u^{(0)}) - F(u^*)} \quad (5.1)$$

of Algorithm 1 under various settings on p , h , H and δ . More precisely, in Figs 5 and 6 we choose p as moderate values 1.5 and 5, in Figs 3 and 4 p is chosen very close to 1 ($p = 1.05, 1.1$) and in Figs 7 and 8 p is chosen large ($p = 10, 20$). In all figures h and H vary such that $H/h = 2^3$, and δ is chosen as $\delta \in \{2^0 h, 2^1 h, 2^2 h\}$.

In every case we observe that the convergence curve of the relative energy error with respect to the number of iterations n appears linear in the x -linear y -log scale plot when n is large enough, consistent with our theoretical result presented in Theorem 2.4. It is noteworthy that even in cases where p is very close to 1 (see Figs 3 and 4), where Theorem 2.4 cannot be applied due to the flat region in the solution u^* as shown in Fig. 2(a, b), the convergence curve still appears linear. However, a theoretical explanation for the linear convergence in these cases is currently lacking.

On the other hand, for each p , we observe that the asymptotic convergence rate of Algorithm 1 shown in Figs 3, 4, 5, 6, 7, 8 remains bounded when h decreases keeping H/δ constant. This behaviour aligns

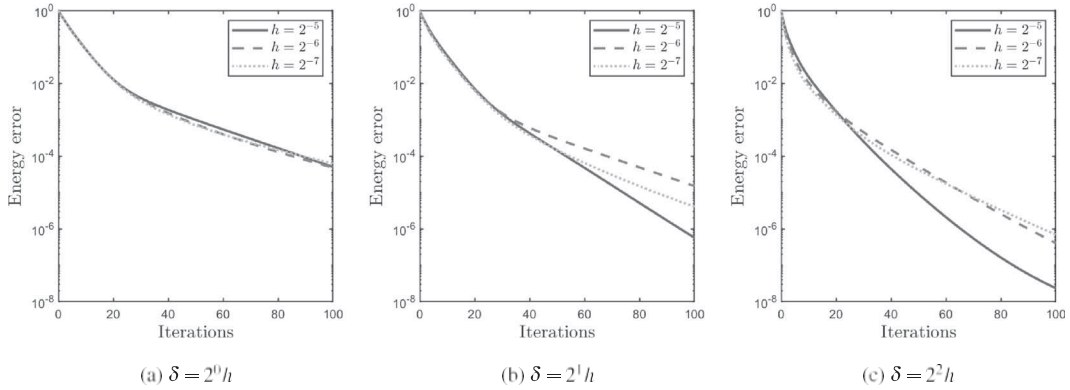


FIG. 3. Decay of the relative energy error (5.1) in the two-level additive Schwarz method (Algorithm 1) for the p -Laplacian problem (2.1) ($p = 1.05$). Parameters h , H and δ stand for the characteristic element size, subdomain size and overlapping width among subdomains, respectively ($H/h = 2^3$).

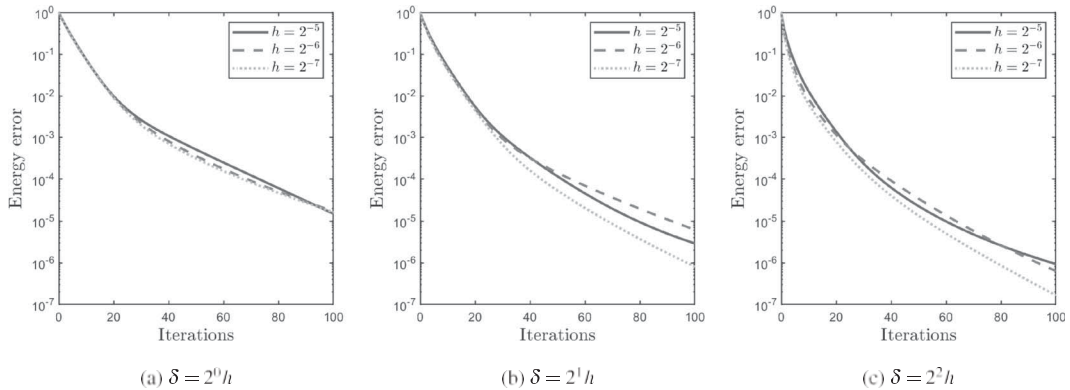


FIG. 4. Decay of the relative energy error (5.1) in the two-level additive Schwarz method (Algorithm 1) for the p -Laplacian problem (2.1) ($p = 1.1$). Parameters h , H and δ stand for the characteristic element size, subdomain size and overlapping width among subdomains, respectively ($H/h = 2^3$).

with the dependence of γ to H/δ explained in Theorem 2.4. Moreover, this observation implies that Algorithm 1 is numerically scalable; the asymptotic linear convergence rate is uniformly bounded when the ratio of the subdomain size to the overlapping width is fixed.

6. Conclusion

In this paper we developed a new convergence theory for additive Schwarz methods for boundary value problems involving the p -Laplacian. To the best of our knowledge our theory is the first theoretical result that explains the asymptotic linear convergence of additive Schwarz methods for the p -Laplacian. Our work successfully bridges the gap between theory and practice by demonstrating that our theoretical findings align well with numerical results.

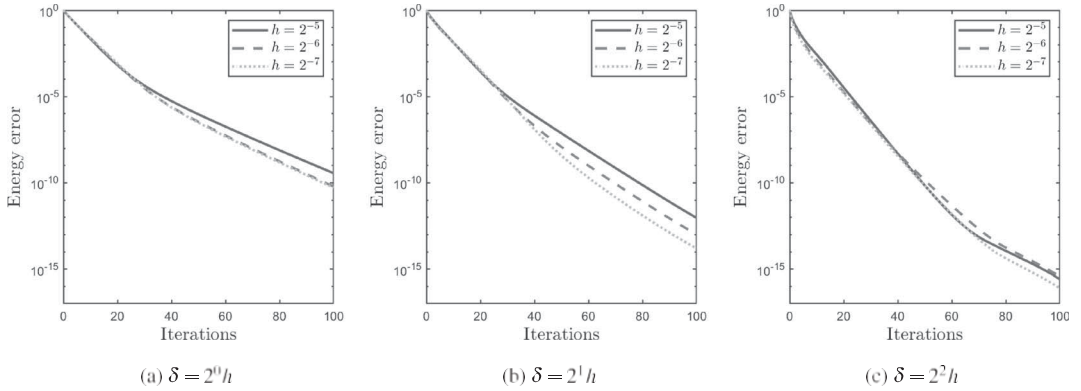


FIG. 5. Decay of the relative energy error (5.1) in the two-level additive Schwarz method (Algorithm 1) for the p -Laplacian problem (2.1) ($p = 1.5$). Parameters h , H and δ stand for the characteristic element size, subdomain size and overlapping width among subdomains, respectively ($H/h = 2^3$).

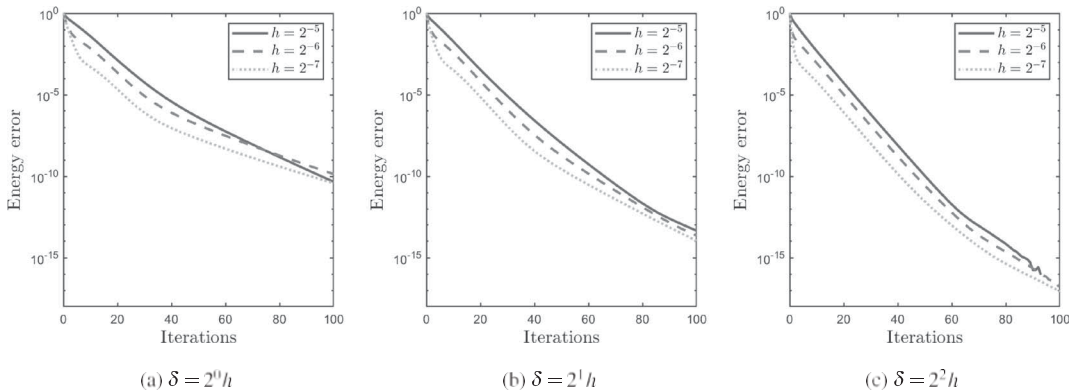


FIG. 6. Decay of the relative energy error (5.1) in the two-level additive Schwarz method (Algorithm 1) for the p -Laplacian problem (2.1) ($p = 5$). Parameters h , H and δ stand for the characteristic element size, subdomain size and overlapping width among subdomains, respectively ($H/h = 2^3$).

While the convergence theory of subspace correction methods for linear problems appears to be well-developed (Xu & Zikatanov, 2002; Lee *et al.*, 2008), there remains a need for further research on the theory of subspace correction methods for nonlinear problems. We believe that our result can serve as a foundation for the sharp convergence theory of general subspace correction methods for complex nonlinear problems.

Funding

NSF-DMS (2208499 to Y.-J.L.); Shapiro Fellowship from Penn State University in the Spring of 2022; REP grant for the year of 2022, from Texas State University; National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2021R1C1C2095193 to J.P.).

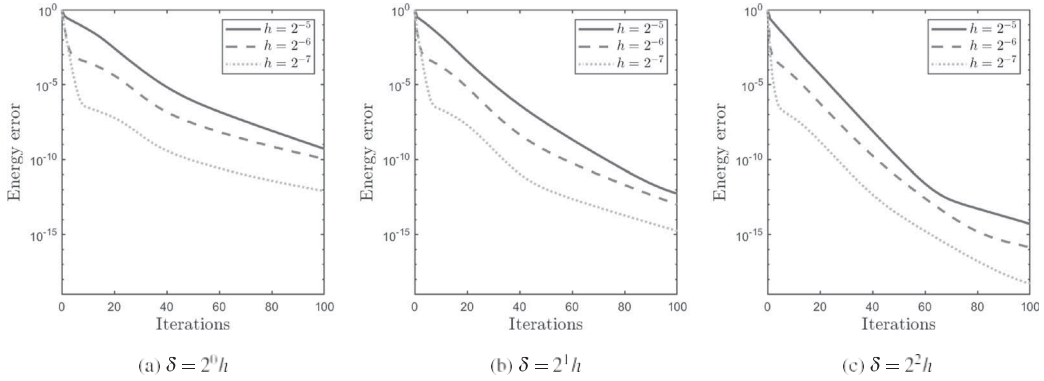


FIG. 7. Decay of the relative energy error (5.1) in the two-level additive Schwarz method (Algorithm 1) for the p -Laplacian problem (2.1) ($p = 10$). Parameters h , H and δ stand for the characteristic element size, subdomain size and overlapping width among subdomains, respectively ($H/h = 2^3$).

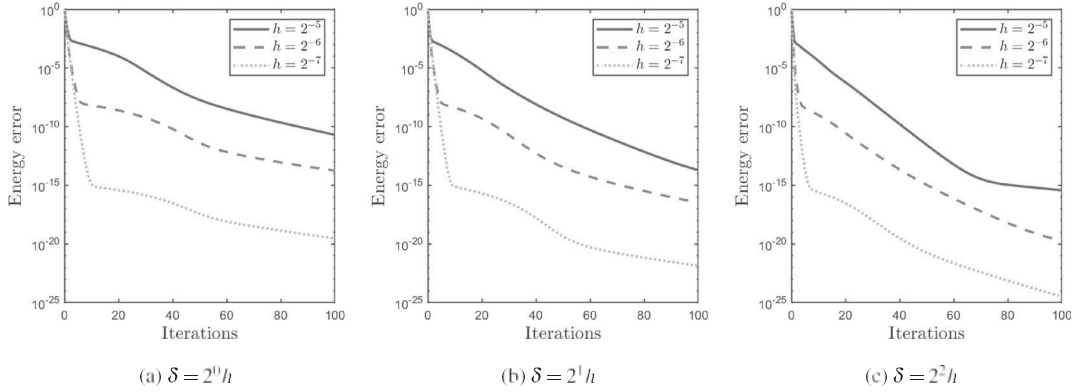


FIG. 8. Decay of the relative energy error (5.1) in the two-level additive Schwarz method (Algorithm 1) for the p -Laplacian problem (2.1) ($p = 20$). Parameters h , H and δ stand for the characteristic element size, subdomain size and overlapping width among subdomains, respectively ($H/h = 2^3$).

REFERENCES

- BADEA, L. (2006) Convergence rate of a Schwarz multilevel method for the constrained minimization of nonquadratic functionals. *SIAM J. Numer. Anal.*, **44**, 449–477.
- BADEA, L. (2019) Additive and restricted additive Schwarz–Richardson methods for inequalities with nonlinear monotone operators. *Comput. Optim. Appl.*, **74**, 345–385.
- BADEA, L. & KRAUSE, R. (2012) One-and two-level Schwarz methods for variational inequalities of the second kind and their application to frictional contact. *Numer. Math.*, **120**, 573–599.
- BARRETT, J. W. & LIU, W. B. (1993) Finite element approximation of the p -Laplacian. *Math. Comp.*, **61**, 523–537.
- BARRETT, J. W. & LIU, W. B. (1994) Finite element approximation of the parabolic p -Laplacian. *SIAM J. Numer. Anal.*, **31**, 413–428.
- BAUSCHKE, H. H., BOLTE, J. & TEBoulLE, M. (2017) A descent lemma beyond Lipschitz gradient continuity: First-order methods revisited and applications. *Math. Oper. Res.*, **42**, 330–348.

- BECK, A. & TEBoulLE, M. (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, **2**, 183–202.
- BENEDIKT, J., GÍRG, P., KOTRLA, L. & TAKÁČ, P. (2018) Origin of the p -Laplacian and A. Missbach. *Electron. J. Differ. Equations*, **16**, 17.
- BERMEJO, R. & INFANTE, J.-A. (2000) A multigrid algorithm for the p -Laplacian. *SIAM J. Sci. Comput.*, **21**, 1774–1789.
- CARSTENSEN, C., LIU, W. & YAN, N. (2006) A posteriori FE error control for p -Laplacian by gradient recovery in quasi-norm. *Math. Comp.*, **75**, 1599–1616.
- CIARLET, P. G. (2002) *The Finite Element Method for Elliptic Problems*. Philadelphia: SIAM.
- DÍAZ, J. I. (1985) *Nonlinear Partial Differential Equations and Free Boundaries*, vol. **I**, volume **106**. Boston, MA: Pitman (Advanced Publishing Program).
- DIENING, L., FORNASIER, M., TOMASI, R. & WANK, M. (2020) A relaxed Kac̣anov iteration for the p -Poisson problem. *Numer. Math.*, **145**, 1–34.
- DIENING, L. & KREUZER, C. (2008) Linear convergence of an adaptive finite element method for the p -Laplacian equation. *SIAM J. Numer. Anal.*, **46**, 614–638.
- DIENING, L. & RUŽIČKA, M. (2007) Interpolation operators in Orlicz–Sobolev spaces. *Numer. Math.*, **107**, 107–129.
- DRYJA, M. & WIDLUND, O. B. (1994) Domain decomposition algorithms with small overlap. *SIAM J. Sci. Comput.*, **15**, 604–620.
- EBMEYER, C. & LIU, W. B. (2005) Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of p -Laplacian problems. *Numer. Math.*, **100**, 233–258.
- ERN, A. & GUERMOND, J.-L. (2021) *Finite Elements I—Approximation and Interpolation*. Cham: Springer.
- GALVIS, J. & EFENDIEV, Y. (2010) Domain decomposition preconditioners for multiscale flows in high-contrast media. *Multiscale Model. Simul.*, **8**, 1461–1483.
- GLOWINSKI, R. & MARROCCO, A. (1975) Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité, d’une classe de problèmes de Dirichlet non linéaires. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér.*, **9**, 41–76.
- HUANG, Y., LI, R. & LIU, W. (2007) Preconditioned descent algorithms for p -Laplacian. *J. Sci. Comput.*, **32**, 343–371.
- LEE, Y.-J., WU, J. & CHEN, J. (2009) Robust multigrid method for the planar linear elasticity problems. *Numer. Math.*, **113**, 473–496.
- LEE, Y.-J., WU, J., XU, J. & ZIKATANOV, L. (2008) A sharp convergence estimate for the method of subspace corrections for singular systems of equations. *Math. Comp.*, **77**, 831–851.
- LIU, W. & YAN, N. (2001) Quasi-norm local error estimators for p -Laplacian. *SIAM J. Numer. Anal.*, **39**, 100–127.
- LIU, W. & YAN, N. (2002) On quasi-norm interpolation error estimation and a posteriori error estimates for p -Laplacian. *SIAM J. Numer. Anal.*, **40**, 1870–1895.
- LIU, X., CHUNG, E. & ZHANG, L. (2021) Iterated numerical homogenization for multiscale elliptic equations with monotone nonlinearity. *Multiscale Model. Simul.*, **19**, 1601–1632.
- MISHCHENKO, K. (2023) Regularized Newton method with global $O(1/k^2)$ convergence. *SIAM J. Optim.*, **33**, 1440–1462.
- O’DONOGHUE, B. & CANDÉS, E. (2015) Adaptive restart for accelerated gradient schemes. *Found. Comput. Math.*, **15**, 715–732.
- PARK, J. (2020) Additive schwarz methods for convex optimization as gradient methods. *SIAM J. Numer. Anal.*, **58**, 1495–1530.
- PARK, J. (2021) Accelerated additive schwarz methods for convex optimization with adaptive restart. *J. Sci. Comput.*, **89**.
- PARK, J. (2022) Additive schwarz methods for convex optimization with backtracking. *Comput. Math. Appl.*, **113**, 332–344.
- PECHSTEIN, C. & SCHEICHL, R. (2013) Weighted Poincaré inequalities. *IMA J. Numer. Anal.*, **33**, 652–686.
- SCHEICHL, R., VASSILEVSKI, P. S. & ZIKATANOV, L. T. (2012) Multilevel methods for elliptic problems with highly varying coefficients on nonaligned coarse grids. *SIAM J. Numer. Anal.*, **50**, 1675–1694.

- SCHEINBERG, K., GOLDFARB, D. & BAI, X. (2014) Fast first-order methods for composite convex optimization with backtracking. *Found. Comput. Math.*, **14**, 389–417.
- SCOTT, L. R. & ZHANG, S. (1990) Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, **54**, 483–493.
- SHAPOVALOV, V. M. (2017) On the applicability of the Ostwald–de Waele model in solving applied problems. *J. Engrg. Phys. Thermophys.*, **90**, 1213–1218.
- TAI, X.-C. & ESPEDAL, M. (1998) Rate of convergence of some space decomposition methods for linear and nonlinear problems. *SIAM J. Numer. Anal.*, **35**, 1558–1570.
- TAI, X.-C. & XU, J. (2002) Global and uniform convergence of subspace correction methods for some convex optimization problems. *Math. Comp.*, **71**, 105–125.
- TEBOULLE, M. (2018) A simplified view of first order methods for optimization. *Math. Program.*, **170**, 67–96.
- TOSELLI, A. & WIDLUND, O. (2005) *Domain Decomposition Methods—Algorithms and Theory*. Berlin: Springer.
- XU, J. (1992) Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, **34**, 581–613.
- XU, J. & ZIKATANOV, L. (2002) The method of alternating projections and the method of subspace corrections in Hilbert space. *J. Am. Math. Soc.*, **15**, 573–597.