Theory of Parallel Subspace Correction Methods for Smooth Convex Optimization

Jongho Park$^{[0000-0002-0496-649X]}$ and Jinchao Xu$^{[0000-0002-8221-4656]}$

1 Introduction

Subspace correction methods [15, 16] offer a unified framework for designing and analyzing modern iterative methods in scientific computing. These algorithms employ a divide, conquer, and combine strategy: they decompose a target problem into smaller local problems defined on subspaces, solve each subproblem independently, and then combine the solutions. In particular, in parallel subspace correction methods, all subproblems can be solved simultaneously, making them well-suited for implementation on parallel computers. Notable examples of parallel subspace correction methods include additive Schwarz methods (see [4, Chapter 7] and references therein), the Bramble–Pasciak–Xu preconditioner [3], and a training algorithm for federated learning [11].

While the theory of subspace correction methods was originally developed for symmetric positive definite linear problems [15], it has been extended to a wider range of problems, including indefinite linear problems [7, 16], variational inequalities [2, 13], and convex optimization problems [1, 8, 14]. Focusing on convex optimization, early results for convex optimization under strong convexity and smoothness assumptions can be found in [5, 12]. The convergence theory under weaker conditions, such as uniform convexity and weak smoothness [10], was first considered in [14], and later extended to constrained and nonsmooth settings in [1, 8, 9].

The aim of this paper is to propose a new convergence theory for parallel subspace correction methods for smooth convex optimization. This work is motivated by the well-known linear theory [15, 16], where the parameter ω, which measures the stability of local problems in subspace correction methods (the precise definition can be found in Assumption 1(c)), is allowed to be in the interval $(0, 2)$; see [15, Corollary 4.2] and [16, Equation (A2)]. However, to the best of our knowledge,

---

Jongho Park · Jinchao Xu
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, e-mail: {jongho.park,jinchao.xu}@kaust.edu.sa

existing works on subspace correction methods for convex optimization do not explain this range of $(0, 2)$. Previous studies either considered only exact local problems [1, 5, 12, 14] or dealt with inexact local problems with $\omega \in (0, 1)$ [8, 9]. In this paper, we propose a new theory for parallel subspace correction methods that accommodates a broader range of local problems. When restricted to linear problems, this corresponds to $\omega \in (0, 2)$, thereby aligning with the linear theory.

This paper is organized as follows. In Section 2, an abstract framework of parallel subspace correction methods for smooth convex optimization. In Section 3, we present the main convergence theorems. In Section 4, we discuss comparisons with existing works on the convergence theory of subspace correction methods for convex optimization. In Section 5, we apply the proposed convergence theory to the numerical solution of a nonlinear partial differential equation. In Section 6, we conclude the paper with remarks.

2 Parallel subspace correction methods

In this section, we present an abstract framework of parallel subspace correction methods for smooth convex optimization. Let V be a reflexive Banach space equipped with a norm $\| \cdot \|$. The topological dual space of V is denoted by V^*, and the duality pairing of V is written as

$$\langle p, v \rangle = p(v), \quad v \in V, \ p \in V^*.$$

We consider the following abstract convex optimization problem on V:

$$\min_{u \in V} F(u), \tag{1}$$

where $F \colon V \to \mathbb{R}$ is a Gâteaux differentiable and coercive convex functional. Due to the coercivity of the energy functional F, (1) admits a solution $u^* \in V$. For each $v \in V$, we denote the Gâteaux derivative of F at v as $F'(v) \in V^*$.

We assume that the solution space V of (1) admits a space decomposition of the form

$$V = \sum_{j=1}^{N} V_j,$$

where each V_j, $j \in [N] = \{1, 2, \dots, N\}$, is a closed subspace of V. By a simple application of the open mapping theorem, we have (cf. [16, Equation (2.15)])

$$\sup_{\|w\|=1} \inf_{\sum_{j=1}^{N} w_j = w} \left(\sum_{j=1}^{N} \|w_j\|^q \right)^{\frac{1}{q}} < \infty, \tag{2}$$

for any $q \in [1, \infty)$, where w and w_j are taken from V and V_j, respectively.

Since the energy functional F is convex, the following inequality, known as the strengthened convexity condition [8, Assumption 4.2], holds for some $\tau > 0$:

$$(1-\tau N)F(v) + \tau \sum_{j=1}^{N} F(v+w_j) \geq F\left(v + \tau \sum_{j=1}^{N} w_j\right), \ v \in V, \ w_j \in V_j. \tag{3}$$

We define the constant $\tau_0 > 0$ as the maximum τ that satisfies the strengthened convexity condition. Namely,

$$\tau_0 = \max\left\{\tau > 0 : \text{The inequality (3) holds}\right\}. \tag{4}$$

Given the convexity and coercivity of F, it follows that $\tau_0 \in [1/N, 1]$. In many applications, better estimates for τ_0, which are usually independent of N, can be obtained; see [8, 14].

Subspace correction methods involve local problems defined in the subspaces $\{V_j\}_{j=1}^N$. For a given $v \in V$, the optimal residual in the subspace V_j that minimizes the energy functional F is determined by a solution of the minimization problem

$$\min_{w_j \in V_j} F(v + w_j). \tag{5}$$

Generalizing (5) to account for scenarios with inexact local problems [8, 9], it is natural to consider the following general local problem:

$$\min_{w_j \in V_j} F_j(w_j; v), \tag{6}$$

where $F_j(\cdot; v) \colon V_j \to \mathbb{R}$ is a Gâteaux differentiable and coercive convex functional for each $v \in V$. The parallel subspace correction method for solving (1) with the inexact local problem (6) is presented in Algorithm 1. Note that the upper bound τ_0 for the step size τ was given in (4).

Algorithm 1 Parallel subspace correction method

```
Given the number of subspaces $N$ and the step size $\tau \in (0, \tau_0]$:
Choose $u^{(0)} \in V$.
for $n = 0, 1, 2, \dots$ do
    for $j \in [N]$ in parallel do
        $w_j^{(n+1)} \in \operatorname*{arg\,min}_{w_j \in V_j} F_j(w_j; u^{(n)})$
    end for
    $u^{(n+1)} = u^{(n)} + \tau \sum_{j=1}^{N} w_j^{(n+1)}$
end for
```

To ensure the convergence of the parallel subspace correction method, we need several assumptions on the local energy functional F_j. These assumptions are summarized in Assumption 1. In what follows, we introduce some notation for the sake of convenience:

$$\begin{aligned} d_F(w;v) &= F(v+w) - F(v) - \langle F'(v), w\rangle, \quad v, w \in V, \\ d_j(w_j;v) &= F_j(w_j;v) - F_j(0;v) - \langle F_j'(0;v), w_j\rangle, \quad v \in V,\ w_j \in V_j. \end{aligned}$$

Assumption 1 (Local problems). For any $j \in [N]$ and $v \in V$, the local energy functional $F_j(\cdot;v)\colon V_j \to \mathbb{R}$ in (6) satisfies the following:

(a) (Convexity) The functional $F_j(\cdot;v)\colon V_j \to \mathbb{R}$ is Gâteaux differentiable, coercive, and convex.

(b) (Consistency) We have

$$F_j(0;v) = F(v),$$

and

$$\langle F_j'(0;v), w_j\rangle = \langle F'(v), w_j\rangle, \quad w_j \in V_j.$$

(c) (Stability) For some $\omega \in (0,1] \cup (1,\rho)$, we have

$$d_F(w_j;v) \leq \omega d_j(w_j;v), \quad w_j \in V_j, \tag{7}$$

where the constant ρ is defined as

$$\rho = \min_{j \in [N]} \inf_{d_j(w_j;v) \neq 0} \frac{\langle d_j'(w_j;v), w_j\rangle}{d_j(w_j;v)}. \tag{8}$$

(d) (Smoothness) For some $q > 1$, each $F_j(w_j;v)$ is v-locally uniformly w_j-locally q-weakly smooth around $w_j = 0$, i.e., for any bounded and convex subsets $K \subset V$ and $K_j \subset V_j$ satisfying $0 \in K_j$, we have

$$\sup_{v \in K, w_j \in K_j \setminus \{0\}} \frac{d_j(w_j;v)}{\|w_j\|^q} < \infty.$$

Under Assumption 1(a, b), one can prove that the constant ρ defined in (8) is always greater than or equal to 1. Moreover, its value can be explicitly determined in many cases. We present some of these examples below.

Example 1 Throughout this example, we assume that the solution space V is a Hilbert space equipped with an inner product $\langle \cdot, \cdot \rangle$. Suppose that the local energy functional F_j is given by

$$F_j(w_j;v) = F(v) + \langle F'(v), w_j\rangle + \frac{1}{2}\langle A_j w_j, w_j\rangle, \quad v \in V,\ w_j \in V_j,$$

where $A_j\colon V_j \to V_j$ is a bounded, symmetric, and positive-definite linear operator. This formulation corresponds to the cases of general subspace correction methods for linear problems [8, Section 4.1] and gradient descent methods for convex problems. It is straightforward to verify that Assumption 1(a, b) holds. Moreover, it readily follows that

$$\frac{\langle d_j'(w_j; v), w_j \rangle}{d_j(w_j; v)} = 2$$

for any $v \in V$ and $w_j \in V_j \setminus \{0\}$. Hence, we deduce that $\rho = 2$.

Example 2 Suppose that the local energy functional F_j is given by

$$F_j(w_j; v) = F(v) + \langle F'(v), w_j \rangle + \frac{M}{s} \|w_j\|^s, \quad v \in V, \ w_j \in V_j,$$

for some $s > 1$ and $M > 0$. We readily observe that Assumption 1(a, b) holds. Moreover, it follows that (see, e.g., [18])

$$\frac{\langle d_j'(w_j; v), w_j \rangle}{d_j(w_j; v)} = s$$

for any $v \in V$ and $w_j \in V_j \setminus \{0\}$, which implies $\rho = s$.

The assumptions on the local problem (6) summarized in Assumption 1 ensure that solving (6) contributes effectively to minimizing the energy functional F. The following lemma establishes a sufficient decrease property for (6).

Lemma 1 (Local sufficient decrease) *For $j \in [N]$ and $v \in V$, let*

$$\hat{w}_j \in \operatorname*{arg\,min}_{w_j \in V_j} F_j(w_j; v). \tag{9}$$

Under Assumption 1(a–c), we have

$$F(v) - F(v + \hat{w}_j) \geq \left(1 - \frac{\omega}{\rho}\right) \langle d_j'(\hat{w}_j; v), \hat{w}_j \rangle \geq 0.$$

Sketch of proof. Using the optimality condition of $\hat{w}_j$, written as

$$\langle F'(v), w_j \rangle + \langle d_j'(\hat{w}_j; v), w_j \rangle = 0, \quad w_j \in V_j,$$

the desired result is obtained by invoking (7) and (8). □

In [8, Lemma 4.5], it was proven that the parallel subspace correction method for solving (1) can be interpreted as a gradient descent method equipped with respect to a certain nonlinear metric-like function. We state this result for our purposes in Lemma 2. Note that this generalizes a fundamental result on the theory for linear problems, often referred to as the additive Schwarz lemma, as given in, e.g., [16, Lemma 2.4].

Lemma 2 (Generalized additive Schwarz lemma) *Suppose that Assumption 1(b) holds. For $v \in V$, we have*

$$\hat{w} := \sum_{j=1}^N \hat{w}_j \in \operatorname*{arg\,min}_{w \in V} \left\{ \langle F'(v), w \rangle + \inf_{w = \sum_{j=1}^N w_j} \sum_{j=1}^N d_j(w_j; v) \right\}, \tag{10}$$

where $\hat{w}_j$ *was given in* (9)*. Moreover, we have*

$$\inf_{\hat{w}=\sum_{j=1}^N w_j} \sum_{j=1}^N d_j(w_j; v) = \sum_{j=1}^N d_j(\hat{w}_j; v). \tag{11}$$

Combining Lemmas 1 and 2 yields the main result of this section: a descent property of the parallel subspace correction method, which is presented in Theorem 1. Due to the page limit, we only provide a sketch of proof of Theorem 1.

Theorem 1 *Suppose that Assumption 1(a–c) holds. In Algorithm 1, we have*

$$F(u^{(n+1)}) \leq F(u^{(n)}) + \tau\theta \min_{w \in V} \left\{ \langle F'(u^{(n)}), w \rangle + \inf_{w=\sum_{j=1}^N w_j} \sum_{j=1}^N d_j(w_j; u^{(n)}) \right\}, \quad n \geq 0,$$

where the constant θ *is given by*

$$\theta = \begin{cases} 1, & \text{if } \omega \in (0,1], \\ \frac{\rho-\omega}{\rho-1}, & \text{if } \omega \in (1,\rho). \end{cases} \tag{12}$$

Sketch of proof. Take any $n \geq 0$. Thanks to (4), it suffices to estimate $\sum_{j=1}^N F(u^{(n)} + w_j^{(n+1)})$. The case $\omega \in (0,1]$ is straightforward by invoking Lemma 2, while the case $\omega \in (1,\rho)$ additionally requires Lemma 1 to obtain a bound for $\langle F'(u^{(n)}), w_j^{(n+1)} \rangle$. □

3 Convergence theorems

In this section, we establish convergence theorems for the parallel subspace correction method in general convex optimization settings. Thanks to Theorem 1, it suffices to estimate

$$\min_{w \in V} \left\{ \langle F'(u^{(n)}), w \rangle + \inf_{w=\sum_{j=1}^N w_j} \sum_{j=1}^N d_j(w_j; u^{(n)}) \right\} \tag{13}$$

to derive a convergence rate of the parallel subspace correction method.

The following lemma provides an important result for estimating the d_j-term in (13), demonstrating that a stable decomposition can be found for the subspaces and the corresponding local problems. Related conditions can be found in [8, Assumption 4.1], [14, Equation (13)], and [16, Equation (2.17)].

Lemma 3 (Stable decomposition) *Suppose that Assumption 1(a, b, d) holds. For any bounded subset* $K \subset V$*, the following holds:*

$$C_K := q \sup_{v, v+w \in K,\, w \neq 0} \inf_{w=\sum_{j=1}^N w_j} \frac{\sum_{j=1}^N d_j(w_j; v)}{\|w\|^q} < \infty. \tag{14}$$

Sketch of proof. The desired result follows from combining (2) and Assumption 1(d). □

Given $u^{(0)} \in V$, we define

$$K_0 = \left\{ u \in V : F(u) \leq F(u^{(0)}) \right\}, \tag{15a}$$

$$R_0 = \inf \left\{ R > 0 : K_0 \subset \overline{B}(u^*; R) \right\}, \tag{15b}$$

where $\overline{B}(u^*; R)$ is the closed ball of radius R centered at u^*. By the convexity and coercivity of F, the set K_0 is bounded and convex, ensuring $R_0 < \infty$. Moreover, Theorem 1 implies that the sequence $\{u^{(n)}\}$ generated by Algorithm 1 is contained in K_0. Consequently, for any $n \geq 0$, by Lemma 3, we get

$$\inf_{w = \sum_{j=1}^N w_j} \sum_{j=1}^N d_j(w_j; u^{(n)}) \leq \frac{C_{K_0}}{q} \|w\|^q. \tag{16}$$

Using a similar argument as in [8, 9] with (13) and (16), we are able to derive the following convergence theorem for the parallel subspace correction method for solving (1).

Theorem 2 *Suppose that Assumption 1 holds. In Algorithm 1, let* $\zeta_n = F(u^{(n)}) - F(u^*)$ *for* $n \geq 0$. *If* $\zeta_0 > C_{K_0} R_0^q$, *then we have*

$$\zeta_1 \leq \left(1 - \tau\theta \left(1 - \frac{1}{q} \right) \right) \zeta_0,$$

where θ, C_{K_0}, *and* R_0 *were given in* (12), (14), *and* (15). *Otherwise, we have*

$$\zeta_n \leq \frac{C}{\left(n + (C/\zeta_0)^{1/\beta} \right)^\beta}, \quad n \geq 0,$$

where

$$\beta = q - 1, \quad C = \left(\frac{q}{\tau\theta} \right)^{q-1} C_{K_0} R_0^q.$$

Meanwhile, in many applications, the energy functional F satisfies the sharpness condition [8, Assumption 3.4], which characterizes the growth rate of F around its minimizer u^*.

Assumption 2 (Sharpness). For some $p > 1$, the following holds: for any bounded and convex subset $K \subset V$ satisfying $u^* \in K$, we have

$$\mu_K := p \inf_{u \in K \setminus \{u^*\}} \frac{F(u) - F(u^*)}{\|u - u^*\|^p} > 0. \tag{17}$$

Example 3 As in Example 1, we assume that the solution space V is a Hilbert space equipped with an inner product $\langle \cdot, \cdot \rangle$. Suppose that the energy functional F is given by

$$F(v) = \frac{1}{2}\langle Av, v\rangle - \langle f, v\rangle, \quad v \in V,$$

where $A\colon V \longrightarrow V$ is a bounded, symmetric, and positive definite linear operator, and $f \in V$. In this case, we have $u^* = A^{-1}f$. For any $u \in V \setminus \{u^*\}$, it follows that

$$\frac{F(u) - F(u^*)}{\|u - u^*\|^2} = \frac{1}{2}\frac{\langle A(u-u^*), u - u^*\rangle}{\|u - u^*\|^2} \geq \frac{1}{2}\lambda_{\min}(A),$$

where $\lambda_{\min}(A)$ denotes the minimum eigenvalue of A. This implies that Assumption 2 holds with $p = 2$ and $\mu_K = \lambda_{\min}(A)$ for any K. In particular, if we equip the space V with the energy norm $\|\cdot\|_A = (\langle A\cdot, \cdot\rangle)^{\frac{1}{2}}$, then we can deduce that $\mu_K = 1$ for any K.

If we additionally assume that Assumption 2 holds, then we are able to derive the following improved convergence theorem for the parallel subspace correction method for solving (1).

Theorem 3 *Suppose that Assumptions 1 and 2 hold. In Algorithm 1, let* $\zeta_n = F(u^{(n)}) - F(u^*)$ *for* $n \geq 0$. *Then we have the following:*

(a) In the case $p = q$, *we have*

$$\zeta_n \leq \left(1 - \tau\theta\left(1 - \frac{1}{q}\right)\min\left\{1, \frac{\mu_{K_0}}{qC_{K_0}}\right\}^{\frac{1}{q-1}}\right)^n \zeta_0, \quad n \geq 0,$$

where θ, C_{K_0}, *and* μ_{K_0} *were given in* (12), (14), (15a), *and* (17).

(b) In the case $p > q$, *if* $\zeta_0 > \left(\frac{p}{\mu_{K_0}}\right)^{\frac{q}{p-q}} C_{K_0}^{\frac{p}{p-q}}$, *then we have*

$$\zeta_1 \leq \left(1 - \tau\theta\left(1 - \frac{1}{q}\right)\right)\zeta_0.$$

Otherwise, we have

$$\zeta_n \leq \frac{C}{\left(n + (C/\zeta_0)^{1/\beta}\right)^\beta}, \quad n \geq 0,$$

where

$$\beta = \frac{p(q-1)}{p-q}, \quad C = \left(\frac{pq}{(p-q)\tau\theta}\right)^{\frac{p(q-1)}{p-q}} \left(\frac{p}{\mu_K}\right)^{\frac{q}{p-q}} C_K^{\frac{p}{p-q}}.$$

4 Comparison to existing works

In this section, we compare the convergence theory proposed in this paper with existing works on the theory of subspace correction methods for convex optimization [1, 8, 9, 14].

In [1, 14], the convergence analyses were based on the assumptions that exact local problems are adopted and that the energy functional F is locally p-uniformly convex and q-weakly smooth [10] for some $p, q > 1$, i.e.,

$$\inf_{v \in K, w \in K \setminus \{0\}} \frac{d_F(w; v)}{\|w\|^p} > 0 \quad \text{and} \quad \sup_{v \in K, w \in K \setminus \{0\}} \frac{d_F(w; v)}{\|w\|^q} < \infty \tag{18}$$

for any bounded subset $K \subset V$; see [1, Equation (2.8)] and [14, Lemma 2.1]. It is straightforward to observe that (18) implies Assumptions 1 and 2 if the exact local problems are used. This means that the proposed convergence theory requires weaker assumptions than [1, 14].

Meanwhile, the convergence rates of the energy error established in [1] and [14] are $O(n^{-\frac{q-1}{p-q}})$ and $O(n^{-\frac{q(q-1)}{(p-q)(p+q-1)}})$, respectively. Since

$$\frac{q(q-1)}{(p-q)(p+q-1)} < \frac{q-1}{p-q} < \frac{p(q-1)}{p-q}$$

for $1 < q < p$, we conclude that Theorem 3 offers a sharper estimate than [1, 14]. We note that a similar discussion was presented in [8, Section 6.1].

Next, we compare the proposed convergence theory with those presented in [8, 9]. The local stability condition introduced in [8, Assumption 4.3] and [8, Assumption 2.4] corresponds to the case $\omega \in (0, 1]$ in Assumption 1(c). Hence, the proposed convergence theory can accommodate a broader class of local problems than those considered in [8, 9]. Furthermore, the proposed theory aligns with the linear theory introduced in [15, 16], where the allowed range of ω is $(0, 2)$; see [15, Corollary 4.2] and [16, Equation (A2)].

In addition, the constants appearing in the convergence theorems are improved compared to [8, 9]. In Theorems 2 and 3, the set associated with the stable decomposition parameter is K_0 (see (15a)). However, in [8, 9], the corresponding set is a superset of K_0 that depends on the step size τ; see [8, Equation (4.14)] and [9, Equation (2.5)]. Consequently, the proposed convergence theory provides better estimates than those in [8, 9].

5 An application

In this section, we present an application of the proposed convergence theory to the numerical solution of the following boundary value problem involving the s-Laplacian ($s > 1$):

$$\begin{aligned} -\operatorname{div}(|\nabla u|^{s-2}\nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded polygonal domain and $f \in W^{-1,s^*}(\Omega)$ with $\frac{1}{s} + \frac{1}{s^*} = 1$. The weak formulation of the above problem reads as

$$\min_{u \in W_0^{1,s}(\Omega)} \left\{ \frac{1}{s} \int_\Omega |\nabla u|^s \, dx - \langle f, u \rangle \right\}. \tag{19}$$

To numerically solve (19), we employ a finite element discretization of (19). Let $\mathcal{T}_h$ be a quasi-uniform triangulation of Ω with h the characteristic element diameter. The lowest-order Lagrangian finite element space with the homogeneous essential boundary condition defined on $\mathcal{T}_h$ is denoted by $S_h(\Omega)$. Then the finite element approximation of (19) defined on $S_h(\Omega)$ is written as

$$\min_{u \in S_h(\Omega)} \left\{ \frac{1}{s} \int_\Omega |\nabla u|^s \, dx - \langle f, u \rangle \right\}. \tag{20}$$

We see that (20) is an instance of (1). Namely, we obtain (20) from (1) if we set

$$V = S_h(\Omega), \quad F(u) = \frac{1}{s} \int_\Omega |\nabla u|^s \, dx - \langle f, u \rangle.$$

In what follows, we analyze a two-level overlapping Schwarz method for solving (20) [6, 14] based on the proposed convergence theory. Let $\mathcal{T}_H$ be a quasi-uniform triangulation of Ω such that $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$, where H stands for the characteristic element diameter of $\mathcal{T}_H$. In addition, let $\{\Omega_j\}_{j=1}^N$ be a quasi-uniform overlapping domain decomposition of Ω, such that each subdomain Ω_j is a union of $\mathcal{T}_h$-elements and it has diameter of order H. The overlap width among the subdomains is measured by a parameter δ. In the two-level overlapping Schwarz method, we define the subspaces as follows:

$$V_0 = I_0 S_H(\Omega), \quad V_j = I_j S_h(\Omega_j), \quad j \in [N], \tag{21}$$

where $S_H(\Omega)$ and each $S_h(\Omega_j)$ are defined similarly to $S_h(\Omega)$. The operators $I_0 \colon S_H(\Omega) \to S_h(\Omega)$ and $I_j \colon S_h(\Omega_j) \to S_h(\Omega)$ are the natural embedding operators. In this setting, we readily get the two-level space decomposition

$$V = V_0 + \sum_{j=1}^N V_j.$$

For the local problem (6), we employ the exact one (5), i.e.,

$$F_j(w_j; v) = F(v + w_j), \quad v \in V, \; w_j \in V_j.$$

Now, we verify the required assumptions and estimate the parameters in the proposed convergence theory. Since we adopt the exact local problems, Assumption 1(a–c) is obvious with $\omega = 1$. Moreover, thanks to [8, Eqations (6.6) and (6.7)], Assumption 1(d) and Assumption 2 hold with $p = \max\{s, 2\}$, $q = \min\{s, 2\}$, and $\mu_K \gtrsim 1$. The strengthened convexity parameter τ_0 in (4) enjoys a lower bound $\tau_0 \geq \frac{1}{5}$ due to a usual coloring argument [8, Section 5.1]. Finally, invoking [14, Lemma 4.1], we deduce that the stable decomposition parameter C_K admits an upper bound whose geometric dependence is only on H/δ.

In conclusion, Theorem 3 implies that the two-level additive Schwarz method for solving (20) satisfies the following convergence rate:

$$F(u^{(n)}) - F(u^*) \lesssim \frac{C_{H/\delta}}{n^{\frac{p(q-1)}{p-q}}},$$

where $C_{H/\delta}$ is a positive constant depending on H/δ.

6 Conclusion

In this paper, we proposed a new convergence theory of parallel subspace correction methods for smooth convex optimization, accommodating a broader range of local problems that align with the linear theory [15, 16]. We demonstrated that, compared to existing works [1, 8, 9, 14], the proposed theory provides better estimates on the convergence rates. Additionally, we applied our theory to a two-level additive Schwarz method for solving a nonlinear boundary value problem.

Several fundamental directions for future research remain. One natural extension is to develop the proposed convergence theory for nonsmooth convex optimization [8]. Another task is to design a corresponding convergence theory for successive subspace correction methods [7, 16], which is challenging given the nonlinearity inherent in convex optimization problems. Establishing such a convergence theory for successive methods would be particularly useful for designing efficient multilevel algorithms [17] for convex optimization.

References

1. Badea, L.: Convergence rate of a Schwarz multilevel method for the constrained minimization of nonquadratic functionals. SIAM J. Numer. Anal. **44**(2), 449–477 (2006)
2. Badea, L., Krause, R.: One-and two-level Schwarz methods for variational inequalities of the second kind and their application to frictional contact. Numer. Math. **120**(4), 573–599 (2012)
3. Bramble, J.H., Pasciak, J.E., Xu, J.: Parallel multilevel preconditioners. Math. Comp. **55**(191), 1–22 (1990)
4. Brenner, S.C., Scott, R.: The Mathematical Theory of Finite Element Methods. Springer, New York (2008)

5. Carstensen, C.: Domain decomposition for a non-smooth convex minimization problem and its application to plasticity. Numer. Linear Algebra Appl. **4**(3), 177–190 (1997)
6. Lee, Y.J., Park, J.: On the linear convergence of additive Schwarz methods for the p-Laplacian. IMA J. Numer. Anal. p. drae068 (2024). DOI 10.1093/imanum/drae068
7. Lee, Y.J., Wu, J., Xu, J., Zikatanov, L.: A sharp convergence estimate for the method of subspace corrections for singular systems of equations. Math. Comp. **77**(262), 831–850 (2008)
8. Park, J.: Additive Schwarz methods for convex optimization as gradient methods. SIAM J. Numer. Anal. **58**(3), 1495–1530 (2020)
9. Park, J.: Additive Schwarz methods for convex optimization with backtracking. Comput. Math. Appl. **113**, 332–344 (2022)
10. Park, J.: Fast gradient methods for uniformly convex and weakly smooth problems. Adv. Comput. Math. **48**(3), Paper No. 34 (2022)
11. Park, J., Xu, J.: DualFL: A duality-based federated learning algorithm with communication acceleration in the general convex regime. arXiv preprint arXiv:2305.10294 (2023)
12. Tai, X.C., Espedal, M.: Rate of convergence of some space decomposition methods for linear and nonlinear problems. SIAM J. Numer. Anal. **35**(4), 1558–1570 (1998)
13. Tai, X.C., Heimsund, B., Xu, J.: Rate of convergence for parallel subspace correction methods for nonlinear variational inequalities. In: Domain Decomposition Methods in Science and Engineering (Lyon, 2000), pp. 127–138 (2002)
14. Tai, X.C., Xu, J.: Global and uniform convergence of subspace correction methods for some convex optimization problems. Math. Comp. **71**(237), 105–124 (2002)
15. Xu, J.: Iterative methods by space decomposition and subspace correction. SIAM Rev. **34**(4), 581–613 (1992)
16. Xu, J., Zikatanov, L.: The method of alternating projections and the method of subspace corrections in Hilbert space. J. Amer. Math. Soc. **15**(3), 573–597 (2002)
17. Xu, J., Zikatanov, L.: Algebraic multigrid methods. Acta Numer. **26**, 591–721 (2017)
18. Xu, Z.B., Roach, G.F.: Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces. J. Math. Anal. Appl. **157**(1), 189–210 (1991)