algorithms

MDPI

Article

Delayed Star Subgradient Methods for Constrained Nondifferentiable Quasi-Convex Optimization

Ontima Pankoon and Nimit Nimana *

Department of Mathematics, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand; ontimapa@kkumail.com
* Correspondence: nimitni@kku.ac.th

Abstract

In this work, we consider the problem of minimizing a quasi-convex function over a nonempty closed convex constrained set. In order to approximate a solution of the considered problem, we propose delayed star subgradient methods. The main feature of the proposed methods is that it allows us to use the stale star subgradients when updating the next iteration rather than computing the new star subgradient in every iteration. We subsequently investigate the convergence results of sequences generated by the proposed methods. Finally, we present some numerical experiments on the Cobb–Douglas production efficiency problem to illustrate the effectiveness of the proposed method.

Keywords: delay; quasi-convex optimization; nonsmooth optimization; quasi-convex function; star subgradient

1. Introduction

In this work, we consider the quasi-convex optimization problem:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in X, \end{array} \tag{1}$$

where the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is quasi-convex continuous (possibly) and nondifferentiable, and the constrained set X is nonempty closed and convex. Throughout this work, we will denote the set of all optimal solutions and the optimal value of the problem (1) as $X^* := \operatorname{argmin}_{\mathbf{x}\in X} f(\mathbf{x})$ and $f^* := \min_{\mathbf{x}\in X} f(\mathbf{x})$, respectively. We also assume that $X^* \neq \emptyset$.

The problem (1) not only forms a fundamental setting for optimization as every convex function is quasi-convex but also generalizes to several practical situations in many fields such as economics, engineering, and others. An important special setting is the problem of minimizing the ratios of two functions [1] (Lemma 4), [2] (Theorem 2.3.8), for example, optimizing the ratio of outputs and inputs for the highest-quality productivity assessments in economics, such as debt-to-equity ratio analysis in corporate financial planning, inventory-to-sales analysis in production planning, and nurse–patient ratio analysis in hospital planning. This kind of problem is known as fractional programming [3–5]. As mentioned above, various kinds of iterative methods have been developed to solve the quasi-convex optimization problem (1); see [1,6–14] and references therein.

check for updates

Academic Editor: Binlin Zhang

Received: 1 July 2025
Revised: 21 July 2025
Accepted: 24 July 2025
Published: 26 July 2025

Citation: Pankoon, O.; Nimana, N. Delayed Star Subgradient Methods for Constrained Nondifferentiable Quasi-Convex Optimization. *Algorithms* **2025**, *18*, 469. https://doi.org/10.3390/a18080469

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

If the function f is convex, it is well-known that the classical method to deal with the constrained convex optimization problem is the projected subgradient method: for an initial point $\mathbf{x}_0 \in X$, compute

$$\mathbf{x}_{k+1} := P_X(\mathbf{x}_k - \alpha_k \widetilde{\nabla} f(\mathbf{x}_k)), \qquad \forall k \in \mathbb{N}_0,$$

where $\widetilde{\nabla} f(\mathbf{x}_k) \in \{\mathbf{g} \in \mathbb{R}^n : \langle \mathbf{g}, \mathbf{y} - \mathbf{x} \rangle \leq f(\mathbf{y}) - f(\mathbf{x}) \text{ for all } \mathbf{y} \in \mathbb{R}^n\}$ is the Fenchel subgradient of f at $\mathbf{x}_k$. Because of the simplicity, this method has been continuously developed in the literature, see for instance [8,15]. However, for the quasi-convex case, it should be noted that the Fenchel subgradient may not exist, in general. In this situation, we need a more general subgradient concept, which is the so-called star subgradient proposed in [16,17]. By utilizing the notion of a star subgradient, Kiwiel [1] proposed the star subgradient method to solve the constrained quasi-convex optimization problem (1): for an initial point $\mathbf{x}_0 \in X$,

$$\mathbf{x}_{k+1} := P_X\left(\mathbf{x}_k - \alpha_k \frac{\widetilde{\nabla}^\star f(\mathbf{x}_k)}{\|\widetilde{\nabla}^\star f(\mathbf{x}_k)\|}\right), \qquad \forall k \in \mathbb{N}_0, \tag{2}$$

where $\widetilde{\nabla}^\star f(\mathbf{x}_k) \in \partial^\star f(\mathbf{x}_k) \setminus \{\mathbf{0}\}$ is the nonzero star subgradient of f at $\mathbf{x}_k$. After that, Hu et al. [6] proposed an inexact version of the star-subgradient method by considering the approximated star subgradient with the presence of noises. By imposing the Hölder condition, they presented the convergence results in various aspects for both constant and diminishing step sizes. Subsequently, several methods have been investigated for solving more general problem settings, for instance, a conditional star subgradient method [8], a stochastic star subgradient method [7], and an incremental star subgradient method [9].

As we can notice, the above-mentioned methods require computing the star subgradient in every single iteration; however, there are some practical situations such as large-scale problems in which the exact star subgradients may be computationally expensive. This shortcoming can be overcome by using the idea of delayed subgradients in which we can retrieve stale subgradients instead of spending time computing a new subgradient in every iteration. This approach is not only useful for dealing with excessive computation but also for handling communication delays in networked systems. Note that various convex methods with delayed gradient or even subgradient updates are developed and studied, for instance, stochastic gradient descent [18,19], incremental type methods [20–23], distributed type methods [24–26], and references therein.

To the best of our knowledge, there is no report about the star subgradient method with delayed updates to solve the problem (1). To fill this gap, in this work, we propose two delayed star subgradient methods (in short, DSSMs) based on the ideas of the star subgradient method (2) and the delayed subgradient updates, called DSSM-I and DSSM-II, respectively. We investigate the convergence properties of the proposed methods in both objective values and iterations, including proving finite convergence.

This paper is organized as follows: Section 2 provides the essential notations, definitions, and facts utilized in this work. Section 3 contains the proposed methods, including their convergence results. This section is divided into two subsections. In Section 4, we then apply DSSM-I to the Cobb–Douglas production efficiency problem. Finally, we provide the summary in Section 5.

2. Preliminaries

This section contains important symbols, definitions, and tools utilized in this work. Throughout this paper, let $\mathbb{R}$ be the set of all real numbers, $\mathbb{N}_0$ the set of all non-negative integers, and $\mathbb{N}$ the set of all positive integers. We denote with $\mathbb{R}^n$ a Euclidean space with an inner product $\langle \cdot, \cdot \rangle$ and an induced norm $\| \cdot \|$. We denote the unit sphere by the set

$\mathcal{S} := \{\mathbf{z} \in \mathbb{R}^n : \|\mathbf{z}\| = 1\}$ and the open ball centered at the $\mathbf{x} \in \mathbb{R}^n$ with radius $\delta > 0$ by the set $\mathcal{B}(\mathbf{x}; \delta) := \{\mathbf{z} \in \mathbb{R}^n : \|\mathbf{z} - \mathbf{x}\| < \delta\}$.

Let $X \subset \mathbb{R}^n$ be a nonempty set. The symbol $\mathrm{cl}(X)$ is the closure of X. The distance (function) between $\mathbf{y} \in \mathbb{R}^n$ and X is the function $\mathrm{dist} : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\mathrm{dist}(\mathbf{y}, X) := \inf_{\mathbf{x} \in X} \|\mathbf{y} - \mathbf{x}\|.$$

The metric projection of $\mathbf{y} \in \mathbb{R}^n$ onto X is the point $P_X(y) \in X$ such that $\|P_X(\mathbf{y}) - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x} \in X$. Note that the metric projection $P_X(\mathbf{y})$ exists and is unique for all $\mathbf{y} \in \mathbb{R}^n$ whenever X is nonempty, closed, and convex [27] (Theorem 1.2.3). The metric projection onto a nonempty closed convex set X has a nonexpansive property, i.e., $\|P_X(\mathbf{x}) - P_X(\mathbf{y})\| \leq \|\mathbf{x} - \mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ [27] (Theorem 2.2.21).

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function, and let $\alpha \in \mathbb{R}$. The strictly sublevel and sublevel set of f corresponding to α are defined by

$$S^{<}_{f,\alpha} := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) < \alpha\} \text{ and } S^{\leq}_{f,\alpha} := \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq \alpha\},$$

respectively. The function f is called upper semicontinuous on $\mathbb{R}^n$ if $S^{<}_{f,\alpha}$ is an open set for all $\alpha \in \mathbb{R}$. The function f is called lower semicontinuous on $\mathbb{R}^n$ if $S^{\leq}_{f,\alpha}$ is a closed set for all $\alpha \in \mathbb{R}$. The function f is continuous on $\mathbb{R}^n$ if it is both upper semicontinuous and lower semicontinuous on $\mathbb{R}^n$. The function f is called quasi-convex on $\mathbb{R}^n$ if $S^{\leq}_{f,\alpha}$ is convex for all $\alpha \in \mathbb{R}$.

For any $\mathbf{x} \in \mathbb{R}^n$, the star subdifferential [16,17] of f at $\mathbf{x}$ is the set

$$\partial^\star f(\mathbf{x}) := \left\{\mathbf{g} \in \mathbb{R}^n : \langle \mathbf{g}, \mathbf{y} - \mathbf{x}\rangle \leq 0 \text{ for all } \mathbf{y} \in S^{<}_{f,f(\mathbf{x})}\right\}.$$

The element $\mathbf{g} \in \partial^\star f(\mathbf{x})$ is called the star subgradient of f at $\mathbf{x}$, and it is denoted by $\widetilde{\nabla}^\star f(\mathbf{x})$. It is obvious that $\mathbf{0} \in \partial^\star f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. The following fact provides the basic properties of the star subgradient, which can be found in [16] (Proposition 30) and [6] (Lemma 2.1).

Fact 1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a quasi-convex, and let $\mathbf{x} \in \mathbb{R}^n$. Then the following statements hold:*

(i) $\partial^\star f(\mathbf{x}) \setminus \{\mathbf{0}\}$ *is a nonempty set;*
(ii) $\partial^\star f(\mathbf{x})$ *is a closed, convex cone.*

Next, we will recall the definition and properties of the quasi-Fejer monotone sequence. Let X be a nonempty subset of $\mathbb{R}^n$. The sequence $\{\mathbf{x}_k\}_{k=0}^{\infty} \subset \mathbb{R}^n$ is said to be quasi-Fejer monotone with respect to X if for any $\mathbf{x} \in X$, there exists a sequence $\{\zeta_k\}_{k=0}^{\infty} \subset (0, \infty)$ such that

$$\sum_{k=0}^{\infty} \zeta_k < \infty \text{ and } \|\mathbf{x}_{k+1} - \mathbf{x}\|^2 \leq \|\mathbf{x}_k - \mathbf{x}\|^2 + \zeta_k \text{ for all } k \in \mathbb{N}_0.$$

Fact 2 ([28]). *(Theorem 5.33) Let $X \subset \mathbb{R}^n$ be a nonempty set and $\{\mathbf{x}_k\}_{k=0}^{\infty}$ be a sequence in $\mathbb{R}^n$. If the sequence $\{\mathbf{x}_k\}_{k=0}^{\infty}$ is quasi-Fejer monotone with respect to X, then $\{\mathbf{x}_k\}_{k=0}^{\infty}$ is bounded and $\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}\|$ exists for all $\mathbf{x} \in X$.*

Finally, we provide the lemma, which is used in the next section.

Lemma 1 ([29]). *(Lemma 2.1) Let $\{a_k\}_{k=0}^{\infty}$ be a scalar sequence, and let $\{b_k\}_{k=0}^{\infty}$ be a positive real number sequence such that $\lim_{k\to\infty} \sum_{i=0}^{k} b_i = \infty$. If $\lim_{k\to\infty} a_k = 0$, then $\lim_{k\to\infty} \frac{\sum_{i=0}^{k} a_i b_i}{\sum_{i=0}^{k} b_i} = 0$.*

3. Algorithms and Convergence Results

In this section, we introduce two delayed star subgradient methods (DSSM-I and DSSM-II) to solve Problem (1) and subsequently investigate the convergence properties of the proposed methods. To deal with delayed-type methods, we need the following assumption of the boundedness of the time-varying delays, as stated below.

Assumption 1. *The sequence of time-varying delays* $\{\tau_k\}_{k=0}^{\infty} \subset \mathbb{N}_0$ *is bounded, that is, there exists a non-negative integer* τ *such that* $0 \leq \tau_k \leq \tau$ *for all* $k \in \mathbb{N}_0$.

3.1. Delayed Star Subgradient Method I

Now, we are ready to propose the delayed star subgradient method of the first kind which is defined as the following Algorithm 1.

Algorithm 1 Delayed Star Subgradient Method I (DSSM-I)

Initialization: Given a stepsize $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$, the delays $\{\tau_k\}_{k=0}^{\infty} \subset \mathbb{N}_0$, and initial points $\mathbf{x}_0, \mathbf{x}_{-1}, \mathbf{x}_{-2}, \ldots, \mathbf{x}_{-\tau} \in X$.
Iterative Step: For a current point $\mathbf{x}_k \in X$, we compute

$$\mathbf{x}_{k+1} := P_X(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k}),$$

where $\mathbf{g}_{k-\tau_k} \in \partial^\star f(\mathbf{x}_{k-\tau_k}) \cap \mathcal{S}$ is a unit star subgradient of f at $\mathbf{x}_{k-\tau_k}$.
Update $k := k + 1$.

Throughout this work, to simplify the analysis, we denote $\mathbf{x}_0 = \mathbf{x}_{-1} = \mathbf{x}_{-2} = \cdots = \mathbf{x}_{-\tau}$.

Remark 1.

(i) *Since the function* f *is quasi-convex, we note that DSSM-I is well defined. In fact, using Fact 1, we have for any* $k \in \mathbb{N}_0$, $\partial^\star f(\mathbf{x}_k) \setminus \{\mathbf{0}\} \neq \emptyset$. *Hence, we are able to select any point* $\widetilde{\nabla}^\star f(\mathbf{x}_k) \in \partial^\star f(\mathbf{x}_k) \setminus \{\mathbf{0}\}$ *and subsequently set* $\mathbf{g}_k := \frac{\widetilde{\nabla}^\star f(\mathbf{x}_k)}{\|\widetilde{\nabla}^\star f(\mathbf{x}_k)\|} \in \partial^\star f(\mathbf{x}_k) \cap \mathcal{S}$. *Moreover, since* $\{\mathbf{x}_k\}_{k=0}^{\infty} \subset X$, *it is clear that* $f(\mathbf{x}_k) \geq f^*$ *for all* $k \in \mathbb{N}_0$.

(ii) *If* $\tau_k = 0$ *for all* $k \in \mathbb{N}_0$, *DSSM-I coincides with the star subgradient method (SSM), which was proposed by Kiwiel [1].*

Remark 2. *Note that methods with time-varying delays allow us to use stale information, which is very helpful when the star subgradients are not easily computed. Assumption 1 is typically assumed for analyzing the convergence results of the delayed-type methods, see [18–20,24–26] and references therein. The presence of the delayed bound* τ *ensures that we need to update the unit star subgradient of* f *at* $\mathbf{x}_{k-\tau_k}$ *in at least every* τ *iteration. Some examples of the delayed sequences are as follows:*

(i) *Constant delay [18], that is,* $\tau_k := \tau$ *for all* $k \in \mathbb{N}_0$.

(ii) *Cyclic delay [19,20,24,26] and references therein. The typical form of this type is* $\tau_k := k \mod (\tau + 1)$ *for all* $k \in \mathbb{N}_0$. *In this case, the delays* τ_k $(k \in \mathbb{N}_0)$ *are chosen with deterministic order from the set* $\{0, 1, 2, \ldots, \tau\}$. *This means that we will use stale information over a consistent period of length* $\tau + 1$ *and update the new star subgradient at every* $\tau + 1$ *iteration.*

(iii) *Random delay [25], that is, the delays* τ_k $(k \in \mathbb{N}_0)$ *are randomly chosen in the set* $\{0, 1, 2, \ldots, \tau\}$.

Assumption 2. *The function* f *satisfies the Hölder condition with order* $p > 0$ *and modulus* $L > 0$ *on* X, *i.e.,* $f(\mathbf{x}) - f^* \leq L\text{dist}(\mathbf{x}, X^*)^p$ *for all* $\mathbf{x} \in X$.

The Hölder condition assumption is typically assumed when investigating the convergence results of methods for solving quasi-convex optimization problems [6,7,9–12]. Note that, for the special case of $p = 1$, the Hölder condition is nothing else than the Lipschitz condition.

The following lemma provides a useful property of the star subgradient and will play an important role in the convergence analysis. The formal proof is due to Konnov [10] (Proposition 2.1).

Lemma 2. *Suppose that Assumption 2 holds. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-I. For any* $\mathbf{x}^* \in X^*$ *and* $k \in \mathbb{N}_0$*, if* $\mathbf{x}_k \notin X^*$*, then*

$$f(\mathbf{x}_k) - f^* \leq L\langle \mathbf{g}_k, \mathbf{x}_k - \mathbf{x}^* \rangle^p \qquad \text{for all } \mathbf{g}_k \in \partial^* f(\mathbf{x}_k) \cap \mathcal{S}.$$

Next, we will provide some lemma, which are important tools for proving the convergence results of DSSM-I. We first provide the basic inequality relating to the sufficient decreasing property of the generated sequence.

Lemma 3. *Suppose that Assumption 2 holds. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-I. For any* $\mathbf{x}^* \in X^*$ *and* $k \in \mathbb{N}_0$*, if* $\mathbf{x}_{k-\tau_k} \notin X^*$*, then*

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k L^{-\frac{1}{p}} \left(f(\mathbf{x}_{k-\tau_k}) - f^*\right)^{\frac{1}{p}}.$$

Proof. Let $\mathbf{x}^* \in X^*$ and $k \in \mathbb{N}_0$ be fixed. Suppose that $\mathbf{x}_{k-\tau_k} \notin X^*$. We note from the definition $\mathbf{x}_k$ in DSSM-I that

$$\begin{aligned}
\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 &= \|P_X(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k}) - P_X(\mathbf{x}^*)\|^2 \\
&\leq \|\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k} - \mathbf{x}^*\|^2 \\
&= \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 - 2\alpha_k \langle \mathbf{g}_{k-\tau_k}, \mathbf{x}_k - \mathbf{x}^* \rangle \\
&\leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k \langle \mathbf{g}_{k-\tau_k}, \mathbf{x}_{k-\tau_k} - \mathbf{x}^* \rangle \\
&\leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k L^{-\frac{1}{p}} \left(f(\mathbf{x}_{k-\tau_k}) - f^*\right)^{\frac{1}{p}},
\end{aligned}$$

where the first inequality follows from the nonexpansiveness of metric projection, and the second one follows from the Cauchy–Schwarz inequality and the fact that $\|\mathbf{g}_{k-\tau_k}\| = 1$ for all $k \in \mathbb{N}_0$, and the last one follows from Lemma 2. This completes the proof. □

Lemma 4. *Suppose that Assumption 1 holds. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-I, and let* $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$ *be a nonincreasing sequence with* $\lim_{k\to\infty} \alpha_k = 0$ *and* $\sum_{k=0}^{\infty} \alpha_k = \infty$*. Then* $\|\mathbf{x}_{k-\tau_k} - \mathbf{x}_k\| \leq (\tau + 1)\alpha_{k-\tau}$ *for all* $k \geq \tau$ *and* $\lim_{k\to\infty} \|\mathbf{x}_{k-\tau_k} - \mathbf{x}_k\| = 0$*.*

Proof. We first note that for each $k \in \mathbb{N}_0$

$$\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| = \left\| \sum_{i=1}^{\tau_k} (\mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}) \right\| \leq \sum_{i=1}^{\tau_k} \|\mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}\| \leq \sum_{i=0}^{\tau} \|\mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}\| \tag{3}$$

and

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| = \|P_X(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k}) - P_X(\mathbf{x}_k)\| \leq \alpha_k \|\mathbf{g}_{k-\tau_k}\| = \alpha_k. \tag{4}$$

Let $k \in \mathbb{N}_0$ be such that $k \geq \tau$. Combining (3) and (4) together with the nonincreasing of $\{\alpha_k\}_{k=0}^{\infty}$ yields

$$\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| \leq \sum_{i=0}^{\tau} \|\mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}\| \leq \sum_{i=0}^{\tau} \alpha_{k-i} \leq \sum_{i=0}^{\tau} \alpha_{k-\tau} = (\tau+1)\alpha_{k-\tau}. \tag{5}$$

Furthermore, since $\lim_{k\to\infty} \alpha_k = 0$, we conclude that $\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| = 0$. The proof is complete. □

Now, we are ready to prove the first convergence result in the sense of the inferior limit of the function values of the generated sequence is equal to the optimal value f^*.

Theorem 1. *Suppose that Assumptions 1 and 2 hold. Let $\{\mathbf{x}_k\}_{k=0}^{\infty}$ be a sequence generated by DSSM-I and $\{\alpha_k\}_{k=0}^{\infty}$ be a nonincreasing sequence with $\lim_{k\to\infty} \alpha_k = 0$ and $\sum_{k=0}^{\infty} \alpha_k = \infty$. Then, we have $\liminf_{k\to\infty} f(\mathbf{x}_k) = f^*$.*

Proof. It is obvious that $\liminf_{k\to\infty} f(\mathbf{x}_k) \geq f^*$ since $f(\mathbf{x}_k) \geq f^*$ for all $k \in \mathbb{N}_0$, so that it is sufficient to prove that $\liminf_{k\to\infty} f(\mathbf{x}_k) \leq f^*$. Now, assume to the contrary that $\liminf_{k\to\infty} f(\mathbf{x}_k) > f^*$. Then there exists $\epsilon > 0$ and $k_1 \in \mathbb{N}_0$ such that $f(\mathbf{x}_k) > f^* + \epsilon$ for all $k \geq k_1$. This yields that

$$f(\mathbf{x}_{k-\tau_k}) > f^* + \epsilon \qquad \text{for all } k \geq k_1 + \tau. \tag{6}$$

On the other hand, since $\lim_{k\to\infty} \alpha_k = 0$ and $\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| = 0$ (from Lemma 4), there exists $k_2 \in \mathbb{N}_0$ such that

$$\alpha_k < \left(\frac{1}{3}\right)\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} \text{ and } \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| < \left(\frac{1}{3}\right)\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} \qquad \text{for all } k \geq k_2. \tag{7}$$

Let $\mathbf{x}^* \in X^*$ and let $k \in \mathbb{N}_0$ be such that $k \geq N_1 := \max\{k_1 + \tau, k_2\}$. Then, we obtain from Lemma 3 together with the relations (6) and (7) that for any $k \geq N_1$,

$$\begin{aligned} \|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 &< \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k\left(\frac{1}{3}\right)\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} + 2\alpha_k\left(\frac{1}{3}\right)\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} - 2\alpha_k\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} \\ &= \|\mathbf{x}_k - \mathbf{x}^*\|^2 - \alpha_k\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}}, \end{aligned}$$

which yields

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 < \|\mathbf{x}_{N_1} - \mathbf{x}^*\|^2 - \left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} \sum_{i=N_1}^{k} \alpha_i. \tag{8}$$

Since $\sum_{k=0}^{\infty} \alpha_k = \infty$, we have $\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} \sum_{k=N_1}^{\infty} \alpha_k = \infty$, and it follows that there exists $N_2 \in \mathbb{N}_0$ such that

$$\left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} \sum_{i=N_1}^{k} \alpha_i > \|\mathbf{x}_{N_1} - \mathbf{x}^*\|^2 \qquad \text{for all } k \geq N_2. \tag{9}$$

Thus, for any $k \geq N := \max\{N_1, N_2\}$, we obtain from (8) and (9) that

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 < \|\mathbf{x}_N - \mathbf{x}^*\|^2 - \left(\frac{\epsilon}{L}\right)^{\frac{1}{p}} \sum_{i=N}^{k} \alpha_i < 0,$$

which is a contradiction. Therefore, we can conclude that $\liminf_{k\to\infty} f(\mathbf{x}_k) \leq f^*$, and hence $\liminf_{k\to\infty} f(\mathbf{x}_k) = f^*$, as desired. □

Corollary 1. *Suppose that Assumptions 1 and 2 hold. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-I, and let* $\{\alpha_k\}_{k=0}^{\infty}$ *be a nonincreasing sequence with* $\lim_{k\to\infty} \alpha_k = 0$ *and* $\sum_{k=0}^{\infty} \alpha_k = \infty$. *If the sequence* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *is bounded, then there exists a subsequence of* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *that converges to an optimal solution in* X^*.

Proof. We first note from the fact that $\liminf_{k\to\infty} f(\mathbf{x}_k) = f^*$ in Theorem 1 together with the boundedness of the sequence $\{\mathbf{x}_k\}_{k=0}^{\infty}$ and the continuity of f that there exists a subsequence $\{\mathbf{x}_{k_i}\}_{i=0}^{\infty}$ of $\{\mathbf{x}_k\}_{k=0}^{\infty}$ such that

$$\lim_{i\to\infty} f(\mathbf{x}_{k_i}) = \liminf_{k\to\infty} f(\mathbf{x}_k) = f^*. \tag{10}$$

Since the sequence $\{\mathbf{x}_{k_i}\}_{i=0}^{\infty}$ is also bounded, there exists a subsequence $\{\mathbf{x}_{k_{i_j}}\}_{j=0}^{\infty}$ of $\{\mathbf{x}_{k_i}\}_{i=0}^{\infty}$ such that $\lim_{j\to\infty} \mathbf{x}_{k_{i_j}} = \bar{\mathbf{x}} \in \mathbb{R}^n$. Again, the continuity of f and using (10) yield

$$f(\bar{\mathbf{x}}) = \lim_{j\to\infty} f(\mathbf{x}_{k_{i_j}}) = \lim_{i\to\infty} f(\mathbf{x}_{k_i}) = f^*.$$

Note that since $\{\mathbf{x}_{k_{i_j}}\}_{j=0}^{\infty} \subset X$, the closedness of X yields $\bar{\mathbf{x}} \in X$, which therefore implies that $\bar{\mathbf{x}}$ is an element in X^*. □

By imposing either the coercivity of the objective function f or the compactness of the constraint X, we can obtain an approximate convergence behavior and some nice convergence results in comparison to the one obtained in Theorem 1.

Theorem 2. *Suppose that Assumptions 1 and 2 hold. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-I, and let* $\{\alpha_k\}_{k=0}^{\infty}$ *be a nonincreasing sequence with* $\lim_{k\to\infty} \alpha_k = 0$ *and* $\sum_{k=0}^{\infty} \alpha_k = \infty$. *If the function* f *is coercive or the constrained set* X *is compact, then the following statements hold:*

(*i*) *For any* $\sigma > 0$, *there exists* $N_\sigma \in \mathbb{N}_0$ *such that*

$$\operatorname{dist}(\mathbf{x}_k, X^*) \leq \rho(\sigma) + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3} \quad \text{for all } k \geq N_\sigma,$$

where $\rho(\sigma) := \max\Big\{\operatorname{dist}(\mathbf{x}, X^*) : \mathbf{x} \in X \cap S^{\leq}_{\bar{f}, f^*+L\sigma}\Big\}$;

(*ii*) $\lim_{k\to\infty} \operatorname{dist}(\mathbf{x}_k, X^*) = 0$, *and* $\lim_{k\to\infty} f(\mathbf{x}_k) = f^*$.

Proof. Let $\sigma > 0$ be arbitrary and denote the notation

$$\rho(\sigma) := \max\Big\{\operatorname{dist}(\mathbf{x}, X^*) : \mathbf{x} \in X \cap S^{\leq}_{\bar{f}, f^*+L\sigma}\Big\}.$$

Since the function f is coercive or the constrained set X is compact, we ensure that $X^* \neq \emptyset$ which together with $X^* \subseteq X \cap S^{\leq}_{\bar{f}, f^*+L\sigma}$ yield that $X \cap S^{\leq}_{\bar{f}, f^*+L\sigma} \neq \emptyset$. Moreover, the continuity and quasi-convexity of f, respectively, imply that $S^{\leq}_{\bar{f}, f^*+L\sigma}$ is closed and convex, and so the set $X \cap S^{\leq}_{\bar{f}, f^*+L\sigma}$ is also closed and convex. Again, with either the compactness of X or the coercivity of f, the intersection $X \cap S^{\leq}_{\bar{f}, f^*+L\sigma}$ is bounded, and hence, $X \cap S^{\leq}_{\bar{f}, f^*+L\sigma}$ is a compact set, which ensures that $\rho(\sigma) < \infty$.

(i) Let $\mathbf{x}^* \in X^*$ be given. Since $\lim_{k\to\infty} \alpha_k = 0$, we also have $\lim_{k\to\infty} \alpha_{k-\tau} = 0$. Consequently, there exists $k_\sigma \in \mathbb{N}_0$ such that $k_\sigma \geq \tau$ and

$$\alpha_{k-\tau} \leq \frac{\sigma^{\frac{1}{p}}}{2\tau+3} \qquad \text{for all } k \geq k_\sigma. \tag{11}$$

This together with the relation that $\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| \leq (\tau+1)\alpha_{k-\tau}$ which is obtained in Lemma 4 yield that

$$\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| \leq \frac{(\tau+1)\sigma^{\frac{1}{p}}}{2\tau+3} \qquad \text{for all } k \geq k_\sigma. \tag{12}$$

Now, we will show that

$$\operatorname{dist}(\mathbf{x}_{k+1}, X^*) \leq \max\left\{\operatorname{dist}(\mathbf{x}_k, X^*), \rho(\sigma) + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3}\right\} \qquad \text{for all } k \geq k_\sigma. \tag{13}$$

Let $k \geq k_\sigma$ be fixed. We divide the proof into two cases according to the behavior of $f(\mathbf{x}_{k-\tau_k})$.

Case 1: Suppose that $f(\mathbf{x}_{k-\tau_k}) > f^* + L\sigma$. Then we have $\mathbf{x}_{k-\tau_k} \notin X^*$. Applying Lemma 3 and the nonincreasing property of the stepsize $\{\alpha_k\}_{k=0}^{\infty}$, we get

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k\alpha_{k-\tau} + 2\alpha_k\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k\sigma^{\frac{1}{p}}. \tag{14}$$

Substituting the inequalities (11) and (12) in (14) gives

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \frac{\sigma^{\frac{1}{p}}\alpha_k}{2\tau+3} + \frac{2(\tau+1)\sigma^{\frac{1}{p}}\alpha_k}{2\tau+3} - 2\alpha_k\sigma^{\frac{1}{p}} \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2.$$

By putting $\mathbf{x}^* = P_{X^*}(\mathbf{x}_k)$, we obtain $\operatorname{dist}(\mathbf{x}_{k+1}, X^*) \leq \operatorname{dist}(\mathbf{x}_k, X^*)$.

Case 2: Suppose that $f(\mathbf{x}_{k-\tau_k}) \leq f^* + L\sigma$. Then we have $\mathbf{x}_{k-\tau_k} \in X \cap S^{\leq}_{f, f^*+L\sigma}$ and $\operatorname{dist}(\mathbf{x}_{k-\tau_k}, X^*) \leq \rho(\sigma)$. We note from the definition of $\mathbf{x}_{k+1} = P_X(\mathbf{x}_k - \alpha_k\mathbf{g}_{k-\tau_k})$ the nonexpansiveness of P_X and the nonincreasing property of $\{\alpha_k\}_{k=0}^{\infty}$ that

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq \|\mathbf{x}_k - \mathbf{x}^*\| + \alpha_k \leq \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| + \|\mathbf{x}_{k-\tau_k} - \mathbf{x}^*\| + \alpha_{k-\tau},$$

which together with the inequality (11) and (12) yield

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq \|\mathbf{x}_{k-\tau_k} - \mathbf{x}^*\| + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3}.$$

By putting $\mathbf{x}^* := P_{X^*}(\mathbf{x}_{k-\tau_k})$ and invoking $\operatorname{dist}(\mathbf{x}_{k-\tau_k}, X^*) \leq \rho(\sigma)$, we obtain

$$\operatorname{dist}(\mathbf{x}_{k+1}, X^*) \leq \rho(\sigma) + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3}.$$

Thus, we proved the relation (13). By using the same analogy of Theorem 1, we obtain that $\liminf_{k\to\infty} f(\mathbf{x}_{k-\tau_k}) = f^*$. Consequently, there exists k'_σ such that $k'_\sigma \geq k_\sigma$ and $f(\mathbf{x}_{k'_\sigma-\tau_{k'_\sigma}}) < f^* + L\sigma^{\frac{1}{p}}$. Therefore, we obtain from **Case** 2 that

$$\operatorname{dist}(\mathbf{x}_{k'_\sigma+1}, X^*) \leq \rho(\sigma) + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3}.$$

Put $N_\sigma := k'_\sigma + 1$. Suppose that $\operatorname{dist}(\mathbf{x}_k, X^*) \leq \rho(\sigma) + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3}$ for all $k \geq N_\sigma$. Invoking (13), we obtain $\operatorname{dist}(\mathbf{x}_{k+1}, X^*) \leq \rho(\sigma) + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3}$. Hence, we conclude by the strong induction that

$$\operatorname{dist}(\mathbf{x}_k, X^*) \leq \rho(\sigma) + \frac{(\tau+2)\sigma^{\frac{1}{p}}}{2\tau+3} \qquad \text{for all } k \geq N_\sigma.$$

(ii) For each $r \in \mathbb{N}$, we denote $X_{\frac{1}{r}} := X \cap S^{<}_{f,f^*+L\left(\frac{1}{r}\right)}$. Observe that $\left\{X_{\frac{1}{r}}\right\}_{r=1}^{\infty}$ is nonincreasing and $\bigcap_{r\in\mathbb{N}} X_{\frac{1}{r}} = X \cap S^{\leq}_{f,f^*} = X^*$. Subsequently, $\left\{\rho\left(\frac{1}{r}\right)\right\}_{r\in\mathbb{N}}$ is also nonincreasing and $\inf_{r\in\mathbb{N}} \rho\left(\frac{1}{r}\right) = \inf_{r\in\mathbb{N}} \max_{\mathbf{x}\in X_{\frac{1}{r}}} \operatorname{dist}(\mathbf{x}, X^*) = \max_{\mathbf{x}\in X^*} \operatorname{dist}(\mathbf{x}, X^*) = 0$. Hence, we obtain that $\lim_{r\to\infty} \rho\left(\frac{1}{r}\right) = 0$.

Now, let $\varepsilon > 0$ be arbitrary. Since $\lim_{r\to\infty}\left(\rho\left(\frac{1}{r}\right) + \frac{(\tau+2)\left(\frac{1}{r}\right)^{\frac{1}{p}}}{2\tau+3}\right) = 0$, there exists $M_\varepsilon \in \mathbb{N}$ such that

$$\rho\left(\frac{1}{r}\right) + \frac{(\tau+2)\left(\frac{1}{r}\right)^{\frac{1}{p}}}{2\tau+3} < \varepsilon \qquad \text{for all } r \geq M_\varepsilon.$$

Since $\frac{1}{M_\varepsilon} > 0$, there exists $N \in \mathbb{N}_0$ such that

$$\operatorname{dist}(\mathbf{x}_k, X^*) \leq \rho\left(\frac{1}{M_\varepsilon}\right) + \frac{(\tau+2)\left(\frac{1}{M_\varepsilon}\right)^{\frac{1}{p}}}{2\tau+3} < \varepsilon \qquad \text{for all } k \geq N.$$

Hence, we conclude that $\lim_{k\to\infty} \operatorname{dist}(\mathbf{x}_k, X^*) = 0$. Furthermore, by Assumption 2, we therefore obtain that $\lim_{k\to\infty} f(\mathbf{x}_k) = f^*$. □

Remark 3. *Under the setting of Theorem 2, we observe that if an optimal solution is unique, that is, $X^* = \{\mathbf{x}^*\}$, we obtain that $\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}^*\| = \lim_{k\to\infty} \operatorname{dist}(\mathbf{x}_k, X^*) = 0$, which means that the generated sequence $\{\mathbf{x}_k\}_{k=0}^{\infty}$ converges to an unique solution $\mathbf{x}^*$.*

3.2. Delayed Star Subgradient Method II

In order to obtain the convergence of the generated sequence, we need a modified version of DDSM-I which is defined as the following Algorithm 2.

Algorithm 2 Delayed Star Subgradient Method II (DSSM-II)

Initialization: Given a stepsize $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$, the delays $\{\tau_k\}_{k=0}^{\infty} \subset \mathbb{N}_0$, and initial points $\mathbf{x}_0, \mathbf{x}_{-1}, \mathbf{x}_{-2}, \ldots, \mathbf{x}_{-\tau} \in X$.
Iterative Step: For a current point $\mathbf{x}_k \in X$, if $f(\mathbf{x}_k) = f^*$, we set

$$\mathbf{x}_{k+1} := \mathbf{x}_k$$

and then stop. If $f(\mathbf{x}_k) > f^*$, we compute

$$\mathbf{x}_{k+1} := P_X(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k}),$$

where $\mathbf{g}_{k-\tau_k} \in \partial^\star f(\mathbf{x}_{k-\tau_k}) \cap \mathcal{S}$ is a unit star subgradient of f at $\mathbf{x}_{k-\tau_k}$.
Update: $k := k + 1$.

Remark 4.

(*i*) *If the delays* $\tau_k = 0$ *for all* $k \in \mathbb{N}_0$, *DSSM-II is relating to the method proposed by Hu et al. [9] (Algorithm* 1*) with the special setting of* $m = 1$.

(*ii*) *Note that DSSM-II involves evaluating the function value at the current iterate* $\mathbf{x}_k \in X$ *and deciding to update the next iteration* $\mathbf{x}_{k+1}$. *If the function value* $f(\mathbf{x}_k)$ *equals the optimal value* f^*, *then no additional calculations are performed, and DSSM-II terminates so that an optimal solution is obtained at the point* $\mathbf{x}_k$.

Next, we will derive some fundamental properties for the convergence of DSSM-II.

Lemma 5. *Suppose that Assumption 2 holds. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-II. For any* $k \in \mathbb{N}_0$ *and* $\mathbf{x}^* \in X^*$, *if* $f(\mathbf{x}_k) > f^*$, *then*

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k L^{-\frac{1}{p}}\left(f(\mathbf{x}_{k-\tau_k}) - f^*\right)^{\frac{1}{p}}.$$

Proof. Let $\mathbf{x}^* \in X^*$ and $k \in \mathbb{N}_0$. Suppose that $f(\mathbf{x}_k) > f^*$. We have from the definition of DSSM-II that $\mathbf{x}_{k+1} = P_X(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k})$. Moreover, since $f(\mathbf{x}_k) > f^*$, we also have $f(\mathbf{x}_{k-\tau_k}) > f^*$. Thus, by following the proving lines of Lemma 3, we obtain

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k L^{-\frac{1}{p}}\left(f(\mathbf{x}_{k-\tau_k}) - f^*\right)^{\frac{1}{p}},$$

as desired. □

Theorem 3. *Suppose that Assumptions 1 and 2 hold. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-II, and let* $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$ *be a nonincreasing sequence with* $\lim_{k\to\infty} \alpha_k = 0$ *and* $\sum_{k=0}^{\infty} \alpha_k = \infty$. *Then, we have* $\|\mathbf{x}_{k-\tau_k} - \mathbf{x}_k\| \leq (\tau + 1)\alpha_{k-\tau}$ *for all* $k \geq \tau$, $\lim_{k\to\infty} \|\mathbf{x}_{k-\tau_k} - \mathbf{x}_k\| = 0$, *and* $\liminf_{k\to\infty} f(\mathbf{x}_k) = f^*$.

Proof. By invoking Lemma 5, we can prove the same analog of Lemma 4 and Theorem 1 to obtain $\liminf_{k\to\infty} f(\mathbf{x}_k) = f^*$. □

Theorem 4. *Suppose that Assumptions 1 and 2 hold. Let* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *be a sequence generated by DSSM-II, and let* $\{\alpha_k\}_{k=0}^{\infty} \subset (0, \infty)$ *be a nonincreasing sequence with* $\sum_{k=0}^{\infty} \alpha_k = \infty$ *and* $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$. *Then the sequence* $\{\mathbf{x}_k\}_{k=0}^{\infty}$ *converges to an optimal solution in* X^* *and, moreover,* $\lim_{k\to\infty} f(\mathbf{x}_k) = f^*$.

Proof. If there exists $K \in \mathbb{N}_0$ such that $f(\mathbf{x}_K) = f^*$, we have that $\mathbf{x}_K$ is an optimal solution and $\mathbf{x}_k = \mathbf{x}_K$ for all $k \geq K$. Consequently, we obtain $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}_K \in X^*$ and $\lim_{k\to\infty} f(\mathbf{x}_k) = f^*$ as required.

On the other hand, we suppose that $f(\mathbf{x}_k) > f^*$ for all $k \in \mathbb{N}_0$. Let $k \in \mathbb{N}_0$ and let $\mathbf{x}^* \in X^*$. We note from Lemma 5 that

$$\begin{aligned}\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 &\leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k L^{-\frac{1}{p}}\left(f(\mathbf{x}_{k-\tau_k}) - f^*\right)^{\frac{1}{p}} \\ &\leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\|.\end{aligned}$$

Let us admit from the inequalities (3) and (4) that $\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| \leq \sum_{i=0}^{\tau} \|\mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}\|$ and $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq \alpha_k$ for all $k \in \mathbb{N}_0$. Thus, for a fixed $N \in \mathbb{N}_0$, we have

$$\begin{aligned}\sum_{k=0}^{N} \alpha_k\|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| &\leq \sum_{k=0}^{N} \alpha_k \sum_{i=0}^{\tau} \|\mathbf{x}_{k-i+1} - \mathbf{x}_{k-i}\| \\ &= \sum_{i=0}^{\tau} \sum_{k=-i}^{N-i} \alpha_{k+i}\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq \sum_{i=0}^{\tau} \sum_{k=0}^{N} \alpha_{k+i}\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq \sum_{i=0}^{\tau} \sum_{k=0}^{N} \alpha_{k+i}\alpha_k,\end{aligned}$$

which together with the nonincreasing property of $\{\alpha_k\}_{k=0}^{\infty}$ implies

$$\sum_{k=0}^{N} \alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| \leq \sum_{i=0}^{\tau} \sum_{k=0}^{N} \alpha_k^2 = (\tau+1) \sum_{k=0}^{N} \alpha_k^2. \tag{15}$$

By approaching $N \to \infty$ in (15) and using the assumption that $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$, we obtain that $\sum_{k=0}^{\infty} \alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| < \infty$, and so

$$\sum_{k=0}^{\infty} \left(\alpha_k^2 + 2\alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| \right) < \infty.$$

Therefore, the sequence $\{x_k\}_{k=0}^{\infty}$ is quasi-Fejer monotone. By applying Fact 2, we obtain that the sequence $\{\mathbf{x}_k\}_{k=0}^{\infty}$ is bounded and $\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}^*\|$ exists.

Since the sequence $\{\mathbf{x}_k\}_{k=0}^{\infty}$ is bounded and $\liminf_{k\to\infty} f(\mathbf{x}_k) = f^*$ (Theorem 3), there exists a subsequence $\{\mathbf{x}_{k_i}\}_{i=0}^{\infty}$ of $\{\mathbf{x}_k\}_{k=0}^{\infty}$ such that

$$\lim_{i\to\infty} f(\mathbf{x}_{k_i}) = \liminf_{k\to\infty} f(\mathbf{x}_k) = f^*.$$

Since $\{\mathbf{x}_{k_i}\}_{i=0}^{\infty}$ is also bounded, there exists a subsequence $\{\mathbf{x}_{k_{i_j}}\}_{j=0}^{\infty}$ of $\{\mathbf{x}_{k_i}\}_{i=0}^{\infty}$ such that $\lim_{j\to\infty} \mathbf{x}_{k_{i_j}} = \bar{\mathbf{x}} \in \mathbb{R}^n$. Thus, the continuity of f yields

$$f(\bar{\mathbf{x}}) = \lim_{j\to\infty} f(\mathbf{x}_{k_{i_j}}) = \lim_{i\to\infty} f(\mathbf{x}_{k_i}) = f^*.$$

Moreover, the closeness of X implies $\bar{\mathbf{x}} \in X$, and hence we obtain that $\bar{\mathbf{x}} \in X^*$. Since $\lim_{k\to\infty} \|\mathbf{x}_k - \bar{\mathbf{x}}\|$ exists and $\lim_{j\to\infty} \|\mathbf{x}_{k_{i_j}} - \bar{\mathbf{x}}\| = 0$, we conclude that $\lim_{k\to\infty} \|\mathbf{x}_k - \bar{\mathbf{x}}\| = 0$. Moreover, invoking the continuity of f, we also have that $\lim_{k\to\infty} f(\mathbf{x}_k) = f^*$. The proof is complete. □

We end the convergence analysis of DSSM-II by investigating the finite convergence of the generated sequence as the following theorem.

Theorem 5. *Suppose that Assumptions 1 and 2 hold and $\mathcal{B}(\mathbf{x}^*;\delta) \subset X^*$ for some $\delta > 0$. Let $\{\mathbf{x}_k\}_{k=0}^{\infty}$ be a sequence generated by DSSM-II. If one of the following statements holds:*

(i) $\alpha_k = \alpha \in \left(0, \frac{2\delta}{2\tau+3}\right)$ *for all* $k \in \mathbb{N}_0$;

(ii) $\{\alpha_k\}_{k=0}^{\infty}$ *is nonincreasing with* $\lim_{k\to\infty} \alpha_k = 0$ *and* $\sum_{k=0}^{\infty} \alpha_k = \infty$,

then there exists $K \in \mathbb{N}_0$ *such that* $\mathbf{x}_K \in X^*$.

Proof. We suppose by contradiction that $\mathbf{x}_k \notin X^*$ for all $k \in \mathbb{N}_0$. This implies $\mathbf{x}_{k-\tau_k} \notin X^*$ for all $k \in \mathbb{N}_0$. Note that $\|\mathbf{x}^* + \delta \mathbf{g}_{k-\tau_k} - \mathbf{x}^*\| \leq \delta$, which implies that $\mathbf{x}^* + \delta \mathbf{g}_{k-\tau_k} \in \mathcal{B}(\mathbf{x}^*;\delta) \subset X^*$. Since $\mathbf{x}_{k-\tau_k} \notin X^*$, we have $f(\mathbf{x}^* + \delta \mathbf{g}_{k-\tau_k}) < f(\mathbf{x}_{k-\tau_k})$. Thus, by the definition of the star subgradient of f at $\mathbf{x}_{k-\tau_k}$, we obtain that

$$\langle \mathbf{g}_{k-\tau_k}, \mathbf{x}^* + \delta \mathbf{g}_{k-\tau_k} - \mathbf{x}_{k-\tau_k} \rangle \leq 0,$$

which yields

$$\langle \mathbf{g}_{k-\tau_k}, \mathbf{x}_{k-\tau_k} - \mathbf{x}^* \rangle \geq \delta. \tag{16}$$

On the other hand, since $\mathbf{x}_k \notin X^*$, we have from the definition of DSSM-II that $\mathbf{x}_{k+1} = P_X(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k})$, and the nonexpansive property of P_X implies that

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k \langle \mathbf{g}_{k-\tau_k}, \mathbf{x}_{k-\tau_k} - \mathbf{x}^* \rangle,$$

which combining (16), leads to

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\|^2 \leq \|\mathbf{x}_k - \mathbf{x}^*\|^2 + \alpha_k^2 + 2\alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\alpha_k \delta. \tag{17}$$

Let N be a fixed non-negative integer. By summing up (17) from $k = 0$ to $k = N$, we have

$$0 \leq \|\mathbf{x}_0 - \mathbf{x}^*\|^2 + \sum_{k=0}^{N} \alpha_k^2 + 2 \sum_{k=0}^{N} \alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| - 2\delta \sum_{k=0}^{N} \alpha_k,$$

which implies

$$\delta \leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2\sum_{k=0}^{N} \alpha_k} + \frac{\sum_{k=0}^{N} \alpha_k^2}{2\sum_{k=0}^{N} \alpha_k} + \frac{\sum_{k=0}^{N} \alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\|}{\sum_{k=0}^{N} \alpha_k}. \tag{18}$$

Based on the conditions (i) and (ii), we will divide our consideration into two cases.
Case 1: Assume that $\alpha_k = \alpha \in (0, \frac{2\delta}{2\tau+3})$ for all $k \in \mathbb{N}_0$.
Let us note from (15) that

$$\sum_{k=0}^{N} \alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| \leq \sum_{i=0}^{\tau} \sum_{k=0}^{N} \alpha_k^2 = (\tau+1)(N+1)\alpha^2.$$

Thus, we have

$$\begin{aligned} & \lim_{N\to\infty} \left(\frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2\sum_{k=0}^{N} \alpha_k} + \frac{\sum_{k=0}^{N} \alpha_k^2}{2\sum_{k=0}^{N} \alpha_k} + \frac{\sum_{k=0}^{N} \alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\|}{\sum_{k=0}^{N} \alpha_k} \right) \\ \leq & \lim_{N\to\infty} \left(\frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2(N+1)\alpha} + \frac{\alpha}{2} + (\tau+1)\alpha \right) \\ = & \frac{(2\tau+3)\alpha}{2} < \delta, \end{aligned}$$

which is a contradiction to the inequality (18).

Case 2: Assume that $\{\alpha_k\}_{k=0}^{\infty}$ is a nonincreasing sequence with $\lim_{k\to\infty} \alpha_k = 0$ and $\sum_{k=0}^{\infty} \alpha_k = \infty$. We obtain from Theorem 3 that $\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| = 0$. Thus, we finally obtain from the fact that $\lim_{k\to\infty} \alpha_k = 0$ and $\sum_{k=0}^{\infty} \alpha_k = \infty$ and $\lim_{k\to\infty} \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\| = 0$ with Lemma 1 that

$$\lim_{N\to\infty} \left(\frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2\sum_{k=0}^{N} \alpha_k} + \frac{\sum_{k=0}^{N} \alpha_k^2}{2\sum_{k=0}^{N} \alpha_k} + \frac{\sum_{k=0}^{N} \alpha_k \|\mathbf{x}_k - \mathbf{x}_{k-\tau_k}\|}{\sum_{n=0}^{N} \alpha_k} \right) = 0,$$

which also contradicts to the inequality (18). □

4. Numerical Experiments

In this section, we present experimental results when we apply DSSM-I to solve the Cobb–Douglas production efficiency problem. We performed all experiments on MATLAB (R2023b) on a MacBook Air 13.3-inch with Apple M1 chip processor and 8GB memory.

The Cobb–Douglas production efficiency problem [3,6,12] aims to maximize the ratio of the total profit represented by the Cobb–Douglas product function and the total cost represented by the linear function on the product factor subject to the funding level limitations constraint. More precisely, let $a_0, c_0 > 0$ and let $a_j, b_{ij}, c_j, p_i \geq 0$ for all $j = 1, 2, \ldots, n$ and $i = 1, 2, \ldots, m$, and the Cobb–Douglas production efficiency problem is defined by

$$\begin{array}{ll} \text{maximize} & f(\mathbf{x}) := \frac{a_0 \prod_{j=1}^{n} x_j^{a_j}}{\sum_{j=1}^{n} c_j x_j + c_0} \\ \text{subject to} & \sum_{j=1}^{n} b_{ij} x_j \geq p_i, i = 1, 2, \ldots, m \\ & x_j \geq 0, j = 1, 2, \ldots, n, \end{array} \tag{19}$$

where $\sum_{j=1}^{n} a_j = 1$. Here, $\mathbf{x} := (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ when x_j is the j-th product factor $(j = 1, 2, \ldots, n)$, and for each $i = 1, 2, \ldots, m$, p_i is the profit to be obtained from project i, and b_{ij} is the support of product factor j to project i to achieve p_i. According to the literature [6,12], we know that the Cobb–Douglas production efficiency problem (19) is a quasi-convex optimization problem. Note that the function f is continuous and quasi-concave [4] (Theorem 2.5.1), satisfying the Hölder condition on $\mathbb{R}^n$ [12] (Appendix B), and the constraint set $C := \bigcap_{i=1}^{m} \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{b}_i, \mathbf{x} \rangle \geq p_i\} \cap [0, \infty)^n$ is the nonempty closed convex set, where $\mathbf{b}_i := (b_{i1}, \ldots, b_{in})^T$ is a vector in $\mathbb{R}^n$. Therefore, we can apply DSSM-I to solve the problem (19). However, since the exact optimal solution to problem (19) is not known, we are unable to apply DSSM-II to solve it. To ensure the boundedness of the constrained set, we set the constrained set to be the set $D := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{b}_i, \mathbf{x} \rangle \geq p_i\} \cap [0.001, 100]^n$ which is a subset of C. Thus, the established convergence result from Theorem 2 guarantees that $\lim_{k\to\infty} f(\mathbf{x}_k) = f^*$. It is worth noting that the computational procedure of the metric projection onto the constrained set D is not an easy task since there is no explicit form of P_D in general. To deal with the situation, we utilized the classical Halpern iteration by performing the following inner loop: for arbitrary initial point $\mathbf{u}_l \in \mathbb{R}^n$ and a sequence $\{\lambda_l\}_{l=0}^{\infty} \subseteq (0, 1)$ such that $\lim_{l\to\infty} \lambda_l = 0, \sum_{l=0}^{\infty} \lambda_l = \infty$, and $\sum_{l=1}^{\infty} |\lambda_{l+1} - \lambda_l| < \infty$, we compute

$$\mathbf{u}_{l+1} := \lambda_l(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k}) + (1 - \lambda_l) P_{D_{m+1}} P_{D_m} P_{D_{m-1}} \cdots P_{D_1} \mathbf{u}_l \text{ for all } l \geq 1,$$

where $D_{m+1} = [0.001, 100]^n$, $D_i := \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{b}_i, \mathbf{x} \rangle \geq p_i\}$ for all $i = 1, 2, \ldots, m$. Then the sequence $\{\mathbf{u}_l\}_{l=0}^{\infty}$ converges to the unique point $P_D(\mathbf{x}_k - \alpha_k \mathbf{g}_{k-\tau_k})$, see [28] (Theorem 30.1) for further details. In our experiment, we put the inner initial point $\mathbf{u}_0$ as the vector whose all coordinates are 1 and set $\lambda_l := \frac{1}{l+2}$ for all $l \in \mathbb{N}_0$. We use the stopping criterion $\frac{\|\mathbf{u}_{l+1} - \mathbf{u}_l\|}{\|\mathbf{u}_l\| + 1} \leq 10^{-6}$ for the inner loop.

To perform the numerical experiments, we randomly set the parameter $a_j, b_{ij} \in (0, 1), a_0, c_0, c_j \in (0, 10)$, and $p_i \in (0, \frac{n}{2})$. We choose the stepsize and time-varying delays to be $\alpha_k := \frac{1}{k+1}$ and $\tau_k := k \mod (\tau + 1)$ for all $k \in \mathbb{N}_0$, respectively. We put the initial point $\mathbf{x}_0$ as the vector whose all coordinates are 1. To explain the influence of the delay bound τ, we perform Algorithm DSSM-I for delay bounds $\tau = 0, 1, 3, 5$, and 10 for various sizes of the problem n and m. Note that DSSM-I with $\tau = 0$ is nothing else than the star subgradient method (SSM) proposed by Kiwiel [1]. Figure 1 illustrates the number of subgradient calculations performed by SSM and DSSM-I with the delay strategies $\tau_k := k \mod (\tau + 1)$ for all $k \in \mathbb{N}_0$.

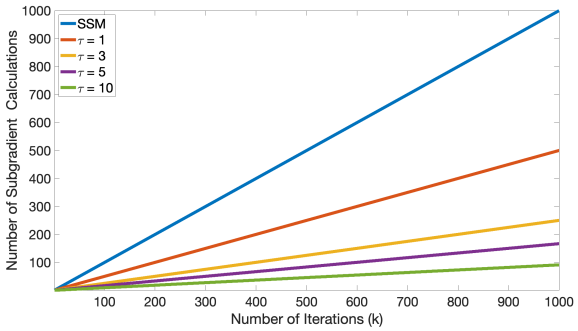
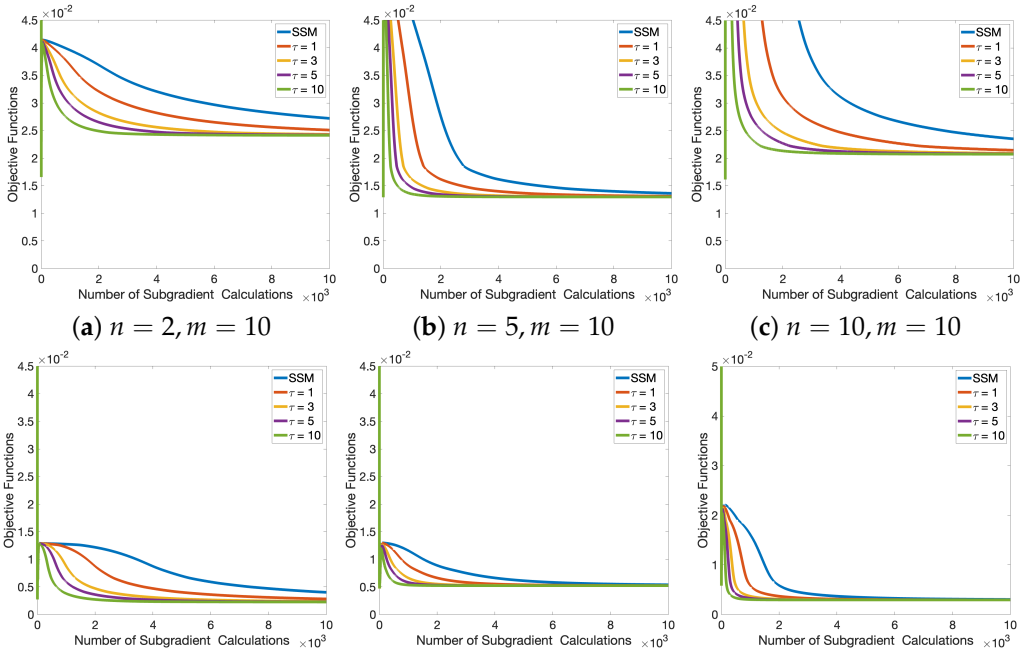


Figure 1. Behaviors of a number of subgradint calculations for various delays bounds τ versus number of iterations.

The average results of the 10 random data sets are presented in Figures 2 and 3 below.



(**a**) $n = 2, m = 10$ (**b**) $n = 5, m = 10$ (**c**) $n = 10, m = 10$

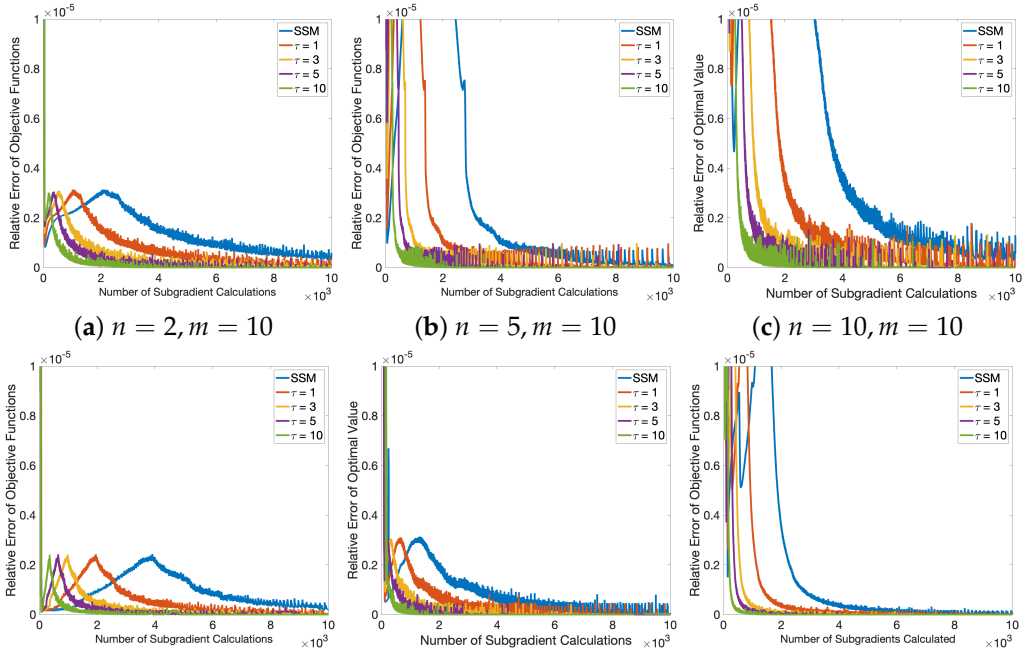
(**d**) $n = 2, m = 100$ (**e**) $n = 5, m = 100$ (**f**) $n = 10, m = 100$

Figure 2. Behaviors of objective function values for various delays bounds τ and problem sizes n and m.

It can be observed from Figure 2 that, overall, all the results show a similar pattern as they approach the (approximated) optimal value as supported by Theorem 2. For each subfigure, we notice that the result of $\tau = 10$ seems stable and decreases faster than other τ.

Next, we plot the relative errors of the objective functions values in the same setting as above in Figure 3.

It can be seen from Figure 3 that all results seem to decrease and stabilize. More precisely, the graph exhibits fluctuations that are alternating decreases and increases in small periods. These fluctuations gradually diminish and vanish to stabilize. Again, this behavior satisfies the convergence result of DSSM-I, and we notice that the results with $\tau = 10$ seems to decrease rapidly to 0 compared to other values of τ.



(**a**) $n = 2, m = 10$ (**b**) $n = 5, m = 10$ (**c**) $n = 10, m = 10$

(**d**) $n = 2, m = 100$ (**e**) $n = 5, m = 100$ (**f**) $n = 10, m = 100$

Figure 3. Behavior of relative errors for various delay bounds τ and problem sizes n and m.

5. Conclusions

In this work, we proposed the so-called projected star subgradient methods with delayed subgradient updates, namely DSSM-I and DSSM-II, for solving the constrained quasi-convex optimization problem. For DSSM-I, we proved that there exists a subsequence of the generated sequence that converges to an optimal solution provided that the boundedness of the generated sequence is imposed. Furthermore, under the compactness or the coercive assumption with the diminishing step size, we proved that the distance between the sequence and the optimal solution set X^* converges to zero, resulting in the convergence of the objective function values to the optimal value. For the modified version as DSSM-II and without requiring the compactness or coercive assumptions, we proved that the generated sequence converges to an optimal solution. We also established the finite convergence provided that the interior of the optimal solution set X^* is nonempty. Finally, we applied DSSM-I to solve the Cobb–Douglas production efficiency problem. It is worth noting that the proposed methods can reduce the computational costs of star subgradients by allowing the method to retrieve stale data instead. It is noted that since we could not indicate the precise upper bound on the number of times the star subgradient could be reused, the open problem with this topic still remains. Moreover, one can notice that we require the closed-form expression of the metric projection onto the constraint set X and overcome this obstacle by solving the sub-problem to approximate P_X at each iteration. This also motivated us to consider a future direction that involves a method that does not require solving a sub-problem.

Author Contributions: Conceptualization, O.P. and N.N.; methodology, O.P. and N.N.; software, O.P.; validation, O.P. and N.N.; formal analysis, O.P. and N.N.; investigation, O.P. and N.N.; writing—original draft preparation, O.P.; writing—review and editing, O.P. and N.N.; visualization, O.P.; supervision, N.N.; project administration, N.N.; funding acquisition, N.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Fundamental Fund of Khon Kaen University. This research has received funding support from the National Science, Research, and Innovation Fund or NSRF. O. Pankoon was supported by the Development and Promotion of Science and Technology Talents Project (DPST).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Dataset available on request from the authors.

Acknowledgments: The authors are thankful to the editor and two anonymous referees for comments and remarks which improved the quality and presentation of this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kiwiel, K.C. Convergence and efficiency of subgradient methods for quasiconvex minimization. *Math. Program.* **2001**, *90*, 1–25. [CrossRef]
2. Cambini, A.; Martein, L. *Generalized Convexity and Optimization*; Springer: Berlin/Heidelberg, Germany, 2008.
3. Bradley, S.P.; Frey, S.C. Fractional programming with homogeneous functions. *Oper. Res.* **1974**, *22*, 350–357. [CrossRef]
4. Stancu-Minasian, I.M. *Fractional Programming*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1997.
5. Schaible, S.; Shi, J. Fractional programming: The sum-of-ratios case. *Optim. Methods Softw.* **2003**, *18*, 219–229. [CrossRef]
6. Hu, Y.; Yang, X.; Sim, C.-K. Inexact subgradient methods for quasi-convex optimization problems. *Eur. J. Oper. Res.* **2015**, *240*, 315–327. [CrossRef]
7. Hu, Y.; Yu, C.K.W.; Li, C. Stochastic subgradient method for quasi-convex optimization problems. *J. Nonlinear Convex Anal.* **2016**, *17*, 711–724.
8. Hu, Y.; Yu, C.K.W.; Li, C.; Yang, X. Conditional subgradient methods for constrained quasi-convex optimization problems. *J. Nonlinear Convex Anal.* **2016**, *17*, 2143–2158.
9. Hu, Y.; Yu, C.K.W.; Yang, X. Incremental quasi-subgradient methods for minimizing the sum of quasi-convex functions. *J. Glob. Optim.* **2019**, *75*, 1003–1028. [CrossRef]
10. Konnov, I.V. On convergence properties of a subgradient method. *Optim. Methods Softw.* **2003**, *18*, 53–62. [CrossRef]
11. Konnov, I.V. On properties of supporting and quasi-supporting vectors. *J. Math. Sci.* **1994**, *71*, 2760–2763. [CrossRef]
12. Hishinuma, K.; Iiduka, H. Fixed point quasiconvex subgradient method. *Eur. J. Oper. Res.* **2020**, *282*, 428–437. [CrossRef]
13. Choque, J.; Lara, F.; Marcavillaca, R.T. A subgradient projection method for quasiconvex minimization. *Positivity* **2024**, *28*, 64. [CrossRef]
14. Zhao, X.; Köbis, M.A.; Yao, Y. A projected subgradient method for nondifferentiable quasiconvex multiobjective optimization problems. *J. Optim. Theory Appl.* **2021**, *190*, 82–107. [CrossRef]
15. Ermol'ev, Y.M. Methods of solution of nonlinear extremal problems. *Cybern* **1966**, *2*, 1–14. [CrossRef]
16. Penot, J.-P. Are generalized derivatives useful for generalized convex functions? In *Generalized Convexity, Generalized Monotonicity: Recent Results*; Crouzeix, J.-P., Martinez-Legaz, J.-E., Volle, M., Eds.; Springer: Boston, MA, USA, 1998; pp. 3–59.
17. Penot, J.-P.; Zălinescu, C. Elements of quasiconvex subdifferential calculus. *J. Convex Anal.* **2000**, *7*, 243–269.
18. Arjevani, Y.; Shamir, O.; Srebro, N. A tight convergence analysis for stochastic gradient descent with delayed updates. In Proceedings of the 31st International Conference on Algorithmic Learning Theory, San Diego, CA, USA, 8–11 February 2020; Kontorovich, A., Neu, G., Eds.; Proceedings of Machine Learning Research, PMLR: Cambridge, MA, USA, 2020; Volume 117, pp. 111–132.
19. Stich, S.U.; Karimireddy, S.P. The error-feedback framework: Better rates for sgd with delayed gradients and compressed updates. *J. Mach. Learn. Res.* **2020**, *21*, 9613–9648.
20. Gürbüzbalaban, M.; Ozdaglar, A.; Parrilo, P.A. On the convergence rate of incremental aggregated gradient algorithms. *SIAM J. Optim.* **2017**, *27*, 1035–1048. [CrossRef]
21. Tseng, P.; Yun, S. Incrementally updated gradient methods for constrained and regularized optimization. *J. Optim. Theory Appl.* **2014**, *160*, 832–853. [CrossRef]
22. Vanli, N.D.; Gürbüzbalaban, M.; Ozdaglar, A. Global convergence rate of proximal incremental aggregated gradient methods. *SIAM J. Optim.* **2008**, *28*, 1282–1300. [CrossRef]
23. Butnariu, D.; Censor, Y.; Reich, S. Distributed asynchronous incremental subgradient methods. In *Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*; Brezinski, C., Wuytack, L., Reich, S., Eds.; Elsevier Science B.V.: Amsterdam, The Netherlands, 2001; Volume 8, p. 381.

24. Namsak, S.; Petrot, N.; Nimana, N. A distributed proximal gradient method with time-varying delays for solving additive convex optimizations. *Results Appl. Math.* **2023**, *18*, 100370. [CrossRef]
25. Deng, X.; Shen, L.; Li, S.; Sun, T.; Li, D.; Tao, D. Towards understanding the generalizability of delayed stochastic gradient descent. *arXiv* **2023**, arXiv:2308.09430. [CrossRef]
26. Arunrat, T.; Namsak, S.; Nimana, N. An asynchronous subgradient-proximal method for solving additive convex optimization problems. *J. Appl. Math. Comput.* **2023**, *69*, 3911–3936. [CrossRef]
27. Cegielski, A. *Iterative Methods for Fixed Point Problems in Hilbert Spaces*; Springer: Berlin/Heidelberg, Germany, 2012.
28. Bauschke, H.H.; Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed.; Springer: Cham, Switzerland, 2017.
29. Kiwiel, K.C. Convergence of Approximate and Incremental Subgradient Methods for Convex Optimization. *SIAM J. Optim.* **2004**, *14*, 807–840. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.