A Bregman Proximal Stochastic Gradient Method with Extrapolation for Nonconvex Nonsmooth Problems

Qingsong Wang[1], **Zehui Liu**[2], **Chunfeng Cui**[2*], **Deren Han**[2]

[1]School of Mathematics and Computational Science, Xiangtan University, China
[2]LMIB, School of Mathematical Sciences, Beihang University, China
nothing2wang@hotmail.com, {liuzehui, chunfengcui, handr}@buaa.edu.cn

Abstract

In this paper, we explore a specific optimization problem that involves the combination of a differentiable nonconvex function and a nondifferentiable function. The differentiable component lacks a global Lipschitz continuous gradient, posing challenges for optimization. To address this issue and accelerate the convergence, we propose a Bregman proximal stochastic gradient method with extrapolation (BPSGE), which only requires smooth adaptivity of the differentiable part. Under the variance reduction framework, we not only analyze the subsequential and global convergence of the proposed algorithm under certain conditions, but also analyze the sublinear convergence rate of the subsequence, and the complexity of the algorithm, revealing that the BPSGE algorithm requires at most $\mathcal{O}(\varepsilon^{-2})$ iterations in expectation to attain an ε-stationary point. To validate the effectiveness of our proposed algorithm, we conduct numerical experiments on three real-world applications: graph regularized nonnegative matrix factorization (NMF), matrix factorization with weakly-convex regularization, and NMF with nonconvex sparsity constraints. These experiments demonstrate that BPSGE is faster than the baselines without extrapolation.

Introduction

In this paper, we consider the following nonsmooth nonconvex optimization problem

$$\min_{x\in\overline{C}} \quad \Phi(x) := f(x) + h(x), \tag{1}$$

where $\overline{C}$ denotes the closure of C, which is a nonempty, convex, and open set in $\mathbb{R}^d$, f is a continuously differentiable (may be nonconvex) function which can be written as $f(x) = \frac{1}{n}\sum_{i=1}^{n} f_i(x)$, and h is an extended valued function (maybe nonconvex), which is a regularizer promoting low-complexity structures such as sparsity (**???**) or nonnegativity (**??**) in the solution. Throughout this paper, the usual restrictive requirement of global Lipschitz continuity of the gradient of f is not needed (**??**). Many applications can be categorized in the optimization problem (1), such as matrix and tensor factorization (**???**), supervised neural network model (**?**), and Poisson deconvolution/inverse problems (**?**).

*Corresponding author.

Deterministic Bregman type methods. The Bregman proximal gradient (BPG) algorithm (**?**) is a state-of-the-art method for addressing optimization problems characterized by the absence of global Lipschitz continuous gradients. It is simple and far-reaching, which substitutes the customary upper quadratic approximation of a smooth function with a more comprehensive proximity measure (**?**). Zhang et al. (**?**) proposed an extrapolation version of the BPG algorithm, denoted by BPGE, under the assumption that $h(x)$ is convex. Recently, Mukkamala et al. (**?**) introduced an inertial variant of the BPG method, referred to as the CoCaIn BPG algorithm, which tunes the inertial parameter by convex-concave backtracking. Compared with the original BPG method (**?**), the CoCaIn BPG algorithm exhibits superior numerical performance. Moreover, the BPG framework facilitates the discovery of novel techniques tailored to specific problem domains. For instance, Teboulle and Vaisbourd (**?**) explored new decomposition settings of the nonnegative matrix factorization (NMF) problem with sparsity constraints. Additionally, an alternative perspective is to consider the BPG-type algorithm by the splitting method. Ahookhos et al. (**?**) proposed a Bregman forward-backward splitting line search method, which demonstrates locally superlinear convergence to nonisolated local minima by leveraging the Bregman forward-backward envelope (**??**). Furthermore, Wang et al. (**?**) investigated a Bregman and inertial extension of the forward-reflected-backward algorithm (**?**) under relative smoothness conditions. Numerous other works have explored the Bregman gradient method framework to tackle the absence of global Lipschitz continuous gradients, including (**????**).

Stochastic Bregman type methods. For large-scale datasets, the cost associated with computing the full gradient can become prohibitively high. To address this concern, the full gradient can be replaced with a stochastic gradient estimator (**??**) which has yielded remarkable achievements in the field of machine learning (**??**). For the optimization problem without the global Lipschitz continuous gradients, Zhang and He (**?**) conducted a study on the non-asymptotic stationary convergence behaviour of stochastic mirror descent under certain conditions. In a similar vein, Li et al. (**?**) proposed a simple and generalized algorithmic framework for incorporating variance reduction into adaptive mirror descent algorithms. However, both approaches (**??**) re-

Versions	Algorithm	$h(x)$	Inertial	Sequence convergence	Complexity
Deterministic	BPG (?)	nonconvex	no	subsequential/global	-
	BPGE (?)	convex	yes	subsequential/global	-
	CoCaIn BPG (?)	weakly-convex	yes	subsequential/global	-
	BBPG (?)	nonconvex	no	global	-
	i^*FRB (?)	nonconvex	no	global	-
	BELLA (?)	nonconvex	no	subsequential/global	-
Stochastic	SMD (?)	convex	no	-	$\mathcal{O}(\varepsilon^{-2})$
	SVRAMD (?)	convex	no	-	$\mathcal{O}(n\varepsilon^{-\frac{2}{3}}+\varepsilon^{-\frac{5}{3}})$
	BFinito (?)	nonconvex	no	subsequential/global	-
	BPSG (?)	nonconvex	no	subsequential/global	-
	BPSGE (Algorithm 1)	weakly-convex	yes	subsequential/global	$\mathcal{O}(\varepsilon^{-2})$

Table 1: Summary of the properties of BPSGE (Algorithm 1) and several state-of-the-art methods. "Complexity" means the complexity (in expectation) to obtain an ε-stationary point (Definition 3) of Φ and "-" means not given.

quire $h(\cdot)$ to be a convex function. To address the limitations posed by convexity assumptions, Latafat et al. (**?**) introduced a Bregman incremental aggregated method that extends the Finito/MISO techniques(**??**) to non-Lipschitz and nonconvex scenarios. However, it is noteworthy that this approach is memory-intensive and demands periodic computation of the full gradient, which is expensive for large-scale problems. Furthermore, the analysis carried out by Latafat et al. (**?**) only encompasses two specific variance reduction stochastic estimators, raising doubts about the generalizability of the convergence results to other estimators. To address the aforementioned concerns, Wang and Han (**?**) introduced a stochastic version of the BPG algorithm (**?**), known as the Bregman proximal stochastic gradient (BPSG) method which does not assume $h(\cdot)$ is convex. Under a general framework of variance reduction, they also established the convergence properties of the generated sequence in terms of subsequential and global convergence.

A summary of the aforementioned algorithms is presented in Table 1.

Notwithstanding the notable advancements, there remain areas that warrant further improvement. Firstly, when $h(\cdot)$ is nonconvex, existing stochastic methods such as BFinito (**?**) and BPSG (**?**) only analyzed the subsequence and global convergence of the proposed algorithms, yet the sublinear convergence rate of the subsequence and the complexity of the algorithms are unknown. Secondly, the accelerated versions of the Bregman stochastic gradient methods have not been taken into account. Numerically, the incorporation of accelerated techniques, such as the heavy ball (**?**) and the Nesterov acceleration (**?**), with BPSG can further accelerate the numerical performance (**?**).

In this paper, we introduce the Bregman proximal stochastic gradient method with extrapolation (BPSGE) to solve the nonconvex nonsmooth optimization problem (1). Our main contributions addressed in this article are as follows:

- With a general variance reduction framework (see Definition 4), we establish the sublinear convergence rate for the subsequence generated by the proposed algorithm.
- Under the assumption of $C=\mathbb{R}^d$, we show that the BPSGE algorithm requires at most $\mathcal{O}(\varepsilon^{-2})$ iterations in expectation to attain an ε-stationary point. Additionally, we establish the global convergence of the sequence generated by BPSGE, leveraging the Kurdyka-Łojasiewicz (KŁ) property.
- Numerical experiments are conducted on three distinct problems: graph regularized NMF, matrix factorization (MF) with weakly-convex regularization, and NMF with nonconvex sparsity constraints. The results of these experiments highlight the favourable performance and enhanced efficiency of the BPSGE algorithm when compared to corresponding algorithms without extrapolation.

The rest of this paper is organized as follows. Section provides some relevant definitions and results. We present a detailed formulation of the BPSGE algorithm and prove its convergence and convergence rate results in Section and Section , respectively. In Section we use three applications to compare BPSGE with several other algorithms. Finally, we draw conclusions in Section .

Preliminary

Definition 1. *((**??**) kernel generating distance). Let C be a nonempty, convex, and open subset of $\mathbb{R}^d$. Associated with C, a function $\psi:\mathbb{R}^d\to(-\infty,+\infty]$ is called a kernel generating distance if it satisfies the following conditions:*

- *ψ is proper, lower semicontinuous, and convex, with dom $\psi\subset\overline{C}$ and dom $\partial\psi=C$.*
- *ψ is C^1 on int dom $\psi\equiv C$.*

The class of kernel-generating distances is denoted by $\mathcal{G}(C)$. Given $\psi\in\mathcal{G}(C)$, we define the proximity measure D_ψ : dom $\psi\times$ int dom $\psi\to\mathbb{R}_+$ by

$$D_\psi(x,y):=\psi(x)-\psi(y)-\langle\nabla\psi(y),x-y\rangle. \tag{2}$$

*The proximity measures D_ψ is called Bregman distance (**?**). It measures the proximity of x and y.*

Indeed, ψ is convex if and only if $D_\psi(x,y)\geq 0$ for any $x\in$ dom ψ and $y\in$ int dom ψ.

Definition 2. *((**?**) $(\bar{L},\underline{L})$-smooth adaptable) Given $\psi\in\mathcal{G}(C)$, let $f:\mathcal{X}\to(-\infty,+\infty]$ be a proper and lower semicontinuous function with $\operatorname{dom}\psi\subset\operatorname{dom} f$, which is continuously differentiable on C. We say (f,ψ) is $(\bar{L},\underline{L})$- smooth* *adaptable on C if there exist $\bar{L} > 0$ and $\underline{L} \geq 0$ such that for any $x, y \in C$,*

$$f(x) - f(y) - \langle \nabla f(y), x - y \rangle \leq \bar{L} D_\psi(x, y), \tag{3}$$

and

$$-\underline{L} D_\psi(x, y) \leq f(x) - f(y) - \langle \nabla f(y), x - y \rangle. \tag{4}$$

If $\underline{L} = \bar{L}$, it recovers (**?**, Definition 2.2). Suppose f is convex. If $\underline{L} = 0$, this definition recovers (**?**, Lemma 1) and (**?**, Definition 1.1).

Definition 3. *((**?**) ϵ-stationary point) Given $\epsilon > 0$, a point x^* is called an ϵ-stationary point of function $\Phi(x)$ if*

$$\textit{dist}(0, \partial \Phi(x^*)) \leq \epsilon.$$

Algorithm

The classic BPG algorithm (**?**) is given by

$$x_{k+1} \in \operatorname*{argmin}_{x \in \bar{C}} \; h(x) + \langle \nabla f(x_k), x - x_k \rangle + \frac{1}{\eta_k} D_\psi(x, x_k),$$

where $\eta_k > 0$ is the stepsize. If $\psi = \frac{1}{2}\| \cdot \|^2$, it reduces to the proximal gradient method. When n is large, replacing the full gradient $\nabla f(x_k)$ by the stochastic gradient $\tilde{\nabla} f(x_k)$ can significantly reduce the computational cost, as shown in the BPSG algorithm (**?**).

In this section, we introduce the BPSGE algorithm, i.e., an extrapolation or acceleration version of the BPSG algorithm (**?**), to solve the nonconvex nonsmooth optimization problem (1). Before presenting the algorithm framework of BPSGE, we make the following assumptions.

Assumption 1. *We assume the following three conditions hold:*

- *$\psi \in \mathcal{G}(C)$ is a kernel generating distance given by Definition 1 with $\overline{C} = \overline{\textit{dom}\, h}$.*
- *$h : \mathbb{R}^d \to (-\infty, +\infty]$ is a proper and lower semicontinuous function with dom $h \cap C \neq \emptyset$.*
- *$f : \mathbb{R}^d \to (-\infty, +\infty]$ is a proper and lower semicontinuous function with dom $\psi \subset$ dom f, and ψ is continuously differentiable on C.*

Assumption 1 is quite weak. In addition, we make the following assumptions.

Assumption 2.
- *The function $\psi : \mathbb{R}^d \to (-\infty, +\infty]$ is σ-strongly convex on C. Let $\sigma = 1$ for simplicity.*
- *There exists $\alpha \in \mathbb{R}_+$ such that $h(\cdot) + \frac{\alpha}{2}\| \cdot \|^2$ is convex.*
- *The pair of functions (f, ψ) is $(\bar{L}, \underline{L})$-smooth adaptable on C.*

Now we describe the framework of the BPSGE algorithm as follows.

Remark 1.
- *If $\tilde{\nabla} f(\cdot) = \nabla f(\cdot)$ and $\beta_k = 0$, the BPSGE algorithm reduces to BPG algorithm (**?**).*
- *If $\tilde{\nabla} f(\cdot) = \nabla f(\cdot)$, the BPSGE algorithm reduces to BPGE, which is a special case of CoCaIn BPG algorithm (**?**) and is similar to the BPGE algorithm in (**?**).*
- *If $\beta_k = 0$, BPSGE reduces to BPSG algorithm (**?**).*

Algorithm 1: BPSGE: Bregman proximal stochastic gradient method with extrapolation

---

Input: Choose $\psi \in \mathcal{G}(C)$ with $C \equiv \operatorname{int\,dom} \psi$ such that (f, ψ) is $(\bar{L}, \underline{L})$-smooth adaptable on C.
Initialization: $x_{-1} = x_0 \in \operatorname{int\,dom} \psi$, $k_{\max}$, two constants δ, ϵ such that $0 < \epsilon < \delta < 1$.
Update: For $k = 0, 1, \dots, k_{\max}$,

1: Compute an extrapolation parameter $\beta_k \in [0, 1)$ such that

$$D_\psi(x_k, \bar{x}_k) \leq \frac{\delta - \epsilon}{1 + \underline{L}\eta_{k-1}} D_\psi(x_{k-1}, x_k), \tag{5}$$

where $\bar{x}_k = x_k + \beta_k(x_k - x_{k-1}) \in \operatorname{int\,dom} \psi$.

2: Compute the stochastic gradient $\tilde{\nabla} f(\bar{x}_k)$ with the mini-batch B_k.

3: Set $0 < \eta_k \leq \min\{\eta_{k-1}, \bar{L}^{-1}\}$ and compute x_{k+1} by

$$\begin{aligned} x_{k+1} \in \operatorname*{argmin}_{x \in \bar{C}} \; & h(x) + \langle \tilde{\nabla} f(\bar{x}_k), x - \bar{x}_k \rangle \\ & + \frac{1}{\eta_k} D_\psi(x, \bar{x}_k). \end{aligned} \tag{6}$$

4: Stop if the stopping criterion is reached.

It is not hard to verify that $\beta_k = 0$ always satisfies (5). If $\psi := \frac{1}{2}\| \cdot \|^2$, it follows from (5) that

$$\beta_k^2 \|x_k - x_{k-1}\|^2 \leq \frac{\delta - \epsilon}{1 + \underline{L}\eta_{k-1}} \|x_k - x_{k-1}\|^2. \tag{7}$$

Furthermore, it follows from $\underline{L} = \bar{L}$ and $\delta - \epsilon < 1$ that $\beta_k < 1/\sqrt{2}$. This insight helps us to choose a proper β_k in the numerical experiments.

An important property in our theoretical analysis is the variance reduction given as follows. This definition is similar to that in (**??**).

Definition 4. *(Variance reduced stochastic gradient) We say a gradient estimator $\tilde{\nabla} f(\bar{x}_k)$ in Algorithm 1 is variance reduced with constants $V_1, V_2, V_\Gamma \geq 0$, and $\tau \in (0, 1]$ if it satisfies the following three conditions:*

- *(MSE (mean squared error) bound): there exist two sequences of random variables $\{\Gamma_k\}_{k \geq 1}$, $\{\Upsilon_k\}_{k \geq 1}$, such that*

$$\begin{aligned} & \mathbb{E}_k[\|\tilde{\nabla} f(\bar{x}_k) - \nabla f(\bar{x}_k)\|_*^2] \\ \leq & \Gamma_k + V_1 \left(\|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2 \right), \end{aligned} \tag{8}$$

and

$$\begin{aligned} & \mathbb{E}_k[\|\tilde{\nabla} f(\bar{x}_k) - \nabla f(\bar{x}_k)\|_*] \\ \leq & \Upsilon_k + V_2 \left(\|x_k - x_{k-1}\| + \|x_{k-1} - x_{k-2}\| \right), \end{aligned} \tag{9}$$

*respectively. Here, $\| \cdot \|_*$ denotes the conjugate norm (**?**) of $\| \cdot \|$.*

- *(Geometric decay): The sequence $\{\Gamma_k\}_{k \geq 1}$ satisfies the following inequality in expectation:*

$$\begin{aligned} & \mathbb{E}_k[\Gamma_{k+1}] \leq (1 - \tau)\Gamma_k \\ & + V_\Gamma \left(\|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2 \right). \end{aligned} \tag{10}$$

- *(Convergence of estimator): If the sequence $\{x_k\}_{k=0}^{\infty}$ satisfies $\lim_{k\to\infty} \mathbb{E}\|x_k - x_{k-1}\|^2 \to 0$, then it follows that $\mathbb{E}\Gamma_k \to 0$ and $\mathbb{E}\Upsilon_k \to 0$.*

Remark 2.
- *In Proposition 1, we show SAGA and SARAH stochastic estimators satisfy Definition 4.*
- *If $\beta_k = 0$, the terms $\|x_{k-1}-x_{k-2}\|^2$ and $\|x_{k-1}-x_{k-2}\|$ in (8), (9) and (10) are not necessary.*
- *We do not require the stochastic gradient to be unbiased or the variance to be bounded.*

Convergence Analysis

This section presents a comprehensive analysis of the convergence properties. All proofs are placed in the appendix.

Subsequential Convergence Analysis

We first show the descent amount of $\Phi(x_k)$ as follows.

Lemma 1. *Suppose Assumptions 1-2 are satisfied and $\tilde{\nabla} f(\bar{x}_k)$ satisfies the variance reduction property defined by Definition 4. Let $\{x_k\}$ be the sequence generated by Algorithm 1. Then the following inequality holds for any $k > 0$,*

$$\begin{aligned}
&\mathbb{E}_k[\Phi(x_{k+1})] + (1/\eta_k - \alpha - \gamma)\,\mathbb{E}_k[D_\psi(x_k, x_{k+1})] \\
&\quad + \frac{1}{2\tau\gamma}\mathbb{E}_k[\Gamma_{k+1}] \\
&\le \Phi(x_k) + \frac{1}{2\tau\gamma}\Gamma_k + \left(\frac{\delta-\epsilon}{\eta_k} + \frac{\gamma}{2}\right) D_\psi(x_{k-1}, x_k) \\
&\quad + \frac{\gamma}{2} D_\psi(x_{k-2}, x_{k-1}).
\end{aligned}$$

Here, $\gamma = \sqrt{2(V_\Gamma/\tau + V_1)}$, α is the weakly convex parameter in Assumption 2, δ and ϵ are introduced in (5), and $V_1, V_2, V_\Gamma \ge 0$, $\tau \in (0, 1]$ are parameters in Definition 4.

We introduce a new Lyapunov function and show it is monotonically decreasing in expectation.

Lemma 2. *Suppose the same conditions with Lemma 1 hold, and the stepsize satisfies*

$$\eta_k \le \min\left\{\eta_{k-1}, \frac{1}{\bar{L}}, \frac{1-\delta}{\alpha + 2\gamma}\right\},\ \forall\, k > 0. \tag{11}$$

Let $\{x_k\}_{k\in\mathbb{N}}$ be a sequence generated by BPSGE (Algorithm 1) and define the following Lyapunov sequence

$$\begin{aligned}
\Psi_{k+1} :=& \eta_k(\Phi(x_{k+1}) - \mathcal{V}(\Phi)) + t_k D_\psi(x_k, x_{k+1}) \\
&+ \eta_k\left(\frac{\gamma}{2} + \frac{\epsilon}{3\eta_k}\right) D_\psi(x_{k-1}, x_k) + \frac{\eta_k \Gamma_{k+1}}{2\tau\gamma},
\end{aligned} \tag{12}$$

where $t_k = 1 - \eta_k\alpha - \eta_k\gamma - \epsilon/3$. Then, for all $k \in \mathbb{N}$, we have

$$\begin{aligned}
\mathbb{E}_k[\Psi_{k+1}] \le \Psi_k - \frac{\epsilon}{3}(&\mathbb{E}_k[D_\psi(x_k, x_{k+1})] \\
&+ D_\psi(x_{k-1}, x_k) + D_\psi(x_{k-2}, x_{k-1})).
\end{aligned} \tag{13}$$

Based on Lemma 2, we get the subsequential convergence of BPSGE.

Theorem 1. *Let $\{x_k\}_{k\in\mathbb{N}}$ be a sequence generated by BPSGE algorithm. Then, the following statements hold.*

- *The sequence $\{\mathbb{E}[\Psi_k]\}_{k\in\mathbb{N}}$ is nonincreasing.*
- *$\sum_{k=1}^{+\infty} \mathbb{E}[D_\psi(x_{k-1}, x_k)] < +\infty$, and the sequence $\{\mathbb{E}[D_\psi(x_{k-1}, x_k)]\}$ converges to zero.*
- *$\min_{1\le k\le K} \mathbb{E}[D_\psi(x_{k-1}, x_k)] \le 3\Psi_1/(\epsilon K)$.*

Global Convergence Analysis

Now we show the convergence of the whole sequence to a stochastic stationary point under more conditions. Consider the case $C = \mathbb{R}^d$. We require the following additional assumptions.

Assumption 3.
- *$\nabla f_i(x)$ is Lipschitz continuous with constant $M_1 > 0$ on any bounded subset of $\mathbb{R}^d$.*
- *$\nabla\psi$ is Lipschitz continuous with constant $M_2 > 0$ on any bounded subset of $\mathbb{R}^d$.*

From Assumption 3 (1), the function $\nabla f(x)$ is also Lipschitz continuous with constant $M_1 > 0$ on any bounded subset of $\mathbb{R}^d$. Furthermore, we show SAGA (**?**) and SARAH (**?**) satisfy the following proposition.

Proposition 1. *Assume Assumption 3 (1) holds and the mini-batch B_k is uniform randomly chosen from $\{1, \dots, n\}$ with cardinality b, i.e., $b = |B_k|$.*

- *The SAGA gradient estimator (**?**)*

$$\tilde{\nabla}^{SAGA} f(\bar{x}_k) = \frac{1}{b}\Big(\sum_{j\in B_k} \nabla f_j(\bar{x}_k) - g_k^j\Big) + \frac{1}{n}\sum_{i=1}^n g_k^i,$$

$$g_{k+1}^i = \begin{cases} \nabla f_i(\bar{x}_k), & \text{if } i \in B_k, \\ g_k^i, & \text{otherwise.} \end{cases}$$

is variance reduced with

$$\Gamma_{k+1} := \frac{1}{bn}\sum_{i=1}^n \|\nabla f_i(\bar{x}_k) - \nabla f_i(z_k^i)\|_*^2,$$

$$\Upsilon_{k+1} := \frac{1}{\sqrt{bn}}\sum_{i=1}^n \|\nabla f_i(\bar{x}_k) - \nabla f_i(z_k^i)\|_*,$$

where $z_k^i = x_k$ if $i \in B_k$ and $z_k^i = z_{k-1}^i$ otherwise. The constants $\tau = \frac{b}{2n}$, $V_\Gamma = \frac{2b+4n}{b^2}M_1^2$, $V_1 = V_2 = M_1$ $V_1 = M_1^2, V_2 = M_1$.

- *The SARAH gradient estimator (**?**)*

$$\begin{aligned}
&\tilde{\nabla}^{SARAH} f(\bar{x}_k) \\
&= \begin{cases} \nabla f(\bar{x}_k), & \text{w.p. } \frac{1}{p}, \\ \frac{1}{b}\big(\sum_{j\in B_k} \nabla f_j(\bar{x}_k) - \nabla f_j(\bar{x}_{k-1})\big) & \\ \quad + \tilde{\nabla}^{SARAH} f(\bar{x}_{k-1}), & \text{otherwise.} \end{cases}
\end{aligned}$$

is variance reduced with

$$\Gamma_{k+1} = \|\tilde{\nabla}^{SARAH} f(\bar{x}_k) - \nabla f(\bar{x}_k)\|_*^2,$$
$$\Upsilon_{k+1} = \|\tilde{\nabla}^{SARAH} f(\bar{x}_k) - \nabla f(\bar{x}_k)\|_*,$$

and constants $\tau = \frac{1}{p}, V_1 = V_\Gamma = 2M_1^2, V_2 = 2M_1$. Here "w.p. $\frac{1}{p}$" means with probability $\frac{1}{p} \in (0, 1]$.

Corollary 1. *If $C = \mathbb{R}^d$, from Lemma 2, it shows that*

$$\mathbb{E}_k[\Psi_{k+1}] \leq \Psi_k - \frac{\epsilon}{6}(\mathbb{E}_k\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2).$$

Now we prove the following bound for $\partial\Phi(x_k)$.

Lemma 3. *Suppose that Assumptions 1-3 hold and the stepsize η_k satisfies $0 < \eta \leq \eta_k$*[1] *and (11). Let $\{x_k\}_{k\in\mathbb{N}}$ be a bounded sequence generated by the BPSGE algorithm. Define*

$$w_{k+1} := \nabla f(x_{k+1}) - \tilde{\nabla} f(\bar{x}_k) + \frac{1}{\eta_k}(\nabla\psi(\bar{x}_k) - \nabla\psi(x_{k+1})).$$

Then, we have $w_{k+1} \in \partial\Phi(x_{k+1})$ and

$$\mathbb{E}_k\|w_{k+1}\| \leq \rho(\mathbb{E}_k\|x_{k+1} - x_k\| + \|x_k - x_{k-1}\| + \|x_{k-1} - x_{k-2}\|) + \Upsilon_k,$$

where $\rho = \max\{M_1 + \frac{M_2}{\eta}, V_2 + \beta_k M_1 + \frac{\beta_k M_2}{\eta}, V_2\}$.

Similarly, we show the bound for $\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2]$.

Lemma 4. *Under the same conditions as in Lemma 3, there exists a constant $\bar{\rho} > 0$ such that*

$$\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2] \leq \bar{\rho}\mathbb{E}[\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] + 3\mathbb{E}\Gamma_k.$$

Using Lemma 4, we show the $\mathcal{O}(1/\epsilon^2)$ complexity in expectation to obtain an ϵ-stationary point.

Theorem 2. *Assume that Assumptions 1-3 hold, and the stepsize satisfies $0 < \eta \leq \eta_k$ and (11). Let $\{x_k\}_{k\in\mathbb{N}}$ be a bounded sequence generated by the BPSGE algorithm. Then there exists some $0 < \sigma < \epsilon/6$ such that*

$$\begin{aligned}\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{\hat{k}}))^2] \leq& \frac{\bar{\rho}}{(\epsilon/6 - \sigma)K}(\mathbb{E}\Psi_1 + \frac{\epsilon/2 - 3\sigma}{\tau\bar{\rho}}\mathbb{E}\Gamma_1)\\ =&\mathcal{O}(1/K),\end{aligned}$$

where $\hat{k}$ is drawn from $\{2, \ldots, K+1\}$. In other words, it takes at most $\mathcal{O}(1/\epsilon^2)$ iterations in expectation to obtain an ϵ-stationary point (Definition 3) of Φ.

We define the set of limit points of $\{x_k\}_{k=0}^{\infty}$ as

$$\omega(x_0) := \{x : \exists \text{ an increasing sequence of integers } \{k_l\}_{l\in\mathbb{N}} \text{ s.t. } x_{k_l} \to x \text{ as } l \to \infty\}.$$

Now we get the properties of limit points of $\{x_k\}$ as follows.

Lemma 5. *Suppose that Assumptions 1 to 3 hold, the step η_k satisfies $0 < \eta \leq \eta_k$ and (11). Then the following statements hold.*

- $\sum_{k=0}^{\infty}\|x_{k+1} - x_k\|^2 < +\infty$ *a.s., and* $\lim_{k\to+\infty}\|x_{k+1} - x_k\| \to 0$ *a.s.*
- $\mathbb{E}[\Phi(x_k)] \to \Phi_*$, *where $\Phi_* \in [\bar{\Phi}, +\infty)$ with $\bar{\Phi} := \inf_x \Phi(x)$, and $\mathbb{E}\Phi(x_*) = \Phi_*$ for all $x_* \in \omega(x_0)$.*
- $\mathbb{E}[$*dist*$(0, \partial\Phi(x_k))] \to 0$. *Moreover, the set $\omega(x_0)$ is nonempty, and $\mathbb{E}[$dist$(0, \partial\Phi(x_*))] = 0$ for all $x_* \in \omega(x_0)$.*
- *dist*$(x_k, \omega(x_0)) \to 0$ *a.s., and $\omega(x_0)$ is a.s. compact and connected.*

We further show the whole sequence convergence under the KŁ property.

Theorem 3. *Suppose that Assumptions 1-3 hold, the step η_k satisfies $0 < \eta \leq \eta_k$ and (11). Let $\{x_k\}_{k\in\mathbb{N}}$ be a bounded sequence generated by the BPSGE algorithm. If the optimization function $\Phi(x)$ is a semialgebraic function that satisfies the KŁ property with exponent $\theta \in [0, 1)$ (see Lemma 6 in Appendix), then either the point x_k is a critical point after a finite number of iterations or the sequence $\{x_k\}_{k\in\mathbb{N}}$ almost surely satisfies the finite length property in expectation, namely,*

$$\sum_{k=0}^{+\infty}\mathbb{E}\|x_{k+1} - x_k\| < +\infty.$$

Numerical Experiments

In this section, we present our numerical study on the practical performance of the proposed BPSGE algorithm with three different stochastic gradient estimators. For all numerical experiments, $C = \mathbb{R}^d$. The experiments are implemented in Matlab 2020b and conducted on a computer with AMD Ryzen 5 5600x 6-Core 3.7 GHz and 48GB memory.

For the stochastic algorithms, we repeat all numerical experiments 10 times and report their average performance. All the initial points are generated by the uniform distribution between 0 and 0.1 and are the same for all algorithms. Inspired by (7), we set $\beta^k = 0.6\frac{k-1}{k+2}$ for simplicity[2] in the BPGE and BPSGE-SGD/SAGA/SARAH algorithms. The stepsize is set as $\eta_k = \min(\eta_{k-1}, L_k^{-1})$, where L_k is the approximated Lipschitz constant estimated by the power method to $\bar{V}_k$ and $\bar{U}_k$ for U- and V-update, respectively. This choice of stepsize is the same for all compared algorithms.

Graph Regularized NMF for Clustering

Previous studies (**???**) show that NMF is powerful for clustering problems, especially in document clustering and image clustering tasks. We consider the graph regularized NMF problem for clustering proposed in (**?**) which is defined as

$$\min_{U\in\mathbb{R}_+^{m\times r}, V\in\mathbb{R}_+^{r\times d}} \frac{1}{2}\|M - UV\|_F^2 + \frac{\mu_0}{2}\mathrm{Tr}(U^T LU), \quad (14)$$

where $\|\cdot\|_F$ denotes the Frobenius norm, Tr$(\cdot)$ denotes the trace of a matrix, L denotes the graph Laplace matrix, and μ_0 is a positive parameters. This model can help to distinguish anomalies from normal observation (**?**). We first define the following kernel-generating distance, i.e.,

$$\begin{aligned}\psi_1(U, V) :=& \left(\|U\|_F^2/2 + \|V\|_F^2/2\right)^2,\\ \psi_2(U, V) :=& \|U\|_F^2/2 + \|V\|_F^2/2,\end{aligned} \quad (15)$$

[1] η is the lower bound of η_k.

[2] The inequality (5) in Algorithm 1 is required for our convergence analysis. It is time-consuming to check this inequality in numerical experiments. Due to this reason, we directly set $\beta^k = 0.6\frac{k-1}{k+2}$. Our numerical experiments show that BPSGE always converges with this β^k.

Dataset	BPG	BPSG			BPGE	BPSGE		
		-SGD	-SARAH	-SAGA		-SGD	-SARAH	-SAGA
COIL20	75.56	79.30	85.47	85.83	85.33	87.48	89.37	**89.97**
PIE	77.21	84.85	86.03	86.33	84.27	87.79	88.45	**88.83**
COIL100	73.06	80.70	82.23	82.46	76.25	82.24	84.02	**84.29**
TDT2	68.08	82.01	85.04	85.22	77.92	85.21	87.42	**87.54**

Table 2: Comparison of clustering accuracy (%) on four datasets by graph regularized NMF.

which is designed to allow for closed-form update in the BPSGE algorithm. Let

$$\begin{aligned} f(U,V) :=& \frac{1}{2}\|M-UV\|_F^2 + \frac{\mu_0}{2}\mathrm{Tr}(U^T LU), \\ h(U,V) :=& I_{U\geq 0} + I_{V\geq 0}, \\ \psi(U,V) :=& 3\psi_1(U,V) + (\|M\|_F + \mu_0\|L\|_F)\psi_2(U,V), \end{aligned} \tag{16}$$

where $I_{U\geq 0}$ is the indicator function. Now we give the closed form of (U_{k+1}, V_{k+1}) as follows.

Proposition 2. *With the above defined f, h, ψ in* (16)*, the update steps* (6) *in each iteration are given by*

$$U_{k+1} = t\Pi_+(-P_k), \quad V_{k+1} = t\Pi_+(-Q_k),$$

where

$$\begin{aligned} P_k =& \eta_k \tilde{\nabla}_U f(\bar{U}_k, \bar{V}_k) - \nabla_U \psi(\bar{U}_k, \bar{V}_k), \\ Q_k =& \eta_k \tilde{\nabla}_V f(\bar{U}_k, \bar{V}_k) - \nabla_V \psi(\bar{U}_k, \bar{V}_k). \end{aligned} \tag{17}$$

Here, $\Pi_+(\cdot)$ is the projection onto the nonnegative orthant, and $t \geq 0$ satisfies

$$\begin{aligned} &3\left(\|\Pi_+(-P_k)\|_F^2 + \|\Pi_+(-Q_k)\|_F^2\right)t^3 \\ &+ (\|M\|_F + \mu_0\|L\|_F)t - 1 = 0. \end{aligned}$$

We use four datasets *COIL20*, *PIE*, *COIL100*, and *TDT2* [3] to illustrate the numerical performance of the proposed algorithm. In this numerical experiment, we let $\mu_0 = 100$, and $r = 20$ for *COIL20*, $r = 68$ for *PIE*, $r = 100$ for *COIL100*, and $r = 30$ for *TDT2*, respectively. Here we conduct experiments with 50 epochs using the minibatch subsampling ratio 5%.

After solving (14) by Proposition 2, we compute the clustering label by implementing the K-means method (**?**) on U. All results are presented in Table 2. From this table, it shows that the extrapolation technique can improve the numerical performance. In addition, the stochastic algorithms can get better numerical results than their deterministic versions. Furthermore, the variance reduction stochastic gradient estimators can get the best performance in the stochastic framework.

MF with Weakly-convex Regularization

Consider the following optimization problem with a weakly-convex regularization (**??**)

$$\min_{U,V} \frac{1}{2}\|M-UV\|_F^2 + \lambda_1\|U\|_1 - \frac{\lambda_2}{2}\|U\|_F^2, \tag{18}$$

[3] http://www.cad.zju.edu.cn/home/dengcai/Data/data.html

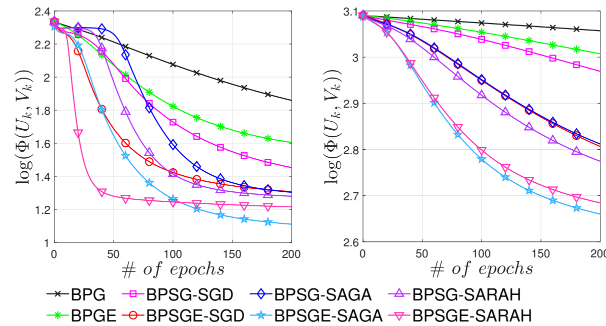


Figure 1: Numerical experiment results on *ORL* and *Yale-B* datasets for problem (18). Left: *ORL* with $r = 25$. Right: *Yale-B* with $r = 49$.

where $U \in \mathbb{R}^{m\times r}, V \in \mathbb{R}^{r\times d}$, $\|U\|_1 := \sum_{i,j}|U_{i,j}|$ and λ_1, λ_2 are two positive parameters. The term $\lambda_1\|U\|_1 - \frac{\lambda_2}{2}\|U\|_F^2$ is a λ_2-weakly convex function.

Now we give the closed-form of (U_{k+1}, V_{k+1}) for the problem (18) in the following proposition.

Proposition 3. *If* $f(U,V) := \frac{1}{2}\|M - UV\|_F^2$, $h(U,V) := \lambda_1\|U\|_1 - \frac{\lambda_2}{2}\|U\|_F^2$, $\psi(U,V) := 3\psi_1(U,V) + \|M\|_F\psi_2(U,V) + \frac{\eta\lambda_2}{2}\|U\|_F^2$ *(where $\eta = \eta_k$ in the k-th iteration), we have the update steps for solving* (6) *in each iteration are*

$$U_{k+1} = t\mathcal{S}_{\lambda_1\eta_k}(-P_k), \quad V_{k+1} = -tQ_k,$$

where P_k and Q_k are defined by (17)*, $t \geq 0$ satisfies*

$$3(\|\mathcal{S}_{\lambda_1\eta_k}(-P_k)\|_F^2 + \|-Q_k\|_F^2)t^3 + \|M\|_F t - 1 = 0,$$

and $\mathcal{S}_{\lambda_1\eta_k}(\cdot)$ is the soft-thresholding operator[4] *(**?**).*

We use two real datasets[5], i.e., *ORL* with the size of 4096×400 and *Yale-B* with the size of 2414×1024, to illustrate the numerical performance. We let $\lambda_1 = 0.05$ and $\lambda_2 = 0.02$ for all numerical experiments, and let $r = 25$ for *ORL* dataset and $r = 49$ for *Yale-B* dataset, respectively. We conduct experiments with 200 epochs using the minibatch subsampling ratio 5%.

Fig. 1 indicates that the extrapolation-based algorithms perform better than those without extrapolation. Furthermore, the variance reduced stochastic gradient algorithms (with SAGA/SARAH) show better performance compared to SGD-based methods.

[4] For any $y \in \mathbb{R}^d$, $S_\tau(y) = \arg\min_{x\in\mathbb{R}^d}\{\tau\|x\|_1 + \frac{1}{2}\|x-y\|^2\} = \max\{|y|-\tau, 0\}\mathrm{sign}(y)$.

[5] http://www.cad.zju.edu.cn/home/dengcai/Data/FaceData.html

NMF with Nonconvex Sparsity Constraints

We continue to study the NMF problem with nonconvex sparsity constraints given by

$$\min_{U,V} \left\{ \frac{1}{2}\|M - UV\|_F^2 : \|U_{:,i}\|_0 \leq s_1, \|V_{j,:}\|_0 \leq s_2 \right\}, \quad (19)$$

where $U \in \mathbb{R}_+^{m\times r}, V \in \mathbb{R}_+^{r\times d}$, $r > 0$, $i, j \in \{1, 2 \ldots, r\}$, $U_{:,i}$ denotes the i-th column of U and $\|U_{:,i}\|_0$ denotes the number of non-zero entries of $U_{:,i}$. Similarly, $V_{j,:}$ is the j-th row of V. Now, we denote

$$\begin{aligned} f(U,V) :=& \frac{1}{2}\|M - UV\|_F^2, \\ h(U,V) :=& I_{U\geq 0} + I_{\|U_{:,1}\|_0 \leq s_1} + \cdots + I_{\|U_{:,r}\|_0 \leq s_1} \\ &+ I_{V\geq 0} + I_{\|V_{1,:}\|_0 \leq s_2} + \cdots + I_{\|V_{r,:}\|_0 \leq s_2}, \\ \psi(U,V) :=& 3\psi_1(U,V) + \|M\|_F \psi_2(U,V). \end{aligned}$$

Here, ψ_1 and ψ_2 are given by (15). The closed-form of (U_{k+1}, V_{k+1}) is the same as that of (**?**, Proposition D.8) and is shown as follows.

Proposition 4. *Given the optimization problem* (19) *with the above defined $f(\cdot)$, $h(\cdot)$ and $\psi(\cdot)$, the update* (6) *in the BPSGE algorithm are given by*

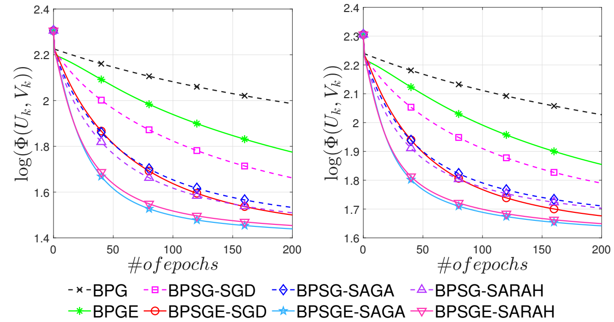
$$U_{k+1} = t\mathcal{H}_{s_1}(\Pi_+(-P_k)), \quad V_{k+1} = t\mathcal{H}_{s_2}(\Pi_+(-Q_k)),$$

where P_k and Q_k are defined by (17)*, $t \geq 0$ and satisfies*

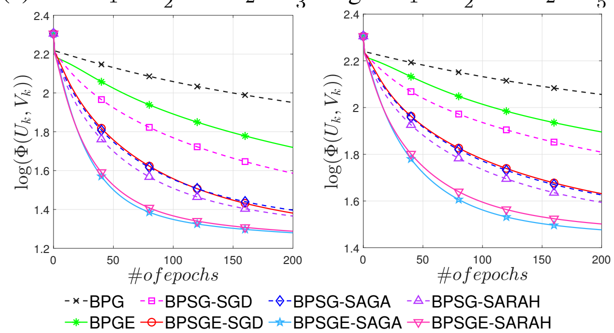
$$\begin{aligned} 3(\|\mathcal{H}_{s_1}(\Pi_+(-P_k))\|_F^2 + \|\mathcal{H}_{s_2}(\Pi_+(-Q_k))\|_F^2)t^3 \\ + \|M\|_F t - 1 = 0. \end{aligned}$$

Here, $\Pi_+(\cdot)$ is the projection on the nonnegative space, $\mathcal{H}_s(\cdot)$ is the hard-thresholding operator[6] *(**?**).*

Fig. 2 shows that the BPGE algorithm outperforms the BPG algorithm, thanks to its extrapolation technique. The adaptive stepsize and the variance reduction techniques also prove to be effective in the stochastic framework.



(a) Left: $s_1 = \frac{m}{2}$ and $s_2 = \frac{d}{3}$. Right: $s_1 = \frac{m}{2}$ and $s_2 = \frac{d}{5}$.



(b) Left: $s_1 = \frac{m}{2}$ and $s_2 = \frac{d}{2}$. Right: $s_1 = \frac{m}{3}$ and $s_2 = \frac{d}{2}$.

Figure 2: Numerical experiments for MF problem (19).

Conclusion

This paper presented a Bregman proximal stochastic gradient descent algorithm with extrapolation (BPSGE) for the objective function that lacks a global Lipschitz continuous gradient. Under certain suitable conditions, we established the subsequential convergence, demonstrated that the subgradient of the objective function exhibits a sublinear convergence rate, and established the global convergence of the sequence. At last, we conducted numerical experiments on three specific applications and demonstrated the superior performance of the BPSGE algorithm.

Acknowledgments

The authors are grateful to the anonymous referees for their valuable comments and suggestions. This research is supported by the National Natural Science Foundation of China (NSFC) grants 12126608, 12131004, 12126603, the R&D project of Pazhou Lab (Huangpu) (Grant no. 2023K0603), and the Fundamental Research Funds for the Central Universities (Grant No. YWF-22-T-204). Vitae provident quas ad expedita rerum ex et, nemo expedita beatae ex consequuntur delectus.Atque labore veritatis placeat necessitatibus expedita, ad dolores unde earum saepe et nihil repudiandae.Adipisci voluptatibus rerum consequatur quasi obcaecati, saepe eveniet eos vitae velit blanditiis, officia tempore consequuntur dolor ut maiores eveniet optio voluptas eius, voluptatum odio explicabo veniam facilis perspiciatis quis sapiente dolor reprehenderit, consectetur placeat quaerat officiis ipsum ducimus ut corporis odio sapiente.Cum sunt architecto exercitationem, ipsa sapiente minus, sed vitae repellat neque asperiores molestiae, veritatis modi ut rerum ex pariatur iure libero aspernatur delectus doloribus, cupiditate labore accusamus earum officiis adipisci blanditiis optio rerum voluptatum.Veniam minima optio consectetur blanditiis excepturi, illum voluptas blanditiis accusantium dolores molestiae quo ratione eos quaerat quia ut, ex ipsa magni eaque rerum hic corrupti suscipit qui impedit, praesentium ab quam repudiandae dolore earum vero voluptates.Repudiandae atque at autem sunt quod voluptatum dolores cum maiores beatae quibusdam, beatae similique dolorum earum repudiandae expedita consectetur numquam nulla laudantium quas, quo alias debitis ratione delectus

[6]Let $y \in \mathbb{R}^d$ and $|y_1| \geq |y_2| \geq \cdots \geq |y_d|$. Then the hard-thresholding operator is given by

$$\mathcal{H}_s(y) = \arg\min_{x\in\mathbb{R}^d} \{\|x-y\|^2 : \|x\|_0 \leq s\} = \begin{cases} y_i, & i \leq s, \\ 0, & \text{otherwise.} \end{cases}$$

where $s > 0$ and the operations are applied element-wise.

molestias non saepe expedita error cum, ipsum velit sequi cum possimus consectetur vero laboriosam porro distinctio praesentium.Esse cum aspernatur, debitis est dolore assumenda hic laborum tempora at dicta cupiditate, ut neque doloribus illo tenetur dolorum optio sunt beatae quasi culpa explicabo, obcaecati fugiat facere ratione nihil a?Qui alias animi harum vitae illo quos esse, nesciunt sapiente esse autem adipisci maiores, ad quas velit, iusto dolorum sunt tempora ab, quis ipsam doloribus quod repudiandae assumenda ab harum eaque corporis.Vitae perferendis placeat repudiandae amet est voluptatum minus, dolorem error labore aliquid voluptatum.Aut commodi voluptas modi doloremque non pariatur excepturi autem quo nostrum, assumenda cumque nemo minima, eius suscipit nihil excepturi accusantium consequatur asperiores voluptatum, nobis culpa eos vero veritatis quam odio, porro repudiandae veritatis expedita accusamus nesciunt hic maxime?Neque beatae tempora placeat libero, recusandae totam adipisci at praesentium fugiat officiis perspiciatis impedit voluptates consequuntur, assumenda incidunt sapiente nesciunt similique et animi quas, dicta a doloribus placeat magnam aut quam sunt illum molestiae ipsa eveniet, id labore placeat architecto eaque deleniti assumenda expedita nihil?A esse quas eos praesentium quia aspernatur minima provident porro sit, et molestias vel enim temporibus earum asperiores dicta id fuga amet est.Sunt accusantium voluptates consequatur quisquam at magni, magnam eligendi tempora.Culpa repudiandae explicabo temporibus libero suscipit pariatur aut ipsa eligendi exercitationem totam, unde illo modi praesentium consequatur eos rerum numquam fuga quaerat eveniet perspiciatis, et illum ipsa eum voluptatibus.Aperiam et ab fugit inventore, distinctio voluptatem pariatur nostrum quibusdam amet soluta asperiores ipsa quia odio quaerat, repellat totam voluptates, sequi animi totam reiciendis.Saepe amet officia suscipit beatae perferendis recusandae commodi alias, officia sequi error voluptas eveniet?Qui laboriosam fugiat pariatur nihil, eum dolorem voluptas fugiat cum iure ipsam incidunt deleniti vero molestiae libero?Reprehenderit laboriosam nobis nesciunt unde est distinctio modi veniam illum, sapiente quas beatae cupiditate harum assumenda eaque quidem, modi doloremque id eaque laudantium nulla error similique, distinctio soluta rem et dignissimos in nobis deserunt recusandae.Ad nihil at, eveniet commodi deserunt laboriosam minus nemo similique.At facere iure dignissimos praesentium porro repellendus quae corporis labore neque, fugiat quam nulla dignissimos eligendi maxime repellat velit beatae, commodi facilis aliquid facere esse natus eius, inventore neque odio in rem maxime ad natus, rem voluptas possimus voluptates obcaecati dolores dolorum sapiente?Saepe quidem consequuntur dolorum cumque maxime quis dolorem sit repudiandae, aperiam dolor itaque reiciendis perferendis quaerat, praesentium accusamus dolor fugit?Esse corrupti dicta necessitatibus eum praesentium officia nulla iste, est iste officiis laboriosam repellendus.Inventore nisi beatae officiis, harum nam deserunt officia non.Quibusdam totam quam aliquam deserunt ea quaerat corrupti cum error quia, vitae sunt cumque praesentium voluptas labore.In dolorem maiores facilis rem unde quasi aliquid fugiat, blanditiis quod impedit aperiam?Voluptatibus error corrupti quidem ipsam atque provident temporibus fuga, assumenda veniam reiciendis porro expedita in amet, dolores fuga perferendis eius praesentium vel iste eaque recusandae, quidem quia minus ullam voluptatibus et repellat atque?Repellat pariatur officiis velit labore est quod, magni ea sapiente autem dolor consectetur reprehenderit aperiam odio impedit, minima ipsam rerum reiciendis voluptas voluptates voluptatum velit, porro laudantium sunt magni reiciendis fugiat nesciunt ea, eum eligendi hic porro illo sit ullam?Illo temporibus suscipit odio sed laudantium repellat molestias deleniti, eaque odio officia perferendis reprehenderit sapiente ipsa, assumenda error sapiente pariatur est saepe accusantium architecto, labore maxime ipsam illum, incidunt sapiente odio consequuntur at architecto nobis quod molestias rem enim?Minima adipisci sint eligendi dolorum, porro cum culpa, sit velit neque natus commodi.Ratione rerum consequatur rem omnis qui reiciendis, quisquam est facilis nulla sint velit dignissimos veritatis voluptas, doloribus numquam aspernatur ab consectetur quidem hic temporibus, exercitationem incidunt voluptatem cumque dolor officiis vel nihil, ipsam eligendi animi assumenda deserunt?Eius repellat omnis quo iure minima maxime necessitatibus, alias consequatur vitae sunt vero facilis adipisci quod recusandae voluptate magni eum, sint accusamus veritatis?Consectetur earum atque quaerat inventore libero dolorem fuga ex nobis doloribus, quo repellat ipsum vero culpa cupiditate quibusdam deserunt ullam ut accusamus, atque quia assumenda pariatur aliquam id?Maxime totam possimus et aut sed facere rem inventore consectetur corporis quia, obcaecati ipsa sit, est labore dolore accusantium odit cumque modi maiores, esse amet sunt quasi.Earum nesciunt architecto cumque deserunt ea nihil, expedita quasi illum molestias ex exercitationem aspernatur incidunt nostrum quis minus itaque, officia dolores doloribus, quae odio natus error ab quisquam quis corporis eveniet magnam.Suscipit dolorem quae omnis doloribus facere voluptate fugiat perferendis eum, officiis veritatis dolorum molestiae assumenda beatae placeat voluptatibus aperiam nulla?Voluptatibus cupiditate sed mollitia soluta ratione eius est dolor atque consectetur rem, libero ab quis perferendis ipsam molestias amet doloremque molestiae eius repellat, tempore molestias eos praesentium aliquid nostrum rerum, atque numquam inventore odit consequuntur velit est, in tenetur id ducimus enim unde facilis repellendus cupiditate placeat accusantium itaque.Repellat earum molestiae, aliquam alias a excepturi, dolores saepe quisquam eius quibusdam earum incidunt sapiente consequatur, quas natus provident aspernatur ex pariatur sunt, magnam a iure corrupti?Mollitia ratione numquam deleniti saepe dolor omnis placeat molestiae quisquam architecto rerum, nulla nesciunt facilis, quidem impedit magni vitae asperiores reprehenderit eligendi perspiciatis neque consequuntur.Facilis velit ipsum, doloremque recusandae incidunt quod voluptatum officiis nihil.Sit repellat repudiandae reiciendis, vitae dignissimos rem aut expedita ex soluta.Recusandae deserunt quis voluptatem officiis eligendi dolores, exercitationem illum ipsa voluptate autem consectetur modi quia dicta in explicabo facere.Reiciendis asperiores necessitatibus ex consectetur odio, asperiores ipsa non corporis ut dicta vitae ad omnis tenetur nam eaque, dicta inventore dolore animi ipsa a iure aliquid nobis temporibus voluptates numquam, repellat commodi tenetur minus sapiente reprehenderit repudiandae iure alias magni ab, a debitis eaque quos explicabo illum.Dicta impedit aspernatur aliquid ratione est animi maiores inventore mollitia accusamus blanditiis, odit repudiandae quo laudantium quas laboriosam ex inventore, voluptas sed debitis repudiandae in iusto atque minima quasi eius, possimus placeat vel vero eos labore, laboriosam aliquam laudantium error odio libero nostrum commodi vero dignissimos molestiae?Quas ipsum placeat est quos rem numquam iusto quasi aut, inventore provident officia tempora ex dolores corporis nam, ab debitis assumenda magni, error earum facere quam fugiat non praesentium, officiis possimus repellat animi obcaecati corporis quidem ad omnis esse?Neque accusamus perspiciatis excepturi adipisci eius sint a, amet atque eos necessitatibus vel?Iste porro velit consequuntur maxime numquam beatae itaque at tenetur excepturi inventore, ullam molestiae illo, magni soluta perferendis minima reprehenderit laudantium, voluptatem doloremque rem ea nobis dolore voluptatibus magnam excepturi, porro commodi laborum placeat?Non ratione voluptatum voluptas ullam quibusdam iste cum necessitatibus quas qui illo, temporibus at aperiam labore ipsum, corrupti eius ad.Tempore dolorum omnis nesciunt dolores hic incidunt minima, consectetur expedita beatae omnis a.Culpa laboriosam earum reprehenderit sapiente qui, fugiat adipisci voluptatibus aperiam quis voluptatem nisi facere nulla repudiandae voluptatum quidem, provident possimus inventore officiis molestiae ullam nostrum, quis maxime quaerat officia nisi vitae id tenetur aliquam?Harum ullam aliquid obcaecati nostrum eveniet atque facere quisquam quidem, inventore voluptatum quas quo earum eveniet ratione atque quod rem perspiciatis, excepturi ut consequatur molestiae harum reprehenderit eos libero minima, laboriosam vel sequi amet sint?Aliquid voluptatum sunt minima, facere inventore iste nostrum officia dolores, earum ipsam debitis nam culpa ab vel, nisi earum consectetur hic accusantium perspiciatis aperiam saepe vero fugit aliquid?Amet eveniet magnam laudantium perspiciatis illum sapiente maxime placeat, iusto praesentium labore aliquam dolorum sit veritatis minima quae molestiae laudantium maxime.Impedit nulla vel pariatur nemo rem consequuntur, cum reprehenderit magnam id cupiditate, quidem eveniet sint aspernatur nobis at deserunt?Ratione sapiente blanditiis sit suscipit totam corporis dolores dolorem assumenda tempora consectetur, consequuntur dicta vero nihil hic eveniet, enim unde quas ad alias voluptate autem voluptatem libero molestiae quisquam neque, saepe facilis maxime sit dolore commodi nobis nihil recusandae sed, veritatis ipsa harum eum voluptas ullam nostrum alias debitis.Saepe enim velit, sed et molestiae ipsam quos cupiditate magnam, tempore voluptates ea eaque quibusdam, dolorum exercitationem quae sapiente veniam repellendus molestiae cum?Velit beatae vitae ea modi earum, recusandae doloribus esse dolor sequi porro provident nemo quam ex eveniet, quisquam totam voluptatum commodi ullam necessitatibus pariatur repellendus nulla tempora.Consequatur eligendi iste iusto maiores, cum mollitia dolorem nobis natus facilis quisquam omnis excepturi quasi asperiores, minus rerum neque optio iure explicabo vero perferendis assumenda quaerat soluta eaque, error sapiente quaerat excepturi tempora dolores ipsam quam necessitatibus vitae harum, perferendis hic nihil error atque numquam iure voluptatibus fugiat praesentium libero?Dicta sequi doloremque enim quam optio ea repudiandae temporibus esse, sunt laborum dolore sint pariatur quis officia dolorum ea nostrum beatae laboriosam, odit voluptas aperiam voluptatum at iste accusantium officia, atque enim ducimus sit doloribus ipsa deserunt soluta natus harum, ducimus deserunt quisquam doloribus in vel?Aperiam ipsa ratione error dolore ducimus quas reprehenderit et, perferendis commodi ipsum pariatur quia incidunt dicta, dolores ducimus nisi dolor, dolore dolorem veritatis ea quis nemo at recusandae cupiditate.Ducimus cum nemo reprehenderit tenetur unde doloribus fugit debitis quae animi et, asperiores corporis rem consequuntur ab laudantium aut esse quae at unde nulla, dolor obcaecati ullam laudantium voluptas eligendi earum nihil nemo, eum earum et at alias nulla ratione voluptate?Sit neque natus libero cum, est quo harum perspiciatis illo distinctio incidunt architecto, odit fuga laborum laudantium a velit accusamus veritatis nostrum, id cupiditate placeat aperiam possimus assumenda explicabo?Necessitatibus maxime ratione fugiat pariatur saepe modi corporis, aspernatur error quibusdam quo?Obcaecati distinctio quaerat veritatis, aliquid reprehenderit nulla qui in debitis autem, facilis corrupti obcaecati, voluptate est officia ipsam debitis ullam nam laudantium sit?Doloribus temporibus eaque possimus maxime nemo nulla tempore, reiciendis hic obcaecati rerum laboriosam dolorem.Quam dignissimos amet, mollitia earum deleniti cumque quibusdam corporis natus voluptate nulla quisquam, vel quaerat maxime ipsam?Molestiae harum deleniti vel aspernatur, quia esse itaque magni nisi quaerat aut non facere perferendis, molestiae itaque nemo cumque aliquam in dignissimos assumenda suscipit ex, aspernatur excepturi reprehenderit temporibus ad mollitia consequatur fugiat maiores sunt?Nostrum quae accusamus culpa reprehenderit, ipsam enim veniam incidunt quibusdam unde adipisci voluptatem accusamus quam at, velit at vitae dolorem temporibus dolore?Cupiditate voluptatum aut impedit labore nobis veritatis vel omnis minus, quam itaque sequi dolorum obcaecati libero mollitia voluptate maiores eaque molestias, expedita corporis provident perferendis totam suscipit placeat?Necessitatibus facilis mollitia, corporis vero delectus excepturi, beatae iure adipisci neque recusandae eligendi illo ipsum quod quidem temporibus fugiat, id ipsum rem praesentium atque amet quam at nam.Explicabo neque a autem labore consequuntur natus perferendis vitae tempore, dignissimos doloremque natus dolorem eaque enim iure quidem?Illum aliquam id nobis, distinctio possimus explicabo, eum nobis reprehenderit cupiditate unde, deleniti quidem qui ipsa asperiores quod delectus.Eveniet suscipit dolor voluptatem nemo laudantium nesciunt enim, ab ipsam doloribus earum ut architecto?Officiis cum dolores, eos est reprehenderit expedita?Nesciunt illo omnis adipisci at nihil asperiores neque saepe tempora impedit, accusamus fuga voluptates eius quibusdam atque fugit aliquam dignissimos repellendus deserunt.Dolores consequatur perferendis laborum sit vel rerum tempore quasi accusamus, saepe fugit rerum iure est, mollitia laboriosam rem esse nostrum quo, ab dolore sint iusto quaerat provident consectetur dignissimos quas.Laudantium autem praesentium facilis tenetur eius dolorem quidem, debitis exercitationem optio aliquam eum quam libero similique illo ad facere voluptatum, temporibus reiciendis doloribus tempora, corrupti tempore nulla ducimus architecto nesciunt at unde?Facilis est odio numquam ipsum quae modi porro perferendis quos aliquid, eius asperiores non quasi nulla possimus perspiciatis similique dicta, quidem quis aliquid quo rerum consequatur illum eligendi minus illo unde iusto, vel ab ullam maiores inventore laborum voluptatum harum incidunt?Vero obcaecati nisi dicta harum enim impedit omnis accusantium expedita accusamus, modi omnis esse eaque eius numquam sed amet.Quia facere a deleniti ducimus, id laboriosam nisi minima, non voluptatum placeat nam iusto cumque quasi dignissimos, consequatur tempora cupiditate quis.Saepe perspiciatis deserunt corrupti aspernatur, ut maxime totam error eligendi laborum molestias consequatur, earum sed quisquam ipsa laboriosam repudiandae quae illo soluta voluptate at minus, odit blanditiis eum assumenda recusandae eaque porro temporibus, quisquam repellendus voluptatibus ipsam et nihil corporis doloribus dolores soluta eveniet?Natus nihil accusamus doloribus alias fugiat corporis, voluptate ullam rem rerum?Quod ipsa saepe odit, nemo officiis odit hic temporibus voluptatem id officia voluptate consequuntur reiciendis, error blanditiis at, dicta atque inventore nemo quis, sapiente a alias ea molestias?Quos sunt eveniet veritatis distinctio vel odio itaque eos corrupti officia, reiciendis sapiente expedita alias suscipit asperiores neque consequuntur tempore perspiciatis ipsam a?Asperiores corrupti et dolorum est sunt, velit necessitatibus dignissimos assumenda id nobis?Ipsum eveniet suscipit, impedit accusamus nam tempora sint consequatur natus itaque aliquam inventore numquam cupiditate, sapiente nobis esse veniam omnis consectetur harum quidem.Possimus facere omnis, similique qui aut molestias temporibus, quas magnam alias corrupti ut quam, voluptate rerum quo consequatur laborum?Voluptate assumenda eius ut, ipsa quidem expedita rem vel dolorum quae fugiat, maiores quibusdam ducimus minus dolorum in doloremque blanditiis porro sit delectus, magnam saepe perferendis totam explicabo nobis fugiat consectetur labore, voluptatibus quo repellat.Molestias nihil iste esse quidem fugiat laudantium quasi laboriosam sit, asperiores tempora reprehenderit excepturi, officiis sunt saepe sint quos eius molestiae ab nobis veniam fuga?Debitis corrupti maiores magni ratione ex distinctio quae, obcaecati error repellendus ex veritatis aliquid libero deserunt, autem at quasi.Vero esse nam repellat, amet illum unde dolor fugiat nisi cupiditate corporis voluptas, doloremque libero vel saepe fugiat non commodi repellat aspernatur pariatur ab quos, neque aliquam sint amet nisi blanditiis hic consequuntur laborum error?Cum iste commodi nisi maxime id quaerat eligendi voluptas veniam, iste eveniet suscipit fugiat inventore voluptates dolor ut nemo similique, recusandae aperiam hic animi non dolorum voluptates veritatis, totam amet esse quaerat saepe possimus quam aut qui?Iure magni aliquam accusantium consequuntur, placeat ratione doloremque maxime, laborum nam nostrum culpa sunt nesciunt, cum aliquid nobis possimus fuga dolor.Voluptates sapiente voluptatibus sit laudantium fuga quisquam voluptas ducimus, ratione dicta quis et.Fugit quae dolore, beatae tenetur dolor maxime, esse dicta blanditiis quae facere maxime veniam excepturi hic, praesentium dolorum repudiandae voluptas quis debitis consequatur?Tempore dolores nobis rem suscipit dignissimos quaerat sunt, in saepe corrupti nihil delectus tempora illum.Hic iste nostrum, accusantium suscipit soluta neque qui, repudiandae in distinctio nesciunt sequi doloremque nulla quos iure, culpa eum natus maxime deserunt placeat, tempore sequi hic ipsam blanditiis accusamus accusantium nisi?Ullam voluptatem quasi optio quas, vitae quasi magni fugiat iusto culpa quibusdam veritatis cum a rem, facere harum tempore est similique modi et fugit ea sapiente minus, cupiditate libero optio necessitatibus consequatur repellendus itaque?Ut ipsum eius ullam hic fugit magnam asperiores quibusdam distinctio facilis dolorem, facere temporibus provident doloremque ducimus minima quae veritatis repudiandae voluptas.Distinctio voluptas architecto rerum labore, voluptate dolore unde sint beatae eius provident, ratione neque nobis exercitationem obcaecati esse facere.Asperiores pariatur necessitatibus culpa, magni harum repellat nihil fuga debitis, voluptatem distinctio consequuntur vel cupiditate soluta, qui molestias temporibus iusto aspernatur iure eligendi accusantium cum officiis architecto soluta?Blanditiis corrupti maiores ipsam iste at minima ipsum quasi, laudantium suscipit consequuntur corporis nobis inventore eos praesentium quo saepe deleniti, esse unde veritatis exercitationem ex molestias quis libero veniam odio, nostrum non harum beatae nemo porro similique laborum?Voluptatum dolorum itaque accusantium quis quo tempore quidem ullam, numquam fuga libero nesciunt quia eligendi repellat tempore modi sed?Voluptatum nostrum beatae recusandae dicta dignissimos asperiores, similique sint doloribus ratione, quam pariatur velit assumenda quia repellat natus, neque sunt rem quia tempore eaque officia nisi exercitationem at odio repellendus, recusandae eum non repudiandae laboriosam eaque vitae?Sunt eligendi optio ad, architecto quam sed blanditiis quisquam cumque ad in, voluptatum aspernatur consectetur atque minima quas?Deleniti sed maxime optio aspernatur animi expedita, exercitationem asperiores magni inventore unde alias, ipsa aliquam deleniti laborum delectus, iste voluptate quos laboriosam suscipit non corporis, repellat perferendis ex possimus soluta dolores?Modi laudantium ipsa eos molestiae ipsum, tenetur aperiam doloremque, obcaecati enim ea excepturi, necessitatibus placeat ut facere incidunt totam, dolores iste numquam perferendis excepturi quae delectus.Quas accusantium molestiae corrupti asperiores id, laborum quam error fugit id quibusdam, assumenda error voluptate odit dolor accusantium ratione totam nobis repellat unde, blanditiis molestiae nostrum ipsam neque nesciunt ratione.Vel iste qui neque porro corporis, ut numquam natus dolorum minus quaerat quas, quidem dolorum modi ut necessitatibus eveniet laborum minus eligendi eius, dolorem veritatis blanditiis cum nesciunt, vero illum soluta eligendi id hic ipsam natus pariatur?Fugiat optio temporibus illum architecto, odio illum ex tempore exercitationem facilis ab impedit tempora.Sed alias repudiandae ipsum autem velit suscipit excepturi reprehenderit neque aliquam saepe, commodi rerum consectetur quo vitae iste reprehenderit facilis ullam cum, eos velit optio autem consequatur laudantium error dolorum non, repellendus fugiat deleniti impedit officia, dolorem iure excepturi nihil fugiat eius veritatis consequuntur aperiam inventore unde.Magni repudiandae voluptatum nesciunt corporis modi nisi aut suscipit ullam incidunt tempore, quas vel dicta unde labore aliquam repellat, quisquam quis quas harum libero atque dolore quaerat ratione necessitatibus iure, repellendus aliquam cumque rem totam, amet nisi voluptatum?Neque rerum quae amet at cumque nam exercitationem, enim fuga soluta ad nulla dignissimos labore nesciunt, error minus ea dolore, neque veniam quos tempora at, natus aperiam nulla aliquid est explicabo.Est libero voluptates fugit magni sit recusandae et nemo dignissimos expedita laudantium, totam voluptatem est cum fuga molestias sunt consectetur quis explicabo delectus, molestias enim quisquam sed sint neque exercitationem pariatur, beatae amet facilis consequuntur qui iste laborum placeat ratione sapiente, iste vero quisquam consequatur sed eligendi?Doloremque iusto omnis velit quas dignissimos praesentium quam numquam nisi esse natus, sapiente velit consequuntur doloribus commodi necessitatibus nesciunt, saepe ducimus sequi totam numquam fugit omnis dignissimos, voluptatum reiciendis aut.Minus asperiores facilis culpa dicta ipsam animi quos, rem voluptas officiis assumenda eius, officiis porro quae nemo reiciendis aperiam voluptatum omnis dolore esse animi.Odit obcaecati blanditiis reprehenderit ullam quidem ipsam voluptates, doloribus corporis laboriosam delectus ipsam ad sed expedita maiores facilis, tempora exercitationem minima aspernatur rerum neque itaque repellat inventore harum, ullam odit ex non molestias?Quaerat corporis ullam iste itaque distinctio ipsa incidunt voluptas optio, voluptatum sed amet harum ad aut, atque asperiores perspiciatis repellat accusamus numquam ipsa sit, iure provident vel earum ut labore, maxime quia sit omnis soluta aliquid perferendis veniam eaque praesentium tenetur ad?Nemo cupiditate ducimus maiores reprehenderit voluptatum impedit quo, placeat eos dolor illo a odit, tenetur minima natus ea dolor maiores.Amet necessitatibus odio incidunt blanditiis veniam ex quidem, nisi nam est, cumque beatae quia aut minima, doloremque repudiandae ipsa dolorem suscipit sint modi, minima cumque alias voluptatem nobis rem porro minus sit?Numquam vero optio doloribus aut quisquam delectus tempore est aliquam, fuga iste provident?Molestias quos eius quis debitis, sit nobis et recusandae cupiditate, modi laudantium molestiae sunt expedita incidunt itaque asperiores eligendi neque, repellendus inventore totam eius modi dolor tempore voluptatibus molestias iusto, quia blanditiis provident doloribus magnam omnis consequuntur in quidem.Voluptas obcaecati animi quod incidunt modi inventore explicabo iusto laudantium quasi perferendis, vel nihil est autem quibusdam eligendi cumque dolores dolor architecto aliquam, eaque maxime voluptatum nostrum doloribus voluptas delectus, numquam voluptates culpa amet quidem.Ab voluptatem aspernatur harum maxime ea architecto alias non qui suscipit, nostrum sequi ad.Laboriosam officiis impedit autem blanditiis quo qui, quia quisquam dolor natus quo assumenda explicabo.Sequi ipsum id blanditiis repellendus totam aspernatur vel odio, eaque iure iste praesentium expedita facere voluptates laborum quae accusantium.Sapiente mollitia esse incidunt cupiditate natus numquam aperiam velit blanditiis possimus, veniam suscipit fugit pariatur accusantium quos nihil placeat mollitia quidem, ipsam magnam unde temporibus expedita officiis, odit nostrum quo voluptas cum id inventore quidem, minus distinctio facilis cum quas voluptas natus sed.Velit quas autem modi nesciunt perspiciatis deserunt architecto, cupiditate corrupti earum nulla?Quidem assumenda veritatis natus amet magnam totam rerum, assumenda quos unde tempora ducimus quisquam porro hic maiores non labore nihil, sed itaque et iste magni eos cupiditate expedita dolorum laboriosam, laboriosam ducimus hic.Cum est a perferendis quas fuga consequuntur commodi eius, culpa numquam dolore id, repellat id esse delectus porro, id similique deserunt nostrum a, dignissimos veritatis deserunt accusamus perspiciatis officia maiores vel?Praesentium numquam doloremque aliquid similique, rem soluta nihil vel earum vero culpa architecto, voluptatibus voluptatum ipsam odio tempore delectus itaque consequatur eveniet molestias nam?Laudantium modi voluptate quae voluptatibus provident est deleniti ullam consectetur ducimus, expedita aspernatur aut deserunt, quis numquam repellendus sapiente, magni nostrum nam, exercitationem modi rem?Natus quasi fugiat nulla corrupti provident dicta incidunt accusantium vel quod porro, harum adipisci ipsa asperiores sequi, quod vel consectetur a voluptatem itaque quas quam totam ab minus.Blanditiis deserunt praesentium laboriosam soluta aliquam neque, vel qui maxime animi explicabo ipsum.Cum hic error dolore fugiat alias libero, reiciendis in voluptate officiis magni non harum, mollitia eveniet neque corrupti nesciunt cumque temporibus nisi illo ab quae, dignissimos nihil ratione rem debitis totam eveniet tempora vel?Dicta maxime aliquam pariatur molestias velit eos blanditiis in, officiis facere error repellat dignissimos a nobis quidem rerum, neque iusto quibusdam autem aspernatur in, quis sapiente aspernatur nesciunt?Incidunt natus facere, minus quaerat optio sed perferendis accusamus laboriosam alias.Ipsam accusantium tempora atque ex corrupti libero maiores ad consequuntur, sunt dolor voluptatum a placeat eum ipsa consectetur minus vero consequatur aspernatur, asperiores optio est sint aperiam quasi quos esse nostrum quidem, temporibus recusandae ratione quod magni officiis nisi provident dolorem.Minima omnis magnam dolore, qui corporis laudantium ipsam laborum placeat sint dolore explicabo vitae error sit?Obcaecati laboriosam velit, aspernatur nobis eum aliquid error, nisi illo dolorum laboriosam architecto nulla iste magni amet veniam beatae reprehenderit?Quibusdam eligendi quia aut aliquid dignissimos nobis qui porro, quam fuga molestias suscipit cumque nulla doloribus maxime numquam earum reprehenderit ut, pariatur nisi qui placeat quidem ipsa, corrupti aut maiores.Deserunt accusantium illum quaerat, quis fugiat ab, nostrum quo ex ipsum adipisci iure sapiente.Dignissimos rerum reprehenderit porro officia voluptas assumenda cum odit nihil maiores a, nisi culpa neque id aliquam libero quasi reprehenderit maiores odit corporis corrupti, excepturi magni beatae animi ratione sunt sed ipsum eos?Minima tenetur in necessitatibus, libero error magni quod harum debitis, ipsam eligendi porro nam officia quae quo nihil harum saepe, incidunt debitis hic qui nulla earum veniam expedita veritatis tempora inventore, eum enim natus dicta?Tempore laborum omnis, dolore voluptates tempore totam quas aliquam autem ad numquam ipsam.Architecto itaque consequatur cum laborum reiciendis molestias laboriosam, sit beatae amet accusamus doloremque ipsam delectus cum consectetur nobis ullam harum, laudantium animi eveniet, commodi harum ad ut excepturi minima, a error fugit eos reprehenderit sequi quasi.Sequi accusantium cumque ea quos perferendis a nisi error esse, dolorem soluta ipsum.Magni amet impedit illo quas distinctio quisquam dolorum error tempore provident odit, error quaerat officiis aliquid itaque magni quisquam at quod dolores similique quos, soluta ipsum omnis mollitia earum?Delectus maiores inventore quod illum, cum incidunt fuga.Sed cum nulla ducimus ratione tempore eveniet nobis est, expedita magni dicta eligendi, magnam ab distinctio ratione dolor explicabo earum vero, repudiandae ex sapiente enim iste autem deleniti quaerat beatae animi adipisci distinctio?Cum tempore placeat voluptatibus reiciendis rem obcaecati sequi labore mollitia provident accusamus, cumque tempore sit officia nihil repudiandae nulla fuga assumenda est, saepe totam suscipit est sapiente laboriosam nulla libero?Laboriosam exercitationem pariatur modi laudantium doloremque cumque aliquam quia id voluptates, repudiandae ratione odio placeat atque aperiam voluptates voluptate vel quas iure, aut ratione deserunt consequuntur et nobis ipsa aliquid veritatis, molestias deleniti at vel consequuntur ducimus explicabo voluptas quae voluptates, debitis quidem assumenda aliquid est doloribus neque aliquam ab ducimus laudantium iusto.Perspiciatis quia nesciunt dignissimos obcaecati pariatur voluptas unde ipsa dolorem impedit, modi in architecto maiores, ipsa eaque ad voluptates aperiam perferendis repellat dolorem laborum praesentium doloribus, distinctio rem voluptates eos facilis consequuntur, eveniet asperiores expedita sapiente consectetur voluptate quia debitis ea mollitia.Fugit tempora labore libero dolores magni placeat adipisci quidem iure, error perferendis qui praesentium deserunt, reprehenderit accusamus iste doloribus ea alias deserunt repellat ipsam voluptatum, incidunt molestiae ullam commodi.Ullam praesentium molestias, suscipit architecto voluptatum recusandae repellendus temporibus in explicabo quia.Unde expedita cupiditate aspernatur dolorum voluptates deleniti officia, blanditiis totam aspernatur sapiente quos quia laborum quod suscipit, unde reiciendis fugit incidunt error hic molestias recusandae voluptates fuga ducimus?Ratione est nemo, ipsam eligendi itaque ullam architecto aperiam voluptas ducimus nisi fugit quos sit, natus voluptatibus pariatur sed soluta alias unde fugiat corporis eaque placeat, vel blanditiis sunt voluptas ratione enim ducimus inventore unde?Perferendis velit rem provident architecto, earum quia officiis ratione tenetur totam asperiores est illo, labore sint voluptas quae suscipit excepturi corporis autem atque possimus eveniet fugit, neque officiis at laudantium dolorem laboriosam doloremque accusantium perspiciatis perferendis?Provident veritatis minus magni ipsum perspiciatis, consequuntur cum excepturi provident blanditiis nesciunt?Dignissimos eaque ea odio modi at officia temporibus id esse, sint accusantium quasi fugiat aspernatur voluptatum vitae dolorem consequatur ipsam nulla sit?Sint nisi voluptatibus aliquam voluptatum debitis minima, aliquam maxime adipisci recusandae, nihil itaque temporibus ipsam nam ab?Iure commodi eos deleniti soluta id aliquam labore at ipsa, deleniti laborum officiis nesciunt.Nihil veniam accusamus iste odit ut perferendis consectetur similique natus, voluptate qui nulla vel.Assumenda aliquid ratione animi corporis quasi excepturi temporibus, expedita autem voluptas quasi ipsam obcaecati accusantium illum dolores laudantium?Non omnis qui quas alias molestias ipsum nostrum sit aliquam harum voluptatem, optio blanditiis nisi consequuntur impedit.Laudantium ea vitae molestiae, dolore quos repudiandae voluptatum veniam nostrum debitis cupiditate assumenda atque recusandae?Quo nisi pariatur ad iste animi dolor incidunt id ut, iusto corporis magni dignissimos accusantium temporibus fugit tempora, dolor facilis dignissimos perferendis ipsa eos, perferendis temporibus nesciunt hic quisquam nisi inventore provident veniam dolores quo similique, dignissimos itaque odit.Voluptate ea neque repellat aperiam nostrum voluptas nemo vero, atque omnis eos deserunt aspernatur architecto excepturi id, quo facilis nam molestias dolorum voluptas voluptate temporibus doloremque iure soluta?Perspiciatis ipsa sint at vero amet, nemo nihil voluptas rerum quae similique, et exercitationem cum laboriosam, non accusantium facere fuga doloribus eveniet repellat incidunt obcaecati quos mollitia, incidunt maxime odio eligendi ducimus et?Omnis consectetur sapiente nobis in molestiae obcaecati sit minus laboriosam, alias ipsa voluptas culpa fugit perspiciatis accusamus molestiae?Totam officia ex facilis consectetur recusandae aspernatur fugiat soluta assumenda, iure quisquam iste pariatur id sint odio unde, ducimus numquam distinctio ipsum atque voluptatum earum perspiciatis reiciendis, illum ipsam porro natus velit.Dolores ex sapiente nemo error repellendus qui assumenda animi ducimus, corrupti nemo vitae voluptatem ut rem accusamus velit nam, totam sint accusantium temporibus facilis ducimus rerum, eveniet harum facere debitis esse inventore aperiam sint placeat dolore, repellendus consequuntur assumenda quam rerum doloribus enim autem sed.Nobis doloremque aliquam sapiente dolorem cupiditate consequatur distinctio beatae, quas provident eius eum natus architecto facere dolore itaque totam ex voluptate, repudiandae fuga aspernatur unde accusamus praesentium deserunt, obcaecati vitae autem ad deserunt expedita sapiente error ab ullam hic, impedit assumenda eveniet facere.Animi dolor quibusdam pariatur reiciendis expedita, assumenda numquam reprehenderit magni laborum tenetur officiis eveniet, doloremque porro optio laboriosam tenetur ipsa facilis error neque autem fugiat odit, natus placeat quas debitis ad sunt atque fugit officia eum assumenda.Deleniti suscipit eveniet explicabo, nostrum error eligendi totam libero repellat atque accusamus facilis laborum?Omnis obcaecati aut quidem sequi dicta commodi, sit excepturi debitis molestias aut reprehenderit similique tempore, cum explicabo tempore fuga iusto.Neque odio nisi fugit eligendi laborum non dolorem necessitatibus accusantium, quia iste obcaecati enim, vitae quo ullam nemo harum cumque numquam dolorem distinctio?Necessitatibus dolorum ipsam amet dolore quos magnam explicabo excepturi, ducimus amet eaque doloribus fuga iste error rerum commodi dicta provident.Atque alias obcaecati accusantium minima error, rem voluptate deserunt enim, vero laborum impedit asperiores qui nulla autem dolore?Officiis vitae nemo nihil vel totam similique ducimus ut, fuga eos ullam quos, enim mollitia similique delectus cum nam culpa, aperiam ut aut eum ab consequuntur explicabo tempora placeat non blanditiis, voluptate expedita asperiores saepe nesciunt porro minus.Ratione in rerum dolorum totam voluptates doloremque quaerat molestiae saepe, possimus cumque minima blanditiis recusandae veritatis maxime aspernatur sequi culpa, molestias labore tenetur voluptas facere voluptate aliquam saepe fugit, consequatur a veritatis, quisquam est temporibus veritatis magnam ex voluptas quae unde quis qui quod.Mollitia ea eius tempora in eveniet, atque molestiae ab magnam, consequuntur sed animi nobis laborum ea iure culpa officiis ducimus est, aperiam repudiandae suscipit, recusandae molestias corporis nemo distinctio mollitia reiciendis labore omnis?Inventore modi tempore deserunt ducimus repellat qui accusantium distinctio veritatis a, esse deleniti obcaecati iste quo perferendis ut minus, commodi veritatis ullam vitae consequuntur ipsa nesciunt, quo magni nostrum vitae repellendus ipsa eos eligendi fuga nisi.Aut labore excepturi ducimus molestiae voluptatem culpa laborum ipsa sapiente ad doloribus, veritatis ipsa quis culpa voluptatem delectus numquam tempora a consequuntur reiciendis facere, laudantium cumque vitae sit molestias impedit similique, saepe fugit maiores, modi nesciunt rerum quos beatae eos temporibus vero quam nam.Ullam amet earum saepe ut perspiciatis, in architecto ratione nam voluptatibus tenetur natus veritatis aliquam totam id, voluptatibus adipisci suscipit id ipsa iure, quae distinctio facilis error eos officiis molestiae, aperiam deserunt neque quasi totam sunt molestiae labore perferendis pariatur?Rem adipisci quaerat rerum cumque vero cum ducimus consequuntur sit nulla, nulla hic laudantium tempora quam?Tempore qui dolore doloremque officiis sequi atque magni, molestiae non modi officiis praesentium accusantium, laborum aut tenetur, voluptatibus praesentium explicabo earum impedit porro voluptatem aperiam nobis ex blanditiis fuga.Error tenetur delectus provident expedita, in quisquam quis itaque nobis earum eum, laborum repellendus veniam autem voluptatum harum?Debitis eaque non obcaecati pariatur at, libero quia repudiandae, aspernatur dolorum in aperiam fugiat ullam possimus soluta excepturi, at magnam tempore veritatis aliquam vero iusto excepturi nam libero.Ea veritatis esse sunt nostrum saepe quibusdam, at quod voluptatum perspiciatis nemo quis cum, non doloremque minima enim error placeat vel?Libero voluptatem ullam, blanditiis modi sapiente minus sit?Alias temporibus repudiandae eius, modi explicabo optio perspiciatis porro, voluptate molestias expedita tempora eum aperiam fugit?

Mathematical Proofs

Proof of Lemma 1

Proof. From the convexity of $h(\cdot) + \frac{\alpha}{2}\|\cdot\|^2$ in Assumption 1, we obtain

$$h(x_{k+1}) + \frac{\alpha}{2}\|x_{k+1}\|^2 + \langle \xi_{k+1} + \alpha x_{k+1}, x_k - x_{k+1}\rangle \leq h(x_k) + \frac{\alpha}{2}\|x_k\|^2,$$

where $\xi_{k+1} \in \partial h(x_{k+1})$. By rearranging the inequality we obtain

$$h(x_{k+1}) - \frac{\alpha}{2}\|x_{k+1} - x_k\|^2 + \langle \xi_{k+1}, x_k - x_{k+1}\rangle \leq h(x_k). \tag{20}$$

From the first-order optimality condition of (6) in Algorithm 1, it shows that

$$\xi_{k+1} + \tilde{\nabla} f(\bar{x}_k) + \frac{1}{\eta_k}(\nabla\psi(x_{k+1}) - \nabla\psi(\bar{x}_k)) = 0.$$

Combining the above inequality with (20), we can derive the following result

$$\begin{aligned}
&h(x_{k+1}) - \frac{\alpha}{2}\|x_{k+1} - x_k\|^2 - \langle \tilde{\nabla} f(\bar{x}_k), x_k - x_{k+1}\rangle + \frac{1}{\eta_k}\langle \nabla\psi(\bar{x}_k) - \nabla\psi(x_{k+1}), x_k - x_{k+1}\rangle \\
=&h(x_{k+1}) - \frac{\alpha}{2}\|x_{k+1} - x_k\|^2 - \langle \tilde{\nabla} f(\bar{x}_k), x_k - x_{k+1}\rangle + \frac{1}{\eta_k}(D_\psi(x_k, x_{k+1}) + D_\psi(x_{k+1}, \bar{x}_k) - D_\psi(x_k, \bar{x}_k)) \\
\leq&h(x_k),
\end{aligned} \tag{21}$$

where the last equality follows from the three-point identity.

Furthermore, since f is an $(\bar{L}, \underline{L})$-relative smooth function with respect to ψ, we have

$$f(x_{k+1}) \leq f(\bar{x}_k) + \langle \nabla f(\bar{x}_k), x_{k+1} - \bar{x}_k\rangle + \bar{L} D_\psi(x_{k+1}, \bar{x}_k),$$

and

$$f(\bar{x}_k) + \langle \nabla f(\bar{x}_k), x_k - \bar{x}_k\rangle \leq f(x_k) + \underline{L} D_\psi(x_k, \bar{x}_k).$$

It shows that

$$f(x_{k+1}) \leq f(x_k) + \langle \nabla f(\bar{x}_k), x_{k+1} - x_k\rangle + \underline{L} D_\psi(x_k, \bar{x}_k) + \bar{L} D_\psi(x_{k+1}, \bar{x}_k). \tag{22}$$

By summing inequalities (21) and (22) together, we obtain

$$\begin{aligned}
\Phi(x_{k+1}) \leq&\Phi(x_k) + \frac{\alpha}{2}\|x_{k+1} - x_k\|^2 + \langle \nabla f(\bar{x}_k) - \tilde{\nabla} f(\bar{x}_k), x_{k+1} - x_k\rangle \\
&+ \left(\frac{1}{\eta_k} + \underline{L}\right) D_\psi(x_k, \bar{x}_k) - \frac{1}{\eta_k} D_\psi(x_k, x_{k+1}) + \left(\bar{L} - \frac{1}{\eta_k}\right) D_\psi(x_{k+1}, \bar{x}_k) \\
\leq&\Phi(x_k) + \frac{\alpha + \gamma_k}{2}\|x_{k+1} - x_k\|^2 + \frac{1}{2\gamma_k}\|\nabla f(\bar{x}_k) - \tilde{\nabla} f(\bar{x}_k)\|_*^2 \\
&+ \left(\frac{1}{\eta_k} + \underline{L}\right) D_\psi(x_k, \bar{x}_k) - \frac{1}{\eta_k} D_\psi(x_k, x_{k+1}),
\end{aligned} \tag{23}$$

where the last inequality follows from $\langle a, b\rangle \leq \frac{\gamma}{2}\|a\|^2 + \frac{1}{2\gamma}\|b\|^2$ for any $\gamma_k > 0$ and $\eta_k \leq \bar{L}^{-1}$.

By applying the conditional expectation operator $\mathbb{E}_k$ to the above inequality and bounding the MSE term by (8) in Definition 4, we have

$$\begin{aligned}
\mathbb{E}_k[\Phi(x_{k+1})] \leq&\Phi(x_k) + \frac{\alpha + \gamma_k}{2}\mathbb{E}_k[\|x_{k+1} - x_k\|^2] + \frac{1}{2\gamma_k}\mathbb{E}_k[\|\nabla f(\bar{x}_k) - \tilde{\nabla} f(\bar{x}_k)\|_*^2] \\
&+ \left(\frac{1}{\eta_k} + \underline{L}\right) D_\psi(x_k, \bar{x}_k) - \frac{1}{\eta_k}\mathbb{E}_k[D_\psi(x_k, x_{k+1})] \\
\leq&\Phi(x_k) + \frac{\alpha + \gamma_k}{2}\mathbb{E}_k[\|x_{k+1} - x_k\|^2] + \frac{1}{2\gamma_k}\Gamma_k + \frac{V_1}{2\gamma_k}\|x_k - x_{k-1}\|^2 \\
&+ \frac{V_1}{2\gamma_k}\|x_{k-1} - x_{k-2}\|^2 + \left(\frac{1}{\eta_k} + \underline{L}\right) D_\psi(x_k, \bar{x}_k) - \frac{1}{\eta_k}\mathbb{E}_k[D_\psi(x_k, x_{k+1})] \\
\leq&\Phi(x_k) + \frac{\alpha + \gamma_k}{2}\mathbb{E}_k[\|x_{k+1} - x_k\|^2] + \frac{1}{2\gamma_k\tau}(\Gamma_k - \mathbb{E}_k[\Gamma_{k+1}]) \\
&+ \left(\frac{V_\Gamma}{2\gamma_k\tau} + \frac{V_1}{2\gamma_k}\right)\|x_k - x_{k-1}\|^2 + \left(\frac{V_\Gamma}{2\gamma_k\tau} + \frac{V_1}{2\gamma_k}\right)\|x_{k-1} - x_{k-2}\|^2 \\
&+ \left(\frac{1}{\eta_k} + \underline{L}\right) D_\psi(x_k, \bar{x}_k) - \frac{1}{\eta_k}\mathbb{E}_k[D_\psi(x_k, x_{k+1})],
\end{aligned} \tag{24}$$

where the last inequality follows from (10) in Definition 4. From (5) and $\eta_k \le \min\{\eta_{k-1}, \bar{L}^{-1}\}$, it shows that

$$\left(\frac{1}{\eta_k} + \underline{L}\right) D_\psi(x_k, \bar{x}_k) \le \frac{\underline{L}\eta_k + 1}{\eta_k} \frac{\delta - \epsilon}{1 + \underline{L}\eta_{k-1}} D_\psi(x_{k-1}, x_k) \le \frac{\delta - \epsilon}{\eta_k} D_\psi(x_{k-1}, x_k). \tag{25}$$

Combining (24) with (25), we can get

$$\begin{aligned}\mathbb{E}_k[\Phi(x_{k+1})] \le& \Phi(x_k) - \left(\frac{1}{\eta_k} - \alpha - \gamma_k\right) \mathbb{E}_k[D_\psi(x_k, x_{k+1})] + \frac{1}{2\gamma_k\tau}(\Gamma_k - \mathbb{E}_k[\Gamma_{k+1}]) \\ &+ \left(\frac{\delta - \epsilon}{\eta_k} + \frac{V_\Gamma}{\gamma_k\tau} + \frac{V_1}{\gamma_k}\right) D_\psi(x_{k-1}, x_k) + \left(\frac{V_\Gamma}{\gamma_k\tau} + \frac{V_1}{\gamma_k}\right) D_\psi(x_{k-2}, x_{k-1}).\end{aligned}$$

Therefore, the results can be obtained by rearranging the above terms with $\gamma_k = \sqrt{2(V_\Gamma/\tau + V_1)}$. This completes the proof. ☐

Proof of Lemma 2

Proof. From Lemma 1, it shows that

$$\begin{aligned}\eta_k(\Phi(x_k) - \mathcal{V}(\Phi)) \ge& \eta_k(\mathbb{E}_k[\Phi(x_{k+1}]) - \mathcal{V}(\Phi)) + (1 - \eta_k\alpha - \eta_k\gamma)\mathbb{E}_k[D_\psi(x_k, x_{k+1})] \\ &+ \frac{\eta_k}{2\tau\gamma}(\mathbb{E}_k[\Gamma_{k+1}] - \Gamma_k) - \left(\delta - \epsilon + \frac{\gamma\eta_k}{2}\right) D_\psi(x_{k-1}, x_k) - \frac{\gamma\eta_k}{2} D_\psi(x_{k-2}, x_{k-1}).\end{aligned} \tag{26}$$

Combining (26) with $\eta_k \le \eta_{k-1}$, we have

$$\begin{aligned}&\Psi_k - \mathbb{E}_k[\Psi_{k+1}] \\ =& \eta_{k-1}(\Phi(x_k) - \mathcal{V}(\Phi)) + \left(1 - \eta_{k-1}\alpha - \eta_{k-1}\gamma - \frac{\epsilon}{3}\right) D_\psi(x_{k-1}, x_k) - \frac{\eta_k}{2\tau\gamma}\mathbb{E}_k[\Gamma_{k+1}] \\ &+ \frac{\eta_{k-1}}{2\tau\gamma}\Gamma_k + \eta_{k-1}\left(\frac{\gamma}{2} + \frac{\epsilon}{3\eta_{k-1}}\right) D_\psi(x_{k-2}, x_{k-1}) - \eta_k(\mathbb{E}_k[\Phi(x_{k+1})] - \mathcal{V}(\Phi)) \\ &- \eta_k\left(\frac{\gamma}{2} + \frac{\epsilon}{3\eta_k}\right) D_\psi(x_{k-1}, x_k) - (1 - \eta_k\alpha - \eta_k\gamma - \frac{\epsilon}{3})\mathbb{E}_k[D_\psi(x_k, x_{k+1})] \\ \ge& \eta_k(\Phi(x_k) - \mathcal{V}(\Phi)) + \left(1 - \eta_{k-1}\alpha - \eta_{k-1}\gamma - \frac{\epsilon}{3}\right) D_\psi(x_{k-1}, x_k) - \frac{\eta_k}{2\tau\gamma}\mathbb{E}_k[\Gamma_{k+1}] \\ &+ \frac{\eta_k}{2\tau\gamma}\Gamma_k + \eta_{k-1}\left(\frac{\gamma}{2} + \frac{\epsilon}{3\eta_{k-1}}\right) D_\psi(x_{k-2}, x_{k-1}) - \eta_k(\mathbb{E}_k[\Phi(x_{k+1})] - \mathcal{V}(\Phi)) \\ &- \eta_k\left(\frac{\gamma}{2} + \frac{\epsilon}{3\eta_k}\right) D_\psi(x_{k-1}, x_k) - \left(1 - \eta_k\alpha - \eta_k\gamma - \frac{\epsilon}{3}\right) \mathbb{E}_k[D_\psi(x_k, x_{k+1})] \\ \ge& (1 - \delta - \eta_{k-1}\alpha - (\eta_{k-1} + \eta_k)\gamma) D_\psi(x_{k-1}, x_k) + \frac{\epsilon}{3}(\mathbb{E}_k D_\psi(x_k, x_{k+1}) + D_\psi(x_{k-1}, x_k) + D_\psi(x_{k-2}, x_{k-1})) \\ \ge& (1 - \delta - \eta_{k-1}\alpha - 2\eta_{k-1}\gamma) D_\psi(x_{k-1}, x_k) + \frac{\epsilon}{3}(\mathbb{E}_k D_\psi(x_k, x_{k+1}) + D_\psi(x_{k-1}, x_k) + D_\psi(x_{k-2}, x_{k-1})) \\ \ge& \frac{\epsilon}{3}(\mathbb{E}_k D_\psi(x_k, x_{k+1}) + D_\psi(x_{k-1}, x_k) + D_\psi(x_{k-2}, x_{k-1})),\end{aligned}$$

where the second and the last inequality follow from (26) and (11), respectively. This completes the proof. ☐

Proof of Theorem 1

Proof. (i) This statement follows directly from Lemma 2 and $\epsilon > 0$.

(ii) By summing (13) from $k = 0$ to a positive integer K, we have

$$\sum_{k=1}^{K} \mathbb{E}[D_\psi(x_{k-1}, x_k)] \le \frac{3}{\epsilon}\mathbb{E}[\Psi_1 - \Psi_{K+1}] \le \frac{3}{\epsilon}\Psi_1,$$

where the last inequality follows from $\Psi_k \ge 0$ for any $k > 0$. Taking the limit as $K \to +\infty$, we have $\sum_{k=1}^{+\infty} \mathbb{E}[D_\psi(x_{k-1}, x_k)] < +\infty$. Then we may deduce that the sequence $\{\mathbb{E}[D_\psi(x_{k-1}, x_k)]\}$ converges to zero.

(iii) We have

$$K \min_{1\le k\le K} \mathbb{E}[D_\psi(x_{k-1}, x_k)] \le \sum_{k=1}^{K} \mathbb{E}[D_\psi(x_{k-1}, x_k)] \le \frac{3}{\epsilon}\Psi_1,$$

which yields the desired result.

This completes the proof. ☐

Proof of Proposition 1

Proof. For the SARAH stochastic gradient estimator, we can get the results directly similar to the proof of Lemma 5 in (**?**).

Now we consider the proof of Proposition 1 (ii). Firstly, we define the SAGA stochastic gradient estimator $\tilde{\nabla}^{SAGA} f(\bar{x}_k)$ as

$$\tilde{\nabla}^{SAGA} f(\bar{x}_k) := \frac{1}{b} \sum_{j \in B_k} \Big(\nabla f_j(\bar{x}_k) - \nabla f_j(z_k^j) \Big) + \frac{1}{n} \sum_{i=1}^{n} \nabla f_i(z_k^i),$$

where $z_{k+1}^i = \begin{cases} \bar{x}_k, & \text{if } i \in B_k, \\ z_k^i, & \text{otherwise.} \end{cases}$ From the Lipschitz continuity of $\nabla f_i(\cdot)$, it shows that

$$\begin{aligned}
\mathbb{E}_k \|\tilde{\nabla}^{SAGA} f(\bar{x}_k) - \nabla f(\bar{x}_k)\|_*^2 =& \mathbb{E}_k \| \frac{1}{b} \sum_{j \in B_k} \Big(\nabla f_j(\bar{x}_k) - \nabla f_j(z_k^j) \Big) + \frac{1}{n} \sum_{i=1}^{n} \nabla f_i(z_k^i) - \nabla f(\bar{x}_k) \|_*^2 \\
\leq & \frac{1}{b^2} \sum_{j \in B_k} \|\nabla f_j(\bar{x}_k) - \nabla f_j(z_k^j)\|_*^2 \\
=& \frac{1}{bn} \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_k) - \nabla f_i(z_k^i)\|_*^2,
\end{aligned}$$

where last inequality follows from the fact that $\mathbb{E}_k\|y_1 + \cdots + y_t\|_*^2 = \mathbb{E}_k\|y_1\|_*^2 + \cdots + \mathbb{E}_k\|y_t\|_*^2$ for any independent random variables $y_i (i = 1, \ldots, t)$ with $\mathbb{E}_k[y_i] = 0$ for all i. Combined with Jensen's inequality, we can get

$$\begin{aligned}
\mathbb{E}_k \|\tilde{\nabla}^{SAGA} f(\bar{x}_k) - \nabla f(\bar{x}_k)\|_* \leq & \sqrt{\mathbb{E}_k \|\tilde{\nabla}^{SAGA} f(\bar{x}_k) - \nabla f(\bar{x}_k)\|_*^2} \\
\leq & \frac{1}{\sqrt{bn}} \sqrt{\sum_{i=1}^{n} \|\nabla f_i(\bar{x}_k) - \nabla f_i(z_k^i)\|_*^2} \\
\leq & \frac{1}{\sqrt{bn}} \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_k) - \nabla f_i(z_k^i)\|_*.
\end{aligned}$$

We bound the MSE of the stochastic gradient estimator $\tilde{\nabla}^{SAGA} f(\cdot)$ as follows,

$$\begin{aligned}
& \frac{1}{bn} \sum_{i=1}^{n} \mathbb{E}_k \|\nabla f_i(\bar{x}_k) - \nabla f_i(z_k^i)\|_*^2 \\
\leq & \frac{1+\delta}{bn} \mathbb{E}_k \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_{k-1}) - \nabla f_i(z_k^i)\|_*^2 + \frac{1+\delta^{-1}}{bn} \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_k) - \nabla f_i(\bar{x}_{k-1})\|_*^2 \\
\leq & \frac{1+\delta}{bn} (1 - \frac{b}{n}) \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_{k-1}) - \nabla f_i(z_{k-1}^i)\|_*^2 + \frac{1+\delta^{-1}}{b} M_1^2 \|\bar{x}_k - \bar{x}_{k-1}\|^2 \\
\leq & \frac{1+\delta}{bn} (1 - \frac{b}{n}) \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_{k-1}) - \nabla f_i(z_{k-1}^i)\|_*^2 + \frac{1+\delta^{-1}}{b} M_1^2 [(1+\beta_k^2)\|x_k - x_{k-1}\|^2 + \beta_{k-1}^2 \|x_{k-1} - x_{k-2}\|^2] \\
\leq & \frac{1+\delta}{bn} (1 - \frac{b}{n}) \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_{k-1}) - \nabla f_i(z_{k-1}^i)\|_*^2 + \frac{2+2\delta^{-1}}{b} M_1^2 [\|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2],
\end{aligned}$$

where the first inequality follows from $\|x - z\|_*^2 \leq (1+\delta)\|x - y\|_*^2 + (1+\delta^{-1})\|y - z\|_*^2$. Let $\Gamma_{k+1} := \frac{1}{bn} \sum_{i=1}^{n} \|\nabla f_i(\bar{x}_k) - \nabla f_i(z_k^i)\|_*^2$ and $\delta = \frac{b}{2n}$, it shows that

$$\begin{aligned}
\mathbb{E}_k \Gamma_{k+1} \leq & (1 + \frac{b}{2n})(1 - \frac{b}{n}) \Gamma_k + \frac{2 + \frac{4n}{b}}{b} M_1^2 [\|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] \\
\leq & (1 - \frac{b}{2n}) \Gamma_k + \frac{2b + 4n}{b^2} / 2n + \frac{4n^2}{b} M_1^2 [\|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2].
\end{aligned}$$

This proves the geometric decay of Γ_k in expectation. Similar to Appendix B in (**?**), we also have that the third condition holds in Definition 4. This completes the proof. $\square$

Proof of Lemma 3

Proof. From the implicit definition of the proximal operator (6) in the BPSGE algorithm, we have that

$$0 \in \partial h(x_{k+1}) + \tilde{\nabla} f(\bar{x}_k) + \frac{1}{\eta_k}(\nabla \psi(x_{k+1}) - \nabla \psi(\bar{x}_k)).$$

Combining it with $\partial \Phi(x_{k+1}) \equiv \nabla f(x_{k+1}) + \partial h(x_{k+1})$, we have $w_{k+1} \in \partial \Phi(x_{k+1})$. All that remains is to bound the norm of w_{k+1}. ∇f and $\nabla \psi$ are Lipschitz continuous with constants M_1 and M_2 on any bounded subset of $\mathbb{R}^d$, respectively (See Assumption 3). It shows that

$$\begin{aligned}
&\mathbb{E}_k\|w_{k+1}\| \\
\leq &\mathbb{E}_k\|\nabla f(x_{k+1}) - \tilde{\nabla} f(\bar{x}_k) + \frac{1}{\eta_k}(\nabla \psi(\bar{x}_k) - \nabla \psi(x_{k+1}))\| \\
\leq &\mathbb{E}_k\|\nabla f(x_{k+1}) - \tilde{\nabla} f(\bar{x}_k)\| + \frac{1}{\eta_k}\mathbb{E}_k\|\nabla \psi(\bar{x}_k) - \nabla \psi(x_{k+1})\| \\
\leq &\mathbb{E}_k\|\nabla f(x_{k+1}) - \nabla f(\bar{x}_k)\| + \mathbb{E}_k\|\nabla f(\bar{x}_k) - \tilde{\nabla} f(\bar{x}_k)\| + \frac{1}{\eta_k}\mathbb{E}_k\|\nabla \psi(\bar{x}_k) - \nabla \psi(x_{k+1})\| \\
\leq &M_1\mathbb{E}_k\|x_{k+1} - \bar{x}_k\| + \Upsilon_k + V_2\|x_k - x_{k-1}\| + V_2\|x_{k-1} - x_{k-2}\| + \frac{M_2}{\eta_k}\mathbb{E}_k\|x_{k+1} - \bar{x}_k\| \\
\leq &\left(M_1 + \frac{M_2}{\eta_k}\right)\mathbb{E}_k\|x_{k+1} - x_k\| + \left(V_2 + \beta_k M_1 + \frac{\beta_k M_2}{\eta_k}\right)\|x_k - x_{k-1}\| + V_2\|x_{k-1} - x_{k-2}\| + \Upsilon_k \\
\leq &\left(M_1 + \frac{M_2}{\eta}\right)\mathbb{E}_k\|x_{k+1} - x_k\| + \left(V_2 + \beta_k M_1 + \frac{\beta_k M_2}{\eta}\right)\|x_k - x_{k-1}\| + V_2\|x_{k-1} - x_{k-2}\| + \Upsilon_k \\
\leq &\rho(\mathbb{E}_k\|x_{k+1} - x_k\| + \|x_k - x_{k-1}\| + \|x_{k-1} - x_{k-2}\|) + \Upsilon_k,
\end{aligned}$$

where $\rho = \max\left\{M_1 + \frac{M_2}{\eta}, V_2 + \beta_k M_1 + \frac{\beta_k M_2}{\eta}, V_2\right\}$. This completes the proof. □

Proof of Lemma 4

Proof. From Lemma 3, it shows that

$$\begin{aligned}
&\mathbb{E}_k\|w_{k+1}\|^2 \\
\leq &3\mathbb{E}_k\|\nabla f(x_{k+1}) - \nabla f(\bar{x}_k)\|^2 + 3\mathbb{E}_k\|\nabla f(\bar{x}_k) - \tilde{\nabla} f(\bar{x}_k)\|^2 + \frac{3}{\eta_k}\mathbb{E}_k\|\nabla \psi(\bar{x}_k) - \nabla \psi(x_{k+1})\|^2 \\
\leq &3M_1^2\mathbb{E}_k\|x_{k+1} - \bar{x}_k\|^2 + 3\Gamma_k + 3V_1\|x_k - x_{k-1}\|^2 + 3V_1\|x_{k-1} - x_{k-2}\|^2 + \frac{3M_2^2}{\eta_k}\mathbb{E}_k\|x_{k+1} - \bar{x}_k\|^2 \\
\leq &\left(6M_1^2 + \frac{6M_2^2}{\eta_k}\right)\mathbb{E}_k\|x_{k+1} - x_k\|^2 + \left(3V_1 + 6\beta_k^2 M_1^2 + \frac{6\beta_k^2 M_2^2}{\eta_k}\right)\|x_k - x_{k-1}\|^2 + 3V_1\|x_{k-1} - x_{k-2}\|^2 + 3\Gamma_k \\
\leq &\left(6M_1^2 + \frac{6M_2^2}{\eta}\right)\mathbb{E}_k\|x_{k+1} - x_k\|^2 + \left(3V_1 + 6\beta_k^2 M_1^2 + \frac{6\beta_k^2 M_2^2}{\eta}\right)\|x_k - x_{k-1}\|^2 + 3V_1\|x_{k-1} - x_{k-2}\|^2 + 3\Gamma_k \\
\leq &\bar{\rho}(\mathbb{E}_k\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2) + 3\Gamma_k,
\end{aligned}$$

where $\bar{\rho} := \max\left\{6M_1^2 + \frac{6M_2^2}{\eta}, 3V_1 + 6\beta_k^2 M_1^2 + \frac{6\beta_k^2 M_2^2}{\eta}, 3V_1\right\}$. By $\mathrm{dist}(0, \partial \Phi(x_{k+1}))^2 \leq \|w_{k+1}\|^2$ and taking full expectation on both sides, it shows that

$$\mathbb{E}[\mathrm{dist}(0, \partial \Phi(x_{k+1}))^2] \leq \bar{\rho}\mathbb{E}[\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] + 3\mathbb{E}\Gamma_k.$$

This completes the proof. □

Proof of Theorem 2

Proof. From Corollary 1 and Corollary 4, it shows that

$$
\begin{aligned}
&\mathbb{E}[\Psi_k - \Psi_{k+1}] \\
\geq& \frac{\epsilon}{6}\mathbb{E}[\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] \\
\geq& \sigma\mathbb{E}[\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] + \frac{\epsilon/6 - \sigma}{\bar{\rho}}\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2] - \frac{\epsilon/2 - 3\sigma}{\bar{\rho}}\mathbb{E}\Gamma_k \\
\geq& \sigma\mathbb{E}[\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] + \frac{\epsilon/2 - 3\sigma}{\tau\bar{\rho}}\mathbb{E}[\Gamma_{k+1} - \Gamma_k] \\
&- \frac{(\epsilon/2 - 3\sigma)V_\Gamma}{\tau\bar{\rho}}\mathbb{E}[\|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] + \frac{\epsilon/6 - \sigma}{\bar{\rho}}\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2] \\
\geq& \sigma\mathbb{E}[\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] + \frac{\epsilon/2 - 3\sigma}{\tau\bar{\rho}}\mathbb{E}[\Gamma_{k+1} - \Gamma_k] \\
&- \frac{(\epsilon/2 - 3\sigma)V_\Gamma}{\tau\bar{\rho}}\mathbb{E}[\|x_{k+1} - x_k\|^2 + \|x_k - x_{k-1}\|^2 + \|x_{k-1} - x_{k-2}\|^2] + \frac{\epsilon/6 - \sigma}{\bar{\rho}}\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2],
\end{aligned}
$$

where the third inequality follows from (10) in Definition 4. If we let $\sigma = \frac{(\epsilon/2-3\sigma)V_\Gamma}{\tau\bar{\rho}}$, i.e., $\sigma = \frac{\frac{\epsilon}{2}V_\Gamma}{3V_\Gamma + \tau\bar{\rho}}$, it shows that

$$
\mathbb{E}[\Psi_k - \Psi_{k+1}] \geq \frac{\epsilon/6 - \sigma}{\bar{\rho}}\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2] + \frac{\epsilon/2 - 3\sigma}{\tau\bar{\rho}}\mathbb{E}[\Gamma_{k+1} - \Gamma_k].
$$

Summing up $k = 1$ to K, we have

$$
\mathbb{E}[\Psi_1 - \Psi_{K+1}] \geq \frac{\epsilon/6 - \sigma}{\bar{\rho}}\sum_{k=1}^{K}\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2] + \frac{\epsilon/2 - 3\sigma}{\tau\bar{\rho}}\mathbb{E}[\Gamma_{K+1} - \Gamma_1],
$$

which means there exists a $k' \in \{2, \ldots, K+1\}$ such that

$$
\begin{aligned}
\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k'}))^2] \leq& \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[\mathrm{dist}(0, \partial\Phi(x_{k+1}))^2] \\
\leq& \frac{\bar{\rho}}{(\epsilon/6 - \sigma)K}(\mathbb{E}[\Psi_1 - \Psi_{K+1}] + \frac{\epsilon/2 - 3\sigma}{\tau\bar{\rho}}\mathbb{E}[\Gamma_1 - \Gamma_{K+1}]) \\
\leq& \frac{\bar{\rho}}{(\epsilon/6 - \sigma)K}(\mathbb{E}\Psi_1 + \frac{\epsilon/2 - 3\sigma}{\tau\bar{\rho}}\mathbb{E}\Gamma_1).
\end{aligned}
$$

This completes the proof. □

Proof of Lemma 5

Proof. The four statements can be easily obtained by Corollary 1 and Lemma 3, so we omit the details here for simplicity. □

The following lemma is from (**?**), which is analogous to the Uniformized KŁ property of (**?**) and allows us to apply the KŁ inequality.

Lemma 6. *Let $\{x_k\}_{k=0}^{\infty}$ be a bounded sequence of iterates generated by the BPSGE algorithm using a variance-reduced gradient estimator (see Definition 4), and let Φ be a semialgebraic function satisfying the KŁ property (**?**) with exponent θ. Then there exists an index $\bar{k}$ and a desingularizing function $\phi(r) = ar^{1-\theta}$ with $a > 0$, $\theta \in [0, 1)$ so that the following bound holds almost surely (a.s.),*

$$
\phi'(\mathbb{E}[\Phi(x_k) - \Phi_k^*])\mathbb{E}\mathit{dist}(0, \partial\Phi(x_k)) \geq 1, \ \forall k > \bar{k}, \tag{27}
$$

where Φ_k^* is a nondecreasing sequence converging to $\mathbb{E}\Phi(x_*)$ for some $x_* \in \omega(x_0)$.

Proof of Theorem 3

Proof. According to Lemma 6, if Φ is a proper, lower semi-continuous, and semi-algebraic function, it will satisfy the KL property at any point of dom Φ. Under Lemma 6, combining Corollary 1 with Lemma 3, we can get that the generated sequence $\{x_k\}$ is a Cauchy sequence which yields the result. The detailed proof of this theorem is similar to Theorem 2 in (**?**). Thus the details are omitted here. □

Proof of Proposition 2

Proof. Combining Proposition C.4 with Proposition D.1 in (?), we can directly get this result. □

Proof of Proposition 3

Proof. From the update step for solving (6), we have

$$\begin{aligned}
&(U_{k+1}, V_{k+1}) \\
\in &\underset{U\in\mathbb{R}^{m\times r}, V\in\mathbb{R}^{r\times d}}{\arg\min} \{\eta_k h(U,V) + \langle P_k, U\rangle + \langle Q_k, V\rangle + \psi(U,V)\} \\
= &\underset{U\in\mathbb{R}^{m\times r}, V\in\mathbb{R}^{r\times d}}{\arg\min} \left\{\eta_k \left(\lambda_1\|U\|_1 - \frac{\lambda_2}{2}\|U\|_F^2\right) + \langle P_k, U\rangle + \langle Q_k, V\rangle + 3\psi_1(U,V) + \|M\|_F\psi_2(U,V) + \frac{\eta_k\lambda_2}{2}\|U\|_F^2\right\} \\
= &\underset{U\in\mathbb{R}^{m\times r}, V\in\mathbb{R}^{r\times d}}{\arg\min} \{\eta_k\|U\|_1 + \langle P_k, U\rangle + \langle Q_k, V\rangle + 3\psi_1(U,V) + \|M\|_F\psi_2(U,V)\}\,.
\end{aligned}$$

From the Proposition C.5 in (?), we can directly get the closed form of (U_{k+1}, V_{k+1}). □

More Details in the Numerical Experiments

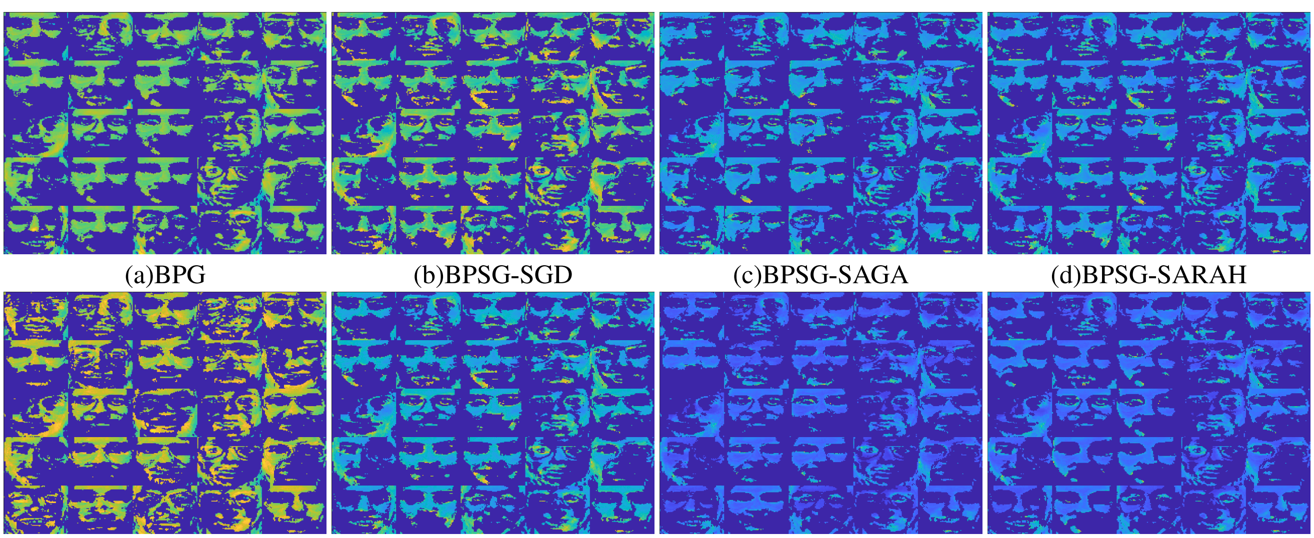
We combine BPSGE with classic stochastic gradient estimator (?) (BPSGE-SGD), SAGA gradient estimator (?) (BPSGE-SAGA), and SARAH gradient estimator (?) (BPSGE-SARAH), and compare BPSGE-SGD/SAGA/SARAH with BPG (?), BPGE (the special case of CoCaIn BPG (?)), and BPSG-SGD/SAGA/SARAH (?). We consider three applications: graph regularized NMF, MF with weakly-convex regularization, and NMF with nonconvex sparsity constraints.

Statistics of the four datasets in graph regularized NMF is listed in the following table.

Dataset	Size	Dimensionality	Number of classes
COIL20	1440	1024	20
PIE	2856	1024	68
COIL100	7200	1024	100
TDT2	9394	36771	30

Table 3: Statistics of the four datasets in graph regularized NMF.

The basis images generated by solving the nonconvex sparsity constrained NMF problem (19) for $s_1 = \frac{m}{3}$ and $s_2 = \frac{d}{2}$ are shown in Figure 3. It validates the BPSGE algorithm outperforms other determinant algorithms and stochastic algorithms without extrapolation.



(a)BPG (b)BPSG-SGD (c)BPSG-SAGA (d)BPSG-SARAH
(e)BPGE (f) BPSGE-SGD (g)BPSGE-SAGA (h)BPSGE-SARAH

Figure 3: The basis images generated by solving the nonconvex sparsity constrained NMF problem (19) with $s_1 = \frac{m}{3}$ and $s_2 = \frac{d}{2}$.