Nonconvex stochastic zeroth-order optimization with decision-dependent distributions: from momentum tracking to coupled sampling

Yuefang Lian* Yongtang Shi† Xiao Wang ‡

September 6, 2026

Abstract

In this paper, we study nonconvex stochastic optimization with decision-dependent distributions, where the decision variable influences the underlying sampling distribution and only stochastic function-value feedback is available. We address two challenges induced by decision-dependent distributions: transport error in momentum-based gradient tracking and variance inflation in zeroth-order estimation. We first develop a Polyak-momentum zeroth-order method that tracks the smoothed gradient under decision-dependent distributions, and establish a sample complexity of $\mathcal{O}(d^2\epsilon^{-6})$ under first-order smoothness. We then introduce implicit gradient transport (IGT) to mitigate the first-order transport error in the tracking recursion, improving the sample complexity to $\mathcal{O}(d^2\epsilon^{-4.5})$ under second-order smoothness. Finally, for a structured class of problems with decision-dependent distributions admitting a common-latent-source representation, we develop a marginal-preserving coupled sampling framework that exploits shared latent randomness between function evaluations to reduce estimator variance. Under this coupled-sampling setting, the proposed methods achieve sample complexities $\mathcal{O}(d^2\epsilon^{-4})$ and $\mathcal{O}(d^2\epsilon^{-3.5})$ under first- and second-order smoothness, respectively. Numerical test results are also reported to showcase the performances of proposed methods.

Keywords Stochastic optimization, decision-dependent distributions, zeroth-order optimization, gradient tracking, variance reduction

1 Introduction

In this paper, we study nonconvex stochastic zeroth-order optimization with decision-dependent distributions:

$$\min_{x\in\mathbb{R}^d} \quad F(x) := \mathbb{E}_{\xi\sim\mathcal{D}(x)}[f(x,\xi)]. \tag{1.1}$$

Here, $f:\mathbb{R}^d\times\Omega\to\mathbb{R}$ is continuously differentiable and $F:\mathbb{R}^d\to\mathbb{R}$ is nonconvex. Unlike classical stochastic optimization with a fixed distribution $\mathcal{D}$, decision-dependent distributions allow the underlying sampling distribution to vary with the decision variable. Specifically, given a query point x, a stochastic oracle returns a stochastic function value $f(x,\xi)$ with $\xi\sim\mathcal{D}(x)$. We consider a sample-only setting in which the density $p_x(\xi)$ of $\mathcal{D}(x)$ is unavailable, and the distribution can only be accessed implicitly through stochastic function evaluations. Such decision-dependent stochastic optimization problems arise in a wide range of applications, including strategic classification [20], reinforcement learning [8], and multiproduct pricing [10].

The study of decision-dependent distributions originates from the performative prediction framework [19, 20], where deployed decisions influence the future data-generating distribution. Early works mainly focused

*School of Mathematical Sciences and LPMC, Nankai University, Tianjin, China. (9820250112@nankai.edu.cn).
†Center for Combinatorics and LPMC, Nankai University, Tianjin, China. (shi@nankai.edu.cn).
‡School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China (wangx936@mail.sysu.edu.cn).

on strongly convex losses and equilibrium-type solution concepts. For example, Perdomo et al. [20] introduced the notion of a performatively stable point, and subsequent studies investigated repeated risk minimization and retraining procedures under strong convexity and distribution regularity assumptions [5, 19]. These ideas have been further extended to various setting with decision-dependent distributions, including online optimization, saddle-point problems, and learning systems [4, 14, 16, 22]. However, the objective $F(x)$ can be nonconvex even when the sample loss $f(x, \xi)$ is convex in x, as the distribution itself changes with the decision. Moreover, exact first-order methods typically require access to the density or score function of $\mathcal{D}(x)$. Under suitable regularity conditions, the gradient of the decision-dependent objective admits the representation

$$\nabla F(x) = \mathbb{E}_{\xi\sim\mathcal{D}(x)}[\nabla_x f(x, \xi) + f(x, \xi)\nabla_x \log p_x(\xi)], \tag{1.2}$$

where $p_x(\xi)$ denotes the density of $\mathcal{D}(x)$. Existing first-order methods for stochastic nonconvex optimization with decision-dependent distributions [8, 10, 15] typically rely on unbiased gradient or score-function information, which is unavailable in the sample-only oracle considered here. This motivates zeroth-order methods that access the distribution only through stochastic function values.

For sample-only decision-dependent problems, existing zeroth-order methods have mainly focused on improving stochastic estimators by exploiting distributional regularity and by developing smoothing or variance-reduction techniques [9, 11, 12, 18, 21]. In particular, Hikima and Takeda [12] provided a unified analysis of multi-direction zeroth-order estimators and showed that two-point Gaussian or spherical estimators achieve sample complexities $\mathcal{O}(d^2\epsilon^{-6})$ and $\mathcal{O}(d^2\epsilon^{-5})$ under first- and second-order smoothness, respectively. They also introduced an adaptive parameter c_k, updated from historical samples, to reduce the variance of one-point estimators without changing their expectation [11]. Recently, Liu et al. [17] developed an online-to-nonconvex (O2NC) reduction framework for sample-only zeroth-order optimization with decision-dependent distributions, achieving sample complexities of $\mathcal{O}(d^2\epsilon^{-6})$ and $\mathcal{O}(d^2\epsilon^{-4.5})$ under first- and second-order smoothness, respectively. Their improved rates are obtained by converting Goldstein-type stationarity guarantees into gradient-stationarity bounds under additional smoothness assumptions. In contrast, we study the interaction between momentum-based zeroth-order methods and decision-dependent distributions, focusing on gradient tracking and variance control under distribution shifts.

The first challenge concerns gradient tracking. Momentum-based methods recursively track gradient information across iterations, and their accuracy depends on how the target gradient varies between consecutive iterates. With decision-dependent distributions, this variation also reflects changes in the induced distribution, introducing an additional first-order transport component that is not handled by standard zeroth-order momentum schemes. To address this issue, we develop a zeroth-order implicit gradient transport (IGT) mechanism that mitigates the the resulting tracking error. The second challenge concerns variance control. In a two-point zeroth-order estimator, the two function evaluations are taken under different distributions induced by the perturbed decisions. Naively sharing the same sample, as in the standard CRN technique, may fail to preserve the correct marginals required by the two perturbed decisions. Independent sampling preserves these marginals, but removes useful correlation and introduces an additional variance component. We therefore propose a common-latent-source coupling framework that preserves the required marginals while exploiting shared latent randomness to reduce the estimator variance.

Table 1 summarizes the comparison with existing zeroth-order methods. Different from the O2NC-based reduction of Liu et al. [17], we directly address momentum tracking under decision-dependent distributions and further exploit common-latent-source coupling to improve the complexity under coupled sampling.

1.1 Contributions

Our main contributions are summarized as follows. We provide a systematic development of zeroth-order methods for stochastic optimization with decision-dependent distributions, moving from momentum-based gradient tracking to implicit transport correction and marginal-preserving coupled sampling.

- *Momentum tracking.* We develop a momentum-based zeroth-order framework for stochastic optimization under a sample-only oracle. Under first-order smoothness, we analyze gradient tracking under decision-dependent sampling distributions and establish a sample complexity of $\mathcal{O}(d^2\epsilon^{-6})$.

Reference	Method	Access model	Smoothness	Gradient Estimator	Distribution	Complexity
Liu et al [18]	ZO-OG	$\mathcal{D}(x)$	first-order	one-point	sensitivity	$\mathcal{O}(d^2G^6\epsilon^{-6})$
Hikima and Takeda [11]	ZO-OGVR	$\mathcal{D}(x)$	first-order	one-point	sensitivity	$\mathcal{O}(d^{4.5}\epsilon^{-6})$
Hikima et al [9]	Guided ZO	$\mathcal{D}(x)$	first-order	two-point	sensitivity	$\mathcal{O}(d^3\epsilon^{-6})$
Hikima and Takeda [12]	ZO-GA	$\mathcal{D}(x)$	first-/second-order	two-point	-	$\mathcal{O}(d^2\epsilon^{-6})/\mathcal{O}(d^2\epsilon^{-5})$
Liu et al. [17]	ZO-O2NC	$\mathcal{D}(x)$	first-/second-order	one-/two-point	-	$\mathcal{O}(d^2\epsilon^{-6})/\mathcal{O}(d^2\epsilon^{-4.5})$
This paper	ZO-PM	$\mathcal{D}(x)$	first-order	two-point	-	$\mathcal{O}(d^2\epsilon^{-6})$
	ZO-PM-IGT	$\mathcal{D}(x)$	second-order	two-point	-	$\mathcal{O}(d^2\epsilon^{-4.5})$
	ZO-PM-CPL	latent simulator $T(x,\zeta)$	first-order	two-point	coupled-sampling assumptions	$\mathcal{O}(d^2\epsilon^{-4})$
	ZO-PM-IGT-CPL	latent simulator $T(x,\zeta)$	second-order	two-point	coupled-sampling assumptions	$\mathcal{O}(d^2\epsilon^{-3.5})$

Table 1: Comparison of existing zeroth-order methods for problem (1.1). Here, $G := \max_{x,\xi} |f(x,\xi)|$ follows [18], and the sensitivity condition refers to the Wasserstein-1 continuity $W_1(\mathcal{D}(x),\mathcal{D}(y)) \le \varepsilon\|x-y\|$. For consistency, we report only Gaussian smoothing-based results; other sampling directions, including coordinate-wise and multiple-unit-sphere directions, are studied in [12].

- *Implicit gradient transport.* We identify a first-order transport component in the momentum tracking recursion that arises from changes in the induced distribution. To address this issue, we develop an implicit gradient transport (IGT) mechanism based on extrapolation. Under second-order smoothness, IGT cancels this first-order transport term and leaves only a higher-order residual, leading to an improved sample complexity of $\mathcal{O}(d^2\epsilon^{-4.5})$.
- *Coupled sampling.* We characterize a structured subclass of problems with decision-dependent distributions that admit a common-latent-source representation. By coupling two-point estimator evaluations through shared latent randomness, the proposed framework preserves the required decision-dependent marginal distributions while reducing the additional variance introduced by independent sampling. Under this additional structure, we establish improved variance bounds and obtain sample complexities of $\mathcal{O}(d^2\epsilon^{-4})$ under first-order smoothness and $\mathcal{O}(d^2\epsilon^{-3.5})$ when combined with IGT under second-order smoothness.

Experiments on two decision-dependent applications and a controlled synthetic benchmark further validate the proposed mechanisms.

2 Preliminaries

In this section, we introduce the basic notation, assumptions, and zeroth-order tools used throughout the paper. Unless otherwise specified, $\|\cdot\|$ denotes the Euclidean norm for vectors and the induced spectral norm for matrices. We denote $F_{\inf} := \inf_{x\in\mathbb{R}^d} F(x) > -\infty$. Throughout the paper, we assume that F is bounded below and G-Lipschitz continuous, A point $\hat{x}$ is called an ϵ-stationary point if $\|\nabla F(\hat{x})\| \le \epsilon$. The following assumptions are used in the convergence analysis.

Assumption 2.1 (Bounded variance)**.** *There exists a constant $\sigma \ge 0$ such that, for all $x \in \mathbb{R}^d$, $\mathbb{E}_{\xi\sim\mathcal{D}(x)}\big[(f(x,\xi)-F(x))^2\big] \le \sigma^2$.*

Assumption 2.2 (First-order smoothness)**.** *Function F is continuously differentiable and has L-Lipschitz continuous gradients on $\mathbb{R}^d$.*

Assumption 2.3 (Second-order smoothness)**.** *Function F is twice continuously differentiable and has H-Lipschitz continuous Hessians on $\mathbb{R}^d$.*

Assumptions 2.1 and 2.2 are the basic conditions required for our zeroth-order analysis, The second-order smoothness assumption is additionally required only for the IGT-based methods to obtain improved complexity bounds.

2.1 Gaussian smoothing and zeroth-order approximation

To exploit zeroth-order information, we introduce a smooth approximation of the objective function F. Let $u \sim \mathcal{N}(0, I_d)$ and $\mu > 0$ be a smoothing parameter. The Gaussian-smoothed objective is defined as

$$F_\mu(x) = \frac{1}{(2\pi)^{d/2}} \int F(x + \mu u) e^{-\frac{1}{2}\|u\|^2} \, du = \mathbb{E}_u[F(x + \mu u)].$$

By construction, $F_\mu(x) \geq F_{\inf}$. Moreover, the Gaussian smoothing preserves the first-order smoothness: if F is L-smooth, then F_μ is L_μ-smooth with $L_\mu = L$. The smoothing bias is controlled as follows.

LEMMA **2.1.** *[2] Assume that F is L-smooth, then $\|\nabla F_\mu(x) - \nabla F(x)\| \leq \sqrt{d}\mu L$, and $\|\nabla F(x)\|^2 \leq 2\|\nabla F_\mu(x)\|^2 + 2d\mu^2 L^2$.*

For a static distribution, the forward and backward evaluations in a two-point estimator may reuse the same sample, e.g. $\frac{f(x+\mu u,\xi)-f(x-\mu u,\xi)}{2\mu} u$. The shared sample induces correlation between the two evaluations, allowing part of the sampling noise to cancel and reducing the estimator variance. In contrast, when the sampling distribution depends on the decision variable, the forward and backward evaluations correspond to different distributions, e.g. $\mathcal{D}(x + \mu u)$ and $\mathcal{D}(x - \mu u)$. To obtain an unbiased estimator of the smoothed gradient, each evaluation must follow its corresponding marginal distribution [18]. Independent sampling preserves these marginals, but removes the correlation needed for noise cancellation and therefore loses the variance-reduction effect available in the static-distribution setting.

2.2 Momentum and implicit gradient transport

Variance-reduced methods such as STORM [3] and SPIDER [6] control stochastic gradient differences at nearby iterates by reusing common randomness. In the decision-dependent setting, however, different iterates induce different sampling distributions, so such estimator differences also contain distribution-shift errors. This makes such recursive constructions difficult to implement and analyze under the sample-only oracle considered here.

Polyak momentum Given a stochastic gradient estimator g_t, Polyak momentum forms the exponential moving average

$$m_t = (1 - \beta_t) m_{t-1} + \beta_t g_t, \qquad \beta_t \in (0, 1].$$

This recursion does not require stochastic gradient differences and can therefore be applied under decision-dependent sampling.

Implicit gradient transport Implicit gradient transport (IGT) [1] evaluates the stochastic estimator at the extrapolated point

$$w_t = x_t + \frac{1 - \alpha_t}{\alpha_t}(x_t - x_{t-1}), \qquad \alpha_t \in (0, 1]. \tag{2.1}$$

While existing analyses of IGT mainly consider stochastic optimization with static distributions, the role of extrapolation in correcting distribution-induced tracking errors remains unexplored.

3 Zeroth-order methods with independent sampling

In this section, we first analyze a Polyak-momentum zeroth-order method under first-order smoothness and then introduce an IGT-based normalized variant under second-order smoothness.

3.1 Two-point zeroth-order estimator with decision-dependent distributions

Motivated by the discussion in Section 2.1, we define a two-point estimator for optimization with decision-dependent distributions. In this setting, the two queried points $x+\mu u$ and $x-\mu u$ correspond to two different distributions, $\mathcal{D}(x+\mu u)$ and $\mathcal{D}(x-\mu u)$. Therefore, to preserve unbiasedness for the smoothed gradient, the forward and backward samples must be drawn from their corresponding distributions. For a queried point x, let $u^i \sim \mathcal{N}(0, I_d)$, $i = 1, \ldots, N$, and randomly generate

$$\xi^{1,i} \sim \mathcal{D}(x + \mu u^i), \qquad \xi^{2,i} \sim \mathcal{D}(x - \mu u^i).$$

Let $U := \{u^i\}_{i=1}^N$, $\Xi^1 := \{\xi^{1,i}\}_{i=1}^N$, and $\Xi^2 := \{\xi^{2,i}\}_{i=1}^N$. The two-point zeroth-order estimator is defined as

$$G_\mu(x, \Xi^1, \Xi^2, U) := \frac{1}{N} \sum_{i=1}^{N} \frac{f(x + \mu u^i, \xi^{1,i}) - f(x - \mu u^i, \xi^{2,i})}{2\mu} u^i. \tag{3.1}$$

This construction avoids the bias issue that may arise if one reuses a single sample drawn from only one perturbed decision-dependent distribution for both function evaluations. Indeed, conditioned on U, the marginal distributions of $\xi^{1,i}$ and $\xi^{2,i}$ are exactly $\mathcal{D}(x + \mu u^i)$ and $\mathcal{D}(x - \mu u^i)$, respectively. Hence, by the tower property,

$$\mathbb{E}_{\Xi^1,\Xi^2,U}\left[G_\mu(x, \Xi^1, \Xi^2, U)\right] = \mathbb{E}_U\left[\frac{F(x + \mu u) - F(x - \mu u)}{2\mu} u\right] = \nabla F_\mu(x),$$

where the last equality follows from the standard Gaussian smoothing identity. Thus, $G_\mu(x, \Xi^1, \Xi^2, U)$ is an unbiased estimator of $\nabla F_\mu(x)$. Since the forward and backward samples are drawn independently from different distributions, the classical noise cancelation effect obtained by reusing the same sample is no longer available. This leads to a sample-noise term amplified by the smoothing parameter μ, producing the unfavorable μ^{-2}-dependence in the variance bound below.

LEMMA **3.1.** *Suppose that Assumptions 2.1 and 2.2 hold. Then*

$$\mathbb{E}_{\Xi^1,\Xi^2,U}\left[\left\|G_\mu(x, \Xi^1, \Xi^2, U) - \nabla F_\mu(x)\right\|^2\right] \le \frac{2\sigma^2 d}{\mu^2 N} + \frac{dL^2\mu^2}{N}(d+2)(d+4) + \frac{12d}{N}\|\nabla F(x)\|^2.$$

Proof. The proof follows from the standard variance decomposition for two-point Gaussian estimators; see, e.g., Lemma 12 of [12]. □

3.2 A Polyak-momentum stochastic zeroth-order algorithm

We first use the estimator in (3.1) with Polyak momentum for Algorithm 3.1, which serves as a baseline for subsequent IGT construction.

To analyze Algorithm 3.1, we define $\varepsilon_t := m_t - \nabla F_\mu(x_t)$, and

$$\bar{\sigma} := \left(\frac{2\sigma^2 d}{\mu^2 N} + \frac{dL^2\mu^2}{N}(d+2)(d+4) + \frac{12dG^2}{N}\right)^{1/2},$$

where G is the Lipschitz constant of F. It is obvious that

$$\mathbb{E}_{\Xi^1,\Xi^2,U}\left[\left\|G_\mu(x, \Xi^1, \Xi^2, U) - \nabla F_\mu(x)\right\|^2\right] \le \bar{\sigma}^2 \tag{3.2}$$

due to $\|\nabla F(x)\| \le G$. The following lemmas give the corresponding momentum-tracking recursion and the descent estimate of the smoothed objective.

Algorithm 3.1 Zeroth-Order SGD with Polyak Momentum (ZO-PM)

Input: Maximum iteration number T, stepsizes $\{\eta_t\}_{t=1}^T$, momentum parameters $\{\beta_t\}_{t=0}^{T-1}$ with $\beta_0 = 1$, smoothing radius μ, batch size N, and initial iterate x_1.

1: Set $m_0 = 0$
2: **for** $t = 1, \ldots, T$ **do**
3: Generate i.i.d. samples $u_t^i \sim \mathcal{N}(0, I_d), i = 1, \ldots, N$
4: Sample i.i.d. samples $\xi_t^{1,i} \sim \mathcal{D}(x_t + \mu u_t^i)$ and $\xi_t^{2,i} \sim \mathcal{D}(x_t - \mu u_t^i), i = 1, \ldots, N$
5: Compute the zeroth-order estimator $G_\mu(x_t, \Xi_t^1, \Xi_t^2, U_t)$ as (3.1)
6: Update $m_t = (1 - \beta_{t-1})m_{t-1} + \beta_{t-1} G_\mu(x_t, \Xi_t^1, \Xi_t^2, U_t)$.
7: Update $x_{t+1} = x_t - \eta_t m_t$
8: **end for**

Output: x_r, chosen uniformly at random from $\{x_t\}_{t=1}^T$.

LEMMA **3.2** (Momentum-tracking recursion)**.** *Suppose that Assumptions 2.1 and 2.2 hold. Then for any $t \geq 1$, we have*

$$\mathbb{E}[\|\varepsilon_{t+1}\|^2] \leq (1 - \beta_t)\mathbb{E}[\|\varepsilon_t\|^2] + \frac{L_\mu^2}{\beta_t}\mathbb{E}\|x_{t+1} - x_t\|^2 + \beta_t^2 \bar{\sigma}^2, \tag{3.3}$$

where $\bar{\sigma} = (\frac{2\sigma^2 d}{\mu^2 N} + \frac{dL^2\mu^2}{N}(d+2)(d+4) + \frac{12dG^2}{N})^{1/2}$.

Proof. Let $\mathcal{F}_t$ denote the filtration generated by all randomness up to iteration t. Since x_{t+1} and m_t are $\mathcal{F}_t$-measurable and the new oracle samples are generated independently at iteration $t+1$, the estimator satisfies

$$\mathbb{E}[G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1})|\mathcal{F}_t] = \nabla F_\mu(x_{t+1}).$$

According to $m_{t+1} = (1 - \beta_t)m_t + \beta_t G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1})$, we obtain

$$\begin{aligned}
&\mathbb{E}[\|\varepsilon_{t+1}\|^2] \\
=&\mathbb{E}[\|m_{t+1} - \nabla F_\mu(x_{t+1})\|^2] \\
=&\mathbb{E}[\big\|(1 - \beta_t)m_t + \beta_t G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1}) - \nabla F_\mu(x_{t+1}))\big\|^2] \\
=&(1 - \beta_t)^2\mathbb{E}[\|m_t - \nabla F_\mu(x_{t+1})\|^2] + \beta_t^2\mathbb{E}[\big\|G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1}) - \nabla F_\mu(x_{t+1}))\big\|^2] \\
&+ (1 - \beta_t)\beta_t\mathbb{E}[\langle m_t - \nabla F_\mu(x_{t+1}), G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1}) - \nabla F_\mu(x_{t+1}))] \\
=&(1 - \beta_t)^2\mathbb{E}[\|m_t - \nabla F_\mu(x_{t+1})\|^2] + \beta_t^2\mathbb{E}[\big\|G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1}) - \nabla F_\mu(x_{t+1}))\big\|^2] \\
=&(1 - \beta_t)^2\mathbb{E}[\|m_t - \nabla F_\mu(x_t) + \nabla F_\mu(x_t) - \nabla F_\mu(x_{t+1})\|^2] \\
&+ \beta_t^2\mathbb{E}[\big\|G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1}) - \nabla F_\mu(x_{t+1}))\big\|^2] \\
\leq&(1 - \beta_t)^2(1 + \alpha)\mathbb{E}[\|m_t - \nabla F_\mu(x_t)\|^2] + (1 - \beta_t)^2(1 + \frac{1}{\alpha})\mathbb{E}[\|\nabla F_\mu(x_t) - \nabla F_\mu(x_{t+1})\|^2] \\
&+ \beta_t^2\mathbb{E}[\big\|G_\mu(x_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1}) - \nabla F_\mu(x_{t+1}))\big\|^2],
\end{aligned}$$

where the cross term vanishes due to the conditional unbiasedness of the zeroth-order estimator, and the last inequality follows from Young's inequality

$$\|a + b\|^2 \leq (1 + \alpha)\|a\|^2 + (1 + \frac{1}{\alpha})\|b\|^2.$$

where $\alpha > 0$. Choosing $\alpha = \frac{\beta_t}{1 - \beta_t}$ and using the L_μ-smoothness of F_μ together with the variance bound in (3.2), we obtain (3.3). The variance bound follows from the G-Lipschitz continuity of F. □

LEMMA **3.3** (Descent estimate)**.** *Suppose that Assumption 2.2 holds, then*

$$F_\mu(x_{t+1}) \le F_\mu(x_t) - \frac{1-L_\mu\eta_t}{4\eta_t}\|x_{t+1}-x_t\|^2 + \frac{\eta_t}{1-L_\mu\eta_t}\|\varepsilon_t\|^2 \quad \forall t \ge 1.$$

Proof. According to the smoothness of the function F_μ, we have

$$\begin{aligned}
F_\mu(x_{t+1}) &\le F_\mu(x_t) + \langle \nabla F_\mu(x_t), x_{t+1}-x_t\rangle + \frac{L_\mu}{2}\|x_{t+1}-x_t\|^2 \\
&= F_\mu(x_t) + \langle \nabla F_\mu(x_t) - m_t, x_{t+1}-x_t\rangle + \langle m_t, x_{t+1}-x_t\rangle + \frac{L_\mu}{2}\|x_{t+1}-x_t\|^2 \\
&= F_\mu(x_t) + \langle \nabla F_\mu(x_t) - m_t, x_{t+1}-x_t\rangle - \langle \frac{1}{\eta_t}(x_{t+1}-x_t), x_{t+1}-x_t\rangle + \frac{L_\mu}{2}\|x_{t+1}-x_t\|^2 \\
&= F_\mu(x_t) + \langle \nabla F_\mu(x_t) - m_t, x_{t+1}-x_t\rangle - (\frac{1}{\eta_t} - \frac{L_\mu}{2})\|x_{t+1}-x_t\|^2 \\
&\le F_\mu(x_t) + \frac{1}{2\alpha}\|\nabla F_\mu(x_t) - m_t\|^2 + \frac{\alpha}{2}\|x_{t+1}-x_t\|^2 - \frac{1-L_\mu\eta_t}{2\eta_t}\|x_{t+1}-x_t\|^2 \\
&= F_\mu(x_t) + \frac{\eta_t}{1-L_\mu\eta_t}\|\nabla F_\mu(x_t) - m_t\|^2 - \frac{1-L_\mu\eta_t}{4\eta_t}\|x_{t+1}-x_t\|^2,
\end{aligned}$$

where the last inequality is from the Young inequality, and we choose the parameter $\alpha = \frac{1-L_\mu\eta_t}{2\eta_t}$ for the last equality. □

The previous two lemmas show that Polyak momentum can be incorporated into the decision-dependent zeroth-order framework without changing the basic structure of the analysis. The following result shows that Algorithm 3.1 matches the best known sample complexity under first-order smoothness.

THEOREM **3.1.** *Suppose that Assumptions 2.1 and 2.2 hold. Let*

$$\eta_t = \eta = \frac{1}{4L_\mu + \sqrt{\frac{TL_\mu}{\Phi(x_1)}\bar{\sigma}^2}}, \qquad \beta_t = \frac{3\eta L_\mu}{1-L_\mu\eta}, \quad t = 1,\dots,T,$$

where $\bar{\sigma}$ is defined in Lemma 3.2. Then the following inequality holds:

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[\|\nabla F(x_t)\|^2\right] \le \frac{C_1 L_\mu \Phi(x_1)}{T} + \frac{C_2\sqrt{L_\mu\Phi(x_1)}}{\sqrt{T}}\bar{\sigma} + 2\mu^2L^2d, \tag{3.4}$$

where $\Phi(x_1) = F_\mu(x_1) - F_{\inf} + \frac{3}{8L_\mu}\mathbb{E}\left[\|m_1 - \nabla F_\mu(x_1)\|^2\right]$ and C_1 and C_2 are positive absolute constants.

Proof. We first establish a relation between the step size and the stationarity measure. Firstly, we have

$$\nabla F_\mu(x_t) = m_t - \varepsilon_t = \frac{x_t - x_{t+1}}{\eta_t} - \varepsilon_t,$$

where $\varepsilon_t := m_t - \nabla F_\mu(x_t)$ and the update rule, i.e., $x_{t+1} = x_t - \eta_t m_t$. Using the inequality $\|a+b\|^2 \le 2\|a\|^2 + 2\|b\|^2$, we obtain

$$\|\nabla F_\mu(x_t)\|^2 \le 2\left\|\frac{x_t - x_{t+1}}{\eta_t}\right\|^2 + 2\|\varepsilon_t\|^2 = \frac{2}{\eta_t^2}\|x_{t+1}-x_t\|^2 + 2\|\varepsilon_t\|^2. \tag{3.5}$$

Taking expectations and defining $\Delta_t := \mathbb{E}[\|\varepsilon_t\|^2]$, we obtain

$$\mathbb{E}[\|x_{t+1}-x_t\|^2] \ge \frac{\eta_t^2}{2}\left(\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] - 2\Delta_t\right). \tag{3.6}$$

Next, we consider the following Lyapunov function

$$\Phi(x_t) := \mathbb{E}[F_\mu(x_t) - F_{\inf}] + c\Delta_t,$$

where $c = \frac{\beta_t(1-L_\mu\eta_t)}{8L_\mu^2\eta_t}$. Combining Lemmas 3.2 and 3.3, we derive

$$\Phi(x_{t+1}) \le \mathbb{E}[F_\mu(x_t) - F_{\inf}] - \left(\frac{1-L_\mu\eta_t}{4\eta_t} - \frac{cL_\mu^2}{\beta_t}\right)\mathbb{E}[\|x_{t+1}-x_t\|^2] + \left(\frac{\eta_t}{1-L_\mu\eta_t} + c(1-\beta_t)\right)\Delta_t + c\beta_t^2\bar{\sigma}^2.$$

Define $H_t := \frac{1-L_\mu\eta_t}{4\eta_t} - \frac{cL_\mu^2}{\beta_t}$. By the choice of c, we have $H_t = \frac{1-L_\mu\eta_t}{8\eta_t}$. Substituting the lower bound (3.6) into above inequality gives

$$\Phi(x_{t+1}) \le \mathbb{E}[F_\mu(x_t) - F_{\inf}] - \frac{H_t\eta_t^2}{2}\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] + \left(\frac{\eta_t}{1-L_\mu\eta_t} + c(1-\beta_t) + H_t\eta_t^2\right)\Delta_t + c\beta_t^2\bar{\sigma}^2. \quad (3.7)$$

We next choose the momentum parameter as $\beta_t = \frac{3\eta_t L_\mu}{1-L_\mu\eta_t}$. Then $c = \frac{3}{8L_\mu}$ and the coefficient of Δ_t in (3.7) satisfies $\frac{\eta_t}{1-L_\mu\eta_t} + c(1-\beta_t) + H_t\eta_t^2 \le 0$, provided that $\eta_t < \frac{1}{4L_\mu}$. Therefore, dropping the non-positive error term and using $H_t = \frac{1-L_\mu\eta_t}{8\eta_t}$ yields

$$\begin{aligned}
\Phi(x_{t+1}) &\le \Phi(x_t) - \frac{H_t\eta_t^2}{2}\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] + c\beta_t^2\bar{\sigma}^2 \\
&\le \Phi(x_t) - \frac{\eta_t(1-L_\mu\eta_t)}{16}\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] + c\beta_t^2\bar{\sigma}^2 \\
&\le \Phi(x_t) - \frac{3\eta_t}{64}\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] + \frac{3}{2}\eta_t^2 L_\mu\bar{\sigma}^2, \qquad (3.8)
\end{aligned}$$

where the last inequality uses $1 - L_\mu\eta_t \ge \frac{3}{4}$ and $c\beta_t^2 = \frac{9\eta_t^2 L_\mu}{8(1-L_\mu\eta_t)} \le \frac{3}{2}\eta_t^2 L_\mu$ by the settings of η_k and β_k. Then summing (3.8) from $t = 1$ to T with $\eta_t = \eta$, we obtain

$$\frac{3\eta}{64}\sum_{t=1}^{T}\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] \le \Phi(x_1) - \Phi(x_{T+1}) + \frac{3}{2}T\eta^2 L_\mu\bar{\sigma}^2.$$

Since $\Phi(x_{T+1}) \ge 0$, we have $\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] \le \frac{64\Phi(x_1)}{3T\eta} + 32\eta L_\mu\bar{\sigma}^2$. From the stepsize choice $\eta = \frac{1}{4L_\mu + \sqrt{\frac{TL_\mu}{\Phi(x_1)}}\bar{\sigma}^2}$, there exist $C_1, C_2 > 0$ such that

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\|\nabla F_\mu(x_t)\|^2] \le C_1\frac{L_\mu\Phi(x_1)}{T} + C_2\frac{\sqrt{L_\mu\Phi(x_1)}}{\sqrt{T}}\bar{\sigma},$$

It then follows from $\|\nabla F(x_t)\|^2 \le 2\|\nabla F_\mu(x_t)\|^2 + 2d\mu^2L^2$ that (3.4) holds. □

Corollary 3.2. *Suppose that the assumptions of Theorem 3.1 hold. If $\mu = \epsilon d^{-1/2}$ and $N = d^2\epsilon^{-4}$, then Algorithm 3.1 outputs an ϵ-stationary point with iteration complexity in order of $\mathcal{O}(\epsilon^{-2})$ and total sample complexity in order of $\mathcal{O}(d^2\epsilon^{-6})$.*

Proof. Note that by $\mu = \epsilon d^{-1/2}$ and $N = d^2\epsilon^{-4}$, the smoothing error satisfies $d\mu^2L^2 = \epsilon^2$, and the variance term satisfies $\bar{\sigma} = \mathcal{O}(\frac{d}{\epsilon\sqrt{N}})$. Substituting these choices into the convergence bound yields $\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\|\nabla F(x_t)\|^2] = \mathcal{O}(\epsilon^2)$ after $T = \mathcal{O}(\varepsilon^{-2})$. Thus, we derive

$$\mathbb{E}[\|\nabla F(x_r)\|] \le \sqrt{\mathbb{E}\|\nabla F(x_r)\|^2} = \mathcal{O}(\epsilon)$$

with the total sample complexity $\mathcal{O}(TN) = \mathcal{O}(d^2\epsilon^{-6})$. □

3.3 A Polyak-momentum stochastic zeroth-order algorithm with IGT

The Polyak-momentum method matches the best known sample complexity $\mathcal{O}(d^2\epsilon^{-6})$, but its tracking recursion still contains a first-order tracking error between consecutive gradients. Under second-order smoothness, we introduce implicit gradient transport to replace this term with a second-order residual and thereby improve the complexity. Under Assumption 2.3, there is

$$\|\nabla F(x) - \nabla F(y) - \nabla^2 F(y)(x-y)\| \leq H\|x-y\|^2, \tag{3.9}$$

for any $x, y \in \mathbb{R}^d$. Furthermore, the Gaussian smoothing bias admits the sharper bound [2], $\|\nabla F_\mu(x) - \nabla F(x)\| \leq d\mu^2 H$, and hence

$$\|\nabla F(x)\| \leq \|\nabla F_\mu(x)\| + d\mu^2 H. \tag{3.10}$$

Define the second-order remainder

$$Z_\mu(x,y) := \nabla F_\mu(x) - \nabla F_\mu(y) - \nabla^2 F_\mu(y)(x-y). \tag{3.11}$$

Since Gaussian smoothing preserves the second-order smoothness constant, the same bound holds for F_μ, so that $\|Z_\mu(x,y)\| \leq H\|x-y\|^2$. The following lemma gives the corresponding variance bound for the two-point zeroth-order estimator.

LEMMA **3.4.** *Suppose that Assumptions 2.1 and 2.3 hold. Then, we have*

$$\mathbb{E}_{U,\Xi^1,\Xi^2}\left[\left\|G_\mu(x,\Xi^1,\Xi^2,U) - \nabla F_\mu(x)\right\|^2\right] \leq \tilde{\sigma}^2(x),$$

where $\tilde{\sigma}(x) = (\frac{2\sigma^2 d}{\mu^2 N} + \frac{H^2\mu^4}{N} d(d+2)(d+4)(d+6) + \frac{6d}{N}\|\nabla F(x)\|^2)^{1/2}$.

Proof. The proof follows from the standard variance decomposition for two-point Gaussian estimators; see, e.g., Lemma 13 of [12]. □

Although Lemma 3.4 provides a sharper μ-dependence in the variance bound than Lemma 3.1, this improvement alone does not enhance the final sample complexity of Algorithm 3.1. The remaining bottleneck arises from the momentum-tracking recursion. Let $\varepsilon_t = m_t - \nabla F_\mu(x_t)$ denote the tracking error. Its recursion is given by

$$\varepsilon_{t+1} = (1-\beta)\varepsilon_t + \beta(G_t - \nabla F_\mu(x_t)) + (1-\beta)(\nabla F_\mu(x_t) - \nabla F_\mu(x_{t+1})).$$

The last term is the transport error caused by the changing evaluation point. With decision-dependent distributions, the gradient variation reflects both the movement of the decision variable and the induced distribution shift, resulting in an additional first-order transport component that is not explicitly handled by standard zeroth-order momentum schemes. The IGT mechanism evaluates the momentum correction at an extrapolated point to cancel this first-order component. Under second-order smoothness, the remaining transport error is reduced to a higher-order residual, leading to an improved complexity bound.

The analysis of Algorithm 3.2 has two ingredients, including momentum tracking recursion and the descent estimate of smoothed objective function. Denote

$$\tilde{\varepsilon}_t := G_\mu(w_t, \Xi_t^1, \Xi_t^2, U_t) - \nabla F_\mu(w_t),$$

which satisfies $\mathbb{E}[\|\tilde{\varepsilon}_t\|^2] \leq \tilde{\sigma}^2(w_t)$ by Lemma 3.4, and

$$\tilde{\sigma}_{\max} := (\frac{2\sigma^2 d}{\mu^2 N} + \frac{H^2\mu^4}{N} d(d+2)(d+4)(d+6) + \frac{6dG^2}{N})^{1/2}, \tag{3.12}$$

which satisfies that $\mathbb{E}[\|\tilde{\varepsilon}_t\|^2] \leq \tilde{\sigma}_{\max}^2$ due to $\|\nabla F(w_t)\| \leq G$. With the choice $\alpha_t = \beta_t$ in (2.1), the first-order Hessian transport terms in the momentum-tracking recursion cancel, leaving only a second-order residual, as formalized in Lemma 3.5.

Algorithm 3.2 Zeroth-Order SGD with Polyak Momentum and Implicit Gradient Transport (ZO-PM-IGT)

Input: Total iterations T, stepsizes $\{\eta_t\}_{t=1}^T$, momentum parameters $\{\beta_t\}_{t=0}^{T-1}$ with $\beta_0 = 1$, extrapolation parameter α, smoothing radius μ, batch size N, and initial iterates $x_0 = x_1$.

1: Initialize $m_0 = 0$
2: **for** $t = 1, \dots, T$ **do**
3: Set $w_t = x_t + \frac{1-\alpha_t}{\alpha_t}(x_t - x_{t-1})$.
4: Generate i.i.d. samples $u_t^i \sim \mathcal{N}(0, I_d), i = 1, \dots, N$
5: Sample i.i.d. $\xi_t^{1,i} \sim \mathcal{D}(w_t + \mu u_t^i)$ and $\xi_t^{2,i} \sim \mathcal{D}(w_t - \mu u_t^i), i = 1, \dots, N$
6: Compute the zeroth-order estimator $G_\mu(w_t, \Xi_t^1, \Xi_t^2, U_t)$ as (3.1)
7: Update $m_t = (1 - \beta_{t-1})m_{t-1} + \beta_{t-1} G_\mu(w_t, \Xi_t^1, \Xi_t^2, U_t)$.
8: Update $x_{t+1} = x_t - \eta_t \frac{m_t}{\|m_t\|}$
9: **end for**

Output: x_r chosen uniformly at random from $\{x_t\}_{t=1}^T$.

LEMMA **3.5** (Momentum tracking recursion)**.** *Suppose that Assumptions 2.1 and 2.3 hold, we set* $\alpha_t = \beta_t = \beta \in (0, 1)$ *and* $\eta_t = \eta$ *for all* $t \geq 1$*. Then it holds that*

$$\mathbb{E}[\|\varepsilon_{t+1}\|] \leq (1-\beta)^t \tilde{\sigma}_{\max} + \sqrt{\beta}\, \tilde{\sigma}_{\max} + \frac{2H\eta^2(1-\beta)}{\beta^2}, \quad t \geq 0. \tag{3.13}$$

Proof. Let $\tilde{\varepsilon}_{t+1} := G_\mu(w_{t+1}, \Xi_{t+1}^1, \Xi_{t+1}^2, U_{t+1}) - \nabla F_\mu(w_{t+1})$. Using the momentum update and $\varepsilon_t = m_t - \nabla F_\mu(x_t)$, we have

$$\begin{aligned} \varepsilon_{t+1} = (1-\beta_t)\varepsilon_t + \beta_t \tilde{\varepsilon}_{t+1} + (1-\beta_t)\big(\nabla F_\mu(x_t) - \nabla F_\mu(x_{t+1})\big) \\ + \beta_t\big(\nabla F_\mu(w_{t+1}) - \nabla F_\mu(x_{t+1})\big). \end{aligned} \tag{3.14}$$

By the definition of Z_μ, obtaining

$$\begin{aligned} \nabla F_\mu(x_t) - \nabla F_\mu(x_{t+1}) &= \nabla^2 F_\mu(x_{t+1})(x_t - x_{t+1}) + Z_\mu(x_t, x_{t+1}), \\ \nabla F_\mu(w_{t+1}) - \nabla F_\mu(x_{t+1}) &= \nabla^2 F_\mu(x_{t+1})(w_{t+1} - x_{t+1}) + Z_\mu(w_{t+1}, x_{t+1}). \end{aligned}$$

Since $w_{t+1} - x_{t+1} = \frac{1-\alpha_t}{\alpha_t}(x_{t+1} - x_t)$, setting $\alpha_t = \beta_t = \beta$ gives

$$(1-\beta)(x_t - x_{t+1}) + \beta(w_{t+1} - x_{t+1}) = 0.$$

Hence, the first-order Hessian terms in (3.14) cancel, and we derive

$$\varepsilon_{t+1} = (1-\beta)\varepsilon_t + \beta\tilde{\varepsilon}_{t+1} + r_{t+1}, \tag{3.15}$$

where $r_{t+1} := (1-\beta)Z_\mu(x_t, x_{t+1}) + \beta Z_\mu(w_{t+1}, x_{t+1})$. Because F_μ has an H-Lipschitz Hessian, $\|Z_\mu(x, y)\| \leq H\|x - y\|^2$. The normalized update gives $\|x_{t+1} - x_t\| \leq \eta$, and therefore $\|w_{t+1} - x_{t+1}\| \leq \frac{1-\beta}{\beta}\eta$. Consequently, we bound the norm of r_{t+1} as follows

$$\|r_{t+1}\| \leq (1-\beta)H\eta^2 + \beta H\left(\frac{1-\beta}{\beta}\right)^2 \eta^2 = \frac{1-\beta}{\beta} H\eta^2 \leq \frac{2(1-\beta)}{\beta} H\eta^2. \tag{3.16}$$

Unrolling (3.15), we yield

$$\varepsilon_{t+1} = (1-\beta)^t \varepsilon_1 + \beta \sum_{s=2}^{t+1} (1-\beta)^{t+1-s} \tilde{\varepsilon}_s + \sum_{s=2}^{t+1} (1-\beta)^{t+1-s} r_s.$$

Let $\mathcal{F}_s$ denote the sigma-field generated by all randomness up to iteration s. Since the query point w_s is $\mathcal{F}_{s-1}$-measurable and the zeroth-order estimator is conditionally unbiased, there is $\mathbb{E}[\tilde{\varepsilon}_s|\mathcal{F}_{s-1}] = 0$. Therefore, $\mathbb{E}\langle\tilde{\varepsilon}_s, \tilde{\varepsilon}_j\rangle = 0, \forall j < s$, as $\tilde{\varepsilon}_j$ is $\mathcal{F}_{s-1}$-measurable. Therefore, we obtain

$$\mathbb{E}[\|\sum_{s=2}^{t+1}(1-\beta)^{t+1-s}\tilde{\varepsilon}_s\|]^2 = \sum_{s=2}^{t+1}(1-\beta)^{2(t+1-s)}\mathbb{E}[\|\tilde{\varepsilon}_s\|^2].$$

Applying Jensen's inequality, we further have

$$\mathbb{E}[\|\sum_{s=2}^{t+1}(1-\beta)^{t+1-s}\tilde{\varepsilon}_s\|] \leq \left(\sum_{s=2}^{t+1}(1-\beta)^{2(t+1-s)}\mathbb{E}[\|\tilde{\varepsilon}_s\|^2]\right)^{1/2} \leq \frac{\tilde{\sigma}_{\max}}{\sqrt{1-(1-\beta)^2}},$$

where the last inequality is from $\max_t \mathbb{E}[\|\tilde{\varepsilon}_t\|^2] \leq \tilde{\sigma}_{\max}^2$. Moreover, it follows from $m_1 = G_\mu(w_1, \Xi_1^1, \Xi_1^2, U_1)$ that $\varepsilon_1 = \tilde{\varepsilon}_1$, thus $\mathbb{E}[\|\varepsilon_1\|] \leq \sqrt{\mathbb{E}[\|\varepsilon_1\|]^2} \leq \tilde{\sigma}_{\max}$. And by (3.16), it holds that $\sum_{s=2}^{t+1}(1-\beta)^{t+1-s}\|r_s\| \leq \frac{2H\eta^2(1-\beta)}{\beta^2}$. Then combining these bounds gives

$$\mathbb{E}[\|\varepsilon_{t+1}\|] \leq (1-\beta)^t\tilde{\sigma}_{\max} + \frac{\beta}{\sqrt{1-(1-\beta)^2}}\tilde{\sigma}_{\max} + \frac{2H\eta^2(1-\beta)}{\beta^2}.$$

Finally, from $\frac{\beta}{\sqrt{1-(1-\beta)^2}} = \sqrt{\frac{\beta}{2-\beta}} \leq \sqrt{\beta}$, we derive the conclusion. □

We focus on the descent estimate of $F_\mu(x)$ under the second-order smoothness.

LEMMA **3.6** (Descent estimate)**.** *Suppose that Assumption 2.2 holds, we have*

$$F_\mu(x_{t+1}) - F_\mu(x_t) \leq -\eta_t\|\nabla F_\mu(x_t)\| + 2\eta_t\|\varepsilon_t\| + \frac{L_\mu\eta_t^2}{2}. \tag{3.17}$$

If $\eta_t = \eta$, then $\sum_{t=1}^T \|\nabla F_\mu(x_t)\| \leq \frac{F_\mu(x_1)-F_{\inf}}{\eta} + \frac{L_\mu T\eta}{2} + 2\sum_{t=1}^T \|\varepsilon_t\|$.

Proof. Let $g_t := \nabla F_\mu(x_t)$, and define $d_t = \begin{cases} m_t/\|m_t\|, & m_t \neq 0, \\ 0, & m_t = 0. \end{cases}$. By the L_μ-smoothness of F_μ, we have

$$F_\mu(x_{t+1}) \leq F_\mu(x_t) - \eta_t\langle g_t, d_t\rangle + \frac{L_\mu\eta_t^2}{2}. \tag{3.18}$$

If $m_t \neq 0$, then, it holds from $g_t = m_t - \varepsilon_t$ that

$$\langle g_t, d_t\rangle = \|m_t\| - \frac{\langle\varepsilon_t, m_t\rangle}{\|m_t\|} \geq \|m_t\| - \|\varepsilon_t\| \geq \|g_t\| - 2\|\varepsilon_t\|.$$

If $m_t = 0$, then $d_t = 0$ and $\|g_t\| = \|\varepsilon_t\|$, so we have $\langle g_t, d_t\rangle = 0 \geq \|g_t\| - 2\|\varepsilon_t\|$. Thus, in both cases, $\langle g_t, d_t\rangle \geq \|g_t\| - 2\|\varepsilon_t\|$. We then derive (3.17) from (3.18). Summing over t and using $F_\mu(x_{T+1}) \geq F_{\inf}$ further proves the second claim. □

Combining Lemmas 3.5 and 3.6, we obtain the following convergence result.

THEOREM **3.3.** *Suppose that Assumptions 2.1, 2.2 and 2.3 hold, and*

$$\eta_t \equiv \eta = \min\Big\{\frac{R^{5/7}}{T^{5/7}H^{1/7}\tilde{\sigma}_{\max}^{4/7}}, \sqrt{\frac{R}{TL_\mu}}\Big\}, \quad \beta_t = \alpha_t = \beta = \min\Big\{\frac{R^{4/7}H^{2/7}}{T^{4/7}\tilde{\sigma}_{\max}^{6/7}}, 1\Big\},$$

where $R := F_\mu(x_1) - F_{\inf}$. Then the iterates generated by Algorithm 3.2 satisfy

$$\frac{1}{T}\sum_{t=1}^T \mathbb{E}[\|\nabla F_\mu(x_t)\|] \leq C_1\frac{\sqrt{RL_\mu}}{\sqrt{T}} + C_2\frac{R^{2/7}H^{1/7}\tilde{\sigma}_{\max}^{4/7}}{T^{2/7}} + C_3\frac{\tilde{\sigma}_{\max}^{13/7}}{R^{4/7}H^{2/7}T^{3/7}}, \tag{3.19}$$

where $C_1, C_2, C_3 > 0$ are numerical constants independent of T, d, μ, and N, and $\tilde{\sigma}_{\max}$ is defined in (3.12).

Proof. By (3.13), summing the inequality over $t = 1, \ldots, T$ and using $\sum_{t=1}^{T}(1-\beta)^t \leq \frac{1}{\beta}$ for $1-\beta < 1$, we obtain

$$\sum_{t=1}^{T} \mathbb{E}[\|\varepsilon_t\|] \leq \frac{\tilde{\sigma}_{\max}}{\beta} + T\sqrt{\beta}\tilde{\sigma}_{\max} + \frac{2TH\eta^2(1-\beta)}{\beta^2}.$$

Combining this bound with Lemma 3.6 gives

$$\sum_{t=1}^{T} \mathbb{E}[\|\nabla F_\mu(x_t)\|] \leq \frac{R}{\eta} + \frac{L_\mu T\eta}{2} + 2\left(\frac{\tilde{\sigma}_{\max}}{\beta} + T\sqrt{\beta}\tilde{\sigma}_{\max} + \frac{2TH\eta^2}{\beta^2}\right). \tag{3.20}$$

By the definition of η, we consider two cases. If $\eta = \sqrt{\frac{R}{TL_\mu}}$, then $\frac{R}{\eta} + \frac{L_\mu T\eta}{2} = \frac{3}{2}\sqrt{RL_\mu T}$. Otherwise, $\eta = \frac{R^{5/7}}{T^{5/7}H^{1/7}\tilde{\sigma}_{\max}^{4/7}}$, and since $\eta \leq \sqrt{R/(TL_\mu)}$, we have $\frac{R}{\eta} + \frac{L_\mu T\eta}{2} \leq \frac{3}{2}R^{2/7}H^{1/7}\tilde{\sigma}_{\max}^{4/7}T^{5/7}$. Thus the following inequality holds:

$$\frac{R}{\eta} + \frac{L_\mu T\eta}{2} \leq \frac{3}{2}\left(\sqrt{RL_\mu T} + R^{2/7}H^{1/7}\tilde{\sigma}_{\max}^{4/7}T^{5/7}\right).$$

By the choice of β, we bound the other terms of (3.20) as follows

$$T\sqrt{\beta}\,\tilde{\sigma}_{\max} \leq R^{2/7}H^{1/7}\tilde{\sigma}_{\max}^{4/7}T^{5/7}, \qquad \frac{TH\eta^2}{\beta^2} \leq R^{2/7}H^{1/7}\tilde{\sigma}_{\max}^{4/7}T^{5/7},$$

and $\frac{\tilde{\sigma}_{\max}}{\beta} \leq \frac{\tilde{\sigma}_{\max}^{13/7}T^{4/7}}{R^{4/7}H^{2/7}}$. Therefore, we arrive at the conclusion as shown in (3.19). □

Corollary 3.4. *Suppose that the assumptions of Theorem 3.3 hold. If $\mu^2 d = \epsilon$ and $N = d^2\epsilon^{-5/2}$, Algorithm 3.2 returns an ϵ-stationary point, i.e., $\mathbb{E}[\|\nabla F(x_r)\|] \leq \epsilon$ with iteration complexity and total sample complexity of order $\mathcal{O}(\epsilon^{-2})$ and $\mathcal{O}(d^2\epsilon^{-4.5})$, respectively.*

Proof. With $\mu^2 d = \epsilon$ and $N = d^2\epsilon^{-5/2}$, the variance bound in (3.12) gives $\tilde{\sigma}_{\max} = \mathcal{O}(\epsilon^{3/4})$. Taking $T = \mathcal{O}(\epsilon^{-2})$, the above inequality yields $\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[\|\nabla F_\mu(x_t)\|] = \mathcal{O}(\epsilon)$. Moreover, using the smoothing error bound $\|\nabla F(x)\| \leq \|\nabla F_\mu(x)\| + d\mu^2 H$ shown in (2.3) and $d\mu^2 = \epsilon$, we obtain $\mathbb{E}[\|\nabla F(x_r)\|] = \mathcal{O}(\epsilon)$. Finally, the total sample complexity is $TN = \mathcal{O}(\epsilon^{-2}) \cdot \mathcal{O}(d^2\epsilon^{-5/2})$ which is in order $\mathcal{O}(d^2\epsilon^{-4.5})$. The proof is completed. □

4 Zeroth-order methods with coupled sampling

Although IGT reduces the momentum-tracking error, the two-point estimator still suffers from an additional variance term caused by independently sampling from different decision-dependent distributions. In particular, independent sampling prevents the two finite-difference evaluations from exploiting shared randomness for variance reduction, as reflected in the variance bound in Lemma 3.4. To address this limitation, we consider a structured subclass of problems with decision-dependent distributions admitting a common-latent-source representation and develop coupled sampling estimators. Such structures naturally arise in simulation-based models with decision-dependent distributions, where decisions affect observations through transformations of shared underlying randomness. Unlike the general sample-only setting considered above, this additional structure enables shared latent randomness across different decisions while preserving their corresponding marginal distributions.

4.1 Coupled sampling with a common latent source

We assume that there exist a probability space $(\mathcal{Z}, \mathcal{F}, P_0)$ and a measurable mapping $T : \mathbb{R}^d \times \mathcal{Z} \to \Xi$ such that, for any decision x, the distribution $\mathcal{D}(x)$ is the law of

$$\xi = T(x, \zeta), \qquad \zeta \sim P_0.$$

This representation allows function evaluations at different decisions to be coupled by sharing the same latent random variable. Unlike classical CRN in fixed-distribution stochastic optimization, the coupling must preserve the decision-dependent marginal distributions at perturbed decisions, while exploiting the correlation induced by shared randomness. For each random direction u^i, we draw $\zeta^i \sim P_0$ and generate

$$\xi^{+,i} = T(x + \mu u^i, \zeta^i), \qquad \xi^{-,i} = T(x - \mu u^i, \zeta^i).$$

Although the marginal distributions remain $\xi^{+,i} \sim \mathcal{D}(x + \mu u^i)$, $\xi^{-,i} \sim \mathcal{D}(x - \mu u^i)$, the two samples are correlated through the shared latent variable ζ^i. Let $U = \{u^i\}_{i=1}^N$. The resulting coupled two-point estimator is defined as

$$G_\mu^{\mathrm{c}}(x, \{\zeta^i\}_{i=1}^N, U) = \frac{1}{N} \sum_{i=1}^N \frac{f(x + \mu u^i, \xi^{+,i}) - f(x - \mu u^i, \xi^{-,i})}{2\mu} u^i. \tag{4.1}$$

Under this coupled sampling oracle, the forward and backward function evaluations can be generated using the same latent randomness. The optimization updates remain identical to Algorithms 3.1 and 3.2; only the estimator construction is modified. Specifically, for a query center z_t, where $z_t = x_t$ for ZO-PM-CPL and $z_t = w_t$ for ZO-PM-IGT-CPL, the coupled two-point estimator $G_\mu^{\mathrm{c}}(x, \{\zeta^i\}_{i=1}^N, U)(z_t)$ is generated by the above procedure.

Algorithm 4.1 Coupled two-point sampling procedure

1: Given query center z_t.
2: **for** $i = 1, \ldots, N$ **do**
3: Draw $u_t^i \sim \mathcal{N}(0, I_d)$ and $\zeta_t^i \sim P_0$.
4: Set $\xi_t^{+,i} = T(z_t + \mu u_t^i, \zeta_t^i), \xi_t^{-,i} = T(z_t - \mu u_t^i, \zeta_t^i)$.
5: **end for**
6: Return $G_\mu^{\mathrm{c}}(z_t, \{\zeta_t^i\}_{i=1}^N, U_t)$ defined in (4.1).

We impose the regularity assumptions for the coupled-sampling framework.

Assumption 4.1. *There exists a measurable function $L_T : \mathcal{Z} \to \mathbb{R}_+$ such that for all $x, y \in \mathbb{R}^d$ and $\zeta \in \mathcal{Z}$,$\|T(x, \zeta) - T(y, \zeta)\| \le L_T(\zeta)\|x - y\|$ and $\mathbb{E}_{\zeta \sim P_0}[L_T(\zeta)^2] \le \sigma_T^2 < \infty$.*

Assumption 4.2. *There exist constants $L_x, L_\xi > 0$ such that for every $x, y \in \mathbb{R}^d$ and $\xi, \xi' \in \Xi$, $|f(x, \xi) - f(y, \xi)| \le L_x\|x - y\|$ and $|f(x, \xi) - f(x, \xi')| \le L_\xi\|\xi - \xi'\|$.*

These assumptions control the variation of individual sample paths under shared latent randomness and the resulting loss values, whereas Wasserstein sensitivity conditions characterize distributional shifts only through marginal distributions. The common-latent-source representation naturally covers location-shift models proposed in [21]. For example, consider the location-shift model $T(x, \zeta) = \zeta + Ax$, then

$$\|T(x, \zeta) - T(y, \zeta)\| \le \|A\|\|x - y\|,$$

with the constant choice $L_T(\zeta) = \|A\|$. Under Assumptions 4.1 and 4.2, the composite sample-path loss $x \mapsto f(x, T(x, \zeta))$ is Lipschitz continuous for each fixed latent realization ζ. More specifically, for every $x, y \in \mathbb{R}^d$, we obtain

$$\begin{aligned} |f(x, T(x, \zeta)) - f(y, T(y, \zeta))| &\le |f(x, T(x, \zeta)) - f(y, T(x, \zeta))| + |f(y, T(x, \zeta)) - f(y, T(y, \zeta))| \\ &\le (L_x + L_\xi L_T(\zeta))\|x - y\|. \end{aligned} \tag{4.2}$$

To compare the coupled estimator with the ideal finite-sample Gaussian finite-difference estimator, we define

$$G_\mu^F(x) := \frac{1}{N} \sum_{i=1}^N \frac{F(x + \mu u^i) - F(x - \mu u^i)}{2\mu} u^i.$$

The following lemma shows that, under the common-latent-source structure, the coupled estimator remains unbiased and its sample noise admits a variance bound without the unfavorable μ^{-2}-dependent term. This improvement is obtained by exploiting stronger sample-path regularity assumptions enabled by the latent representation, rather than relying only on distribution-level variance control.

LEMMA **4.1.** *Under Assumptions 4.1 and 4.2, we have*

$$\mathbb{E}_{\{\zeta^i\}}\left[G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)\;\middle|\;U\right]=G_\mu^F(x),$$
$$\mathbb{E}_{U,\{\zeta^i\}}[\|G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-G_\mu^F(x)\|^2]\le\frac{\bar{L}^2\,d(d+2)}{N},$$

where $\bar{L}^2:=\mathbb{E}_{\zeta\sim P_0}[(L_x+L_\xi L_T(\zeta))^2]$.

Proof. The marginal distributions of $\xi^{+,i}$ and $\xi^{-,i}$ are given by $\mathcal{D}(x+\mu u^i)$ and $\mathcal{D}(x-\mu u^i)$, respectively. Since the coupling preserves these marginal distributions, the conditional expectation of the coupled estimator coincides with that of the ideal finite-difference estimator. Therefore, the coupled estimator is conditionally unbiased given U. For the variance bound, we define

$$H(x,u,\zeta):=f(x+\mu u,T(x+\mu u,\zeta))-f(x-\mu u,T(x-\mu u,\zeta)).$$

Conditioned on U, we have

$$G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-G_\mu^F(x)=\frac{1}{N}\sum_{i=1}^N\frac{H(x,u^i,\zeta^i)-\mathbb{E}_\zeta H(x,u^i,\zeta)}{2\mu}u^i.$$

Since the samples $\{\zeta^i\}$ are independent and the centered terms have zero conditional mean given U, expanding the squared norm and eliminating the cross terms yields

$$\begin{aligned}\mathbb{E}[\|G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-G_\mu^F(x)\|^2]&=\frac{1}{4\mu^2N^2}\sum_{i=1}^N\mathbb{E}\left[\|u^i\|^2\,\mathrm{Var}_\zeta(H(x,u^i,\zeta))\right]\\&\le\frac{\bar{L}^2}{N^2}\sum_{i=1}^N\mathbb{E}[\|u^i\|^4],\end{aligned}$$

where $\bar{L}^2:=\mathbb{E}[(L_x+L_\xi L_T(\zeta))^2]$. The inequality follows from the path-Lipschitz condition (4.2), which gives $|H(x,u,\zeta)|\le 2\mu(L_x+L_\xi L_T(\zeta))\|u\|$ and $\mathrm{Var}_\zeta(H(x,u,\zeta))\le\mathbb{E}_\zeta[H(x,u,\zeta)^2]$. Thus, it holds that

$$\mathrm{Var}_\zeta(H(x,u,\zeta))\le\mathbb{E}_\zeta[H(x,u,\zeta)^2]\le 4\mu^2\bar{L}^2\|u\|^2.$$

The result then follows from $\mathbb{E}[\|u^i\|^4]=d(d+2)$. □

By the conditional unbiasedness established above, the coupled estimator error satisfies the same conditional zero-mean property used in Lemma 3.5. Therefore, the tracking analysis can be directly extended to the coupled setting.

4.2 Convergence of ZO-PM-CPL

We replace the independent estimator G_μ in Algorithm 3.1 with the coupled estimator returned by Procedure 4.1, resulting in the ZO-PM-CPL method. The following result establishes the conditional unbiasedness and variance bound of the coupled estimator used in ZO-PM-CPL. Under the common-latent-source structure, this variance bound does not contain the unfavorable μ^{-2}-dependent term that arises under independent sampling.

LEMMA **4.2.** *Suppose Assumptions 2.2, 4.1 and 4.2 hold. Then the coupled estimator* (4.1) *satisfies* $\mathbb{E}[G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)|x] = \nabla F_\mu(x)$ *and*

$$\mathbb{E}\left[\left\|G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-\nabla F_\mu(x)\right\|^2\right] \le \sigma_{\mathrm{c}}^2(x),$$

where $\sigma_{\mathrm{c}}(x) = (\frac{\bar{L}^2 d(d+2)}{N} + \frac{dL^2\mu^2}{2N}(d+2)(d+4) + \frac{6d}{N}\|\nabla F(x)\|^2)^{1/2}$.

Proof. Using the decomposition

$$G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-\nabla F_\mu(x) = \big(G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-G_\mu^F(x)\big)+\big(G_\mu^F(x)-\nabla F_\mu(x)\big),$$

we obtain

$$\mathbb{E}\left[\|G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-\nabla F_\mu(x)\|^2\right] = \mathbb{E}\left[\|G_\mu^{\mathrm{c}}(x,\{\zeta^i\}_{i=1}^N,U)-G_\mu^F(x)\|^2\right] + \mathbb{E}\left[\|G_\mu^F(x)-\nabla F_\mu(x)\|^2\right],$$

where the cross term vanishes due to the conditional unbiasedness of the coupled estimator given U. The first term is bounded by Lemma 4.1, while the second term follows from the standard Gaussian finite-difference variance bound [12]. Combining the two bounds gives the desired result. □

We note that

$$\bar{\sigma}_c := (\frac{2\bar{L}^2 d(d+2)}{N} + \frac{dL^2\mu^2}{N}(d+2)(d+4) + \frac{12dG^2}{N})^{1/2} \ge \sigma_c(x_t), \quad \forall t.$$

Similar to the analysis of Theorem 3.1, The following result shows the sample complexity ZO-PM-CPL under first-order smoothness.

THEOREM **4.1.** *Suppose Assumptions 2.2, 4.1 and 4.2 hold, and let*

$$\eta = \frac{1}{4L_\mu + \sqrt{TL_\mu/\Phi(x_1)}\bar{\sigma}_c}, \qquad \beta = \frac{3\eta L_\mu}{1-L_\mu\eta}, \tag{4.3}$$

where $\Phi(x_1) = F_\mu(x_1) - F_{\inf} + \frac{3}{8L_\mu}\mathbb{E}[\|m_1 - \nabla F_\mu(x_1)\|^2]$. *If* $\mu = \epsilon d^{-1/2}$ *and* $N = d^2\epsilon^{-2}$, *then ZO-PM-CPL outputs* x_r *satisfying* $\mathbb{E}[\|\nabla F(x_r)\|] \le \epsilon$ *with iteration complexity of order* $\mathcal{O}(\epsilon^{-2})$ *and total sample complexity of order* $\mathcal{O}(d^2\epsilon^{-4})$.

Proof. The proof follows the Lyapunov analysis of Theorem 3.1. Define the Lyapunov function as follows

$$\Phi(x_t) = \mathbb{E}[F_\mu(x_t) - F_{\inf}] + \frac{3}{8L_\mu}\mathbb{E}[\|\epsilon_t\|^2],$$

where $\epsilon_t = m_t - \nabla F_\mu(x_t)$. By Lemma 4.2, the coupled estimator satisfies that

$$\mathbb{E}[\|G_\mu^c(x_t) - \nabla F_\mu(x_t)\|^2] \le \bar{\sigma}_c^2,$$

which replaces the variance bound used in Lemma 3.2. Therefore, applying the same momentum tracking recursion as Lemma 3.2, we obtain

$$\mathbb{E}[\|\epsilon_{t+1}\|^2] \le (1-\beta)^2\mathbb{E}[\|\epsilon_t\|^2] + \beta^2\bar{\sigma}_c^2 + C_1\eta^2L_\mu^2\mathbb{E}[\|\nabla F_\mu(x_t)\|^2].$$

Combining the above inequality with the descent inequality in Lemma 3.3 gives

$$\frac{1}{T}\sum_{t=1}^T \mathbb{E}[\|\nabla F_\mu(x_t)\|^2] \le C_1\frac{L_\mu\Phi(x_1)}{T} + C_2\sqrt{\frac{L_\mu\Phi(x_1)}{T}}\,\bar{\sigma}_c.$$

Using the choice of parameters as (4.3), we derive the desired convergence bound. Finally, choosing $\mu = \epsilon d^{-1/2}$ and $N = d^2\epsilon^{-2}$ gives $\bar{\sigma}_c = \mathcal{O}(\epsilon)$ and therefore $T = \mathcal{O}(\epsilon^{-2})$, and the total sample complexity is $TN = \mathcal{O}(d^2\epsilon^{-4})$. □

4.3 Convergence of ZO-PM-IGT-CPL

ZO-PM-IGT-CPL is obtained by applying the same coupled sampling procedure at the extrapolated query point w_t in Algorithm 3.2. The following lemma gives the corresponding variance bound for the two-point zeroth-order estimator under second-order smoothness and coupled-sampling regularity. The proof is similar to Lemma 4.2, thus we omit the proof here.

LEMMA **4.3.** *Suppose Assumptions 2.2, 2.3, 4.1 and 4.2 hold, then the coupled estimator* (4.1) *satisfies* $\mathbb{E}[G_\mu^{\mathrm{c}}(x, \{\zeta^i\}_{i=1}^N, U)|x] = \nabla F_\mu(x)$ *and*

$$\mathbb{E}\left[\left\|G_\mu^{\mathrm{c}}(x, \{\zeta^i\}_{i=1}^N, U|x) - \nabla F_\mu(x)\right\|^2\right] \leq \tilde{\sigma}_{\mathrm{c}}^2(x)$$

where $\tilde{\sigma}_{\mathrm{c}}(x) = (\frac{2\bar{L}^2 d(d+2)}{N} + \frac{H^2\mu^4}{N} d(d+2)(d+4)(d+6) + \frac{6d}{N}\|\nabla F(x)\|^2)^{1/2}$ *and* $\bar{L}^2$ *is defined in Lemma 4.1.*

Proof. The proof follows the same decomposition as Lemma 4.2. Specifically,

$$G_\mu^{\mathrm{c}}(x, \{\zeta^i\}_{i=1}^N, U) - \nabla F_\mu(x) = \left(G_\mu^{\mathrm{c}}(x, \{\zeta^i\}_{i=1}^N, U) - G_\mu^F(x)\right) + \left(G_\mu^F(x) - \nabla F_\mu(x)\right).$$

The first term is bounded by Lemma 4.1, while the second term follows from the standard Gaussian finite-difference variance bound under second-order smoothness. Combining the two bounds yields the desired result. □

Define the coupled estimator error as follows

$$\tilde{\varepsilon}_t^{\mathrm{c}} := G_\mu^{\mathrm{c}}\big(w_t, \{\zeta_t^i\}_{i=1}^N, U_t\big) - \nabla F_\mu(w_t).$$

Then, we have $\mathbb{E}[\|\tilde{\varepsilon}_t^{\mathrm{c}}\|^2] \leq \tilde{\sigma}_c^2(w_t)$ according to Lemma 4.3. We note that

$$\tilde{\sigma}_{\max}^c := \left(\frac{2\bar{L}^2 d(d+2)}{N} + \frac{H^2\mu^4}{N} d(d+2)(d+4)(d+6) + \frac{6dG^2}{N}\right)^{1/2} \geq \tilde{\sigma}_c(w_t), \forall t,$$

Similar to the analysis of Theorem 3.3, The following result shows the improved sample complexity under second-order smoothness.

THEOREM **4.2.** *Suppose that Assumptions 2.2, 2.3, 4.1 and 4.2 hold, and choose*

$$\eta = \min\Big\{\frac{R^{5/7}}{T^{5/7}H^{1/7}(\tilde{\sigma}_{\max}^{\mathrm{c}})^{4/7}}, \sqrt{\frac{R}{TL}}\Big\}, \qquad \beta = \min\Big\{\frac{R^{4/7}H^{2/7}}{T^{4/7}(\tilde{\sigma}_{\max}^{\mathrm{c}})^{6/7}}, 1\Big\},$$

where $R := F_\mu(x_1) - F_{\inf}$. *If* $\mu = \epsilon d^{-1/2}$ *and* $N = d^2\epsilon^{-3/2}$, *then the output of* ZO-PM-IGT-CPL *satisfies* $\mathbb{E}[\|\nabla F(x_r)\|] \leq \epsilon$ *with iteration complexity in order* $\mathcal{O}(\epsilon^{-2})$ *and total sample complexity in order* $\mathcal{O}(d^2\epsilon^{-3.5})$.

Proof. The proof follows the IGT-based Lyapunov analysis of Theorem 3.3. By Lemma 4.3, the coupled estimator satisfies $\mathbb{E}[\|\tilde{\epsilon}_t^c\|^2] \leq (\tilde{\sigma}_{\max}^c)^2$. Compared with the analysis of Theorem 3.3, the IGT update remains unchanged, and hence the first-order transport cancellation argument is identical. The only difference is that the stochastic estimator error term is controlled by $\tilde{\sigma}_{\max}^c$ instead of $\tilde{\sigma}_{\max}$. Following the tracking recursion in Lemma 3.5, we obtain

$$\mathbb{E}[\|\epsilon_{t+1}^c\|] \leq (1-\beta)^t\mathbb{E}[\|\epsilon_1^c\|] + C_1\sqrt{\beta}\tilde{\sigma}_{\max}^c + C_2\frac{H\eta^2}{\beta^2}.$$

Combining this bound with the descent inequality of Lemma 3.6 yields

$$\frac{1}{T}\sum_{t=1}^T \mathbb{E}[\|\nabla F_\mu(x_t)\|] \leq C_1\sqrt{\frac{RL_\mu}{T}} + C_2\frac{R^{2/7}H^{1/7}(\tilde{\sigma}_{\max}^c)^{4/7}}{T^{2/7}} + C_3\frac{(\tilde{\sigma}_{\max}^c)^{13/7}}{R^{4/7}H^{2/7}T^{3/7}}.$$

With $\mu = \epsilon d^{-1/2}, N = d^2\epsilon^{-3/2}$, we have $\tilde{\sigma}_{\max}^c = O(\epsilon^{3/4})$. The above inequality gives $\mathbb{E}[\|\nabla F(x_r)\|] \leq \epsilon$. With the iteration complexity $T = \mathcal{O}(\epsilon^{-2})$ and the total sample complexity $TN = \mathcal{O}(d^2\epsilon^{-3.5})$. □

Remark 4.1. *Coupled sampling reduces estimator variance and yields the complexity $\mathcal{O}(d^2\epsilon^{-4})$ for ZO-PM-CPL. However, it does not address the first-order tracking error in the momentum-tracking recursion. The further improvement of ZO-PM-IGT-CPL is therefore attributed to IGT, which replaces the first-order transport term with a second-order residual.*

5 Experiments

Following [9], we conduct experiments on multiproduct pricing [11] and strategic classification [13] to evaluate the proposed methods in realistic decision-dependent applications. We further include a controlled continuous location-shift synthetic experiment to isolate the effect of coupled sampling under the exact common-latent-source structure. All methods use only stochastic function-value feedback and do not explicitly use the closed-form distribution $\mathcal{D}(x)$. Detailed problem settings, parameter choices, and evaluation procedures are provided in Appendix B.

5.1 Setup and compared methods

We implemented the following methods.

- **ZO-OG**: a zeroth-order method with a one-point gradient estimator, which is analogous to the zeroth-order method proposed in [18].
- **ZO-OGVR**: a zeroth-order method using a one-point gradient estimator with a variance reduction parameter [11].
- **ZO-GA**: a zeroth-order method with a Gaussian-smoothed two-point gradient estimator, which is analogous to the method proposed in [12].
- **ZO-O2NC-I & ZO-O2NC-II**: the two variants of the O2NC-based zeroth-order method proposed by Liu et al. [17]. ZO-O2NC-I uses a two-point estimator, whereas ZO-O2NC-II uses a one-point residual estimator that reuses the previous function evaluation.
- **ZO-PM & ZO-PM-IGT**: the momentum-based zeroth-order methods corresponding to Algorithms 3.1 and 3.2, respectively.
- **ZO-PM-CPL**: the coupled variant of ZO-PM, obtained by replacing the independent two-point estimator in Algorithm 3.1 with the coupled estimator generated by Procedure 4.1.
- **ZO-PM-IGT-CPL**: the coupled variant of ZO-PM-IGT, obtained by applying Procedure 4.1 at the extrapolated query point w_t in Algorithm 3.2.

The CPL variants exploit common-latent sampling by reusing the same latent realization across paired perturbed evaluations, whereas the remaining methods follow their original sampling constructions without common-latent coupling.

5.2 Multiproduct pricing application

We consider a multiproduct pricing problem following [7, 11], where a seller sets the prices of ten products for forty buyers. The demand ξ follows a multinomial logit model whose distribution depends on the price vector $x \in \mathbb{R}^{10}$, and the objective is the negative expected profit.

5.2.1 Setting and metrics

All methods are initialized at $x_0 = 0.5 \cdot \mathbf{1}$, where $\mathbf{1}$ denotes the all-ones vector, and terminated after $10,000$ function evaluations. For the final output $\hat{x}$, we estimate the objective using $1,000$ independent samples from $\mathcal{D}(\hat{x})$. Results are averaged over 20 random seeds, and lower values are better. We consider two regimes. The standard regime uses $\mu = 0.19$ and $N = 10$ for overall performance comparison, while the small-smoothing regime uses $\mu = 0.02$ and $N = 2$ to emphasize the variance sensitivity of independently sampled two-point estimators and assess the benefit of common-latent coupling.

Estimator variance We examine the variance-reduction effect of common-latent coupling using an offline stage-wise diagnostic. Along a representative ZO-PM-IGT-CPL trajectory, we report the average conditional variance over multiple query points from the early, middle, and final stages. We compare PM-IND, PM-CPL, and ZO-O2NC-I; PM-IND/PM-CPL isolates the coupling effect, while ZO-O2NC-I serves as an independent-sampling reference in the small-μ regime.

5.2.2 Results

We evaluate the proposed methods from three perspectives: the benefit of IGT over plain momentum, the effect of coupled sampling under small smoothing, and the resulting variance reduction of the zeroth-order estimator.

Effect of IGT Table 2 compares the proposed momentum-based methods with existing zeroth-order baselines. Both ZO-PM and ZO-PM-IGT consistently outperform the one-point methods ZO-OG and ZO-OGVR. Compared with ZO-PM, ZO-PM-IGT achieves lower objective values on seven of the eight weekly instances, suggesting that the IGT mechanism improves the finite-budget performance of momentum-based zeroth-order optimization. ZO-O2NC-I also performs strongly and achieves the best result on several instances, providing a competitive baseline in the standard regime.

Table 2: Standard-regime results on the multiproduct pricing problem. Values report the mean objective and standard deviation over 20 matched random seeds; lower objective values are better.

date	ZO-PM		ZO-PM-IGT		ZO-GA		ZO-OG		ZO-OGVR		ZO-O2NC-I		ZO-O2NC-II	
	obj	sd	obj	sd	obj	sd	obj	sd	obj	sd	obj	sd	obj	sd
2/21–2/27	-12.17	0.84	**-12.85**	0.76	-12.11	0.84	1.70	5.21	-6.66	1.77	-12.73	0.75	-12.24	0.76
3/21–3/27	-11.84	0.88	-12.53	0.74	-11.76	0.96	-0.24	2.00	-7.45	1.58	**-12.54**	0.76	-12.08	0.77
5/23–5/29	**-16.07**	0.77	-15.68	0.86	-15.93	0.81	-0.44	1.64	-4.92	1.48	-15.98	0.79	-15.76	0.94
6/20–6/26	-8.82	1.08	-9.09	0.97	-8.79	1.07	2.24	3.18	-6.83	1.82	**-10.00**	0.79	-9.57	1.10
7/18–7/24	-8.00	1.02	-8.19	0.94	-7.97	1.03	6.85	9.19	-4.11	1.84	**-8.96**	0.91	-8.64	0.88
8/08–8/14	-15.72	0.72	**-15.87**	0.76	-15.64	0.71	-1.23	1.12	-6.94	1.54	-15.79	0.76	-15.33	0.79
9/19–9/25	-7.17	1.21	-7.26	1.16	-7.15	1.21	7.46	8.15	-3.91	2.63	**-7.92**	1.03	-7.59	1.16
12/05–12/11	-7.53	1.05	-7.69	1.00	-7.50	1.08	4.60	6.41	-4.11	2.25	**-8.51**	0.98	-8.19	1.00

Effect of coupled sampling Table 3 evaluates coupled sampling in the small-smoothing regime, where estimator variance becomes more pronounced. Both ZO-PM-CPL and ZO-PM-IGT-CPL consistently improve upon their uncoupled counterparts, with lower objective values and generally smaller standard deviations across all weekly instances. They also outperform the ZO-O2NC baselines. Notably, ZO-PM-IGT-CPL achieves the best objective value on all eight instances, showing that coupling and IGT provide complementary gains under small smoothing.

Table 3: Small-smoothing coupling results on the multiproduct pricing problem. Values report the mean objective and standard deviation over 20 matched random seeds; lower objective values are better.

date	ZO-PM		ZO-PM-CPL		ZO-PM-IGT		ZO-PM-IGT-CPL		ZO-O2NC-I		ZO-O2NC-II	
	obj	sd	obj	sd	obj	sd	obj	sd	obj	sd	obj	sd
2/21–2/27	-8.73	1.93	-12.55	0.93	-13.05	0.83	**-13.24**	0.79	-10.32	1.69	-9.13	2.38
3/21–3/27	-7.77	2.42	-12.39	1.21	-12.84	0.87	**-12.99**	0.85	-9.60	2.01	-8.00	3.32
5/23–5/29	-9.70	2.60	-16.02	0.85	-16.11	0.78	**-16.65**	0.74	-7.43	2.62	-7.70	2.54
6/20–6/26	-6.37	2.51	-9.92	0.89	-10.28	0.82	**-10.40**	0.83	-9.45	1.26	-8.83	1.78
7/18–7/24	-5.90	1.92	-9.07	0.81	-9.18	0.80	**-9.29**	0.79	-7.96	1.90	-7.27	2.24
8/08–8/14	-10.66	2.22	-15.86	0.69	-15.96	0.79	**-16.33**	0.74	-9.64	2.11	-8.05	3.12
9/19–9/25	-6.42	1.86	-7.97	1.19	-8.24	1.01	**-8.37**	0.95	-7.50	1.99	-7.32	1.52
12/05–12/11	-6.78	1.87	-8.48	0.93	-8.80	0.93	**-8.83**	0.89	-8.03	1.49	-7.59	1.47

Estimator variance Figure 1 reports the stage-wise estimator variance along a representative optimization trajectory. PM-CPL consistently reduces variance relative to PM-IND across stages and smoothing radii, with larger gains as μ decreases, confirming the benefit of common-latent coupling in the small-smoothing regime. ZO-O2NC-I, as an independent-sampling reference, exhibits variance sensitivity for small μ, indicating that independent sampling remains a source of variance amplification.

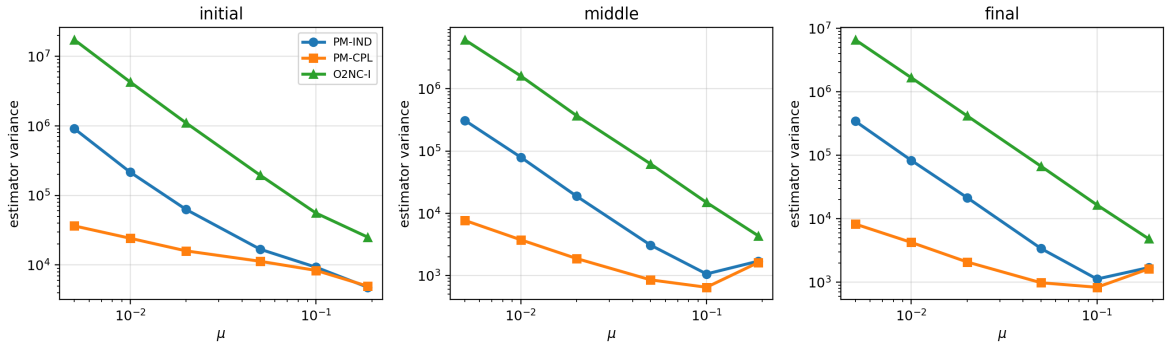

Figure 1: Stage-wise estimator variance of PM-IND, PM-CPL, and ZO-O2NC-I under different smoothing radii in the multiproduct pricing problem. Each value is averaged over multiple fixed query points within the corresponding optimization stage.

5.3 Strategic classification

We further evaluate the proposed methods on strategic classification, following [9, 13]. In this setting, a decision maker deploys a linear classifier, while agents strategically modify their observable features after observing the classifier. Consequently, the distribution of reported features depends on x. Given a reported feature vector ξ_F, the predicted probability is

$$p_x(\xi_F) = 1/(1 + \exp\Big(-x_{[11]}^\top \xi_F - x_{12}\Big)).$$

The objective is to minimize the expected logistic loss:

$$\min_x F(x) = \mathbb{E}_{(\xi_F, L)\sim\mathcal{D}(x)}[-L \log p_x(\xi_F) - (1 - L)\log(1 - p_x(\xi_F))].$$

where $\mathcal{D}(x)$ is induced by the strategic response to the deployed classifier.

5.3.1 Setting and metrics

We use the processed credit-card default dataset considered in [9, 13], containing 11 features and a binary default label. All methods are initialized at $x_0 = \mathbf{1} \in \mathbb{R}^{12}$ and use the same stochastic function-evaluation budget. We report training loss, test loss, test AUC, and test accuracy, with lower losses and higher AUC/accuracy indicating better performance. We consider two strategic-response settings. The standard setting uses $\tau = 1.0$ for overall performance comparison, while the $\tau = 5.0$ setting is used to evaluate coupled sampling under an alternative strategic-response regime. All methods based on random-direction zeroth-order estimators use $N = 10$ directions. The smoothing radius μ is method-dependent and follows the tuned or sensitivity-validated choices reported in Appendix B.2.

5.3.2 Results

We evaluate the proposed methods from two perspectives: the effect of implicit gradient transport under the standard strategic classification setting and the effect of coupled sampling under a high-variance strategic-response setting.

Effect of IGT Table 4 reports the results in the standard setting. The proposed momentum-based methods achieve lower losses than the existing zeroth-order baselines. ZO-PM-IGT attains the best mean value on all four metrics, while ZO-O2NC-I remains less competitive in test loss under its sensitivity-validated smoothing radius. These results support the finite-budget benefit of IGT in this application.

Table 4: Results on strategic classification with $\tau = 1.0$. The reported values are mean and standard deviation over 20 random seeds.

metric	ZO-PM-IGT		ZO-PM		ZO-GA		ZO-OG		ZO-OGVR		ZO-O2NC-I		ZO-O2NC-II	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
train loss (↓)	**0.742**	0.050	0.748	0.047	0.795	0.071	1.069	0.159	0.890	0.075	0.878	0.281	2.506	0.972
test loss (↓)	**0.741**	0.073	0.754	0.073	0.799	0.094	1.085	0.143	0.899	0.088	0.895	0.328	2.508	0.997
test accuracy (↑)	**0.605**	0.032	0.599	0.043	0.601	0.041	0.506	0.022	0.585	0.042	0.589	0.043	0.537	0.066
test AUC (↑)	**0.655**	0.037	0.653	0.045	0.652	0.050	0.518	0.048	0.643	0.039	0.644	0.050	0.550	0.106

Effect of coupled sampling Table 5 reports the results under the high-variance strategic-response setting. Coupled sampling substantially improves both momentum-based methods. The coupled variants also outperform ZO-O2NC-I with its sensitivity-validated smoothing radius across the reported metrics, with ZO-PM-IGT-CPL achieving the best overall performance.

Table 5: Results of coupled sampling under the high-variance strategic-response setting. The reported values are mean and standard deviation over 20 random seeds.

metric	ZO-PM-IGT		ZO-PM-IGT-CPL		ZO-PM		ZO-PM-CPL		ZO-O2NC-I		ZO-O2NC-II	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
train loss (↓)	0.958	0.176	**0.701**	0.039	2.902	1.734	0.745	0.052	1.535	0.477	2.018	0.250
test loss (↓)	0.973	0.190	**0.703**	0.048	2.870	1.750	0.753	0.060	1.550	0.487	2.028	0.284
test accuracy (↑)	0.593	0.032	**0.639**	0.042	0.560	0.063	0.638	0.031	0.527	0.066	0.444	0.027
test AUC (↑)	0.638	0.045	**0.692**	0.039	0.586	0.090	0.689	0.030	0.530	0.108	0.397	0.043

5.4 Ablation study on coupled sampling

To isolate the effect of coupled sampling, we construct a synthetic decision-dependent phase retrieval benchmark with an explicit common-latent representation. Specifically, the sample associated with decision x is

generated as $\xi = T(x, \zeta) = (a(x, \zeta), y(\zeta))$, where the same latent variable ζ can be shared by the forward and backward evaluations in the two-point estimator while preserving the correct decision-dependent marginals. We use the phase retrieval loss

$$f(x, \xi) = \frac{1}{4}\big((a^\top x)^2 - y\big)^2, \qquad F(x) = \mathbb{E}_\zeta[f(x, T(x, \zeta))].$$

This benchmark is therefore well suited for isolating the variance-reduction effect of common-latent coupling. The detailed construction is provided in Appendix B.3.

5.4.1 Setting and metrics

We use $d = 10$ and initialize all methods at $x_0 = 0.3 \cdot \mathbf{1}$. Each method is run for $10,000$ stochastic function evaluations, and results are averaged over 50 independent random seeds. We report the excess objective $F(x) - F(x^*_{\text{ref}})$ and the gradient norm $\|\nabla F(x)\|$, where x^*_{ref} is a high-accuracy numerical reference solution; lower values are better. Details on the parameter generation and reference-solution computation are provided in Appendix B.3.

5.4.2 Results

Figure 2 compares the coupled and uncoupled variants under the same function-evaluation budget. Common-latent coupling improves both ZO-PM and ZO-PM-IGT, yielding lower excess objective values and gradient norms than their uncoupled counterparts. The gain is most pronounced for ZO-PM-IGT-CPL, which achieves the lowest final objective gap and stationarity error among the four methods. These results support the role of common-latent coupling in reducing sampling-induced variability.

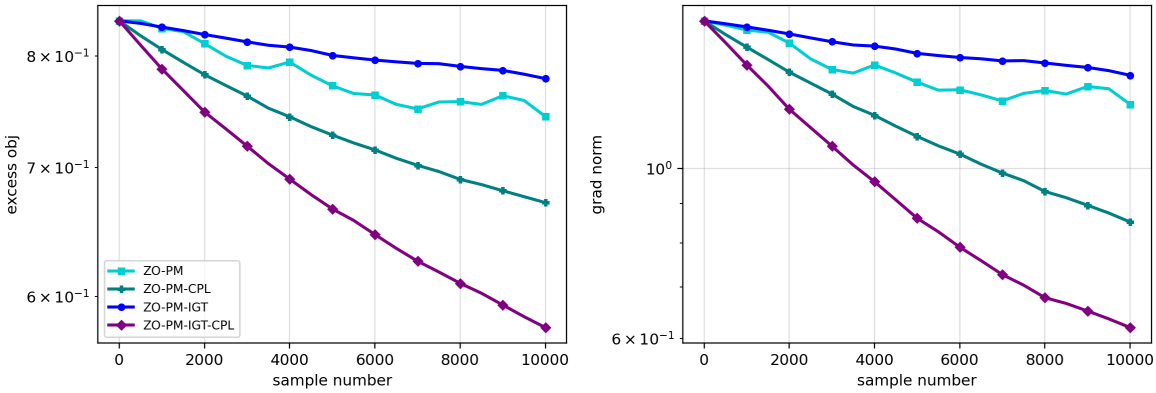

Figure 2: Convergence curves for the decision-dependent phase retrieval ablation study. The plots show the mean excess objective value and mean gradient norm over 50 random seeds as functions of cumulative stochastic function evaluations. Lower values are better.

6 Conclusions

We studied nonconvex stochastic zeroth-order optimization with decision-dependent distributions under a sample-only function-value oracle. The proposed framework provides a systematic development from momentum-based gradient tracking to implicit transport correction and coupled sampling. We first developed a Polyak-momentum zeroth-order method with sample complexity $\mathcal{O}(d^2\epsilon^{-6})$ under first-order smoothness, and showed that decision-dependent sampling distributions introduce a first-order transport component in the momentum-tracking recursion. We then introduced implicit gradient transport (IGT), which reduces

this component to a higher-order residual and improves the complexity to $\mathcal{O}(d^2\epsilon^{-4.5})$ under second-order smoothness. For structured problems with decision-dependent distributions admitting a common-latent-source representation, we further developed a marginal-preserving coupled-sampling framework that reduces the variance caused by independent two-point sampling, yielding complexities $\mathcal{O}(d^2\epsilon^{-4})$ and $\mathcal{O}(d^2\epsilon^{-3.5})$ under first- and second-order smoothness, respectively. Experiments on multiproduct pricing, strategic classification, and a synthetic benchmark support the effectiveness of the proposed mechanisms. Future work includes relaxing the coupled-sampling assumptions, reducing the dependence on unknown problem parameters, and extending the framework to broader settings with decision-dependent distributions.

A Gradient representation under density access

LEMMA **A.1** (Differentiability and gradient representation)**.** *Suppose that, for every $x \in \mathbb{R}^d$, the distribution $\mathcal{D}(x)$ admits a density $p_x(\xi) > 0$ with respect to a reference measure $d\xi$. Let $F(x) = \int_\Xi f(x,\xi)p_x(\xi)\,d\xi$. Assume that $f(x,\xi)$ and $p_x(\xi)$ are differentiable in x for almost every ξ. In addition, assume that for every compact set $\mathcal{K} \subset \mathbb{R}^d$, there exists an integrable function $M_\mathcal{K} : \Xi \to [0,\infty)$ such that*

$$\|\nabla_x f(x,\xi)p_x(\xi) + f(x,\xi)\nabla_x p_x(\xi)\| \le M_\mathcal{K}(\xi), \qquad \forall x \in \mathcal{K},$$

for almost every ξ. Then F is differentiable, and its gradient is given by

$$\nabla F(x) = \mathbb{E}_{\xi\sim\mathcal{D}(x)}\left[\nabla_x f(x,\xi) + f(x,\xi)\nabla_x \log p_x(\xi)\right].$$

Proof. Define $H(x,\xi) := f(x,\xi)p_x(\xi)$. Then $F(x) = \int_\Xi H(x,\xi)\,d\xi$. By assumption, $H(x,\xi)$ is differentiable in x for almost every ξ, and we have

$$\nabla_x H(x,\xi) = \nabla_x f(x,\xi)p_x(\xi) + f(x,\xi)\nabla_x p_x(\xi).$$

Moreover, for every compact set $\mathcal{K} \subset \mathbb{R}^d$, we have $\|\nabla_x H(x,\xi)\| \le M_\mathcal{K}(\xi), \forall x \in \mathcal{K}$, where $M_\mathcal{K} \in L^1(d\xi)$. Therefore, by the dominated convergence theorem, differentiation can be exchanged with integration. Hence, F is differentiable and there is

$$\nabla F(x) = \nabla_x \int_\Xi H(x,\xi)\,d\xi = \int_\Xi \nabla_x H(x,\xi)\,d\xi.$$

Substituting the expression of $\nabla_x H(x,\xi)$, we derive

$$\nabla F(x) = \int_\Xi \left[\nabla_x f(x,\xi)p_x(\xi) + f(x,\xi)\nabla_x p_x(\xi)\right] d\xi.$$

Since $p_x(\xi) > 0$, we have $\nabla_x p_x(\xi) = p_x(\xi)\nabla_x \log p_x(\xi)$. Therefore, we conclude

$$\nabla F(x) = \mathbb{E}_{\xi\sim\mathcal{D}(x)}\left[\nabla_x f(x,\xi) + f(x,\xi)\nabla_x \log p_x(\xi)\right].$$

This proves both the differentiability of F and the claimed gradient representation. □

B Additional experimental details

B.1 Multiproduct Pricing

The stochastic objective is defined through the negative profit $f(x,\xi) = -\sum_{i=1}^n x_i\xi_i + \sum_{i=1}^n c_i(\xi_i)$, where x_i is the price of product i and ξ_i is its realized demand. The production cost is modeled as

$$c_i(\xi_i) = \begin{cases} 2w_i\xi_i, & \xi_i < l_i, \\ w_i(\xi_i - l_i) + 2w_i l_i, & l_i \le \xi_i < u_i, \\ 3w_i(\xi_i - u_i) + w_i(u_i - l_i) + 2w_i l_i, & \xi_i \ge u_i. \end{cases}$$

We set $l_i = 0.5m/n$, $u_i = 1.5m/n$, and $w_i = \rho_i\theta_i$, where $\rho_i \sim \text{Unif}[0.25, 0.5]$ and θ_i is the normalized historical average selling price.

Decision-dependent demand model The demand distribution $\mathcal{D}(x)$ is induced by a multinomial logit model. Each buyer chooses product i with probability

$$p_i(x) = \frac{\exp(\gamma_i(\theta_i - x_i))}{a_0 + \sum_{j=1}^n \exp(\gamma_j(\theta_j - x_j))},\ p_0(x) = \frac{a_0}{a_0 + \sum_{j=1}^n \exp(\gamma_j(\theta_j - x_j))},$$

where $p_0(x)$ denotes the probability of no purchase. We use $\gamma_i = \frac{\pi}{0.5\sqrt{6}\theta_i}, a_0 = 0.1n$. Conditioned on x, the demand vector follows a multinomial distribution with parameters $(m, p_0(x), p_1(x), \ldots, p_n(x))$.

Estimator variance evaluation For the multiproduct pricing variance diagnostic, we use a representative ZO-PM-IGT-CPL trajectory in the small-μ regime. We select $K = 5$ query points from each of the early, middle, and final stages, uniformly over $[0, 2000]$, $[4000, 6000]$, and $[8000, 10000]$ in cumulative function evaluations. At each point $x_{s,k}$, we generate $R = 200$ independent estimator realizations and compute

$$\widehat{\mathrm{Var}}(G_\mu(x_{s,k})) = \frac{1}{R-1}\sum_{r=1}^{R} \|G_\mu^{(r)}(x_{s,k}) - \bar{G}_\mu(x_{s,k})\|^2,$$

where $\bar{G}_\mu(x_{s,k})$ is the sample mean. The reported stage-level variance is averaged over the K points, with standard errors across them. We compare PM-IND, PM-CPL, and ZO-O2NC-I for $\mu \in \{0.005, 0.01, 0.02, 0.05, 0.10, 0.19\}$. PM-IND and PM-CPL differ only in independent versus shared latent realizations, while ZO-O2NC-I is an independent-sampling reference.

B.2 Strategic classification

We follow the strategic response model in [9,13]. Each agent has a true feature vector $\xi_{\text{true}} \in \mathbb{R}^{11}$ and label $L \in \{0, 1\}$. Given classifier parameter $x \in \mathbb{R}^{12}$, the reported feature vector ξ_F is obtained by solving

$$\xi_F \in \arg\max_{\xi \in \mathbb{R}^{11}} \left\{ r(\xi, x) - \tau \|\xi - \xi_{\text{true}}\|_2^2 \right\},$$

where

$$r(\xi, x) = \begin{cases} 2, & x_{[11]}^\top \xi + x_{12} \geq 0, \\ 0, & x_{[11]}^\top \xi + x_{12} < 0. \end{cases}$$

Here, $\tau > 0$ controls the cost of strategic manipulation; larger τ discourages feature modification. The pair (ξ_F, L) induces the decision-dependent distribution $\mathcal{D}(x)$.

Parameter settings We consider $\tau = 1.0$ for the standard setting and $\tau = 5.0$ for the coupled-sampling experiment. In the standard setting, ZO-PM and ZO-PM-IGT use $\mu = 0.7$, ZO-GA and ZO-OGVR use $\mu = 1.0$, and ZO-OG uses $\mu = 0.3$. In the $\tau = 5.0$ setting, ZO-PM and ZO-PM-CPL use $\mu = 0.12$, whereas ZO-PM-IGT and ZO-PM-IGT-CPL use $\mu = 0.175$. For the O2NC baselines, we performed a lightweight pre-specified sensitivity check over μ. ZO-O2NC-I uses the validated radii $\mu = 1.0$ and $\mu = 0.35$ in the two settings, respectively. ZO-O2NC-II retains the matched radius because alternative choices either degrade AUC/accuracy or do not significantly improve test loss. For the CPL variants, the two perturbed evaluations share the same latent realization $\zeta = (\xi_{\text{true}}, L)$, whereas the uncoupled variants use independent latent realizations for the two evaluations.

B.3 Decision-dependent phase retrieval

Let $\zeta = (a_0, \varepsilon)$, $a_0 \sim \mathcal{N}(0, I_d)$ and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. Given a decision x, define

$$a(x, \zeta) = a_0 + \rho_{\text{dd}} A x, \qquad y(\zeta) = (a_0^\top x^{\text{sig}})^2 + \varepsilon,$$

where $A \in \mathbb{R}^{d\times d}$ is fixed, $\rho_{\mathrm{dd}} \geq 0$ controls the strength of decision dependence, and $\|x^{\mathrm{sig}}\| = 1$. The induced sample is $\xi = T(x,\zeta) = (a(x,\zeta), y(\zeta))$, and the loss is $f(x,\xi) = \frac{1}{4}\big((a^\top x)^2 - y\big)^2$. The case $\rho_{\mathrm{dd}} = 0$ reduces to standard stochastic phase retrieval, whereas $\rho_{\mathrm{dd}} > 0$ introduces decision dependence through the sensing distribution. The uncoupled variants use independent latent realizations for the two perturbed evaluations, whereas the coupled variants reuse the same latent realization. To isolate the coupling effect, we compare only paired methods that differ in their sampling construction; O2NC is therefore omitted.

Parameter settings. We set $\rho_{\mathrm{dd}} = 0.25$, $\sigma = 0.1$, $\mu = 0.02$, and $N = 2$. For each random seed, $A = U\,\mathrm{diag}(0.2,\ldots,1.0)V^\top$, where U, V are obtained from the SVD of a Gaussian random matrix, and x^{sig} is sampled from a standard Gaussian distribution and normalized. ZO-PM and ZO-PM-CPL use $(\eta,\beta) = (5\times10^{-5}, 0.2)$, while ZO-PM-IGT and ZO-PM-IGT-CPL use $(\eta,\beta,\alpha) = (6\times10^{-4}, 0.2, 0.5)$.

Evaluation details. The population objective $F(x)$ and gradient $\nabla F(x)$ are evaluated using Gaussian moment identities. For each instance, x^*_{ref} is computed by BFGS using the exact objective and gradient from multiple deterministic starting points and is used only for numerical evaluation. The convergence curves record $F(x) - F(x^*_{\mathrm{ref}})$ and $\|\nabla F(x)\|$ every 500 stochastic function evaluations up to a total budget of 10,000, using the same algorithmic parameters and random seeds across methods.

References

[1] S. Arnold, P.-A. Manzagol, R. Babanezhad Harikandeh, I. Mitliagkas, and N. Le Roux. Reducing the variance in online optimization by transporting past gradients. *Advances in Neural Information Processing Systems*, 32, 2019.

[2] A. S. Berahas, L. Cao, K. Choromanski, and K. Scheinberg. A theoretical and empirical comparison of gradient approximations in derivative-free optimization. *Foundations of Computational Mathematics*, 22(2):507–560, 2022.

[3] A. Cutkosky and F. Orabona. Momentum-based variance reduction in non-convex sgd. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[4] J. Cutler, D. Drusvyatskiy, and Z. Harchaoui. Stochastic optimization under distributional drift. *Journal of machine learning research*, 24(147):1–56, 2023.

[5] D. Drusvyatskiy and L. Xiao. Stochastic optimization with decision-dependent distributions. *Mathematics of Operations Research*, 48(2):954–998, 2023.

[6] C. Fang, C. J. Li, Z. Lin, and T. Zhang. Spider: Near-optimal non-convex optimization via stochastic path-integrated differential estimator. *Adv. Neural Inf. Process. Syst.*, 31:689–699, 2018.

[7] G. Gallego and R. Wang. Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research*, 62(2):450–461, 2014.

[8] H. Hassani, A. Karbasi, A. Mokhtari, and Z. Shen. Stochastic conditional gradient++: (non) convex minimization and continuous submodular maximization. *SIAM Journal on Optimization*, 30(4):3315–3344, 2020.

[9] Y. Hikima, H. Sawada, and A. Fujino. Guided zeroth-order methods for stochastic non-convex problems with decision-dependent distributions. In *Forty-second International Conference on Machine Learning*, 2025b.

[10] Y. Hikima and A. Takeda. Stochastic approach for price optimization problems with decision-dependent uncertainty. *European Journal of Operational Research*, 322(2):541–553, 2025.

[11] Y. Hikima and A. Takeda. Zeroth-order methods for nonconvex stochastic problems with decision-dependent distributions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 17195–17203, 2025a.

[12] Y. Hikima and A. Takeda. Zeroth-order gradient estimators for stochastic problems with decision-dependent distributions. *arXiv preprint arXiv:2510.24929*, 2025c.

[13] S. Levanon and N. Rosenfeld. Strategic classification made practical. In *International Conference on Machine Learning*, pages 6243–6253. PMLR, 2021.

[14] Q. Li and H.-T. Wai. State dependent performative prediction with stochastic approximation. In *International Conference on Artificial Intelligence and Statistics*, pages 3164–3186. PMLR, 2022.

[15] Q. Li and H.-T. Wai. Stochastic optimization schemes for performative prediction with nonconvex loss. *Advances in Neural Information Processing Systems*, 37:8673–8697, 2024.

[16] Q. Li, C.-Y. Yau, and H.-T. Wai. Multi-agent performative prediction with greedy deployment and consensus seeking agents. *Advances in Neural Information Processing Systems*, 35:38449–38460, 2022.

[17] C. Liu, Z. Wan, H. Ye, and J. Lui. Stochastic non-smooth non-convex optimization with decision-dependent distributions. *arXiv preprint arXiv:2605.06549*, 2026.

[18] H. Liu, Q. Li, and H. T. Wai. Two-timescale derivative free optimization for performative prediction with markovian data. In *International Conference on Machine Learning*, pages 31425–31450. PMLR, 2024.

[19] C. Mendler-Dünner, J. Perdomo, T. Zrnic, and M. Hardt. Stochastic optimization for performative prediction. *Advances in Neural Information Processing Systems*, 33:4929–4939, 2020.

[20] J. Perdomo, T. Zrnic, C. Mendler-Dünner, and M. Hardt. Performative prediction. In *International Conference on Machine Learning*, pages 7599–7609. PMLR, 2020.

[21] M. Ray, L. J. Ratliff, D. Drusvyatskiy, and M. Fazel. Decision-dependent risk minimization in geometrically decaying dynamic environments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 8081–8088, 2022.

[22] K. Wood and E. Dall'Anese. Stochastic saddle point problems with decision-dependent distributions. *SIAM Journal on Optimization*, 33(3):1943–1967, 2023.