Online Learning on Hidden-Convex Losses via Algorithmic Equivalence: Optimal Regret, Geometric Barrier, and Bandit Feedback

Anas Barakat[1] **Andreas Kontogiannis**[2,5] **Vasilis Pollatos**[3,5]
Ioannis Panageas[4] **Antonios Varvitsiotis**[1,5,6]

Abstract

We study adversarial online learning with hidden-convex losses, i.e., nonconvex losses that become convex after a nonlinear reparameterization. Ghai, Lu, and Hazan [17] proved that, under geometric and smoothness assumptions, online gradient descent (OGD) on such nonconvex losses approximately simulates online mirror descent (OMD) on the underlying convex losses with a suitable regularizer, yielding $\mathcal{O}(T^{2/3})$ regret. They left open whether the optimal $\Theta(\sqrt{T})$ regret from online convex optimization can be recovered in this hidden-convex setting. We answer this question affirmatively. More specifically, via a sharper discrete-time algorithmic equivalence argument, we prove that OGD achieves $\mathcal{O}(\sqrt{T})$ regret under the same assumptions, matching the optimal worst-case rate for adversarial online convex optimization. We also address another open question of [17] by clarifying the geometry required for this algorithmic equivalence. We replace the diagonal-Jacobian sufficient condition with a necessary-and-sufficient Hessian compatibility condition, thereby expanding the class of admissible reparameterizations. We complement our tight regret bound with a lower bound showing that OGD suffers $\Omega(T)$ regret even with a stepsize of order $T^{-1/2}$, for a smooth Hessian incompatible reparameterization between convex domains and an adversarial sequence of hidden-convex losses. Finally, we extend our analysis to one-point bandit feedback and prove a $\mathcal{O}(T^{3/4})$ expected regret bound for bandit OGD with spherical smoothing, matching its classical rate on convex losses.

1 Introduction

Online convex optimization (OCO) provides a theoretical framework for adversarial sequential decision-making: for convex Lipschitz losses, simple first-order methods such as online gradient descent (OGD) achieve the optimal $\Theta(\sqrt{T})$ regret rate [44, 19, 36]. Beyond convexity, the situation is substantially more delicate. For arbitrary nonconvex losses, sublinear regret against the best fixed decision in hindsight is generally unattainable by simple first-order methods without additional structure or oracle access [25, 1, 37, 22]. For instance, it is known that no deterministic algorithm can achieve sublinear regret in online nonconvex learning in general [37, Proposition 3]. This motivates the search for structured classes of nonconvex online problems that retain enough geometry to permit convex-like guarantees with simple first-order algorithms.

Affiliations: [1] Singapore University of Technology and Design [2] National Technical University of Athens [3] National and Kapodistrian University of Athens [4] University of California, Irvine [5] Archimedes, Athena Research Center, Greece [6] National University of Singapore, Centre for Quantum Technologies

Contact: barakat9anas@gmail.com, andr.kontog@gmail.com, vaspoll97@gmail.com, ipanagea@ics.uci.edu,antonios@sutd.edu.sg.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

A prominent form of benign nonconvexity is *hidden convexity* [7, 29, 12, 28, 18]. This structural property appears in several modern optimization problems, including neural-network training [39, 34, 11, 41], reinforcement learning [20, 42, 40, 5, 4, 10], and nonconvex games [38, 32, 35, 24, 16, 23, 6]. In these settings, the objective is nonconvex in the algorithm's native parameters but convex after an appropriate nonlinear change of variables. This structure has been exploited in deterministic and stochastic optimization to obtain global convergence guarantees for nonconvex problems [9, 12, 13, 8]. In the online setting, however, the learner faces a distinct challenge: the loss functions which are revealed sequentially over time may even be chosen adversarially and the goal becomes to achieve low regret against the best fixed comparator in hindsight rather than convergence to the minimizer of a fixed function.

Ghai, Lu, and Hazan [17] initiated the study of such a structured nonconvex online learning setting through algorithmic equivalence. They showed that, under suitable geometric and smoothness assumptions on the reparameterization, OGD applied to the nonconvex losses approximately simulates OMD applied to the corresponding convex losses in the reparameterized space. This connection led to a $\mathcal{O}(T^{2/3})$ regret bound. Their result left open whether the slower-than-convex rate is inherent to the nonconvex online learning setting. Since the identity reparameterization recovers ordinary OCO, the best possible regret one could hope for in this class with exact gradient feedback is $\Theta(\sqrt{T})$. This raises the central question:

When does OGD recover the optimal regret under exact gradient feedback for nonconvex hidden-convex losses, and does the same algorithmic equivalence principle extend to the bandit feedback setting?

In this paper, we answer the exact gradient feedback question affirmatively and show that the framework extends to one-point bandit feedback. Specifically, we prove a $\mathcal{O}(\sqrt{T})$ regret for OGD, discuss the required compatibility conditions needed for the algorithmic equivalence, show a linear-regret lower bound without it, and sublinear expected regret guarantees in the bandit setting. We elaborate on our contributions in more detail in the next section.

1.1 Main contributions

We study online nonconvex learning for losses that are convex under a nonlinear reparameterization which is *unknown* to the learner. Our results address several open questions formulated by Ghai, Lu, and Hazan [17], refining and extending their algorithmic equivalence analysis between OGD and OMD. Our main contributions are as follows:

- **Optimal regret under exact gradient feedback.** In the exact gradient setting, we prove that OGD achieves $\mathcal{O}(\sqrt{T})$ regret for hidden-convex losses (Theorem 1) under the geometric compatibility and smoothness assumptions considered in [17]. This improves their $\mathcal{O}(T^{2/3})$ bound and matches the optimal regret for adversarial online *convex* optimization. Since classical OCO is recovered by the identity reparameterization, the rate is optimal for this problem class in general. Our analysis is based on a sharper algorithmic equivalence between OGD and online mirror descent (OMD), improving the perturbation analysis underlying the previous $\mathcal{O}(T^{2/3})$ guarantee. Specifically, our proof shows that OGD on the nonconvex losses tracks OMD in the convex parameterization with sufficiently small discretization error to preserve the classical $\sqrt{T}$ regret.
- **Geometry of the reparameterization.** Prior work [17] imposed a diagonal-Jacobian condition to guarantee the Hessian compatibility needed for the OGD–OMD equivalence. We relax this structural assumption and give necessary and sufficient conditions for compatibility (Proposition 1), thereby enlarging the class of reparameterizations covered by the analysis.
- **Geometric barrier for OGD.** We show that the Hessian compatibility condition is not merely a proof artifact. We prove that there exists a smooth reparameterization violating compatibility and an adversarial sequence of hidden-convex losses for which OGD incurs $\Omega(T)$ regret (Theorem 2).
- **Bandit feedback.** We extend our analysis to the one-point bandit setting, where the learner only observes a single loss value per round. Against oblivious adversaries, we prove an expected regret bound of $\mathcal{O}(T^{3/4})$ under standard additional Lipschitzness and boundedness assumptions (Theorem 3). This matches the classical guarantee for OGD with one-point spherical smoothing on convex losses [14].

Together, our results establish new regret guarantees for OGD for a structured class of *online nonconvex* optimization problems, extending classical results from online *convex* optimization.

1.2 Related Work

Online nonconvex optimization. Online convex optimization is a standard framework for adversarial sequential decision-making; see, e.g., [19, 36, 33]. A separate line of work studies online learning with nonconvex losses [25, 1, 37, 22]. Sublinear global regret in such settings typically requires access to strong oracles, such as sampling oracles that allow continuous-domain exponential weights [31, 25], or offline optimization oracles for the nonconvex losses [1, 37]. Our work takes a different route: rather than assuming oracle access for general nonconvex losses, we exploit hidden convexity of the loss sequence and analyze first-order algorithms through the algorithmic equivalence between OGD in the nonconvex parameterization and OMD in the convex parameterization [2, 17, 30]. A few works in the literature also consider structured nonconvex losses with special structure such as the composition of a non-increasing function with a linear function [43], or weaker notions of convexity such as strictly locally quasi-convexity [21] and weak pseudo-convexity [15].

Hidden-convex optimization. Hidden convexity has recently been studied in several offline deterministic and stochastic optimization settings [9, 12, 13, 28]. These works consider nonconvex problems that admit a convex reformulation under a nonlinear transformation, often under assumptions weaker than those used in the online algorithmic-equivalence framework. Their focus, however, is offline optimization: the objective is fixed, or arises as an expectation over a fixed data-generating distribution. By contrast, we study an adversarial online setting in which the loss functions may change arbitrarily over time, and the performance criterion is regret rather than convergence to a global optimizer. This distinction requires a different analysis, since the learner must compete with the best fixed decision in hindsight while only observing feedback sequentially.

Online learning under bandit feedback. The literature on online bandit learning is extensive; we refer to [27, 26] for modern treatments. For bandit online convex optimization, Flaxman, Kalai and McMahan [14] introduced a one-point gradient estimator based on spherical smoothing, showing that gradient-based online methods can be implemented when only function values are observed. The bandit setting has also been studied for general Lipschitz nonconvex losses and for special nonconvex classes [43, 15]. Our bandit results are different in focus: we consider hidden-convex losses and use smoothing together with the OGD–OMD equivalence to transfer regret guarantees from the convex parameterization to the original nonconvex one. Thus, while our algorithmic template is inspired by bandit OCO, the main challenge is controlling the interaction between bandit gradient estimation, nonlinear reparameterization, and the approximation error in algorithmic equivalence.

2 Online Learning on Hidden-Convex Losses

In this section, we introduce the online *hidden*-convex optimization (OHCO) problem which extends the celebrated online convex optimization (OCO) setting [36, 21, 33].

Hidden-convex losses. Let $\mathcal{X}, \mathcal{Y} \subset \mathbb{R}^d$ be compact convex sets with nonempty interiors. Consider a sequence of loss functions (ℓ_t) where for any $t \geq 1$, $\ell_t : \mathcal{X} \to \mathbb{R}$ has the following hidden-convexity structure:

$$\ell_t(x) = h_t(q(x)), \quad \forall x \in \mathcal{X}, \tag{1}$$

where $q : \mathcal{X} \to \mathcal{Y}$ is a smooth bijective reparameterization function and $h_t : \mathcal{Y} \to \mathbb{R}$ is a convex function. In this work, we will focus on a subclass of hidden-convex functions satisfying some geometric and smoothness assumptions made in [17]. We defer the statement of these assumptions on q and (ℓ_t) to Sections 3-5. While ℓ_t is nonconvex in general, it is convex under the reparameterization q and we then say that the function ℓ_t is *hidden-convex*. When the reparameterization q is simply the identity, the function ℓ_t is a (truly) convex function.

Online regret minimization. We are interested in the online optimization setting where a learner seeks to minimize an unknown sequence of hidden-convex functions (ℓ_t) by choosing a sequence (x_t) where each x_t is based only on the information observed before round t. At each round t, a learner chooses $x_t \in \mathcal{X}$ using an online algorithm $\mathcal{A}$, then an adversary reveals a loss function ℓ_t which is hidden-convex (see (1)) and the player suffers the loss $\ell_t(x_t)$. The goal of the learner is to achieve a good performance with respect to the best possible fixed decision in hindsight, minimizing

the (external) regret of the algorithm $\mathcal{A}$ defined for any time horizon $T \geq 1$ as follows:

$$R_T := \sum_{t=1}^{T} \ell_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(x) \,.$$

Since $q : \mathcal{X} \to \mathcal{Y}$ is bijective, the regret can equivalently be written in hidden coordinates as $R_T = \sum_{t=1}^{T} h_t(z_t) - \min_{z \in \mathcal{Y}} \sum_{t=1}^{T} h_t(z)$ where $z_t = q(x_t)$. Unlike in the classical OCO setting, the sequence (ℓ_t) is not supposed to be convex. We recover the standard OCO setting by choosing q to be the identity function in (1).

Information settings. The reparameterization function q is unknown to the learner who cannot run an online algorithm on the sequence of losses (h_t) and recover a decision in the original space $\mathcal{X}$ by inverting q. Preconditioning by the Jacobian of q (which is unknown) is not possible either. We consider two different settings depending on the feedback available to the learner:

(i) *Exact Gradient Feedback:* In this setting, we suppose that the loss function ℓ_t is differentiable and the learner has access to the gradient $\nabla \ell_t(x_t)$ of the loss function ℓ_t at each time step of the online interaction.

(ii) *Bandit Feedback:* In this setting, the learner has only access to a single function value call to the loss function ℓ_t at each round.

Online Gradient Descent. In this work, we will focus on the most famous and classical algorithm to address this problem: Online Gradient Descent (OGD) [44]. In the exact gradient feedback setting, the learner performs the following update for all time steps t:

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \eta \nabla \ell_t(x_t)) \,, \tag{OGD}$$

where η is a positive step size and $\Pi_{\mathcal{X}}$ is the Euclidean projection over the convex compact set $\mathcal{X}$. In the bandit setting, the learner will use the one-point estimate of the gradient resulting in the bandit gradient descent algorithm proposed by [14]. We will discuss this case in more detail in Section 5.

We conclude this section by introducing some general notation that will be used throughout the paper.

Additional notation. For a nonzero integer d, we denote by $[d] := \{1, \dots, d\}$. For any function $f : \mathbb{R}^d \to \mathbb{R}^d$, we denote by $J_f(x) \in \mathbb{R}^{d \times d}$ the Jacobian of f at x. Given a strictly convex and differentiable function $R : \mathbb{R}^d \to \mathbb{R}$, the Bregman divergence induced by R is defined as $D_R(x \| y) := R(x) - R(y) - \nabla R(y)^\top (x - y)$. For a closed convex set $\mathcal{K}$ and a strictly convex regularizer R, we use $\Pi_{\mathcal{K}}^R(x) := \arg\min_{y \in \mathcal{K}} D_R(y \| x)$ to denote the Bregman projection, and we use the shorthand notation $\Pi_{\mathcal{K}} := \Pi_{\mathcal{K}}^{\|\cdot\|^2}$ for the Euclidean projection. Given a positive-definite matrix $M \in \mathbb{R}^{d \times d}$, we define the norm $\|x\|_M := \sqrt{x^\top M x}$ and we denote by $M^{-\top}$ the inverse transpose matrix. We use the notation $B_p := \{x \in \mathbb{R}^d : \|x\|_p \leq 1\}$ for an ℓ_p ball and B_p^+ to denote the intersection of the ℓ_p ball and the positive orthant. Differentiability on a compact domain is understood through extensions of the stated regularity to an open neighborhood. Bounds on derivatives are in the induced Euclidean operator norms.

3 Regret Analysis of OGD for OHCO under Exact Gradient Feedback

In this section, we establish a $\mathcal{O}(\sqrt{T})$ regret bound for OGD in the OHCO setting described in Section 2 under exact gradient feedback, i.e. when the learner has access to gradients of the loss functions ℓ_t. We start by introducing our main assumptions in Section 3.1 before stating our first main result in Section 3.2 and providing a proof sketch in Section 3.3.

3.1 Assumptions

We consider the same set of assumptions on the reparameterization and the loss functions as in [17].

Assumption 1 (Hessian-compatible reparameterization)**.** *The map $q : \mathcal{X} \to \mathcal{Y}$ is a smooth bijection onto the convex set $\mathcal{Y}$ and there exists a twice continuously differentiable, strictly convex regularizer $R : \mathcal{Y} \to \mathbb{R}$ such that, for all $x \in \mathcal{X}$, with $z = q(x)$, $\left[\nabla^2 R(z)\right]^{-1} = J_q(x) J_q(x)^\top$.*

Assumption 1 is a crucial assumption which restricts the class of parameterization functions q. Examples of such parameterizations include the quadratic reparameterization with the entropy regularizer, the exponential reparameterization with the log-barrier regularizer and the power reparameterization with tempered Bregman divergences [3]. We refer the reader to Section 3.1 in [17] for a more detailed discussion on these examples. We will elaborate more on Assumption 1 and its usefulness in Section 3.3 when discussing the proof of our main results in the exact gradient feedback setting, and in Section 4 where we will discuss both verifiable necessary and sufficient conditions under which it is satisfied, and its necessity.

Assumption 2 (Regularizer and reparameterization smoothness)**.** *There exists $G > 1$ such that:*

(i) The map q is G-Lipschitz on $\mathcal{X}$, and the 1st and 2nd derivatives of q^{-1} are bounded by G on $\mathcal{Y}$.

(ii) The regularizer R is 1-strongly convex and smooth with 1st and 3rd derivatives bounded by G.

(iii) For every $z \in \mathcal{Y}$, the map $w \mapsto D_R(z\|w)$ is G-Lipschitz on $\mathcal{Y}$.

Our last assumption asks that the loss functions ℓ_t have uniformly bounded gradients and that the diameter of $\mathcal{X}$ is bounded. These assumptions are also standard in OCO.

Assumption 3 (Boundedness of gradients and diameter)**.** *There exists $G_F > 0$ s.t. $\|\nabla \ell_t(x)\|_2 \leq G_F$ for all $x \in \mathcal{X}$. Moreover, there exists $D_1 > 0$ such that $\sup_{z,z' \in \mathcal{Y}} D_R(z\|z') \leq D_1$.*

3.2 Regret Bound

We are now ready to state the main result of this section.

Theorem 1 (OGD regret under full-information feedback)**.** *Let Assumptions 1-3 hold and let $T \geq 1$. Define $A := G_F^2 \left(5G^6 + \frac{1}{2}G^2\right)$. Setting $\eta = \sqrt{\frac{D_1}{AT}}$, the regret of (OGD) on the sequence of hidden-convex losses (ℓ_t) defined in (1) using stepsize η is bounded as follows:*

$$R_T \leq 2\sqrt{AD_1T} + \frac{G^6 G_F^3 D_1}{A} = \mathcal{O}(\sqrt{T}).$$

This regret bound improves over the $\mathcal{O}(T^{2/3})$ regret bound established in [17]. The leading $\sqrt{T}$ coefficient is $O(G^3 G_F)$, where G_F bounds the original loss gradients. For comparison, the bound in [17] has dependence $O(G^{10/3} G_{F,h})$, where $G_{F,h}$ bounds the hidden loss gradients (under our assumptions, $G_{F,h} \leq G G_F$).

3.3 Proof Sketch of Theorem 1

In this section, we provide a proof sketch of Theorem 1. We provide a complete proof in Section A. The main strategy of the proof is inspired by [17] which shows that OGD applied to non-convex functions is an approximation of online mirror descent applied to convex functions under reparameterization. We briefly recall the intuition behind this idea for the convenience of the reader to motivate Assumption 1 before discussing our new key technical steps driving our tighter analysis.

Implicit OMD via OGD reparameterization. In continuous time, Amid and Warmuth [2] showed that the mirror-flow trajectory in the hidden space and the gradient-flow trajectory in the original space coincide exactly under the reparameterization. More precisely, if $z(t) = q(x(t))$, they showed that both trajectories of the following ordinary differential equations:

$$\frac{d}{dt}\nabla R(z(t)) = -\nabla h(z(t))\,, \quad \frac{dx(t)}{dt} = -\nabla(h \circ q)(x(t)) \tag{2}$$

coincide if the mirror descent regularizer R and the reparameterization function q satisfy $[\nabla^2 R(q(x))]^{-1} = J_q(x) J_q(x)^\top$. In discrete time, OGD iterates on the losses (ℓ_t) are given by (OGD) and Online Mirror Descent (OMD) iterates on the convex losses (h_t) are defined as follows:

$$y_{t+1} = \arg\min_{y \in \mathcal{Y}} \left\{ \langle \nabla h_t(y_t), y - y_t \rangle + \frac{1}{\eta} D_R(y\|y_t) \right\}, \tag{OMD}$$

with $y_0 = q(x_0)$. In discrete time, the exact equivalence in the continuous dynamics in (2) no longer holds because of the discretization error. Starting from a coupled state $y_t = q(x_t)$, one step of OGD

on (ℓ_t) in the original space produces x_{t+1} and hence $z_{t+1} := q(x_{t+1})$, while one step of OMD on (h_t) in the hidden space produces y_{t+1}. Ghai, Lu, and Hazan [17] showed that these two hidden-space iterates are $O(\eta^{3/2})$-close. Throughout the coupling argument, the ghost OMD step is initialized anew at $z_t = q(x_t)$ on each round; it is not a separately evolving OMD trajectory.

Our key improvement. To obtain Theorem 1, our key contribution is to show a sharper $O(\eta^2)$ estimate of the distance between two hidden-space iterates, which is small enough to preserve the classical $\sqrt{T}$ regret rate after applying a perturbed-OMD regret bound.

Lemma 1 (One-step OGD-OMD coupling)**.** *Suppose Assumptions 1-3 hold. If $y_t = z_t = q(x_t)$, x_{t+1} is computed via one step of* (OGD)*, and y_{t+1} via one step of* (OMD)*, then*

$$\|z_{t+1} - y_{t+1}\|_2 \le C_\eta := G^5(5G_F^2 + G_F^3\eta)\eta^2.$$

The proof of Lemma 1 is the main new step. Starting from the coupled state $y_t = z_t = q(x_t)$, we first localize both y_{t+1} and z_{t+1} in an $O(\eta)$-neighborhood $\mathcal{Y}_t$ of y_t. On this local set, both updates can be expressed as minimizers of the same strongly convex quadratic model, up to different higher-order perturbations:

$$y_{t+1} = \arg\min_{y\in\mathcal{Y}_t} \left\{\Phi_t(y) + \varepsilon_t^{\mathrm{OMD}}(y)\right\}, \qquad z_{t+1} = \arg\min_{y\in\mathcal{Y}_t} \left\{\Phi_t(y) + \varepsilon_t^{\mathrm{OGD}}(y)\right\},$$

$$\Phi_t(y) := \langle \nabla h_t(y_t), y - y_t\rangle + \frac{1}{2\eta}\|y - y_t\|^2_{\nabla^2 R(y_t)}.$$

Here $\varepsilon_t^{\mathrm{OMD}}$ is the Taylor remainder obtained by replacing $D_R(y\|y_t)$ with its second-order expansion at y_t, while $\varepsilon_t^{\mathrm{OGD}}$ is the Taylor remainder obtained after changing variables $y = q(x)$ in the OGD update and expanding q^{-1} at y_t (see Lemma 2 and Lemma 3 in App. A.1 for their explicit definitions). The leading quadratic terms agree exactly because Assumption 1 identifies the OGD Euclidean metric in x-space with the mirror metric $\nabla^2 R(y_t)$ in y-space.

The crucial departure from [17] is how we compare these two perturbed minimization problems. Rather than controlling the sup-norm difference between the perturbed objectives and then translating this into a distance between minimizers, we compare the first-order optimality conditions directly. Since Φ_t is $1/\eta$-strongly convex and the normal cone of $\mathcal{Y}_t$ is monotone, Lemma 4 in App. A.1 gives

$$\|y_{t+1} - z_{t+1}\| \le \eta \left\|\nabla\varepsilon_t^{\mathrm{OGD}}(z_{t+1}) - \nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1})\right\|.$$

Since both iterates lie in an $\mathcal{O}(\eta)$-neighborhood of y_t, we prove that $\|\nabla\varepsilon_t^{\mathrm{OGD}}(z_{t+1})\| = \mathcal{O}(\eta)$ and $\|\nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1})\| = \mathcal{O}(\eta)$ (in Lemma 6-7 in App. A.1). Hence $\|y_{t+1} - z_{t+1}\| = \mathcal{O}(\eta^2)$. This sharper one-step coupling improves the $\mathcal{O}(\eta^{3/2})$ discrepancy used in [17] to $\mathcal{O}(\eta^2)$. Finally, the mapped OGD iterates $z_t = q(x_t)$ can be viewed as perturbed OMD iterates, with each ghost step initialized at the current point z_t. The perturbation errors $r_{t+1} = z_{t+1} - y_{t+1}$ satisfy $\|r_{t+1}\|_2 \le C_\eta := G^5(5G_F^2 + G_F^3\eta)\eta^2$. Applying the perturbed-OMD regret bound (Lemma 8) yields:

$$R_T \le \frac{D_1}{\eta} + \frac{\eta G^2 G_F^2 T}{2} + \frac{GC_\eta T}{\eta} = \frac{D_1}{\eta} + A\eta T + G^6 G_F^3 \eta^2 T.$$

The choice $\eta = \sqrt{D_1/(AT)}$ gives Theorem 1. See App. A for the full proof.

4 Discussion of the Hessian-Compatibility Hidden Geometry Assumption 1

Our analysis relies on Assumption 1 which allows to obtain an implicit OMD reparameterization and exploit the connection with OMD analysis over convex functions to obtain our tight regret analysis. In this section, we discuss this assumption in more depth. First, we answer an open question in [17] regarding whether we can relax their sufficient diagonal Jacobian condition (Assumption 4 therein) to verify Assumption 1. In section 4.1, we provide necessary and sufficient conditions under which Assumption 1 is satisfied. Second, we address the question of the necessity of this Hessian-compatibility hidden-geometry assumption, which is also mentioned as an open question in [17]. In Section 4.2, we show that Assumption 1 cannot in general be removed from a uniform sublinear-regret guarantee for Euclidean OGD.

4.1 Relaxation of the diagonal Jacobian assumption in [17]

Define the inverse metric field $M : \mathcal{Y} \to \mathbb{R}^{d\times d}$ by

$$M(z) := \left[J_q(q^{-1}(z))J_q(q^{-1}(z))^\top\right]^{-1}. \tag{3}$$

Recall that Assumption 1 states that there exists a scalar function $R : \mathcal{Y} \to \mathbb{R}$ s.t. $M(z) = \nabla^2 R(z)$, for all $z \in \mathcal{Y}$. Theorem 7 in [17] showed that if J_q is diagonal then there exists such R. They used this diagonal separable structure to construct a regularizer R explicitly as a solution to second-order ODEs with coefficients depending on the parameterization function q and its derivatives. In the following result, we relax this condition and provide more general conditions under which the matrix field M is a Hessian field.

Proposition 1. *Let $q : \mathcal{X} \to \mathcal{Y}$ be a C^3 diffeomorphism, with $\mathcal{Y}$ convex, and let M be defined by (3). Then there exists a C^3 scalar function R such that $M = \nabla^2 R$ if and only if*

$$\partial_{z_k} M_{ij}(z) = \partial_{z_j} M_{ik}(z), \qquad \forall i, j, k \in [d], \quad z \in \mathcal{Y}.$$

Such R is unique up to an affine function. Moreover, since $M(z) \succ 0$, R is strictly convex; on the compact domain $\mathcal{Y}$ of our OHCO setting, R is uniformly strongly convex.

This result follows as a consequence of the fundamental result from differential analysis stating that a vector field with matching cross-partial coordinate function derivatives is a conservative field, i.e., a gradient field deriving from a scalar potential. See Proposition 3 for a precise statement of that result and a complete proof of Proposition 1.

We provide two examples where the Jacobian of the parameterization q is not diagonal in general, yet the matrix field M defined in (3) satisfies the cross-derivatives conditions $\partial_{z_k} M_{ij} = \partial_{z_j} M_{ik}$ of Proposition 1. Moreover, the compatibility Assumption 1 is satisfied with explicit regularizers R. Detailed derivations for both examples can be found in App. B.2.

Example 1 (Affine mixing of a separable parameterization)**.** Let $q(x) = As(x) + b$ and $s(x) = \big(s_1(x_1), \dots, s_d(x_d)\big)$ where $A \in \mathbb{R}^{d\times d}$ is invertible, $b \in \mathbb{R}^d$, and each s_i is a C^2 diffeomorphism from an open interval onto its image. Then $J_q(x) = A\,D(x)$ with $D(x) := \operatorname{diag}\big(s_1'(x_1), \dots, s_d'(x_d)\big)$. Hence $M(z) = A^{-T}D(x)^{-2}A^{-1}, z = q(x)$. Moreover, M is explicitly the Hessian of the following regularizer: $R(z) := \sum_{i=1}^d r_i\big((A^{-1}(z-b))_i\big)$ and $r_i''(t) = \big(s_i'(s_i^{-1}(t))\big)^{-2}$. For example, if $s_i(u_i) = \exp(u_i), i \in [d]$, $r_i(t) = -\log t$ satisfies the conditions.

Example 2 (Rank-one nonlinear mixing)**.** Let $a \in \mathbb{R}^d$ and let $q(x) = x + h(a^\top x)\,a$ where $h : \mathbb{R} \to \mathbb{R}$ is C^2. We have $J_q(x) = I + h'(a^\top x)\,aa^\top$. This Jacobian is typically non-diagonal whenever at least two components of a are nonzero.

Both examples also satisfy Assumption 2 on suitable compact domains; we give the corresponding domain conditions and verification in App. B.2.

These examples enlarge the class of reparameterizations certified in the coordinates in which the problem is presented, but we do not claim that they are intrinsically nonseparable. In particular, the affine mixing in Example 1 can sometimes be absorbed into the hidden loss, while Example 2 becomes coordinatewise after a rotation aligning a with a coordinate axis. Identifying standard applications that are provably irreducibly nonseparable remains open.

4.2 OGD can incur linear regret without Hessian-compatible hidden geometry

In this section, we discuss how crucial is Assumption 1 to obtain sublinear regret for OGD in our nonconvex online learning setting. Perhaps surprisingly, we show that OGD can incur linear regret if the Hessian compatibility assumption does not hold; i.e., if the metric induced by the symmetric positive definite matrix M in (3) does not derive from a Hessian metric such that $M(z) = \nabla^2 R(z)$ for some regularization function R.

Theorem 2 (Linear regret without Hessian-compatible geometry)**.** *There exist a smooth diffeomorphism $q : \mathcal{X} \to \mathcal{Y}$ with $\mathcal{X} = \mathcal{Y} = [0,1]^2$ and $J_q(x)J_q(x)^\top \succeq \mu I$ for some $\mu > 0$, whose inverse metric field is not the Hessian of any regularizer, an initialization $x_1 \in \operatorname{int} \mathcal{X}$, and constants $c, C, S, \eta_0 > 0$ with the following property. For every $T \geq 1$ and every constant stepsize $0 < \eta \leq \eta_0$, there exists an adaptive sequence of linear hidden losses $h_t(z) = \langle s_t, z\rangle$, with $\|s_t\|_2 \leq S$, such that projected OGD on $\ell_t = h_t \circ q$ satisfies $R_T \geq cT - \frac{C}{\eta}$. All constants, q, and x_1 are independent of T and η. In particular, choosing $\eta = \Theta(T^{-1/2})$, held constant during the run, gives $R_T = \Omega(T)$ for sufficiently large T.*

The theorem shows that smoothness and uniform positive definiteness of the induced hidden geometry are not sufficient for sublinear regret of OGD. This means that Assumption 1 is not merely a technical condition for our analysis of OGD in Theorem 1. The obstruction is geometric: when the metric induced by the reparameterization is not Hessian-compatible, OGD may fail to achieve sublinear regret even when the reparameterization is smooth and its induced metric is uniformly positive definite. The hidden-space dynamics of OGD need not arise from an integrable potential. An adversary can then exploit this fact by forcing the iterates to cycle around a loop with nonzero circulation and force linear regret. We elaborate more on this in the next proof sketch.

Proof sketch of Theorem 2. We provide a proof sketch to highlight the intuitions. The full proof is provided in App. B.3. The construction uses the two-dimensional reparameterization $q(x_1, x_2) = \big(x_1,\, x_2 - x_1(1-x_1)x_2(1-x_2)\big), \mathcal{X} = \mathcal{Y} = [0,1]^2$. For each fixed x_1, the second coordinate is strictly increasing in x_2 and maps $[0,1]$ onto itself. Thus q is a smooth diffeomorphism between convex compact domains, with a uniformly positive definite induced metric. Its inverse metric M, defined in (3), is not Hessian-compatible, as reflected in the nonzero circulation below.

We consider linear hidden losses $h_t(z) = \langle s_t, z\rangle$ and original losses $\ell_t(x) = h_t(q(x))$, where the adversary chooses s_t. Writing $z_t = q(x_t)$, we have $\nabla \ell_t(x_t) = J_q(x_t)^\top s_t$. Whenever projection is inactive, a Taylor expansion of q shows that OGD in the original coordinates induces the hidden-space dynamics

$$z_{t+1} = z_t - \eta M(z_t)^{-1} s_t + O(\eta^2), \tag{4}$$

for uniformly bounded s_t. This suggests choosing $s_t = -M(z_t)v_t$ to cancel the induced metric and move the hidden iterate approximately in a prescribed direction v_t: $z_{t+1} = z_t + \eta v_t + O(\eta^2)$. However, these local errors need not cancel after a complete circuit. To repeat the same circuit without accumulating drift, we implement this steering idea exactly rather than only to first order.

Specifically, suppose that the desired next hidden point is $z_t^+ = z_t + \Delta v_t \in \operatorname{int} \mathcal{Y}$, where $v_t \in \{\pm e_1, \pm e_2\}$ and $0 < \Delta \leq \eta$. The corresponding target in the original space is $q^{-1}(z_t^+)$. We therefore choose s_t by solving $x_t - \eta J_q(x_t)^\top s_t = q^{-1}(z_t^+)$, which gives $s_t = \frac{1}{\eta} J_q(x_t)^{-\top}\big[x_t - q^{-1}(z_t^+)\big]$. With this choice, the OGD pre-projection point is exactly $q^{-1}(z_t^+) \in \operatorname{int} \mathcal{X}$. Hence projection is inactive and $z_{t+1} = z_t + \Delta v_t$ holds exactly. Only the adversary uses q; the learner still performs ordinary OGD. Moreover, the vectors s_t are uniformly bounded because $\Delta \leq \eta$ and q^{-1} has bounded derivatives on $\mathcal{Y}$. This exact choice retains the same inverse-metric cancellation as its first-order counterpart. Indeed, Taylor expansion of q^{-1} gives $s_t = -\frac{\Delta}{\eta} M(z_t)v_t + O(\Delta^2/\eta)$. Fixing a comparator $u \in \mathcal{X}$ and writing $z_u = q(u)$, define $F_u(z) := M(z)(z - z_u)$. The instantaneous regret against u can then be written as

$$\ell_t(x_t) - \ell_t(u) = \langle s_t, z_t - z_u\rangle = -\frac{1}{\eta}\langle F_u(z_t), z_{t+1} - z_t\rangle + O(\Delta^2/\eta).$$

Hence cumulative regret along the prescribed trajectory is, up to controlled error terms, a rescaled discrete line integral of the vector field F_u. We now choose a closed path along which this integral is negative. For $u = (0, 1/2)$, we have $z_u = (0, 1/2)$, and a direct calculation at $z^0 = q(1/2, 1/2) = (1/2, 7/16)$ gives $\partial_{z_1}(F_u)_2(z^0) - \partial_{z_2}(F_u)_1(z^0) = -\frac{1}{4}$. By continuity, there is a small square $\mathcal{Q} \subset \operatorname{int} \mathcal{Y}$ on which the curl is uniformly negative and Green's theorem therefore gives a negative counterclockwise circulation $I_{\mathcal{Q}} := \oint_{\partial \mathcal{Q}} F_u(z) \cdot dz < 0$.

Let L be the side length of $\mathcal{Q}$. Set $n = \lceil L/\eta \rceil$ and $\Delta = L/n$, initialize z_1 at its lower-left corner, and prescribe n steps in each of the directions $e_1, e_2, -e_1, -e_2$, in that order. The hidden iterates then traverse the boundary counterclockwise and return exactly to the starting point after $N = 4n = \Theta(1/\eta)$ rounds. The Riemann-sum error is $O(\Delta)$, and the cumulative Taylor error in the

regret is $O(N\Delta^2/\eta) = O(1)$. Consequently, one complete cycle satisfies

$$\sum_{t=1}^{N} \big(\ell_t(x_t) - \ell_t(u)\big) = -\frac{I_{\mathcal{Q}}}{\eta} + O(1) = \Omega(1/\eta)$$

for sufficiently small η. Thus each cycle contributes a positive constant amount of regret per round, with constants independent of η. Because the trajectory closes exactly, the same cycle can be repeated without drift. Completing as many cycles as fit within the horizon T and setting all remaining losses to zero yields $R_T \geq cT - C/\eta$ for constants $c, C > 0$ independent of T and η, since external regret is at least regret against the fixed comparator u. In particular, a constant stepsize within the run of order $\eta = \Theta(T^{-1/2})$ gives $R_T = \Omega(T)$.

Remark 1. Theorem 2 is an existential lower bound for Euclidean OGD: it does not assert that every incompatible reparameterization leads to linear regret. More generally, the same exact steering construction applies to a smooth diffeomorphism between compact convex domains in $\mathbb{R}^2$ whenever, for some comparator u, the field $M(z)(z - q(u))$ has nonzero curl at an interior point. A sufficiently small square and an appropriate orientation then give a negative circulation.

Open question. The lower bound of Theorem 2 only holds for OGD. Whether Assumption 1 is necessary to obtain sublinear regret in general, for *any* online algorithm, remains open.

5 Regret Analysis of OGD for OHCO under Bandit Feedback

5.1 Expected Regret Bound in the Bandit Setting

We now consider the one-point bandit feedback, where the learner observes only one function value per round. We analyze the classical Bandit OGD (BOGD) algorithm with spherical smoothing (Algorithm 1, [14]) in the hidden-convex setting, against an oblivious adversary. Let U_{sp} denote the uniform distribution over the Euclidean unit sphere. For an exploration radius $\delta \in (0, 1)$, the algorithm maintains its iterates in the shrunken domain $\mathcal{X}_\delta := (1 - \delta)\mathcal{X}$, so that the perturbed query points remain feasible (under the standard normalization $0 \in \mathcal{X}$ and $B_2 \subseteq \mathcal{X}$ as in [14]). We denote by $\mathcal{Y}_\delta := q(\mathcal{X}_\delta)$ the corresponding hidden-space domain, which is only used in the analysis to compare the mapped BOGD iterates with mirror descent on the convex losses.

Before stating the result, we introduce a few extra standard assumptions on the smoothness and boundedness of the losses (ℓ_t), and the diameter of $\mathcal{X}$.

Assumption 4 (Loss boundedness)**.** *There exists $M > 0$ s.t. $|\ell_t(x)| \leq M$, for all $x \in \mathcal{X}, t \geq 1$.*

Assumption 5 (Smoothness)**.** *For any t, the loss function ℓ_t is smooth, i.e., there exists a constant $H > 0$, such that for any $x, x' \in \mathcal{X}$ and any $t \geq 1$, $\|\nabla\ell_t(x) - \nabla\ell_t(x')\| \leq H\|x - x'\|$.*

Algorithm 1 Bandit OGD (BOGD) [14]

1: Input: $x_1 \in \mathcal{X}_\delta$
2: **for** $t = 1, \ldots, T$ **do**
3: Sample a random direction $\zeta_t \sim U_{sp}$
4: Choose $\hat{x}_t = x_t + \delta\zeta_t$
5: Observe the bandit loss $\ell_t(\hat{x}_t)$
6: Construct $g_t = \frac{d}{\delta}\ell_t(\hat{x}_t)\zeta_t$
7: Update $x_{t+1} = \Pi_{\mathcal{X}_\delta}(x_t - \eta g_t)$
8: **end for**

Assumption 6 (Feasible set regularity)**.** *There exists $D > 0$ s.t. $\|x\|_2 \leq D, \|x - x'\|_2 \leq D$ for all $x, x' \in \mathcal{X}$. Moreover, $0 \in \mathcal{X}$, $B_2 \subseteq \mathcal{X}$, and the hidden shrunken domain $\mathcal{Y}_\delta := q((1 - \delta)\mathcal{X})$ is convex for the exploration radius $\delta \in (0, 1)$ used by Algorithm 1.*

In the following theorem, we present the no-regret guarantees of Bandit OGD.

Theorem 3 (Expected regret under bandit feedback)**.** *Fix an oblivious sequence of losses and a horizon $T \geq 1$. Let Assumptions 1–6 hold for the exploration radius specified below. Let Assumptions 1 to 6 hold. Let $K_1 := G_F D + G^2 H D + G_F, K_2 := 6d^2G^6M^2, K_3 := d^3G^6M^3$. Then, the expected regret of BOGD (Algorithm 1) is bounded as follows:*

$$\mathbb{E}[R_T] \leq C_{\text{band}}\, T^{3/4}, \quad C_{\text{band}} := 2\sqrt{2}\,(K_1^2 K_2 D_1)^{1/4} + K_1^{1/2} K_2^{-5/4} K_3 D_1^{3/4}/\sqrt{2}\,,$$

when using the stepsize $\eta = \left(\frac{4D_1^3}{K_1^2 K_2 T^3}\right)^{1/4}$, and the radius $\delta = \left(\frac{4K_2 D_1}{K_1^2 T}\right)^{1/4}$ with $\delta \in (0, 1)$.

Thus, under the stated assumptions, BOGD recovers the classical $\mathcal{O}(T^{3/4})$ guarantee of one-point spherical-smoothing OGD for convex losses [14].

5.2 Proof Sketch of Theorem 3

We give a proof sketch of Theorem 3; the complete proof is in App. C. The proof follows the same high-level strategy as in the exact gradient setting with a few similar steps to the analysis of OBGD in OCO: after mapping the BOGD iterates to the hidden space, $z_t = q(x_t)$, we compare them to mirror-descent steps on the convex losses h_t. The bandit setting introduces three additional errors: the domain must be shrunk to allow perturbations, the one-point estimator is biased for the true gradient, and the algorithm suffers the loss at the perturbed point $\hat{x}_t = x_t + \delta\zeta_t$, rather than at x_t. To define the ghost OMD iterate in the hidden space (which is only used in the analysis and never computed), we introduce the hidden-space gradient estimator $\tilde{g}_t := J_q(x_t)^{-\top} g_t$ which is the natural hidden-space analogue of the bandit estimator g_t by the chain rule: $\nabla h_t(z_t) = J_q(x_t)^{-\top} \nabla \ell_t(x_t), z_t = q(x_t)$. Note that it is not computed by the algorithm but introduced only to analyze the mapped BOGD trajectory as a perturbed OMD trajectory in $\mathcal{Y}_\delta = q(\mathcal{X}_\delta)$. Since $|\ell_t| \le M$ and $\|\zeta_t\| = 1$, we have $\|g_t\| \le dM/\delta$. Let $z^\star \in \arg\min_{z\in\mathcal{Y}} \sum_{t=1}^T h_t(z)$ and $z_\delta^\star \in \arg\min_{z\in\mathcal{Y}_\delta} \sum_{t=1}^T h_t(z)$ be two comparators, and let $\hat{z}_t = q(\hat{x}_t)$. The starting point is the following regret decomposition:

$$\mathbb{E}[R_T] \le \underbrace{\mathbb{E}\left[\sum_{t=1}^T \langle \tilde{g}_t, z_t - z_\delta^\star\rangle\right]}_{R_T^{(1)}\text{: perturbed OMD regret for random linear losses}} + \underbrace{\sum_{t=1}^T h_t(z_\delta^\star) - h_t(z^\star)}_{R_T^{(2)}\text{: domain shrinkage error}} + \underbrace{\mathbb{E}\left[\sum_{t=1}^T B_t\right]}_{R_T^{(3)}\text{: gradient estimation bias}} + \underbrace{\mathbb{E}\left[\sum_{t=1}^T h_t(\hat{z}_t) - h_t(z_t)\right]}_{R_T^{(4)}\text{: smoothing error}},$$

where $B_t := \langle b_t, z_\delta^\star - z_t\rangle$ and $b_t := \mathbb{E}[\tilde{g}_t \mid \mathcal{F}_{t-1}] - \nabla h_t(z_t)$. The four terms isolate the distinct sources of error in the bandit setting. The shrinkage term is controlled by comparing optimization over $\mathcal{X}$ and $\mathcal{X}_\delta = (1-\delta)\mathcal{X}$, giving $R_T^{(2)} \le \delta G_F D T$. The smoothing term is controlled directly by Lipschitzness of ℓ_t: $|R_T^{(4)}| \le \delta G_F T$. The bias term comes from the fact that the one-point spherical estimator is unbiased for the gradient of the ball-smoothed loss, not for $\nabla\ell_t(x_t)$. Using smoothness of ℓ_t and the chain-rule transformation above yields $\|b_t\| \le \delta G H$ and $\|z_t - z_\delta^\star\| \le GD$, and hence $|R_T^{(3)}| \le \delta G^2 H D T$. It remains to control $R_T^{(1)}$, which is the term most closely related to the exact gradient setting proof. For each round, define the ghost OMD iterate, only used in the analysis:

$$y_{t+1} = \arg\min_{y\in\mathcal{Y}_\delta} \left\{ \langle \tilde{g}_t, y - z_t\rangle + \frac{1}{\eta} D_R(y\|z_t) \right\}.$$

The actual mapped BOGD iterate satisfies $z_{t+1} = q(x_{t+1}) = y_{t+1} + r_{t+1}$. As in the exact gradient feedback setting, the OGD–OMD one-step coupling applies pathwise because the proof only requires a bounded vector g_t, not an exact gradient. With $\hat{G}_F = dM/\delta$, Corollary 5 gives $\|r_{t+1}\| \le \hat{C}_\eta := G^5\left(5\hat{G}_F^2 + \hat{G}_F^3\eta\right)\eta^2$. As in the exact-gradient setting, the perturbed OMD bound for the random linear losses $\langle \tilde{g}_t, \cdot\rangle$ gives $R_T^{(1)} \le \frac{D_1}{\eta} + \frac{\eta T}{2}\tilde{G}_F^2 + \frac{G\hat{C}_\eta T}{\eta}$, $\tilde{G}_F := G\hat{G}_F = \frac{GdM}{\delta}$. Combining the four bounds yields: $\mathbb{E}[R_T] \le K_1\delta T + \frac{D_1}{\eta} + K_2\frac{\eta T}{\delta^2} + K_3\frac{\eta^2 T}{\delta^3}$. Balancing the first three terms gives the values of η, δ reported in Theorem 3 and the desired $\mathcal{O}(T^{3/4})$ bound follows.

6 Conclusion and Future Work

In this paper, we studied online nonconvex learning for a structured class of hidden-convex losses. We showed that OGD achieves $\mathcal{O}(\sqrt{T})$ regret in the exact-gradient setting, sharpened the geometric conditions required for the underlying algorithmic equivalence, and proved that OGD can suffer linear regret when these conditions fail. We also extended the framework to one-point bandit feedback, establishing sublinear expected regret guarantees against oblivious adversaries.

Several questions remain open. First, our lower bound shows that OGD can suffer linear regret without Hessian compatibility, but it does not rule out sublinear regret for other algorithms. An important direction is therefore to determine whether online hidden-convex optimization admits sublinear, or even $\mathcal{O}(\sqrt{T})$, regret without this geometric assumption. Second, our bandit guarantees leave open the possibility of sharper rates, high-probability bounds, and algorithms matching the best-known guarantees from bandit online convex optimization. More broadly, it would be interesting to understand for which classes of hidden-convex functions online algorithms incur a regret that matches that of their convex counterparts.

Acknowledgments and Disclosure of Funding

The authors acknowledge the use of ChatGPT throughout the development of this work, particularly for constructing the instance for the linear regret lower bound of Theorem 2, developing and refining its proof, and formulating the examples in Section 4.1. The manuscript was primarily written by the authors, with ChatGPT providing assistance with selected drafts and revisions. The authors critically evaluated and revised the AI-assisted material, checked the mathematical arguments, and assessed the novelty of the results against the relevant literature. The authors take full responsibility for the content of the manuscript.

Ioannis Panageas is supported by NSF grant CCF-2454115. This work is supported by the MOE Tier 2 Grant (MOE-T2EP20223-0018), the CQT++ Core Research Funding Grant (SUTD) (RS-NRCQT-00002), the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (Award Number: AISG2-RP-2020-016), and partially by Project MIS 5154714 of the National Recovery and Resilience Plan, Greece 2.0, funded by the European Union under the NextGenerationEU Program.

References

[1] Naman Agarwal, Alon Gonen, and Elad Hazan. Learning in non-convex games with an optimization oracle. In *Conference on Learning Theory*, pages 18–29. PMLR, 2019. 1, 3

[2] Ehsan Amid and Manfred KK Warmuth. Reparameterizing mirror descent as gradient descent. *Advances in Neural Information Processing Systems*, 33:8430–8439, 2020. 3, 5

[3] Ehsan Amid, Manfred KK Warmuth, Rohan Anil, and Tomer Koren. Robust bi-tempered logistic loss based on bregman divergences. *Advances in Neural Information Processing Systems*, 32, 2019. 5

[4] Anas Barakat, Souradip Chakraborty, Peihong Yu, Pratap Tokekar, and Amrit Singh Bedi. On the global optimality of policy gradient methods in general utility reinforcement learning. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. 2

[5] Anas Barakat, Ilyas Fatkhullin, and Niao He. Reinforcement learning with general utilities: Simpler variance reduction and large state-action space. In *International Conference on Machine Learning*, 2023. 2

[6] Anas Barakat, Ioannis Panageas, and Antonios Varvitsiotis. Convex markov games and beyond: New proof of existence, characterization and learning algorithms for nash equilibria. In *International Conference on Artificial Intelligence and Statistics*, 2026. 2

[7] Aharon Ben-Tal and Marc Teboulle. Hidden convexity in some nonconvex quadratically constrained quadratic programming. *Mathematical Programming*, 72(1):51–63, 1996. 2

[8] Aditya Bhaskara, Ashok Cutkosky, Ravi Kumar, and Manish Purohit. Descent with misaligned gradients and applications to hidden convexity. In *International Conference on Learning Representations*, 2025. 2

[9] Xin Chen, Niao He, Yifan Hu, and Zikun Ye. Efficient algorithms for a class of stochastic hidden convex optimization and its applications in network revenue management. *Operations Research*, 73(2):704–719, 2025. 2, 3

[10] Ryan D'Orazio, Danilo Vucetic, Zichu Liu, Junhyung Lyle Kim, Ioannis Mitliagkas, and Gauthier Gidel. Solving hidden monotone variational inequalities with surrogate losses. In *The Thirteenth International Conference on Learning Representations*, 2025. 2

[11] Tolga Ergen and Mert Pilanci. The convex landscape of neural networks: Characterizing global optima and stationary points via lasso models. *IEEE Transactions on Information Theory*, 71(5):3854–3870, 2025. 2

[12] Ilyas Fatkhullin, Niao He, and Yifan Hu. Stochastic optimization under hidden convexity. *SIAM Journal on Optimization*, 35(4):2544–2571, 2025. 2, 3

[13] Ilyas Fatkhullin, Niao He, Guanghui Lan, and Florian Wolf. Global solutions to non-convex functional constrained problems with hidden convexity. *arXiv preprint arXiv:2511.10626*, 2025. 2, 3

[14] Abraham D. Flaxman, Adam Tauman Kalai, and H. Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, page 385–394, USA, 2005. Society for Industrial and Applied Mathematics. 2, 3, 4, 9, 32

[15] Xiand Gao, Xiaobo Li, and Shuzhong Zhang. Online learning with non-convex losses and non-stationary regret. In *International Conference on Artificial Intelligence and Statistics*, 2018. 3

[16] Ian Gemp, Andreas Alexander Haupt, Luke Marris, Siqi Liu, and Georgios Piliouras. Convex Markov games: A new frontier for multi-agent reinforcement learning. In *International Conference on Machine Learning*, 2025. 2

[17] Udaya Ghai, Zhou Lu, and Elad Hazan. Non-convex online learning via algorithmic equivalence. *Advances in Neural Information Processing Systems*, 35:22161–22172, 2022. 1, 2, 3, 4, 5, 6, 7, 14, 18, 21

[18] Bram L Gorissen, Dick den Hertog, and Meike Reusken. Hidden convexity in a class of optimization problems with bilinear terms. *Operations Research*, 74(2):1126–1152, 2026. 2

[19] Elad Hazan. Introduction to online convex optimization. *Foundations and Trends in Optimization*, 2(3-4):157–325, 2016. 1, 3

[20] Elad Hazan, Sham Kakade, Karan Singh, and Abby Van Soest. Provably efficient maximum entropy exploration. In *International Conference on Machine Learning*, 2019. 2

[21] Elad Hazan, Kfir Levy, and Shai Shalev-Shwartz. Beyond convexity: Stochastic quasi-convex optimization. *Advances in Neural Information Processing Systems*, 2015. 3

[22] Amélie Héliou, Matthieu Martin, Panayotis Mertikopoulos, and Thibaud Rahier. Online non-convex optimization with imperfect feedback. *Advances in Neural Information Processing Systems*, 33:17224–17235, 2020. 1, 3

[23] Fivos Kalogiannis, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Ian Gemp, and Georgios Piliouras. Solving zero-sum convex Markov games. In *International Conference on Machine Learning*, 2025. 2

[24] Fivos Kalogiannis, Jingming Yan, and Ioannis Panageas. Learning equilibria in adversarial team markov games: A nonconvex-hidden-concave min-max optimization problem. *Advances in Neural Information Processing Systems*, 37:92832–92890, 2024. 2

[25] Walid Krichene, Maximilian Balandat, Claire Tomlin, and Alexandre Bayen. The hedge algorithm on a continuum. In *International Conference on Machine Learning*, 2015. 1, 3

[26] Tor Lattimore. Bandit convex optimisation. *arXiv preprint arXiv:2402.06535*, 2024. 3

[27] Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020. 3

[28] Eitan Levin, Joe Kileel, and Nicolas Boumal. The effect of smooth parametrizations on nonconvex optimization landscapes. *Mathematical Programming*, 209(1):63–111, 2025. 2, 3

[29] Duan Li, Zhi-You Wu, Heung-Wing Joseph Lee, Xin-Min Yang, and Lian-Sheng Zhang. Hidden convex minimization. *Journal of Global Optimization*, 31(2):211–233, 2005. 2

[30] Zhiyuan Li, Tianhao Wang, Jason D Lee, and Sanjeev Arora. Implicit bias of gradient descent on reparametrized models: On equivalence to mirror descent. *Advances in Neural Information Processing Systems*, 35:34626–34640, 2022. 3

[31] Odalric-Ambrym Maillard and Rémi Munos. Online learning in adversarial lipschitz environments. In *Joint european conference on machine learning and knowledge discovery in databases*, pages 305–320. Springer, 2010. 3

[32] Andjela Mladenovic, Iosif Sakos, Gauthier Gidel, and Georgios Piliouras. Generalized natural gradient flows in hidden convex-concave games and GANs. In *International Conference on Learning Representations*, 2022. 2

[33] Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019. 3

[34] Deep Patel and Emmanouil-Vasileios Vlatakis-Gkaragkounis. Solving neural min-max games: The role of architecture, initialization & dynamics. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*. 2

[35] Iosif Sakos, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Panayotis Mertikopoulos, and Georgios Piliouras. Exploiting hidden structures in non-convex games for convergence to nash equilibrium. *Advances in Neural Information Processing Systems*, 2023. 2

[36] Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012. 1, 3

[37] Arun Sai Suggala and Praneeth Netrapalli. Online non-convex learning: Following the perturbed leader is optimal. In *Algorithmic Learning Theory*, pages 845–861. PMLR, 2020. 1, 3

[38] Emmanouil-Vasileios Vlatakis-Gkaragkounis, Lampros Flokas, and Georgios Piliouras. Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games. *Advances in Neural Information Processing Systems*, 32, 2019. 2

[39] Yifei Wang, Jonathan Lacotte, and Mert Pilanci. The hidden convex optimization landscape of regularized two-layer reLU networks: an exact characterization of optimal solutions. In *International Conference on Learning Representations*, 2022. 2

[40] Tom Zahavy, Brendan O'Donoghue, Guillaume Desjardins, and Satinder Singh. Reward is enough for convex mdps. In *Advances in Neural Information Processing Systems*, 2021. 2

[41] Emi Zeger and Mert Pilanci. Unveiling hidden convexity in deep learning: A sparse signal processing perspective. *arXiv preprint arXiv:2603.23831*, 2026. 2

[42] Junyu Zhang, Alec Koppel, Amrit Singh Bedi, Csaba Szepesvari, and Mengdi Wang. Variational policy gradient method for reinforcement learning with general utilities. In *Advances in Neural Information Processing Systems*, 2020. 2

[43] Lijun Zhang, Tianbao Yang, Rong Jin, and Zhi-Hua Zhou. Online bandit learning for a special class of non-convex losses. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(1), Feb. 2015. 3

[44] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *International Conference on Machine Learning*, pages 928–936, 2003. 1, 4

Contents

1 Introduction **1**

1.1 Main contributions 2

1.2 Related Work 3

2 Online Learning on Hidden-Convex Losses **3**

3 Regret Analysis of OGD for OHCO under Exact Gradient Feedback **4**

3.1 Assumptions 4

3.2 Regret Bound 5

3.3 Proof Sketch of Theorem 1 5

4 Discussion of the Hessian-Compatibility Hidden Geometry Assumption 1 **6**

4.1 Relaxation of the diagonal Jacobian assumption in [17] 7

4.2 OGD can incur linear regret without Hessian-compatible hidden geometry 7

5 Regret Analysis of OGD for OHCO under Bandit Feedback **9**

5.1 Expected Regret Bound in the Bandit Setting 9

5.2 Proof Sketch of Theorem 3 10

6 Conclusion and Future Work **10**

A Proofs for Section 3: Proof of Theorem 1 **15**

A.1 Proof of Lemma 1 15

A.2 Proof of Theorem 1 using Lemma 1 21

B Proofs for Section 4 **22**

B.1 Proof of Proposition 1 22

B.2 Detailed derivations for examples 23

B.3 Proof of Theorem 2: Linear Regret Lower Bound for OGD 24

C Proof of Theorem 3: Expected Regret Bound in the Bandit Setting **29**

A Proofs for Section 3: Proof of Theorem 1

In this section, we provide a complete proof of Theorem 1. The proof strategy consists of showing that the OGD iterates mapped to $\mathcal{Y}$ via q and the OMD iterates are close to each other in the hidden space $\mathcal{Y}$ after a single update step starting from the same initial point. Then we can view the OGD update as a perturbed version of OMD, and combine it with the fact that the OMD algorithm can tolerate bounded noise per trial.

A.1 Proof of Lemma 1

We begin with the following key lemma showing that the updates y_{t+1} and $z_{t+1} = q(x_{t+1})$ created by OMD and OGD respectively, are close to each other when starting from the same initial point $y_t = z_t = q(x_t)$. We show that the two iterates are minimizers of approximately the same strongly-convex objective, hence the minimizers must be close. In preparation of the bandit setting, we prove a general result that holds pathwise when we use any bounded vector instead of the exact gradients of both the original loss function ℓ_t and the hidden function h_t.

Proposition 2. *Suppose Assumptions 1-3 hold and assume $y_t = q(x_t)$. Consider the following OMD and OGD-like update rules:*

$$y_{t+1} = \arg\min_{y \in \mathcal{Y}} \tilde{g}_t^\top (y - y_t) + \frac{1}{\eta} D_R(y \| y_t)\,, \tag{5}$$

$$x_{t+1} = \arg\min_{x \in \mathcal{X}} g_t^\top (x - x_t) + \frac{1}{2\eta} \|x - x_t\|_2^2\,, \tag{6}$$

where $\tilde{g}_t = J_q(x_t)^{-\top} g_t$ and $g_t \in \mathbb{R}^d$ is any vector[a]. If in addition there exists $\hat{G}_F > 0$ such that $\|g_t\| \le \hat{G}_F$ for any t, then,

$$\|y_{t+1} - q(x_{t+1})\|_2 \le G^5 (5\hat{G}_F^2 + \hat{G}_F^3 \eta)\eta^2\,. \tag{7}$$

[a] $\tilde{g}_t$ is used in the analysis but never actually computed, in contrast to g_t which is computed and used in BOGD. Note that we suppose that we do not have access to the parameterization q.

Before providing the proof, we state an immediate corollary of this result in the exact gradient information setting of Theorem 1.

Corollary 4 (Proposition 2 in the exact gradient setting)**.** *In Proposition 2, set $g_t = \nabla \ell_t(x_t)$ and $\tilde{g}_t = \nabla h_t(z_t)$ in (5)-(6), then the approximation inequality (7) holds with $\hat{G}_F = G_F$ where G_F is the uniform loss gradient ($\nabla \ell_t$) bound in Assumption 3, i.e.,*

$$\|y_{t+1} - q(x_{t+1})\|_2 \le C_\eta := G^5 (5G_F^2 + G_F^3 \eta)\eta^2.$$

Proof of Proposition 2. Note first that $\|\tilde{g}_t\| \le \tilde{G}_F := G\hat{G}_F$. Both objectives in (5)-(6) can be written as the sum of a linear function and a strongly convex function. By 1-strong convexity of R and uniform boundedness of g_t by $\hat{G}_F$, we have for any $y \in \mathcal{Y}$,

$$\tilde{g}_t^\top (y - y_t) + \frac{1}{\eta} D_R(y \| y_t) \ge -\tilde{G}_F \|y - y_t\| + \frac{1}{2\eta} \|y - y_t\|^2\,.$$

Therefore, if $\|y - y_t\|_2 > 2\eta \tilde{G}_F$, then the objective takes a positive value which cannot be minimal since $y = y_t$ gives zero objective value in (5). Define the set:

$$\mathcal{Y}_{r, y_t} = \mathcal{Y} \cap \{y \in \mathbb{R}^d : \|y - y_t\|_2 \le r\}\,.$$

Then it follows that the OMD iterate in (5) satisfies:

$$y_{t+1} \in \mathcal{Y}_t := \mathcal{Y}_{2\eta \tilde{G}_F, y_t}\,. \tag{8}$$

As we want to localize both iterates (5)-(6) in a set of size controlled by η, we now show that we also have $q(x_{t+1}) \in \mathcal{Y}_t$ when $y_t = q(x_t)$. Observe for this that:

$$\|q(x_{t+1}) - y_t\| = \|q(x_{t+1}) - q(x_t)\| \le G\|x_{t+1} - x_t\| \le \eta G \|g_t\| \le \eta G \hat{G}_F \le 2\eta G \hat{G}_F = 2\eta \tilde{G}_F\,.$$

We now write both update rules under a similar form. Define the function $\Phi_t : \mathcal{Y} \to \mathbb{R}$ by:

$$\Phi_t(y) := \tilde{g}_t^\top (y - y_t) + \frac{1}{2\eta} \|y - y_t\|^2_{\nabla^2 R(y_t)} \,. \tag{9}$$

The function Φ_t is $1/\eta$-strongly convex because $\nabla^2 \Phi_t(y) = \frac{1}{\eta} \nabla^2 R(y_t) \succeq \frac{1}{\eta} I$ as R is supposed to be 1-strongly convex.

Lemma 2. *For any $t \geq 1$, the iterates of (OMD)-(5) can be rewritten as follows:*

$$y_{t+1} = \operatorname*{arg\,min}_{y \in \mathcal{Y}_t} \left\{ \Phi_t(y) + \varepsilon_t^{\mathrm{OMD}}(y) \right\} , \quad \varepsilon_t^{\mathrm{OMD}}(y) := \frac{1}{\eta} \varepsilon_t(y) \,,$$

where $\varepsilon_t(y) := D_R(y \| y_t) - \frac{1}{2} \|y - y_t\|^2_{\nabla^2 R(y_t)}$. Moreover, for $y \in \mathcal{Y}_t$, $|\varepsilon_t(y)| \leq \frac{4}{3} G \tilde{G}_F^3 \eta^3$.

Proof. We have proved in (8) that $y_{t+1} \in \mathcal{Y}_t$. Hence, we can rewrite (5) by replacing $\mathcal{Y}$ by $\mathcal{Y}_t$:

$$y_{t+1} = \operatorname*{arg\,min}_{y \in \mathcal{Y}_t} \tilde{g}_t^\top (y - y_t) + \frac{1}{\eta} D_R(y \| y_t). \tag{10}$$

Let $\Delta_t := y - y_t$. By Taylor's theorem with integral remainder, we have:

$$\begin{aligned} D_R(y \| y_t) &= R(y) - R(y_t) - \nabla R(y_t)^\top (y - y_t) \\ &= \frac{1}{2} \Delta_t^\top \nabla^2 R(y_t) \Delta_t + \varepsilon_t(y) \\ &= \frac{1}{2} \|\Delta_t\|^2_{\nabla^2 R(y_t)} + \varepsilon_t(y), \end{aligned} \tag{11}$$

where $\varepsilon_t(y)$ is the integral remainder defined as follows:

$$\varepsilon_t(y) := \frac{1}{2} \int_0^1 (1 - s)^2 \nabla^3 R(y_t + s\Delta_t)[\Delta_t, \Delta_t, \Delta_t] \, ds \,,$$

where $\nabla^3 R$ is the third-order derivative tensor of R. Since this third derivative is uniformly bounded by G over $z \in \mathcal{Y}$ by Assumption 2, we have:

$$|\varepsilon_t(y)| \leq \frac{G}{6} \|\Delta_t\|_2^3.$$

In particular, for $y \in \mathcal{Y}_t$, i.e. $\|y - y_t\|_2 \leq 2\eta \tilde{G}_F$, we obtain the following estimate:

$$|\varepsilon_t(y)| \leq \frac{G}{6} (2\eta \tilde{G}_F)^3 = \frac{4}{3} G \tilde{G}_F^3 \eta^3 \,.$$

Combining the Bregman approximation of (11) with (10), we obtain the desired identity using the definition of the strongly-convex function Φ_t introduced in (9). □

In the following lemma, we rewrite the OGD iterates mapped to $\mathcal{Y}$ via q under a similar form to the update rule of OMD obtained in Lemma 2.

Lemma 3. *For the update (6) with a bounded vector g_t, suppose $y_t = q(x_t)$. Then, for $t \geq 1$,*

$$z_{t+1} = q(x_{t+1}) = \operatorname*{arg\,min}_{y \in \mathcal{Y}_t} \left\{ \Phi_t(y) + \varepsilon_t^{\mathrm{OGD}}(y) \right\} , \tag{12}$$

$$\varepsilon_t^{\mathrm{OGD}}(y) := \tilde{\varepsilon}_t(y) + \frac{1}{\eta} \left(\frac{1}{2} \|\varepsilon_t^q(y)\|_2^2 + \left\langle J_q(x_t)^{-1}(y - y_t), \, \varepsilon_t^q(y) \right\rangle \right) , \tag{13}$$

where $\tilde{\varepsilon}_t(y) := g_t^\top \varepsilon_t^q(y)$ and $\varepsilon_t^q(y) := q^{-1}(y) - q^{-1}(y_t) - J_q^{-1}(x_t)(y - y_t)$. Moreover, for every $y \in \mathcal{Y}_t$, we have $\|\varepsilon_t^q(y)\|_2 \leq 2G\tilde{G}_F^2 \eta^2$ and $|\tilde{\varepsilon}_t(y)| \leq 2\tilde{G}_F^3 \eta^2$

Proof. Recall that the update (6) in the original space $\mathcal{X}$ can be written as follows:

$$x_{t+1} = \arg\min_{x \in \mathcal{X}} \left\{ g_t^\top (x - x_t) + \frac{1}{2\eta} \|x - x_t\|_2^2 \right\}. \tag{14}$$

Recall from (9) that for any $y \in \mathcal{Y}$:

$$\Phi_t(y) := \tilde{g}_t^\top (y - y_t) + \frac{1}{2\eta} \|y - y_t\|^2_{\nabla^2 R(y_t)} .$$

To prove (12), we relate the linear term $g_t^\top (x - x_t)$ to $\tilde{g}_t^\top (y - y_t)$ and the quadratic term $\frac{1}{2\eta}\|x - x_t\|_2^2$ to $\frac{1}{2\eta}\|y - y_t\|^2_{\nabla^2 R(y_t)}$ respectively in **(i)** and **(ii)** below.

(i) Relating $g_t^\top (x - x_t)$ to $\tilde{g}_t^\top (y - y_t)$. Recalling that $g_t = J_q(x_t)^\top \tilde{g}_t$, we have:

$$\begin{aligned} g_t^\top (x - x_t) &= \tilde{g}_t^\top J_q(x_t)(x - x_t) \\ &= \tilde{g}_t^\top J_q(x_t)(q^{-1}(y) - q^{-1}(y_t)) , \end{aligned} \tag{15}$$

where the last equality follows from using the identities $y = q(x)$ and $y_t = q(x_t)$.

Under Assumption 2, applying Taylor's theorem to q^{-1} around $y_t = q(x_t)$ gives, for $y = q(x)$,

$$q^{-1}(y) - q^{-1}(y_t) = J_{q^{-1}}(y_t)(y - y_t) + \varepsilon_t^q(y) , \tag{16}$$

where the remainder is given by:

$$\varepsilon_t^q(y) = \int_0^1 (1 - s) \, (y - y_t)^\top \nabla^2 q^{-1}(y_t + s(y - y_t))(y - y_t) \, ds. \tag{17}$$

Since $J_{q^{-1}}(y_t) = J_q(x_t)^{-1}$, it follows that:

$$q^{-1}(y) - q^{-1}(y_t) = J_q(x_t)^{-1}(y - y_t) + \varepsilon_t^q(y). \tag{18}$$

Since the second derivative of q^{-1} is bounded by G, we have:

$$\|\varepsilon_t^q(y)\|_2 \le \frac{G}{2}\|y - y_t\|_2^2.$$

In particular, for $y \in \mathcal{Y}_t = \mathcal{Y}_{2\eta\tilde{G}_F, y_t}$, i.e. $y \in \mathcal{Y}$ s.t. $\|y - y_t\|_2 \le 2\eta\tilde{G}_F$, we obtain:

$$\|\varepsilon_t^q(y)\|_2 \;\le\; \frac{G}{2}(2\eta\tilde{G}_F)^2 = 2G\tilde{G}_F^2\eta^2 . \tag{19}$$

Using (18) in (15) yields:

$$\begin{aligned} g_t^\top (x - x_t) &= \tilde{g}_t^\top J_q(x_t)(q^{-1}(y) - q^{-1}(y_t)) \\ &= \tilde{g}_t^\top (y - y_t) + \tilde{g}_t^\top J_q(x_t) \, \varepsilon_t^q(y) \\ &= \tilde{g}_t^\top (y - y_t) + \tilde{\varepsilon}_t(y), \end{aligned} \tag{20}$$

where $\tilde{\varepsilon}_t(y) := \tilde{g}_t^\top J_q(x_t) \, \varepsilon_t^q(y) = g_t^\top \varepsilon_t^q(y)$. Therefore, using (19), we have for any $y \in \mathcal{Y}_t$,

$$|\tilde{\varepsilon}_t(y)| \;\le\; \|g_t\| \cdot \|\varepsilon_t^q(y)\| \;\le\; G(2\hat{G}_F\tilde{G}_F^2\eta^2) = 2\tilde{G}_F^3\eta^2 ,$$

where the last equality stems from recalling that $\tilde{G}_F = G\hat{G}_F$.

(ii) Relating $\frac{1}{2\eta}\|x - x_t\|_2^2$ to $\frac{1}{2\eta}\|y - y_t\|^2_{\nabla^2 R(y_t)}$. For any $x \in \mathcal{X}$, $y = q(x)$, given that $y_t = q(x_t)$,

$$\begin{aligned} \frac{1}{2\eta}\|x - x_t\|_2^2 &= \frac{1}{2\eta}\|q^{-1}(y) - q^{-1}(y_t)\|_2^2 \\ &= \frac{1}{2\eta}\|J_q(x_t)^{-1}(y - y_t) + \varepsilon_t^q(y)\|_2^2 \\ &= \frac{1}{2\eta}\|J_q(x_t)^{-1}(y - y_t)\|_2^2 + \frac{1}{\eta}\left(\frac{1}{2}\|\varepsilon_t^q(y)\|_2^2 + \left\langle J_q(x_t)^{-1}(y - y_t), \, \varepsilon_t^q(y)\right\rangle\right). \end{aligned} \tag{21}$$

Using Assumption 1 relating J_q and $\nabla^2 R$ and the initial coupling $y_t = z_t = q(x_t)$, we have:

$$\left[\nabla^2 R(y_t)\right]^{-1} = J_q(x_t) J_q(x_t)^\top, \qquad \nabla^2 R(y_t) = [J_q(x_t)^\top]^{-1} J_q(x_t)^{-1}.$$

Therefore we can write:

$$\begin{aligned}\|J_q(x_t)^{-1}(y - y_t)\|_2^2 &= (y - y_t)^\top [J_q(x_t)^{-1}]^\top J_q(x_t)^{-1}(y - y_t) \\ &= (y - y_t)^\top \nabla^2 R(y_t)(y - y_t) \\ &= \|y - y_t\|^2_{\nabla^2 R(y_t)}. \end{aligned} \tag{22}$$

Combining (21) and (22) yields:

$$\frac{1}{2\eta}\|x - x_t\|_2^2 = \frac{1}{2\eta}\|y - y_t\|^2_{\nabla^2 R(y_t)} + \frac{1}{\eta}\left(\frac{1}{2}\|\varepsilon_t^q(y)\|_2^2 + \left\langle J_q(x_t)^{-1}(y - y_t),\, \varepsilon_t^q(y)\right\rangle\right). \tag{23}$$

Given the OGD update rule (14), the change of variable $y = q(x)$ and (20), (23) proven in **(i)** and **(ii)** respectively, we have shown that

$$q(x_{t+1}) = \operatorname*{arg\,min}_{y \in \mathcal{Y}_t} \left\{\Phi_t(y) + \varepsilon_t^{\mathrm{OGD}}(y)\right\},$$

where the error term $\varepsilon_t^{\mathrm{OGD}}(y)$ is defined as follows:

$$\varepsilon_t^{\mathrm{OGD}}(y) := \tilde{\varepsilon}_t(y) + \frac{1}{\eta}\left(\frac{1}{2}\|\varepsilon_t^q(y)\|_2^2 + \left\langle J_q(x_t)^{-1}(y - y_t),\, \varepsilon_t^q(y)\right\rangle\right). \tag{24}$$

□

So far, we have shown in Lemmas 2 and 3 respectively the following similar update rule forms:

$$y_{t+1} = \operatorname*{arg\,min}_{y \in \mathcal{Y}_t}\{\Phi_t(y) + \varepsilon_t^{\mathrm{OMD}}(y)\}, \tag{25}$$

$$z_{t+1} = q(x_{t+1}) = \operatorname*{arg\,min}_{y \in \mathcal{Y}_t}\{\Phi_t(y) + \varepsilon_t^{\mathrm{OGD}}(y)\}. \tag{26}$$

In the next lemma, we use first-order conditions together with strong convexity of the function Φ_t to show that the distance between the minimizers can be upper bounded by η times the distance between the gradients of the errors in each one of the updates (for OMD and OGD). Note that our proof technique here is different from the strategy adopted in [17] (Lemma 4) as we do not control the distance between objectives and translate it to the distance between minimizers using strong convexity.

Lemma 4. *For any $t \geq 1$, if $y_t = q(x_t)$, then*

$$\|y_{t+1} - q(x_{t+1})\| \leq \eta \left\|\nabla \varepsilon_t^{\mathrm{OGD}}(q(x_{t+1})) - \nabla \varepsilon_t^{\mathrm{OMD}}(y_{t+1})\right\|.$$

To prove this result, we will need the following technical lemma stating monotonicity of normal cones of nonempty closed convex sets, which is a standard result in convex analysis.

Lemma 5 (Monotonicity of the normal cone)**.** *Let $K \subset \mathbb{R}^d$ be a nonempty closed convex set. For any $x, y \in K$, let*

$$N_K(x) := \{v \in \mathbb{R}^d : \langle v, z - x\rangle \leq 0, \quad \forall z \in K\}$$

denote the normal cone of K at x. Then, for any $v_x \in N_K(x)$ and $v_y \in N_K(y)$,

$$\langle v_x - v_y, x - y\rangle \geq 0.$$

Proof. Since $v_x \in N_K(x)$ and $y \in K$, the definition of the normal cone gives $\langle v_x, y - x\rangle \leq 0$, or equivalently, $\langle v_x, x - y\rangle \geq 0$. Similarly, since $v_y \in N_K(y)$ and $x \in K$, we have $\langle v_y, x - y\rangle \leq 0$. Therefore, it follows from the above inequalities that:

$$\langle v_x - v_y, x - y\rangle = \langle v_x, x - y\rangle - \langle v_y, x - y\rangle \geq 0.$$

□

Proof of Lemma 4. Let $\delta_{t+1} := y_{t+1} - q(x_{t+1})$ as a shorthand notation. From first-order optimality conditions for (25) and (26), there exist $v_{t+1}^{\mathrm{OMD}} \in N_{\mathcal{Y}_t}(y_{t+1}), v_{t+1}^{\mathrm{OGD}} \in N_{\mathcal{Y}_t}(q(x_{t+1}))$ s.t.

$$\nabla\Phi_t(y_{t+1}) + \nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1}) + v_{t+1}^{\mathrm{OMD}} = 0\,,$$
$$\nabla\Phi_t(q(x_{t+1})) + \nabla\varepsilon_t^{\mathrm{OGD}}(q(x_{t+1})) + v_{t+1}^{\mathrm{OGD}} = 0\,.$$

Subtracting both identities and taking inner product with δ_{t+1} gives

$$\left\langle \nabla\Phi_t(y_{t+1}) - \nabla\Phi_t(q(x_{t+1})),\, \delta_{t+1} \right\rangle + \left\langle \nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1}) - \nabla\varepsilon_t^{\mathrm{OGD}}(q(x_{t+1})),\, \delta_{t+1} \right\rangle + \left\langle v_{t+1}^{\mathrm{OMD}} - v_{t+1}^{\mathrm{OGD}},\, \delta_{t+1} \right\rangle = 0. \quad (27)$$

By $\frac{1}{\eta}$-strong convexity of Φ_t, we have:

$$\left\langle \nabla\Phi_t(y_{t+1}) - \nabla\Phi_t(q(x_{t+1})),\, \delta_{t+1} \right\rangle \geq \frac{1}{\eta}\|\delta_{t+1}\|^2. \quad (28)$$

By monotonicity of normal cones of convex sets (Lemma 5 above), we can also control the term involving the normal cone directions as follows:

$$\left\langle v_{t+1}^{\mathrm{OMD}} - v_{t+1}^{\mathrm{OGD}},\, \delta_{t+1} \right\rangle \geq 0. \quad (29)$$

Plugging (28) and (29) into (27), we obtain:

$$\frac{1}{\eta}\|\delta_{t+1}\|^2 \leq \left\|\nabla\varepsilon_t^{\mathrm{OGD}}(q(x_{t+1})) - \nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1})\right\| \cdot \|\delta_{t+1}\|.$$

Therefore, it follows that the distance δ_{t+1} between the two minimizers is upper bounded as follows:

$$\|\delta_{t+1}\| \leq \eta \left\|\nabla\varepsilon_t^{\mathrm{OGD}}(q(x_{t+1})) - \nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1})\right\|.$$

□

It remains to bound the last norm in the right-hand side of Lemma 4 and show it is also of order η. We show that each one of the norms of the gradients $\|\nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1})\|$ and $\|\nabla\varepsilon_t^{\mathrm{OGD}}(q(x_{t+1}))\|$ are of order η in the following two lemmas.

Lemma 6. $\|\nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1})\| \leq 2G\tilde{G}_F^2\eta\,.$

Proof. Recall that $\varepsilon_t^{\mathrm{OMD}}(y) = \frac{1}{\eta}\varepsilon_t(y)$, where:

$$\varepsilon_t(y) = D_R(y\|y_t) - \frac{1}{2}(y - y_t)^\top\nabla^2 R(y_t)(y - y_t)\,.$$

Differentiating the above error function w.r.t. y yields:

$$\nabla\varepsilon_t(y) = \nabla R(y) - \nabla R(y_t) - \nabla^2 R(y_t)(y - y_t).$$

Using third-order smoothness of R (bounded third derivative) as supposed in Assumption 2,

$$\|\nabla\varepsilon_t(y_{t+1})\| \leq \frac{G}{2}\|y_{t+1} - y_t\|^2. \quad (30)$$

Since $y_{t+1} \in \mathcal{Y}_t$, we have $\|y_{t+1} - y_t\| \leq 2\eta\tilde{G}_F$ and it follows from (30) that:

$$\|\nabla\varepsilon_t(y_{t+1})\| \leq \frac{G}{2}(2\eta\tilde{G}_F)^2 = 2G\tilde{G}_F^2\eta^2\,.$$

Since $\varepsilon_t^{\mathrm{OMD}}(y) = \frac{1}{\eta}\varepsilon_t(y)$, we conclude that:

$$\|\nabla\varepsilon_t^{\mathrm{OMD}}(y_{t+1})\| = \frac{1}{\eta}\|\nabla\varepsilon_t(y_{t+1})\| \leq 2G\tilde{G}_F^2\eta\,.$$

□

Lemma 7. $\|\nabla\varepsilon_t^{\mathrm{OGD}}(q(x_{t+1}))\| \;\le\; G^5(3\hat{G}_F^2+\hat{G}_F^3\eta)\eta.$

Proof. Recall from (24) that:

$$\varepsilon_t^{\mathrm{OGD}}(y)=\tilde{\varepsilon}_t(y)+\frac{1}{\eta}\left(\frac{1}{2}\|\varepsilon_t^q(y)\|_2^2+\left\langle J_q(x_t)^{-1}(y-y_t),\,\varepsilon_t^q(y)\right\rangle\right).$$

Differentiating this error term w.r.t. y gives:

$$\nabla\varepsilon_t^{\mathrm{OGD}}(y)=\nabla\tilde{\varepsilon}_t(y)+\frac{1}{\eta}J_{\varepsilon_t^q}(y)^\top\varepsilon_t^q(y)+\frac{1}{\eta}\Big([J_q(x_t)^{-1}]^\top\varepsilon_t^q(y)+J_{\varepsilon_t^q}(y)^\top J_q(x_t)^{-1}(y-y_t)\Big). \quad (31)$$

Since $\tilde{\varepsilon}_t(y):=g_t^\top\varepsilon_t^q(y)$, we have:

$$\nabla\tilde{\varepsilon}_t(y)=J_{\varepsilon_t^q}(y)^\top g_t\,,\quad \|\nabla\tilde{\varepsilon}_t(y)\|\;\le\;\hat{G}_F\|J_{\varepsilon_t^q}(y)\|\,.$$

We now prove the following estimates:

(i) $\|q(x_{t+1})-y_t\|\;\le\;G\hat{G}_F\eta\,,$

(ii) $\|J_{\varepsilon_t^q}(q(x_{t+1}))\|\;\le\;G^2\hat{G}_F\eta,$

(iii) $\|\varepsilon_t^q(q(x_{t+1}))\|\;\le\;\frac{G^3\hat{G}_F^2}{2}\eta^2\,.$

Proof of (i). Since $y_t=q(x_t)$, $x_t\in\mathcal{X}$, and $x_{t+1}=\Pi_{\mathcal{X}}(x_t-\eta g_t)$,

$$\begin{aligned}\|q(x_{t+1})-y_t\|_2&=\|q(x_{t+1})-q(x_t)\|_2\\&\le G\|x_{t+1}-x_t\|_2\\&=G\|\Pi_{\mathcal{X}}(x_t-\eta g_t)-\Pi_{\mathcal{X}}(x_t)\|_2\\&\le\eta G\|g_t\|_2\\&\le\eta G\hat{G}_F=\eta\tilde{G}_F.\end{aligned}$$

Here we used Lipschitzness of q, nonexpansiveness of Euclidean projection, and the assumed bound $\|g_t\|_2\le\hat{G}_F$.

Proof of (ii). Recall from (16) that for any $y=q(x),x\in\mathcal{X}$,

$$\varepsilon_t^q(y)=q^{-1}(y)-q^{-1}(y_t)-J_{q^{-1}}(y_t)(y-y_t)\,. \quad (32)$$

Differentiating w.r.t. y yields:

$$J_{\varepsilon_t^q}(q(x_{t+1}))=J_{q^{-1}}(q(x_{t+1}))-J_{q^{-1}}(y_t)\,.$$

Using Assumption 2, we obtain:

$$\|J_{\varepsilon_t^q}(q(x_{t+1}))\|=\|J_{q^{-1}}(q(x_{t+1}))-J_{q^{-1}}(y_t)\|\;\le\;G\|q(x_{t+1})-y_t\|\;\le\;\eta G^2\hat{G}_F\,,$$

where the last step uses the first estimate (i) proved above.

Proof of (iii). Recalling (32) and using boundedness of the second-order derivative of q^{-1} (Assumption 2), we immediately obtain:

$$\|\varepsilon_t^q(q(x_{t+1}))\|\;\le\;\frac{G}{2}\cdot\|q(x_{t+1})-y_t\|^2\;\le\;\frac{G^3\hat{G}_F^2}{2}\eta^2\,,$$

where the last estimate follows from using the first proven estimate (i): $\|q(x_{t+1})-y_t\|\;\le\;G\hat{G}_F\eta\,.$

We conclude the proof of Lemma 7 by using all the estimates (i) to (iii) to bound the norm of the gradient $\nabla\varepsilon_t^{\mathrm{OGD}}(x)$ in (31) as follows:

$$\begin{aligned}\|\nabla\varepsilon_t^{\mathrm{OGD}}(q(x_{t+1}))\|\;&\le\;\hat{G}_F\|J_{\varepsilon_t^q}(q(x_{t+1}))\|+\frac{1}{\eta}\|J_{\varepsilon_t^q}(q(x_{t+1}))\|\cdot\|\varepsilon_t^q(q(x_{t+1}))\|\\&\quad+\frac{1}{\eta}G\|\varepsilon_t^q(q(x_{t+1}))\|+\frac{1}{\eta}G\|J_{\varepsilon_t^q}(q(x_{t+1}))\|\cdot\|q(x_{t+1})-y_t\|\\&\le\;G^2\hat{G}_F^2\eta+\frac{1}{2}G^5\hat{G}_F^3\eta^2+\frac{1}{2}G^4\hat{G}_F^2\eta+G^4\hat{G}_F^2\eta\\&\le\;G^5(3\hat{G}_F^2+\hat{G}_F^3\eta)\eta\,,\end{aligned} \quad (33)$$

where the last inequality follows from using the assumption $G \geq 1$ which is without loss of generality (otherwise replace the constant by 1). □

End of Proof of Proposition 2. Using Lemma 4, we have:

$$\|y_{t+1} - q(x_{t+1})\| \leq \eta \left\|\nabla \varepsilon_t^{\mathrm{OGD}}(q(x_{t+1})) - \nabla \varepsilon_t^{\mathrm{OMD}}(y_{t+1})\right\| .$$

Then it follows from using Lemma 6 and Lemma 7 that:

$$\begin{aligned}
\|y_{t+1} - q(x_{t+1})\| = \|\delta_{t+1}\| &\leq \eta \cdot \left(\|\nabla \varepsilon_t^{\mathrm{OGD}}(q(x_{t+1}))\| + \|\nabla \varepsilon_t^{\mathrm{OMD}}(y_{t+1})\|\right) \\
&\leq (G^5(3\hat{G}_F^2 + \hat{G}_F^3\eta) + 2G\tilde{G}_F^2)\eta^2 \\
&= (G^5(3\hat{G}_F^2 + \hat{G}_F^3\eta) + 2G^3\hat{G}_F^2)\eta^2 \\
&\leq G^5(5\hat{G}_F^2 + \hat{G}_F^3\eta)\eta^2 , \qquad (34)
\end{aligned}$$

which concludes the proof. □

A.2 Proof of Theorem 1 using Lemma 1

The following lemma controls the regret of an approximate OMD algorithm.

Lemma 8 (Perturbed OMD regret bound; cf. Lemma 6 in [17])**.** *Let Assumptions 1–3 hold and let $\eta > 0$. Suppose $(z_t)_{t=1}^{T+1} \subseteq \mathcal{Y}$ satisfies*

$$y_{t+1} = \arg\min_{y \in \mathcal{Y}} \left\{ \langle \nabla h_t(z_t), y - z_t \rangle + \frac{1}{\eta} D_R(y \| z_t) \right\}, \qquad z_{t+1} = y_{t+1} + r_{t+1},$$

where $\|r_{t+1}\|_2 \leq C_\eta$ for all t. Then

$$\sum_{t=1}^{T} h_t(z_t) - \min_{z \in \mathcal{Y}} \sum_{t=1}^{T} h_t(z) \leq \frac{D_1}{\eta} + \frac{\eta G^2 G_F^2 T}{2} + \frac{G C_\eta T}{\eta}.$$

Proof. Fix any comparator $z \in \mathcal{Y}$. The optimality condition for the ghost OMD step, the Bregman three-point identity, and 1-strong convexity of R give:

$$\langle \nabla h_t(z_t), z_t - z \rangle \leq \frac{D_R(z\|z_t) - D_R(z\|y_{t+1})}{\eta} + \frac{\eta}{2}\|\nabla h_t(z_t)\|_2^2.$$

By the chain rule and Assumptions 2–3,

$$\|\nabla h_t(z_t)\|_2 = \|J_q(q^{-1}(z_t))^{-\top} \nabla \ell_t(q^{-1}(z_t))\|_2 \leq G G_F.$$

Moreover, both z_{t+1} and y_{t+1} belong to $\mathcal{Y}$, so Assumption 2(iii) implies

$$D_R(z\|z_{t+1}) - D_R(z\|y_{t+1}) \leq G\|z_{t+1} - y_{t+1}\|_2 \leq G C_\eta.$$

Using convexity of h_t and inserting z_{t+1} into the Bregman difference, we obtain

$$h_t(z_t) - h_t(z) \leq \frac{D_R(z\|z_t) - D_R(z\|z_{t+1})}{\eta} + \frac{\eta G^2 G_F^2}{2} + \frac{G C_\eta}{\eta}.$$

Summing over t, the Bregman terms telescope. The result follows from $D_R(z\|z_1) \leq D_1$, $D_R(z\|z_{T+1}) \geq 0$, and choosing a best fixed comparator. □

End of Proof of Theorem 1. The rest of the proof follows the same lines as the proof of [17] using our tighter estimate obtained in Corollary 4. We provide a full proof for completeness.

Proof. Let $z_t = q(x_t)$. At each round, define the ghost OMD step from the current mapped OGD iterate by:

$$y_{t+1} = \arg\min_{y \in \mathcal{Y}} \left\{ \langle \nabla h_t(z_t), y - z_t \rangle + \frac{1}{\eta} D_R(y \| z_t) \right\}.$$

Corollary 4, applied with the common starting point $z_t = q(x_t)$, gives:

$$z_{t+1} = y_{t+1} + r_{t+1}, \qquad \|r_{t+1}\|_2 \le C_\eta := G^5(5G_F^2 + G_F^3\eta)\eta^2.$$

Since $z_t = q(x_t) \in \mathcal{Y}$ for every t, Lemma 8 applies:

$$\begin{aligned} R_T &\le \frac{D_1}{\eta} + \frac{\eta G^2 G_F^2 T}{2} + \frac{G C_\eta T}{\eta} \\ &= \frac{D_1}{\eta} + G_F^2\left(5G^6 + \frac{1}{2}G^2\right)\eta T + G^6 G_F^3 \eta^2 T \\ &= \frac{D_1}{\eta} + A\eta T + G^6 G_F^3 \eta^2 T\,, \end{aligned}$$

where $A := G_F^2\left(5G^6 + \frac{1}{2}G^2\right)$. Substituting $\eta = \sqrt{D_1/(AT)}$ yields:

$$R_T \le 2\sqrt{AD_1T} + \frac{G^6 G_F^3 D_1}{A},$$

which proves Theorem 1. □

B Proofs for Section 4

B.1 Proof of Proposition 1

The proof relies on the following classical result from vector calculus.

Proposition 3. *Let $\mathcal{U} \subseteq \mathbb{R}^d$ be convex[a] and let $F : \mathcal{U} \to \mathbb{R}^d$ be a C^1 vector field. Denoting $F = (F_1, \cdots, F_d)$ using its coordinate functions, if $\partial_{x_j} F_k(x) = \partial_{x_k} F_j(x)$ for all $j, k \in [d]$ and all $x \in \mathcal{U}$, then there exists a scalar function $\phi : \mathcal{U} \to \mathbb{R}$ such that $F = \nabla\phi$.*

[a] A simple connected domain suffices here. Convexity is enough for our purpose.

Proof of Prop. 1. Necessity is immediate: if $M = \nabla^2 R$, then

$$\partial_{z_k} M_{ij} = \partial^3_{z_k z_i z_j} R = \partial^3_{z_j z_i z_k} R = \partial_{z_j} M_{ik}.$$

For each $i \in [d]$, define the vector field $M^{(i)}$ of the ith row of the matrix field M:

$$M^{(i)}(z) := (M_{i,1}(z), \cdots, M_{i,d}(z))\,.$$

The cross-derivatives conditions $\partial_{z_k} M_{ij}(z) = \partial_{z_j} M_{ik}(z)\,, \forall i, j, k \in [d], \forall z \in \mathcal{Y}$, guarantee that $M^{(i)}$ has also matching cross-partial derivatives in the sense of Proposition 3. Therefore, it follows from Proposition 3 that there exists a scalar function $g_i : \mathcal{Y} \to \mathbb{R}$ such that $\nabla g_i = M^{(i)}$ for all $i \in [d]$.

We obtain a vector field $g = (g_1, \cdots, g_d)$ whose Jacobian J_g coincides with M, i.e. $J_g = M$ since $M^{(i)}$ are the rows of the matrix field M. Then, observe that M is symmetric because $J_q J_q^\top$ is symmetric and its inverse is also symmetric. Hence, the Jacobian $J_g = M$ is also symmetric and it follows that g itself has matching cross partial derivatives. Applying Proposition 3 again to g yields the existence of a scalar function R such that $g = \nabla R$. We have proved that $M = J_g$ and $g = \nabla R$ which implies that $M = \nabla^2 R$.

If R_1 and R_2 have the same Hessian on $\mathcal{Y}$, then $\nabla^2(R_1 - R_2) = 0$. Since $\mathcal{Y}$ is convex, $R_1 - R_2$ is affine, which proves uniqueness up to an affine function.

Finally, $M(z) \succ 0$ for every $z \in \mathcal{Y}$. Hence R is strictly convex. Since $\mathcal{Y}$ is compact and M is continuous, $\inf_{z\in\mathcal{Y}} \lambda_{\min}(M(z)) > 0$ (where $\lambda_{\min}$ stands for the smallest eigenvalue), which gives uniform strong convexity. □

B.2 Detailed derivations for examples

Example 1: Affine mixing of a separable reparameterization. Let

$$q(x) = A\,s(x) + b, \qquad s(x) = \big(s_1(x_1), \dots, s_d(x_d)\big),$$

where $A \in \mathbb{R}^{d\times d}$ is invertible, $b \in \mathbb{R}^d$, and each $s_i : \mathbb{R} \to \mathbb{R}$ is a C^2 diffeomorphism. Writing

$$D(x) := \operatorname{diag}\big(s_1'(x_1), \dots, s_d'(x_d)\big),$$

it follows that the Jacobian of q at x and its outer product can be written as:

$$J_q(x) = AD(x), \qquad J_q(x)J_q(x)^\top = AD(x)^2A^\top.$$

Therefore, for $z = q(x)$,

$$M(z) = \big[J_q(q^{-1}(z))J_q(q^{-1}(z))^\top\big]^{-1} = A^{-T}D(x)^{-2}A^{-1}, \qquad x = q^{-1}(z).$$

Let $\xi := A^{-1}(z-b)$. Since $z = As(x) + b$, we have $\xi = s(x)$, hence $x_i = s_i^{-1}(\xi_i)$. Define

$$m_i(\xi_i) := \frac{1}{\big(s_i'(s_i^{-1}(\xi_i))\big)^2}.$$

Using this notation, we can rewrite $M(z)$ as follows:

$$M(z) = A^{-T}\operatorname{diag}\big(m_1(\xi_1), \dots, m_d(\xi_d)\big)A^{-1}, \qquad \xi = A^{-1}(z-b).$$

This matrix field is explicitly Hessian. Indeed, define r_i by

$$r_i''(t) = m_i(t) = \frac{1}{\big(s_i'(s_i^{-1}(t))\big)^2},$$

and set

$$R(z) := \sum_{i=1}^{d} r_i\big((A^{-1}(z-b))_i\big).$$

Since $A^{-1}(z-b)$ is affine in z,

$$\nabla^2 R(z) = A^{-T}\operatorname{diag}\big(r_1''(\xi_1), \dots, r_d''(\xi_d)\big)A^{-1} = M(z).$$

Thus M satisfies the Hessian-compatibility condition. In particular, the compatibility holds although $J_q(x) = AD(x)$ is not diagonal unless A is diagonal.

Example 2: Rank-one nonlinear mixing. Fix $a \in \mathbb{R}^d \setminus \{0\}$ and let

$$q(x) = x + h(a^\top x)a,$$

where $h : \mathbb{R} \to \mathbb{R}$ is C^2. Writing $r := a^\top x$ and $s := \|a\|^2$, we have

$$J_q(x) = I + h'(r)aa^\top.$$

This Jacobian is generally non-diagonal whenever at least two components of a are nonzero.

The matrix $J_q(x)$ acts as multiplication by $1 + sh'(r)$ in the direction a, and as the identity on $a^\perp$. Assume $1 + sh'(r) \neq 0$ on the domain, so that q is locally invertible. Then $J_q(x)J_q(x)^\top$ has eigenvalue $(1 + sh'(r))^2$ along a and eigenvalue 1 on $a^\perp$. Consequently, the inverse metric field has the form

$$M(z) = I + \beta(a^\top z)\,aa^\top, \qquad z = q(x),$$

for a scalar function β. To identify it, note that

$$a^\top z = a^\top q(x) = r + sh(r).$$

Let $g(r) := r + sh(r)$, and assume g is invertible on the relevant interval. Then

$$r = g^{-1}(a^\top z),$$

and the eigenvalue of $M(z)$ along a gives

$$1 + s\,\beta(a^\top z) = \big(1 + sh'(r)\big)^{-2}.$$

Thus we can deduce an expression for β as follows:

$$\beta(t) = \frac{\big(1 + sh'(g^{-1}(t))\big)^{-2} - 1}{s}.$$

Again M is explicitly Hessian. Let $\psi : \mathbb{R} \to \mathbb{R}$ satisfy $\psi''(t) = \beta(t)$, and define:

$$R(z) := \frac{1}{2}\|z\|^2 + \psi(a^\top z).$$

Then, it follows from the above definition that the Hessian can be written as follows:

$$\nabla^2 R(z) = I + \psi''(a^\top z)aa^\top = I + \beta(a^\top z)aa^\top = M(z).$$

Hence the Hessian-compatibility condition holds, despite the generally non-diagonal Jacobian $J_q(x)$.

Verification of Assumption 2. For Example 1, restrict $\mathcal{X}$ to a compact box $\prod_{i=1}^d I_i$ on which each s_i' is bounded away from zero and assume that the resulting inverse metric satisfies $M(z) \succeq I$. Then $s(\mathcal{X})$ is a compact box and $\mathcal{Y} = As(\mathcal{X}) + b$ is compact and convex. Smoothness and nonsingularity imply uniform bounds on the first and second derivatives of q^{-1}, while the explicit formula for R gives bounded first and third derivatives on $\mathcal{Y}$.

For Example 2, write $s = \|a\|^2$ and choose

$$\mathcal{X} = \left\{ w + \frac{r}{s}a : w \in W,\ r \in I \right\},$$

where $W \subset a^\perp$ is compact and convex and I is a compact interval. If $g(r) = r + sh(r)$ satisfies $\inf_{r\in I} |g'(r)| > 0$, then

$$\mathcal{Y} = \left\{ w + \frac{g(r)}{s}a : w \in W,\ r \in I \right\}$$

is also compact and convex and q is a diffeomorphism from $\mathcal{X}$ onto $\mathcal{Y}$. Assuming again $M(z) \succeq I$, the required derivative bounds follow from compactness and nonsingularity. Finally,

$$\nabla_w D_R(z\|w) = \nabla^2 R(w)(w - z),$$

so Assumption 2(iii) follows from compactness and boundedness of $\nabla^2 R$.

B.3 Proof of Theorem 2: Linear Regret Lower Bound for OGD

We provide a complete proof of Theorem 2 following the proof sketch given in Section 4.2.

Proof. **Step 1: A smooth square-to-square diffeomorphism.** Let $\mathcal{X} = \mathcal{Y} = [0, 1]^2$ and define

$$q(x_1, x_2) = \big(x_1,\ x_2 - x_1(1 - x_1)x_2(1 - x_2)\big). \tag{35}$$

Its Jacobian is

$$J_q(x) = \begin{pmatrix} 1 & 0 \\ p(x) & r(x) \end{pmatrix}, \qquad \begin{aligned} p(x) &:= -(1 - 2x_1)x_2(1 - x_2), \\ r(x) &:= 1 - x_1(1 - x_1)(1 - 2x_2). \end{aligned}$$

For every $x \in \mathcal{X}$, $r(x) \geq 3/4$. For each fixed x_1, the function $x_2 \mapsto q_2(x_1, x_2)$ is therefore strictly increasing, with $q_2(x_1, 0) = 0$ and $q_2(x_1, 1) = 1$. Thus q maps the square bijectively onto itself. The inverse function theorem gives a smooth inverse, including smooth local extensions at the boundary, since r remains positive on a neighborhood of $\mathcal{X}$. By compactness and nonsingularity, there exists $\mu > 0$ such that $J_q(x)J_q(x)^\top \succeq \mu I$ uniformly on $\mathcal{X}$.

Throughout this proof, M denotes the inverse metric in (3). For $z = q(x)$,

$$M(z) = J_q(x)^{-\top} J_q(x)^{-1} = \begin{pmatrix} 1 + p(x)^2/r(x)^2 & -p(x)/r(x)^2 \\ -p(x)/r(x)^2 & 1/r(x)^2 \end{pmatrix}. \tag{36}$$

Step 2: Failure of Hessian compatibility and negative circulation. We first compute the derivatives of the inverse metric at an interior point, then use them to exhibit a comparator field with negative curl. Set $x^0 = (1/2, 1/2)$. Then $z^0 = q(x^0) = (1/2, 7/16)$. Recall from Step 1 that

$$p(x) = -(1-2x_1)x_2(1-x_2), \qquad r(x) = 1 - x_1(1-x_1)(1-2x_2).$$

At x^0, we have $p(x^0) = 0$, $r(x^0) = 1$, and hence $J_q(x^0) = I$ and $M(z^0) = I$. However, M is not constant near z^0, so its derivatives need not vanish.

To distinguish derivatives in the original and hidden coordinates, write $\widehat{M}(x) := M(q(x))$. By (36),

$$\widehat{M}(x) = \begin{pmatrix} 1 + p(x)^2/r(x)^2 & -p(x)/r(x)^2 \\ -p(x)/r(x)^2 & 1/r(x)^2 \end{pmatrix}.$$

Direct differentiation of p and r gives:

$$\begin{aligned} \partial_{x_1} p(x) &= 2x_2(1-x_2), & \partial_{x_2} p(x) &= -(1-2x_1)(1-2x_2), \\ \partial_{x_1} r(x) &= -(1-2x_1)(1-2x_2), & \partial_{x_2} r(x) &= 2x_1(1-x_1). \end{aligned}$$

Therefore,

$$\partial_{x_1} p(x^0) = \frac{1}{2}, \qquad \partial_{x_2} p(x^0) = 0, \qquad \partial_{x_1} r(x^0) = 0, \qquad \partial_{x_2} r(x^0) = \frac{1}{2}.$$

For $j \in \{1, 2\}$, differentiating the entries of $\widehat{M}$ and then evaluating at $p(x^0) = 0$ and $r(x^0) = 1$ yields

$$\begin{aligned} \partial_{x_j} \widehat{M}_{11}(x^0) &= \left[\frac{2p}{r^2}\, \partial_{x_j} p - \frac{2p^2}{r^3}\, \partial_{x_j} r \right]_{x=x^0} = 0, \\ \partial_{x_j} \widehat{M}_{12}(x^0) &= \partial_{x_j} \widehat{M}_{21}(x^0) \\ &= \left[-\frac{1}{r^2}\, \partial_{x_j} p + \frac{2p}{r^3}\, \partial_{x_j} r \right]_{x=x^0} = -\partial_{x_j} p(x^0), \\ \partial_{x_j} \widehat{M}_{22}(x^0) &= \left[-\frac{2}{r^3}\, \partial_{x_j} r \right]_{x=x^0} = -2\partial_{x_j} r(x^0). \end{aligned}$$

We now convert these derivatives into hidden coordinates. Since $M(z) = \widehat{M}(q^{-1}(z))$, the chain rule gives

$$\partial_{z_j} M(z^0) = \sum_{k=1}^{2} \partial_{x_k} \widehat{M}(x^0) [J_{q^{-1}}(z^0)]_{kj}.$$

Here $J_{q^{-1}}(z^0) = J_q(x^0)^{-1} = I$, so $\partial_{z_j} M(z^0) = \partial_{x_j} \widehat{M}(x^0)$. Combining the preceding calculations, we obtain

$$\partial_{z_1} M(z^0) = \begin{pmatrix} 0 & -1/2 \\ -1/2 & 0 \end{pmatrix}, \qquad \partial_{z_2} M(z^0) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}. \tag{37}$$

If $M = \nabla^2 R$ for a regularizer R, equality of mixed third derivatives would require $\partial_{z_2} M_{11} = \partial_{z_1} M_{12}$. But (37) gives

$$\partial_{z_2} M_{11}(z^0) = 0 \neq -\frac{1}{2} = \partial_{z_1} M_{12}(z^0).$$

Thus M is not Hessian-compatible.

Next, fix the comparator $u = (0, 1/2)$. Since the nonlinear term in q vanishes when $x_1 = 0$,

$$z_u := q(u) = (0, 1/2), \qquad z^0 - z_u = (1/2, -1/16).$$

Define

$$F_u(z) := M(z)(z - z_u) = \begin{pmatrix} P_u(z) \\ Q_u(z) \end{pmatrix}, \tag{38}$$

where P_u and Q_u are the first and second scalar components of F_u, respectively. Explicitly,

$$P_u(z) = M_{11}(z)z_1 + M_{12}(z)\left(z_2 - \frac{1}{2}\right),$$
$$Q_u(z) = M_{21}(z)z_1 + M_{22}(z)\left(z_2 - \frac{1}{2}\right).$$

Holding the comparator fixed and applying the product rule,

$$\partial_{z_1} Q_u(z) = \big(\partial_{z_1} M_{21}(z)\big)z_1 + M_{21}(z) + \big(\partial_{z_1} M_{22}(z)\big)\left(z_2 - \frac{1}{2}\right),$$
$$\partial_{z_2} P_u(z) = \big(\partial_{z_2} M_{11}(z)\big)z_1 + \big(\partial_{z_2} M_{12}(z)\big)\left(z_2 - \frac{1}{2}\right) + M_{12}(z).$$

Using $M(z^0) = I$ and (37), we obtain

$$\partial_{z_1} Q_u(z^0) = \left(-\frac{1}{2}\right)\left(\frac{1}{2}\right) + 0 + 0\left(-\frac{1}{16}\right) = -\frac{1}{4},$$
$$\partial_{z_2} P_u(z^0) = 0\left(\frac{1}{2}\right) + 0\left(-\frac{1}{16}\right) + 0 = 0.$$

Consequently, the scalar curl of F_u at z^0 is strictly negative:

$$\partial_{z_1} Q_u(z^0) - \partial_{z_2} P_u(z^0) = -\frac{1}{4}.$$

Since $z^0 \in \operatorname{int} \mathcal{Y}$ and the curl is continuous, we can choose a fixed square

$$\mathcal{Q} = [a, a+L] \times [b, b+L] \subset \operatorname{int} \mathcal{Y}, \qquad z^0 \in \operatorname{int} \mathcal{Q},$$

with $L > 0$, and a constant $\gamma > 0$ such that

$$\partial_{z_1} Q_u(z) - \partial_{z_2} P_u(z) \le -\gamma \qquad \text{for every } z \in \mathcal{Q}.$$

The field F_u is continuously differentiable on a neighborhood of $\mathcal{Q}$. Green's theorem therefore converts the circulation around its counterclockwise-oriented boundary into the area integral of its curl:

$$\begin{aligned} I_{\mathcal{Q}} &:= \oint_{\partial \mathcal{Q}} F_u(z) \cdot dz \\ &= \oint_{\partial \mathcal{Q}} \big(P_u(z)\, dz_1 + Q_u(z)\, dz_2\big) \\ &= \int_a^{a+L} \int_b^{b+L} \big(\partial_{z_1} Q_u - \partial_{z_2} P_u\big)(z_1, z_2)\, dz_2\, dz_1 \\ &\le -\gamma L^2 < 0. \end{aligned} \tag{39}$$

Thus the uniformly negative curl yields a strictly negative circulation along this fixed closed path.

Write $\kappa := -I_{\mathcal{Q}} > 0$ and initialize OGD at

$$x_1 = q^{-1}(a, b) \in \operatorname{int} \mathcal{X},$$

so that its initial hidden iterate is the lower-left corner $z_1 = (a, b)$ of $\mathcal{Q}$. The map, square, comparator, initialization, and constants γ, κ are independent of T and η.

Step 3: Exact steering around a closed grid. Fix $0 < \eta \le L$ where we recall here that L is the width of the square $\mathcal{Q}$ and define

$$n := \left\lceil \frac{L}{\eta} \right\rceil, \qquad \Delta := \frac{L}{n}, \qquad N := 4n.$$

Then $\eta/2 \le \Delta \le \eta$ and $N \le 8L/\eta$. Starting from the point (a, b), prescribe n steps in direction e_1, then n in direction e_2, then n in direction $-e_1$, and finally n in direction $-e_2$, each of length Δ. This

grid traverses $\partial\mathcal{Q}$ and returns exactly to (a, b). For each such round, let $v_t \in \{\pm e_1, \pm e_2\}$ denote the prescribed direction and set:

$$z_t^+ := z_t + \Delta v_t \,.$$

Then, choose:

$$s_t := \frac{1}{\eta} J_q(x_t)^{-\top} \big[x_t - q^{-1}(z_t^+)\big], \qquad h_t(z) := \langle s_t, z\rangle, \qquad \ell_t := h_t \circ q. \tag{40}$$

Once selected, s_t is held fixed in the definition of h_t. Thus $\nabla\ell_t(x) = J_q(x)^\top s_t$, and

$$x_t - \eta\nabla\ell_t(x_t) = x_t - \eta J_q(x_t)^\top s_t = q^{-1}(z_t^+) \in \operatorname{int}\mathcal{X}.$$

We briefly justify why $q^{-1}(z_t^+) \in \operatorname{int}\mathcal{X}$. Since the nonlinear term in q vanishes on the boundary of the square, $q(x) = x$ for every $x \in \partial\mathcal{X} = \partial\mathcal{Y}$. Thus the preimage of $z_t^+ \in \operatorname{int}\mathcal{Y}$ cannot lie on $\partial\mathcal{X}$. The unprojected update is therefore already feasible. Since Euclidean projection leaves every feasible point unchanged,

$$x_{t+1} = \Pi_{\mathcal{X}}\big(q^{-1}(z_t^+)\big) = q^{-1}(z_t^+), \qquad z_{t+1} = q(x_{t+1}) = z_t^+.$$

Hence projection is inactive and the prescribed hidden-space update is followed exactly. Induction proves that the prescribed grid is followed exactly and can be repeated without any accumulated discretization drift. Only the adversary uses q; the learner still performs ordinary projected OGD.

The loss vectors are bounded uniformly in η and T. Indeed, let $K := \sup_{z\in\mathcal{Y}} \|J_{q^{-1}}(z)\|_{\mathrm{op}} < \infty$. Since the segment from z_t to z_t^+ lies in $\mathcal{Y}$,

$$\|s_t\|_2 \le \frac{K}{\eta}\|q^{-1}(z_t) - q^{-1}(z_t^+)\|_2 \le K^2\frac{\Delta}{\eta} \le K^2.$$

Consequently, both the losses and their original-space gradients are uniformly bounded on $\mathcal{X}$, independently of η and T.

Step 4: Regret as a discrete line integral. We now express the instantaneous regret against the fixed comparator u as a discrete line integral of the field $F_u(z) = M(z)(z - z_u)$, where $z_u = q(u)$.

Recall that the exact-steering construction gives

$$z_{t+1} = z_t^+ = z_t + \Delta v_t, \qquad s_t = \frac{1}{\eta} J_q(x_t)^{-\top}\big[x_t - q^{-1}(z_t + \Delta v_t)\big].$$

To relate this exact choice of s_t to the inverse metric $M(z_t)$, we expand q^{-1} around z_t. Let $B < \infty$ be a uniform bound on the second derivative of q^{-1} on $\mathcal{Y}$, in operator norm. Since $q^{-1}(z_t) = x_t$, $J_{q^{-1}}(z_t) = J_q(x_t)^{-1}$, and $\|v_t\|_2 = 1$, Taylor's theorem along the grid segment gives

$$q^{-1}(z_t + \Delta v_t) = x_t + \Delta J_q(x_t)^{-1} v_t + \rho_t, \qquad \|\rho_t\|_2 \le \frac{B}{2}\Delta^2.$$

Substituting this expansion into the definition of s_t, we obtain

$$s_t = -\frac{\Delta}{\eta} J_q(x_t)^{-\top} J_q(x_t)^{-1} v_t - \frac{1}{\eta} J_q(x_t)^{-\top}\rho_t = -\frac{\Delta}{\eta} M(z_t) v_t + e_t,$$

where $e_t := -\frac{1}{\eta} J_q(x_t)^{-\top}\rho_t$.

Recall that $K := \sup_{z\in\mathcal{Y}} \|J_{q^{-1}}(z)\|_{\mathrm{op}} < \infty$. Since $\|J_q(x_t)^{-\top}\|_{\mathrm{op}} = \|J_{q^{-1}}(z_t)\|_{\mathrm{op}} \le K$, the vector error satisfies

$$\|e_t\|_2 \le \frac{1}{\eta}\|J_q(x_t)^{-\top}\|_{\mathrm{op}} \cdot \|\rho_t\|_2 \le \frac{KB}{2}\frac{\Delta^2}{\eta}.$$

We next use this decomposition of s_t to compute the loss difference against u. By the definitions $\ell_t = h_t \circ q$ and $h_t(z) = \langle s_t, z\rangle$,

$$\begin{aligned}
\ell_t(x_t) - \ell_t(u) &= h_t(q(x_t)) - h_t(q(u)) \\
&= h_t(z_t) - h_t(z_u) \\
&= \langle s_t, z_t - z_u\rangle \\
&= -\frac{\Delta}{\eta}\langle M(z_t)v_t, z_t - z_u\rangle + \langle e_t, z_t - z_u\rangle \\
&= -\frac{\Delta}{\eta}\langle M(z_t)(z_t - z_u), v_t\rangle + \langle e_t, z_t - z_u\rangle \\
&= -\frac{\Delta}{\eta}\langle F_u(z_t), v_t\rangle + \langle e_t, z_t - z_u\rangle,
\end{aligned}$$

where the penultimate equality uses symmetry of $M(z_t)$.

Define the scalar error

$$\epsilon_t := \langle e_t, z_t - z_u\rangle.$$

Both z_t and z_u belong to $\mathcal{Y} = [0,1]^2$, whose Euclidean diameter is $\sqrt{2}$. Hence, by Cauchy–Schwarz,

$$|\epsilon_t| \le \|e_t\|_2 \cdot \|z_t - z_u\|_2 \le \frac{KB\sqrt{2}}{2}\frac{\Delta^2}{\eta}.$$

Finally, using the exact displacement $z_{t+1} - z_t = \Delta v_t$, we conclude that

$$\ell_t(x_t) - \ell_t(u) = -\frac{1}{\eta}\langle F_u(z_t), z_{t+1} - z_t\rangle + \epsilon_t, \qquad |\epsilon_t| \le C_0\frac{\Delta^2}{\eta}, \tag{41}$$

where $C_0 := KB\sqrt{2}/2$ is independent of η and T.

Thus, over any prefix of τ steering rounds, the cumulative regret against u satisfies

$$R_\tau(u) := \sum_{t=1}^{\tau}\big(\ell_t(x_t) - \ell_t(u)\big) = -\frac{1}{\eta}\sum_{t=1}^{\tau}\langle F_u(z_t), z_{t+1} - z_t\rangle + E_\tau,$$

where

$$E_\tau := \sum_{t=1}^{\tau}\epsilon_t, \qquad |E_\tau| \le C_0\tau\frac{\Delta^2}{\eta} \le C_0\eta\tau,$$

using $\Delta \le \eta$. In particular, one complete cycle has $N = 4n = 4L/\Delta$ rounds, so its cumulative Taylor error is bounded uniformly:

$$|E_N| \le C_0 N\frac{\Delta^2}{\eta} = 4LC_0\frac{\Delta}{\eta} \le 4LC_0.$$

The remaining task is to compare the discrete line integral over one cycle with the negative circulation $I_{\mathcal{Q}}$.

Step 5: A uniform positive regret per cycle. We first compare the discrete line integral over one cycle with the continuous circulation $I_{\mathcal{Q}}$. Since F_u is continuously differentiable on a neighborhood of the fixed compact square $\mathcal{Q}$, it has a Lipschitz constant $L_F < \infty$ on $\mathcal{Q}$, independent of η and T.

Consider one grid segment from z_t to z_{t+1}, and write

$$d_t := z_{t+1} - z_t = \Delta v_t, \qquad \|d_t\|_2 = \Delta.$$

Parameterizing this oriented segment by $z_t + sd_t$, with $s \in [0,1]$, its exact line integral is

$$I_t := \int_{[z_t, z_{t+1}]} F_u(z)\cdot dz = \int_0^1 \langle F_u(z_t + sd_t), d_t\rangle\, ds.$$

The corresponding term in our discrete line integral is $\langle F_u(z_t), d_t\rangle$: it replaces the vector field along the segment by its value at the starting point. Since this term is constant with respect to s,

$$I_t - \langle F_u(z_t), d_t\rangle = \int_0^1 \langle F_u(z_t + sd_t) - F_u(z_t), d_t\rangle\, ds.$$

Using Cauchy–Schwarz's inequality and Lipschitzness of F_u, we obtain

$$|I_t - \langle F_u(z_t), d_t\rangle| \le \int_0^1 \|F_u(z_t + sd_t) - F_u(z_t)\|_2 \cdot \|d_t\|_2 \, ds \le \int_0^1 L_F s \|d_t\|_2^2 \, ds = \frac{L_F}{2}\Delta^2.$$

Thus the approximation error on each segment is of order Δ^2.

The $N = 4n$ grid segments traverse the boundary of $\mathcal{Q}$ exactly once in the counterclockwise direction, with $z_{N+1} = z_1$. By additivity of line integrals,

$$\sum_{t=1}^{N} I_t = \oint_{\partial\mathcal{Q}} F_u(z) \cdot dz = I_{\mathcal{Q}}.$$

Consequently, the triangle inequality and the preceding segment-wise bound give

$$\begin{aligned}\left|\sum_{t=1}^{N}\langle F_u(z_t), z_{t+1} - z_t\rangle - I_{\mathcal{Q}}\right| &= \left|\sum_{t=1}^{N}\big(\langle F_u(z_t), d_t\rangle - I_t\big)\right| \\ &\le \sum_{t=1}^{N}|\langle F_u(z_t), d_t\rangle - I_t| \\ &\le \frac{NL_F}{2}\Delta^2 \\ &= 2LL_F\Delta,\end{aligned}$$

where the last equality uses $N = 4n$ and $n\Delta = L$. Hence the discrete line integral approximates the circulation with an $O(\Delta)$ error, with a constant independent of η and T. Combining this with (41) and $n\Delta = L$,

$$\begin{aligned}\sum_{t=1}^{N}\big(\ell_t(x_t) - \ell_t(u)\big) &\ge \frac{\kappa}{\eta} - \frac{2LL_F\Delta}{\eta} - \frac{4nC_0\Delta^2}{\eta} \\ &= \frac{\kappa}{\eta} - (2LL_F + 4LC_0)\frac{\Delta}{\eta} \\ &\ge \frac{\kappa}{\eta} - C_1,\end{aligned}$$

where $C_1 := 2LL_F + 4LC_0$ is independent of η and T. Choose

$$\eta_0 := \min\left\{L, \frac{\kappa}{2(C_1+1)}\right\} > 0.$$

For every $0 < \eta \le \eta_0$, each cycle therefore contributes at least $\kappa/(2\eta)$ regret against u.

Step 6: Repetition over the horizon. For a given horizon T, perform $m := \lfloor T/N \rfloor$ complete cycles and set $s_t = 0$ on all remaining rounds. Each complete cycle starts at the same point and obeys the same lower bound. Since external regret is at least regret against the fixed comparator u,

$$R_T \ge \sum_{t=1}^{T}\big(\ell_t(x_t) - \ell_t(u)\big) \ge m\frac{\kappa}{2\eta} \ge \left(\frac{T}{N} - 1\right)\frac{\kappa}{2\eta} \ge \frac{\kappa}{16L}T - \frac{\kappa}{2\eta},$$

where the last inequality uses $N \le 8L/\eta$. Thus the theorem holds with $c = \kappa/(16L)$, $C = \kappa/2$, and $S = K^2$. In particular, a horizon-dependent constant stepsize $\eta = \Theta(T^{-1/2})$ gives $R_T \ge cT - O(\sqrt{T}) = \Omega(T)$ for sufficiently large T. □

C Proof of Theorem 3: Expected Regret Bound in the Bandit Setting

Recall first that the regret of bandit OGD (Algorithm 1) is defined for any time horizon $T \ge 1$ by:

$$R_T := \sum_{t=1}^{T}\ell_t(\hat{x}_t) - \min_{x\in\mathcal{X}}\sum_{t=1}^{T}\ell_t(x) \;= \sum_{t=1}^{T}h_t(\hat{z}_t) - \min_{z\in\mathcal{Y}}\sum_{t=1}^{T}h_t(z)\,, \tag{42}$$

where the last equality follows from using the notation $\hat{z}_t := q(\hat{x}_t)$ for any $t \geq 1$ and recalling that $\ell_t = h_t \circ q$ with q being a diffeomorphism.

To prove Theorem 3, we first decompose the expected regret into four different error terms.

Lemma 9 (Expected regret bound decomposition)**.** *Fix a time horizon $T \geq 1$. Let $z^\star \in \arg\min_{z\in\mathcal{Y}} \sum_{t=1}^T h_t(z)$, $z_\delta^\star \in \arg\min_{z\in\mathcal{Y}_\delta} \sum_{t=1}^T h_t(z)$ and let $\mathcal{F}_t := \sigma(\zeta_1, \dots, \zeta_t)$ be the filtration of σ-algebras induced by the random variables (ζ_t) used in bandit OGD. Then the expected regret $\mathbb{E}[R_T]$ of bandit OGD (Algorithm 1) can be upper-bounded as follows:*

$$\mathbb{E}[R_T] \;\leq\; R_T^{(1)} + R_T^{(2)} + R_T^{(3)} + R_T^{(4)}\,,$$

where each error term in the right-hand side above is defined as follows:

$$R_T^{(1)} := \mathbb{E}\left[\sum_{t=1}^T \langle \tilde{g}_t, z_t - z_\delta^\star\rangle\right], \qquad \textit{(perturbed OMD regret for stochastic linearized losses)} \tag{43}$$

$$R_T^{(2)} := \sum_{t=1}^T h_t(z_\delta^\star) - h_t(z^\star)\,, \qquad \textit{(domain shrinkage error)} \tag{44}$$

$$R_T^{(3)} := \mathbb{E}\left[\sum_{t=1}^T \langle b_t, z_\delta^\star - z_t\rangle\right], \qquad \textit{(cumulative gradient estimation bias)} \tag{45}$$

$$b_t := \mathbb{E}[\tilde{g}_t|\mathcal{F}_{t-1}] - \nabla h_t(z_t)\,, \qquad \textit{(gradient estimation bias)} \tag{46}$$

$$R_T^{(4)} := \mathbb{E}\left[\sum_{t=1}^T h_t(\hat{z}_t) - h_t(z_t)\right] \qquad \textit{(smoothing error)}\,. \tag{47}$$

Proof. The loss sequence is fixed in advance, and the initialization is deterministic. We choose fixed minimizers $z^\star$ and $z_\delta^\star$, independently of the random directions. In particular, $z_t - z_\delta^\star$ is $\mathcal{F}_{t-1}$-measurable.

Using the regret definition (42), we have the following decomposition:

$$\mathbb{E}[R_T] = \mathbb{E}\left[\sum_{t=1}^T h_t(\hat{z}_t) - h_t(z_\delta^\star)\right] + \sum_{t=1}^T h_t(z_\delta^\star) - h_t(z^\star) = \mathbb{E}\left[\sum_{t=1}^T h_t(\hat{z}_t) - h_t(z_\delta^\star)\right] + R_T^{(2)}\,. \tag{48}$$

To control regret against the shrunk comparator $z_\delta^\star$, we start by decomposing the instantaneous regret. Since z_t is $\mathcal{F}_{t-1}$-measurable, we have:

$$\mathbb{E}[h_t(\hat{z}_t) - h_t(z_\delta^\star)|\mathcal{F}_{t-1}] = \mathbb{E}[h_t(\hat{z}_t) - h_t(z_t)|\mathcal{F}_{t-1}] + h_t(z_t) - h_t(z_\delta^\star)\,. \tag{49}$$

Then by convexity of h_t, we have:

$$h_t(z_t) - h_t(z_\delta^\star) \;\leq\; \langle \nabla h_t(z_t), z_t - z_\delta^\star\rangle\,. \tag{50}$$

Recall now that we denote by b_t the bias due to one-point smoothing defined by:

$$b_t = \mathbb{E}[\tilde{g}_t|\mathcal{F}_{t-1}] - \nabla h_t(z_t)\,.$$

Note here that $\tilde{g}_t := J_q(x_t)^{-\top} g_t$ is not computed and just used here in the analysis whereas g_t is the one-point estimator actually computed in BOGD. It follows then that:

$$\langle \nabla h_t(z_t), z_t - z_\delta^\star\rangle = \langle \mathbb{E}[\tilde{g}_t|\mathcal{F}_{t-1}], z_t - z_\delta^\star\rangle - \langle b_t, z_t - z_\delta^\star\rangle\,. \tag{51}$$

Combining (49), (50) and (51), taking expectations and noting that $z_t - z_\delta^\star$ is $\mathcal{F}_{t-1}$-measurable, we obtain:

$$\mathbb{E}[h_t(\hat{z}_t) - h_t(z_\delta^\star)] \leq \mathbb{E}[\langle \tilde{g}_t, z_t - z_\delta^\star\rangle] + \mathbb{E}[\langle b_t, z_\delta^\star - z_t\rangle] + \mathbb{E}[h_t(\hat{z}_t) - h_t(z_t)].$$

Summing the above inequality over t yields:

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^T \big(h_t(\hat{z}_t) - h_t(z_\delta^\star)\big)\right] &\leq \mathbb{E}\left[\sum_{t=1}^T \langle \tilde{g}_t, z_t - z_\delta^\star\rangle\right] + \mathbb{E}\left[\sum_{t=1}^T \langle b_t, z_\delta^\star - z_t\rangle\right] + \mathbb{E}\left[\sum_{t=1}^T \big(h_t(\hat{z}_t) - h_t(z_t)\big)\right] \\ &= R_T^{(1)} + R_T^{(3)} + R_T^{(4)}.\end{aligned} \tag{52}$$

Finally, combining (52) with (48) gives the desired bound:

$$\mathbb{E}[R_T] \;\leq\; R_T^{(1)} + R_T^{(2)} + R_T^{(3)} + R_T^{(4)}\,.$$

□

Given the regret decomposition of Lemma 9, we control each of the error terms $R_T^{(i)}, i \in \{1,2,3,4\}$ defined in (43) to (47) respectively, in a dedicated lemma.

Lemma 10 (Domain shrinkage error)**.** *Let $T \;\geq\; 1$, let $z^\star \in \arg\min_{z\in\mathcal{Y}} \sum_{t=1}^T h_t(z)$, and $z_\delta^\star \in \arg\min_{z\in\mathcal{Y}_\delta} \sum_{t=1}^T h_t(z)$. Then we have:*

$$R_T^{(2)} = \sum_{t=1}^T h_t(z_\delta^\star) - h_t(z^\star) \;\leq\; \delta G_F D T\,,$$

where we recall that δ is the smoothing parameter in the one-point zeroth-order gradient estimator, G_F is the uniform Lipschitz constant of ℓ_t and D is the finite diameter of the compact set $\mathcal{X}$.

Proof. Since $\mathcal{Y}_\delta = q(\mathcal{X}_\delta)$, q is a diffeomorphism, $\mathcal{X}_\delta = (1-\delta)\mathcal{X}$ and $\ell_t = h_t \circ q$, the comparator term over the shrunk set $\mathcal{X}_\delta$ can be rewritten as follows:

$$\min_{z\in\mathcal{Y}_\delta} \sum_{t=1}^T h_t(z) = \min_{x\in\mathcal{X}_\delta} \sum_{t=1}^T \ell_t(x) = \min_{x\in\mathcal{X}} \sum_{t=1}^T \ell_t((1-\delta)x)\,. \tag{53}$$

Then for any $x \in \mathcal{X}$, we have:

$$\sum_{t=1}^T \ell_t((1-\delta)x) \;\leq\; \sum_{t=1}^T |\ell_t((1-\delta)x) - \ell_t(x)| + \sum_{t=1}^T \ell_t(x) \;\leq\; \delta G_F D T + \sum_{t=1}^T h_t(q(x))\,,$$

where the last inequality uses uniform G_F-Lipschitzness of the loss functions ℓ_t and boundedness of the set $\mathcal{X}$ which has finite diameter D.

Taking the minimum over $x \in \mathcal{X}$, using (53) and recalling that $z^\star \in \arg\min_{z\in\mathcal{Y}} \sum_{t=1}^T h_t(z)$ and $z_\delta^\star \in \arg\min_{z\in\mathcal{Y}_\delta} \sum_{t=1}^T h_t(z)$ yields the desired inequality:

$$R_T^{(2)} = \sum_{t=1}^T h_t(z_\delta^\star) - h_t(z^\star) \;\leq\; \delta G_F D T\,.$$

□

Lemma 11 (Gradient estimation bias)**.** *For any $T \;\geq\; 1$, we have:*

$$|R_T^{(3)}| = \left|\mathbb{E}\left[\sum_{t=1}^T \langle b_t, z_\delta^\star - z_t\rangle\right]\right| \;\leq\; \delta G^2 H D T\,,$$

where we recall that H is the uniform smoothness constant of the loss ℓ_t.

Proof. First, using the Cauchy–Schwarz inequality, it follows that:

$$|R_T^{(3)}| = \left|\mathbb{E}\left[\sum_{t=1}^T \langle b_t, z_\delta^\star - z_t\rangle\right]\right| \;\leq\; \mathbb{E}\left[\sum_{t=1}^T \|b_t\| \cdot \|z_t - z_\delta^\star\|\right]\,. \tag{54}$$

We control each one of the terms in the right-hand side separately.

Since $z_t = q(x_t)$ and $z_\delta^\star = q(x_\delta^\star)$ for some $x_t \in \mathcal{X}$ and $x_\delta^\star \in \mathcal{X}_\delta \subset \mathcal{X}$, we have:

$$\|z_t - z_\delta^\star\| = \|q(x_t) - q(x_\delta^\star)\| \;\leq\; G\|x_t - x_\delta^\star\| \;\leq\; GD\,, \tag{55}$$

where the first inequality follows from using G-Lipschitzness of q (Assumption 2) and the second one uses the definition of the diameter of $\mathcal{X}$ supposed to be finite.

We now control the norm of the bias $\|b_t\|$ for any t. For $x \in \mathcal{X}_\delta$, define the ball-smoothed loss:

$$\bar{\ell}_t(x) := \mathbb{E}_{v\sim\mathcal{U}(\mathbb{B})}[\ell_t(x+\delta v)]$$

where v is a random variable with uniform distribution $\mathcal{U}(\mathbb{B})$ on the unit ball $\mathbb{B}$. This loss is well defined: if $x = (1-\delta)x'$ with $x' \in \mathcal{X}$ and $v \in \mathbb{B} = B_2 \subseteq \mathcal{X}$, then $x + \delta v = (1-\delta)x' + \delta v \in \mathcal{X}$ by convexity. It is known from [14, Lemma 1] that the conditional expectation of the spherical estimator g_t is equal to the gradient of the ball-smoothed loss:

$$\mathbb{E}[g_t|\mathcal{F}_{t-1}] = \nabla\bar{\ell}_t(x_t)\,.$$

The (hidden-space ghost) gradient estimator of $\nabla h_t(z_t)$ is

$$\tilde{g}_t = J_q(x_t)^{-\top} g_t\,.$$

(Recall that $\nabla\ell_t(x_t) = J_q(x_t)^\top \nabla h_t(z_t)$ by the chain rule.) Since x_t is $\mathcal{F}_{t-1}$-measurable, it follows that:

$$\mathbb{E}[\tilde{g}_t|\mathcal{F}_{t-1}] = J_q(x_t)^{-\top}\mathbb{E}[g_t|\mathcal{F}_{t-1}] = J_q(x_t)^{-\top}\nabla\bar{\ell}_t(x_t)\,. \tag{56}$$

Moreover, we have by the chain rule $\nabla\ell_t(x_t) = J_q(x_t)^\top\nabla h_t(z_t)$ which implies that $\nabla h_t(z_t) = J_q(x_t)^{-\top}\nabla\ell_t(x_t)$. Therefore, using this identity together with (56), we can express the bias b_t as follows:

$$\begin{aligned} b_t &= \mathbb{E}[\tilde{g}_t|\mathcal{F}_{t-1}] - \nabla h_t(z_t) \\ &= J_q(x_t)^{-\top}(\nabla\bar{\ell}_t(x_t) - \nabla\ell_t(x_t)) \\ &= J_q(x_t)^{-\top}\mathbb{E}_{v\sim\mathcal{U}(\mathbb{B})}[\nabla\ell_t(x_t+\delta v) - \nabla\ell_t(x_t)]\,. \end{aligned}$$

Taking the norm, using boundedness of the inverse Jacobian of q (Assumption 2) and smoothness of the loss function ℓ_t (Assumption 5), we have:

$$\begin{aligned} \|b_t\|_2 &\le \|J_q(x_t)^{-\top}\|_{\mathrm{op}}\,\mathbb{E}_{v\sim\mathcal{U}(\mathbb{B})}\left[\|\nabla\ell_t(x_t+\delta v) - \nabla\ell_t(x_t)\|_2\right] \\ &\le \delta GH\,\mathbb{E}_{v\sim\mathcal{U}(\mathbb{B})}[\|v\|_2] \\ &\le \delta GH. \end{aligned} \tag{57}$$

Combining (57) and (55) in (54), we obtain the desired inequality:

$$|R_T^{(3)}| \;\le\; \delta G^2 HDT\,.$$

□

Lemma 12 (Smoothing error)**.** *For any $T \ge 1$, we have:*

$$|R_T^{(4)}| \;\le\; \left|\mathbb{E}\left[\sum_{t=1}^T h_t(\hat{z}_t) - h_t(z_t)\right]\right| \;\le\; \delta G_F T\,.$$

Proof. The desired bound follows from the following:

$$\begin{aligned} |R_T^{(4)}| &= \left|\mathbb{E}\left[\sum_{t=1}^T h_t(\hat{z}_t) - h_t(z_t)\right]\right| \\ &= \left|\mathbb{E}\left[\sum_{t=1}^T \ell_t(\hat{x}_t) - \ell_t(x_t)\right]\right| && (\ell_t = h_t\circ q,\ \hat{z}_t = q(\hat{x}_t), z_t = q(x_t)) \\ &\le\; G_F\mathbb{E}\left[\sum_{t=1}^T \|\hat{x}_t - x_t\|\right] && (\text{by } G_F\text{-Lipschitzness of } \ell_t \text{ (Assumption 3)}) \\ &= G_F\mathbb{E}\left[\sum_{t=1}^T \delta\|\zeta_t\|\right] && (\hat{x}_t = x_t + \delta\zeta_t \text{ in BOGD (Algorithm 1)}) \\ &= \delta G_F T && (\|\zeta_t\| = 1)\,. \end{aligned} \tag{58}$$

□

Lemma 13 (Perturbed ghost OMD regret for random linear losses)**.** *Suppose* $(z_t)_{t=1}^{T+1} \subseteq \mathcal{Y}_\delta$ *satisfies:*

$$z_{t+1} = r_{t+1} + y_{t+1}\,, \quad y_{t+1} = \operatorname*{arg\,min}_{y \in \mathcal{Y}_\delta} \{\tilde{g}_t^\top (y - z_t) + \frac{1}{\eta} D_R(y||z_t)\}\,.$$

Then for every $z_\delta^\star \in \mathcal{Y}_\delta$,

$$\sum_{t=1}^{T} \langle \tilde{g}_t, z_t - z_\delta^\star \rangle \;\leq\; \frac{D_R(z_\delta^\star||z_1)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T} \|\tilde{g}_t\|_2^2 + \frac{G}{\eta}\sum_{t=1}^{T} \|r_{t+1}\|_2\,.$$

If in addition $\|\tilde{g}_t\| \;\leq\; \tilde{G}_F$ *and* $\|r_{t+1}\| \;\leq\; \hat{C}_\eta$*, then:*

$$R_T^{(1)} = \mathbb{E}\left[\sum_{t=1}^{T} \langle \tilde{g}_t, z_t - z_\delta^\star \rangle\right] \;\leq\; \frac{D_1}{\eta} + \frac{\eta \tilde{G}_F^2 T}{2} + \frac{\hat{C}_\eta G T}{\eta}\,. \tag{59}$$

Proof. First, we prove the following inequality which follows from using the first-order optimality condition (from the ghost OMD update) together with Bregman three-point identity. For any comparator $z_\delta^\star \in \mathcal{Y}_\delta$, we have:

$$\langle \tilde{g}_t, z_t - z_\delta^\star \rangle \;\leq\; \frac{D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||y_{t+1})}{\eta} + \frac{\eta}{2}\|\tilde{g}_t\|_2^2\,. \tag{60}$$

We now provide the proof of (60). By definition of y_{t+1}, using the first-order optimality condition gives for every $z_\delta^\star \in \mathcal{Y}_\delta$,

$$\left\langle \tilde{g}_t + \frac{1}{\eta}(\nabla R(y_{t+1}) - \nabla R(z_t)), z_\delta^\star - y_{t+1} \right\rangle \;\geq\; 0\,.$$

Rearranging this inequality, we obtain:

$$\langle \tilde{g}_t, y_{t+1} - z_\delta^\star \rangle \;\leq\; \frac{1}{\eta}\langle \nabla R(y_{t+1}) - \nabla R(z_t), z_\delta^\star - y_{t+1} \rangle\,. \tag{61}$$

Using the three-point identity, we have:

$$\langle \nabla R(y_{t+1}) - \nabla R(z_t), z_\delta^\star - y_{t+1} \rangle = D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||y_{t+1}) - D_R(y_{t+1}||z_t)\,. \tag{62}$$

Combining (61) and (62) yields:

$$\langle \tilde{g}_t, y_{t+1} - z_\delta^\star \rangle \;\leq\; \frac{1}{\eta}\left[D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||y_{t+1}) - D_R(y_{t+1}||z_t)\right]\,.$$

Using the above inequality, we obtain:

$$\begin{aligned}
\langle \tilde{g}_t, z_t - z_\delta^\star \rangle &= \langle \tilde{g}_t, z_t - y_{t+1} \rangle + \langle \tilde{g}_t, y_{t+1} - z_\delta^\star \rangle \\
&\leq \langle \tilde{g}_t, z_t - y_{t+1} \rangle + \frac{1}{\eta}\left[D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||y_{t+1}) - D_R(y_{t+1}||z_t)\right] \\
&= \langle \tilde{g}_t, z_t - y_{t+1} \rangle - \frac{1}{\eta} D_R(y_{t+1}||z_t) + \frac{1}{\eta}\left[D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||y_{t+1})\right]\,.
\end{aligned}$$

Using 1-strong convexity of R and Cauchy-Schwarz inequality yields:

$$\langle \tilde{g}_t, z_t - y_{t+1} \rangle - \frac{1}{\eta} D_R(y_{t+1}||z_t) \;\leq\; \|\tilde{g}_t\|_2 \cdot \|z_t - y_{t+1}\|_2 - \frac{1}{2\eta}\|z_t - y_{t+1}\|_2^2\,.$$

Note now that for any $a \;\geq\; 0, r \;\geq\; 0, ar - \frac{1}{2\eta} r^2 \;\leq\; \frac{\eta}{2} a^2$. Hence with $a = \|\tilde{g}_t\|_2$ and $r = \|z_t - y_{t+1}\|$, we obtain:

$$\langle \tilde{g}_t, z_t - y_{t+1} \rangle - \frac{1}{\eta} D_R(y_{t+1}||z_t) \;\leq\; \frac{\eta}{2}\|\tilde{g}_t\|_2^2\,.$$

We have proved that:

$$\langle \tilde{g}_t, z_t - z_\delta^\star \rangle \;\leq\; \frac{D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||y_{t+1})}{\eta} + \frac{\eta}{2}\|\tilde{g}_t\|_2^2 \,.$$

Using the previous inequality, we have:

$$\langle \tilde{g}_t, z_t - z_\delta^\star \rangle \;\leq\; \frac{D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||z_{t+1})}{\eta} + \frac{D_R(z_\delta^\star||z_{t+1}) - D_R(z_\delta^\star||y_{t+1})}{\eta} + \frac{\eta}{2}\|\tilde{g}_t\|_2^2 \,.$$

By G-Lipschitzness of $z \mapsto D_R(z_\delta^\star||z)$, we have:

$$D_R(z_\delta^\star||z_{t+1}) - D_R(z_\delta^\star||y_{t+1}) \;\leq\; G\|z_{t+1} - y_{t+1}\| \,,$$

which implies that:

$$\langle \tilde{g}_t, z_t - z_\delta^\star \rangle \;\leq\; \frac{D_R(z_\delta^\star||z_t) - D_R(z_\delta^\star||z_{t+1})}{\eta} + \frac{G}{\eta}\|z_{t+1} - y_{t+1}\| + \frac{\eta}{2}\|\tilde{g}_t\|_2^2 \,.$$

Summing up the above inequality, the first term telescopes and we obtain:

$$\sum_{t=1}^{T} \langle \tilde{g}_t, z_t - z_\delta^\star \rangle \;\leq\; \frac{D_R(z_\delta^\star||z_1)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T}\|\tilde{g}_t\|_2^2 + \frac{G}{\eta}\sum_{t=1}^{T}\|z_{t+1} - y_{t+1}\|_2 \,,$$

which concludes the proof upon recalling the notation $r_{t+1} = z_{t+1} - y_{t+1}$. As for the second bound, it follows immediately from using $\|\tilde{g}_t\| \;\leq\; \tilde{G}_F$ and $\|r_{t+1}\| \;\leq\; \hat{C}_\eta$. □

In the next lemma, we show how to control the error sequence (r_t) in Lemma 13 by providing a precise constant $\hat{C}_\eta$ which quantifies the approximation error in the bandit setting, as a corollary of Proposition 2.

Lemma 14 (Gradient estimators norm bounds)**.** *Suppose that ℓ_t is uniformly bounded by M, i.e. $|\ell_t(x)| \;\leq\; M, \forall x \in \mathcal{X}$, and the inverse Jacobian of q is bounded by G at any point $x \in \mathcal{X}$ (as in Assumption 2). Then for any $t \;\geq\; 1$, we have:*

$$\|g_t\| \;\leq\; \hat{G}_F := \frac{dM}{\delta} \,, \quad \|\tilde{g}_t\| \;\leq\; \tilde{G}_F := \frac{dMG}{\delta} \,. \tag{63}$$

Proof. Recall from BOGD (Algorithm 1) that $g_t = \frac{d}{\delta}\ell_t(\hat{x}_t)\zeta_t$. Then, since $\|\zeta_t\| = 1$, we have:

$$\|g_t\| = \frac{d}{\delta}|\ell_t(\hat{x}_t)| \;\leq\; \frac{dM}{\delta} = \hat{G}_F \,.$$

Then, since $\tilde{g}_t = J_q(x_t)^{-\top} g_t$, we have:

$$\|\tilde{g}_t\| \;\leq\; G\|g_t\| \;\leq\; \frac{dMG}{\delta} = \tilde{G}_F \,.$$

□

Corollary 5 (Proposition 2 in the bandit setting)**.** *Under Assumptions 1–6, let x_{t+1} be the BOGD update, set $z_t = q(x_t)$, and define:*

$$y_{t+1} = \operatorname*{arg\,min}_{y \in \mathcal{Y}_\delta} \left\{ \langle \tilde{g}_t, y - z_t \rangle + \frac{1}{\eta} D_R(y\|z_t) \right\}, \qquad \tilde{g}_t = J_q(x_t)^{-\top} g_t.$$

Then, with $\hat{G}_F = dM/\delta$ (as defined in Lemma 14), we have for any $t \;\geq\; 1$,

$$\|y_{t+1} - q(x_{t+1})\|_2 \leq \hat{C}_\eta := G^5(5\hat{G}_F^2 + \hat{G}_F^3\eta)\eta^2.$$

Proof. Apply Proposition 2 to the restricted map $q|_{\mathcal{X}_\delta} : \mathcal{X}_\delta \to \mathcal{Y}_\delta$, with common starting point $z_t = q(x_t)$. The set $\mathcal{X}_\delta$ is convex, and $\mathcal{Y}_\delta$ is convex by Assumption 6. The geometric identities and derivative bounds are inherited from $\mathcal{X}$ and $\mathcal{Y}$, and Lemma 14 gives $\|g_t\|_2 \leq dM/\delta$. □

End of Proof of Theorem 3. We are ready to conclude the proof of the theorem using the results established so far. First, it follows from Lemma 9 that:

$$\mathbb{E}[R_T] \;\leq\; R_T^{(1)} + R_T^{(2)} + R_T^{(3)} + R_T^{(4)}\,. \tag{64}$$

We have proved a bound for each one of the error terms in the above bound:

(i) By Lemma 10, we have $R_T^{(2)} \;\leq\; \delta G_F D T\,.$

(ii) By Lemma 11, we get: $|R_T^{(3)}| \;\leq\; \delta G^2 H D T\,.$

(iii) By Lemma 12, we obtain: $|R_T^{(4)}| \;\leq\; \delta G_F T\,.$

(iv) Lemma 13-(59) together with Corollary 5 yield:

$$R_T^{(1)} \;\leq\; \frac{D_1}{\eta} + \frac{\eta \tilde{G}_F^2 T}{2} + \frac{\hat{C}_\eta G T}{\eta} \;\leq\; \frac{D_1}{\eta} + (6\hat{G}_F^2 + \hat{G}_F^3\eta)G^6\eta T\,,$$

where the last inequality follows from using $\hat{C}_\eta = G^5(5\hat{G}_F^2 + \hat{G}_F^3\eta)\eta^2$ (Corollary 5) and upper bounding the resulting constant under the assumptions $G > 1$ (which is without loss of generality), similarly to the deterministic setting. Note here that $\hat{G}_F = \frac{dM}{\delta}$ as shown in Lemma 14.

Combining all the above bounds in (64), we obtain:

$$\mathbb{E}[R_T] \;\leq\; K_1\delta T + \frac{D_1}{\eta} + K_2\frac{\eta T}{\delta^2} + K_3\frac{\eta^2 T}{\delta^3}\,, \tag{65}$$

where $K_1 := G_F D + G^2 H D + G_F, K_2 := 6d^2G^6M^2, K_3 = d^3G^6M^3\,.$

We choose (η, δ) by balancing the first three terms of (65). The remaining term will be evaluated at the same parameter values and retained in the final bound. Define then for any (η, δ),

$$f(\eta, \delta) := K_1\delta T + \frac{D_1}{\eta} + K_2\frac{\eta T}{\delta^2}\,,$$

which consists of the three first terms in the upperbound (65). We now optimize this function w.r.t (η, δ). By the first order optimality conditions, the optimal pair $(\eta_\star, \delta_\star)$ minimizing $f(\eta, \delta)$ satisfies:

$$\frac{\partial f}{\partial \eta}(\eta_\star, \delta_\star) = -\frac{D_1}{\eta_\star^2} + \frac{K_2 T}{\delta_\star^2} = 0\,, \quad \frac{\partial f}{\partial \delta}(\eta_\star, \delta_\star) = K_1 T - \frac{2K_2\eta_\star T}{\delta_\star^3} = 0\,.$$

Rewriting these conditions yields:

$$\eta_\star = \sqrt{\frac{D_1\delta_\star^2}{K_2 T}}\,, \quad \eta_\star = \frac{K_1\delta_\star^3}{2K_2}\,.$$

Solving for $\delta_\star$ as a function of T (eliminating η) and then plugging back $\delta_\star$ in the first expression of $\eta_\star$ yields:

$$\delta_\star = \left(\frac{4K_2D_1}{K_1^2 T}\right)^{1/4}\,, \quad \eta_\star = \left(\frac{4D_1^3}{K_1^2K_2T^3}\right)^{1/4}\,.$$

Now, plugging back these values of $\eta_\star$ and $\delta_\star$ in (65), we obtain:

$$\mathbb{E}[R_T] \;\leq\; 2\sqrt{2}(K_1^2K_2D_1)^{1/4}T^{3/4} + \frac{1}{\sqrt{2}}\left(\frac{K_1^{1/2}K_3D_1^{3/4}}{K_2^{5/4}}\right)T^{1/4}\,.$$

which concludes the proof of Theorem 3 since $T^{1/4} \;\leq\; T^{3/4}$ for $T \;\geq\; 1$.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Our contributions are stated in the abstract as well as in the contributions section and the core of the paper consistently.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss opportunities for future work in the conclusion, highlight open questions in the main part, and discuss our assumptions.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: Each result is stated with the corresponding assumptions needed. Assumptions are introduced using the Assumption environment. All the proofs are detailed in the appendix.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [N/A]

Justification: [N/A]

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
   - The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: [N/A]

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: This is a theoretical paper and our research complies with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: This work is theoretical and we do not foresee any potentially immediate and direct societal impact.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: [N/A]

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: [N/A]

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.

- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: [N/A]

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: [N/A]

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: [N/A]

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [Yes]

Justification: This paper is theoretical. LLMs were used throughout the development of this project. See acknowledgements for more details.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.