IEEE Control Systems Letters paper presented at
2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Variational Principles for Mirror Descent and Mirror Langevin Dynamics

Belinda Tzen*, Anant Raj†§, Maxim Raginsky§, and Francis Bach†

***Abstract*— Mirror descent, introduced by Nemirovski and Yudin in the 1970s, is a primal-dual convex optimization method that can be tailored to the geometry of the optimization problem at hand through the choice of a strongly convex potential function. It arises as a basic primitive in a variety of applications, including large-scale optimization, machine learning, and control. This paper proposes a variational formulation of mirror descent and of its stochastic variant, mirror Langevin dynamics. The main idea, inspired by the classic work of Brezis and Ekeland on variational principles for gradient flows, is to show that mirror descent emerges as a closed-loop solution for a certain optimal control problem, and the Bellman value function is given by the Bregman divergence between the initial condition and the global minimizer of the objective function.**

I. Introduction

The continuous-time gradient flow

$$\dot{x}(t) = -\nabla f(x(t)), \qquad x(0) = x_0 \tag{1}$$

for a C^1 objective function $f : \mathbb{R}^n \to \mathbb{R}$ is a basic primitive in continuous-time optimization and control. Under appropriate assumptions, the trajectory of (1) converges to a minimizer of f, which justifies thinking of the gradient flow as a method for asymptotically solving the optimization problem

$$\text{minimize } f(x), \quad x \in \mathbb{R}^n. \tag{2}$$

However, apart from the local characterization of $-\nabla f(x)$ as the "direction of steepest descent," there appears to be little discussion of the sense, if any, in which (1) is "optimal" among all dynamical systems that asymptotically solve (2).

One of the few exceptions is the variational principle of Brezis and Ekeland [1], [2]: Fix a time horizon $T > 0$. Then, among all absolutely continuous curves $x : [0, T] \to \mathbb{R}^n$ with $x(0) = x_0$, the trajectory of (1) on $[0, T]$ minimizes the action functional

$$S(x(\cdot)) := \int_0^T \{f(x(t)) + f^*(-\dot{x}(t))\}\,\mathrm{d}t + \frac{1}{2}|x(T)|^2, \tag{3}$$

where

$$f^*(v) := \sup_{x\in\mathbb{R}^n} \{\langle v, x\rangle - f(x)\}$$

is the Legendre–Fenchel conjugate of f and $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$. The minimum value of S over all such $x(\cdot)$ is equal to $\frac{1}{2}|x_0|^2$. The underlying idea is simple and boils down to a careful analysis of the equality cases of the Fenchel–Young inequality

$$f(x) + f^*(v) \geq \langle v, x\rangle.$$

However, the Brezis–Ekeland variational principle does not say anything about the asymptotic behavior of the extremal trajectories or about the curious fact that a finite-horizon problem of minimizing the action (3) has a solution given by the flow of a time-invariant dynamical system.

In this paper, we revisit this problem from a control-theoretic point of view and provide a new variational interpretation of the gradient flow as an *infinite-horizon stabilizing optimal control* [3, §8.5]. Moreover, we consider a more general method of *mirror descent*. This method, introduced in the 1970s by Nemirovski and Yudin [4, Ch. 3], can be tailored to the geometry of the optimization problem at hand through the choice of a strongly convex *potential function*. In continuous time, mirror descent is implemented by a time-invariant dynamical system whose state $x(t) \in \mathbb{R}^n$ and output $y(t) \in \mathbb{R}^n$ evolve according to

$$\begin{aligned} &\dot{x}(t) = -\nabla f(\nabla\varphi^*(x(t))), \qquad x(0) = x_0 \\ &y(t) = \nabla\varphi^*(x(t)) \end{aligned} \tag{4}$$

where $\varphi : \mathbb{R}^n \to \mathbb{R}$ the potential function and φ^* is its Legendre–Fenchel conjugate. In the literature, $x(t)$ and $y(t)$ are referred to as the *dual-space* and the *primal-space* trajectories, respectively, and $y(t)$ is the candidate minimizer at time t. (The Euclidean gradient flow (1) is a special case of (4) with $\varphi(x) = \frac{1}{2}|x|^2$ and $y(t) = x(t)$ for all t.)

A. Brief summary of contributions

We show the following: Suppose that the objective f is strictly convex. For the controlled system $\dot{x}(t) = u(t), y(t) = \nabla\varphi^*(x(t))$, we consider the class of all *stabilizing controls* [3, §8.5], i.e., appropriately well-behaved functions $u : [0, \infty) \to \mathbb{R}^n$ such that the output $y(t)$ converges, as $t \to \infty$, to the unique minimizer of f. This class contains, among others, sufficiently smooth state feedback controls of the form $u(t) = k(x(t))$. We then identify an instantaneous cost function $q(x, u)$, closely related to the Lagrangian in (3), such that the state feedback law $k(x) = -\nabla f(\nabla\varphi^*(x))$

This work was supported in part by the NSF under award CCF-2106358 ("Analysis and Geometry of Neural Dynamical Systems") and in part by the Illinois Institute for Data Science and Dynamical Systems (iDS^2), an NSF HDR TRIPODS institute, under award CCF-1934986. A. Raj was supported by the Marie Sklodowska-Curie Fellowship (project NN-OVEROPT 101030817).

*Columbia University, New York, NY, USA.
§University of Illinois, Urbana, IL, USA.
†INRIA, Ecole Normale Supérieure, PSL Research University, Paris, France.
Emails: bt2314@columbia.edu, anant.raj@inria.fr, maxim@illinois.edu, francis.bach@inria.fr

[or, equivalently, the output feedback law $\tilde{k}(y) = -\nabla f(y)$] gives a control that minimizes the infinite-horizon cost

$$\int_0^\infty q(x(t), u(t))\,\mathrm{d}t \tag{5}$$

over all stabilizing controls. The value function $V(x_0)$, i.e., the minimum value of this cost as a function of the initial state x_0, is given by a certain "distance," induced by the potential φ, between the initial output $y_0 = \nabla\varphi^*(x_0)$ and the unique minimizer of f. At the same time, $V(x)$ is the Lyapunov function for the closed-loop system $\dot{x}(t) = -\nabla f(\nabla\varphi^*(x(t)))$.

We also consider a stochastic variant of mirror descent, the so-called *mirror Langevin dynamics* [5]–[7], which is implemented by an Itô stochastic differential equation

$$\begin{aligned} \mathrm{d}X_t &= -\nabla f(\nabla\varphi^*(X_t))\,\mathrm{d}t + \sqrt{2\varepsilon(\nabla^2\varphi^*(X_t))^{-1}}\,\mathrm{d}W_t \\ Y_t &= \nabla\varphi^*(X_t) \end{aligned} \tag{6}$$

driven by a standard n-dimensional Brownian motion $(W_t)_{t\ge 0}$, where $\varepsilon > 0$ is a small "temperature" parameter. In this setting, we consider a finite-horizon optimal control problem of minimizing the expected cost

$$\mathbf{E}\left[\int_0^T q(X_t, u_t)\,\mathrm{d}t + r(X_T)\middle| X_0 = x_0\right],$$

over all admissible control processes $(u_t)_{0\le t\le T}$ entering into the controlled SDE

$$\mathrm{d}X_t = u_t\,\mathrm{d}t + \sqrt{2\varepsilon(\nabla^2\varphi^*(X_t))^{-1}}\,\mathrm{d}W_t.$$

Here, q is the same cost as in (5) and r is an appropriately chosen terminal cost. As in the deterministic case, the mirror Langevin dynamics (6) emerges as the closed-loop system corresponding to the optimal control, although the value function is now time-dependent.

II. The deterministic problem

A. Some preliminaries

We assume that the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is C^1 and strictly convex and that the potential function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is C^2 and strictly convex. We let $\bar{y}$ denote the unique global minimizer of f. Both f and φ are assumed to be of *Legendre type*, i.e., $|\nabla f(x)|, |\nabla\varphi(x)| \to +\infty$ as $|x| \to +\infty$. As a consequence (see, e.g., [8, Thm. 26.5]), the gradient map $\nabla\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is bijective, with $(\nabla\varphi)^{-1} = \nabla\varphi^*$. The potential φ and its conjugate φ^* induce the *Bregman divergences*

$$\begin{aligned} D_\varphi(y, y') &:= \varphi(y) - \varphi(y') - \langle\nabla\varphi(y'), y - y'\rangle \\ D_{\varphi^*}(x, x') &:= \varphi^*(x) - \varphi^*(x') - \langle\nabla\varphi^*(x'), x - x'\rangle \end{aligned} \tag{7}$$

and the following relation holds for any pair of points $y, y' \in \mathbb{R}^n$ and their "mirror images" $x = \nabla\varphi(y), x' = \nabla\varphi(y')$:

$$D_\varphi(y, y') = D_{\varphi^*}(x', x)$$

(cf. [9, §11.2] for details). We also require that, for each fixed $x' \in \mathbb{R}^n$, the map $x \mapsto D_{\varphi^*}(x, x')$ is *radially unbounded*, i.e., $D_{\varphi^*}(x, x') \to +\infty$ as $|x| \to +\infty$. This will hold, e.g., if φ is strongly convex, i.e., there exists some $\alpha > 0$, such that

$$\varphi(y') \ge \varphi(y) + \langle\nabla\varphi(y), y' - y\rangle + \frac{\alpha}{2}|y' - y|^2 \tag{8}$$

for all $y, y' \in \mathbb{R}^n$. Radial unboundedness is needed for the invocation of the Lyapunov criterion for global asymptotic stability [3, Sec. 5.7].

Remark 1. We assume that φ is finite on all of $\mathbb{R}^n$ mainly to keep the exposition simple. It is not hard to adapt the analysis to the case when the potential function φ is defined on a closed convex set $\mathsf{X} \subseteq \mathbb{R}^n$ with nonempty interior, and $|\nabla\varphi(x)| \to +\infty$ as x approaches any point on the boundary of X. This corresponds to the problem of minimizing $f(x)$ subject to the constraint $x \in \mathsf{X}$.

B. Infinite-horizon optimal stabilizing controls

Consider the time-invariant controlled dynamical system

$$\dot{x}(t) = u(t). \tag{9}$$

For $x_0 \in \mathbb{R}^n$, let $\mathcal{U}_{x_0}$ be the class of all *stabilizing controls* at x_0, i.e., all locally essentially bounded maps $u : [0, \infty) \to \mathbb{R}^n$, such that the trajectory $x(t)$ of (9) with $x(0) = x_0$ is defined for all $t \ge 0$ and $x(t) \to \bar{x}$ as $t \to \infty$, where $\bar{x} := \nabla\varphi(\bar{y})$. We would like to minimize the cost

$$J_\infty(x_0, u(\cdot)) := \int_0^\infty q(x(t), u(t))\,\mathrm{d}t,$$

over all $u(\cdot) \in \mathcal{U}_{x_0}$, where

$$q(x, u) := f(\nabla\varphi^*(x)) + f^*(-u) + \langle u, \bar{y}\rangle. \tag{10}$$

The class $\mathcal{U}_{x_0}$ is nonempty since $u(t) = (\bar{x} - x_0)\mathbf{1}_{\{0\le t\le 1\}}$ is a stabilizing control at x_0. We denote by $V(x_0)$ the *value function*, i.e., infimum of $J_\infty(x_0, u(\cdot))$ over $u(\cdot) \in \mathcal{U}_{x_0}$.

C. The main result

Let $V(x) := D_{\varphi^*}(x, \bar{x})$. Theorem 1, stated and proved below, states that V is the value function for the above infinite-horizon optimal control problem, and that the mirror descent dynamics (4) is the closed-loop system corresponding to an optimal stabilizing control. Moreover, the value function V is also a global Lyapunov function for (4), and the point $\bar{x}$ is its global asymptotically stable equilibrium. The proof makes essential use of the following lemma:

Lemma 1. *The function V has the following properties:*

1) *It is C^2 and strictly convex.*
2) *$V(\bar{x}) = 0$, and $V(x) > 0$ for $x \ne \bar{x}$.*
3) *$V(x) \to +\infty$ as $|x| \to +\infty$.*

Moreover, the following inequality holds for $\dot{V}(x, u) := \langle\nabla V(x), u\rangle$:

$$\dot{V}(x, u) + q(x, u) \ge 0, \qquad x, u \in \mathbb{R}^n \tag{11}$$

and equality is attained iff $u = -\nabla f(\nabla\varphi^*(x))$.

Proof. Items 1)–3) are immediate consequences of our assumptions on φ. Moreover, a simple computation shows that

$$\dot{V}(x, u) + q(x, u) = f(\nabla\varphi^*(x)) + f^*(-u) + \langle u, \nabla\varphi^*(x)\rangle,$$

which is nonnegative by the Fenchel–Young inequality. The equality condition in (11) follows by [8, Thm. 23.5]. □

Theorem 1. *We have the following:*

1) *For any stabilizing control $u(\cdot) \in \mathcal{U}_{x_0}$ and for all $t \geq 0$,*
$$\int_0^t q(x(t), u(t))\,\mathrm{d}t \geq V(x_0) - V(x(t)). \tag{12}$$
In particular, $J_\infty(x_0, u(\cdot)) \geq V(x_0)$.

2) *For each x_0, the closed-loop system $\dot{x}(t) = -\nabla f(\nabla\varphi^*(x(t)))$ gives rise to an optimal stabilizing control $u(t) = -\nabla f(\nabla\varphi^*(x(t)))$, such that*
$$J_\infty(x_0, u(\cdot)) = V(x_0) = D_\varphi(x_0, \bar{x}).$$
Moreover, $V(x)$ is a global Lyapunov function for the closed-loop system.

Proof. Let $x_0 \in \mathbb{R}^n$ be given and consider an arbitrary stabilizing control $u(\cdot) \in \mathcal{U}_{x_0}$. Then, for $\dot{x}(t) = u(t)$ with $x(0) = x_0$ we have
$$\begin{aligned} V(x(t)) - V(x_0) &= \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} V(x(s))\,\mathrm{d}s \\ &= \int_0^t \dot{V}(x(s), u(s))\,\mathrm{d}s \\ &\geq -\int_0^t q(x(s), u(s))\,\mathrm{d}s, \end{aligned}$$
where the last step follows from (11). Rearranging gives (12). Moreover, taking the limit as $t \to \infty$ and using the fact that $V(x(t)) \to 0$ as $t \to \infty$ since $u(\cdot)$ is stabilizing, we get the inequality $J_\infty(x_0, u(\cdot)) \geq V(x)$.

Next, consider the closed-loop system $\dot{x}(t) = -\nabla f(\nabla\varphi^*(x(t)))$, $x(0) = x_0$, that generates the mirror descent flow. Then $\bar{x}$ is clearly an equilibrium point since $\nabla f(\nabla\varphi^*(\bar{x})) = \nabla f(\bar{y}) = 0$. Letting $y(t) := \nabla\varphi^*(x(t))$ and using (11), we have
$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} V(x(t)) &= \dot{V}(x(t), -\nabla f(y(t))) \\ &= -q(x(t), -\nabla f(y(t))) \\ &= \langle \nabla f(y(t)), \bar{y}\rangle - f^*(\nabla f(y(t))) - f(y(t)) \\ &\leq f(\bar{y}) - f(y(t)), \end{aligned}$$
which is strictly negative whenever $y(t) \neq \bar{y}$ by the strict convexity of f, or, equivalently, whenever $x(t) \neq \bar{x}$ since $\nabla\varphi^*$ is a bijection. Together with Lemma 1, this shows that V is a global Lyapunov function [3, Def. 5.7.1] for the above closed-loop system, so $\bar{x}$ is a globally asymptotically stable equilibrium [3, Thm. 17]. Thus, the control $u(t) = -\nabla f(\nabla\varphi^*(x(t)))$ is stabilizing at x_0, and $J_\infty(x_0, u(\cdot)) = V(x_0)$ from the equality condition in (11). □

D. Quantitative estimates

Theorem 1 allows us to obtain quantitative estimates on the approach of the trajectory of (4) to equilibrium. While similar estimates have been given in some earlier works [10], [11], the appeal of our optimal control perspective is that it allows to obtain such guarantees in a unified manner. It will be useful to introduce the following definition [12]: The function f is *μ-strongly convex* ($\mu \geq 0$) w.r.t. the potential function φ if
$$f(y') \geq f(y) + \langle \nabla f(y), y' - y\rangle + \mu D_\varphi(y', y),\ \ y, y' \in \mathbb{R}^n.$$
If $\mu = 0$, this is simply convexity; when $\mu > 0$, f has some nonzero "curvature" in some neighborhood of each point x, where the "geometry" is determined by the potential φ.

Theorem 2. *Let $(x(t), y(t))$, $t \geq 0$, be the state and the output trajectories of the mirror descent dynamics (4) starting from $x(0) = x_0$ and $y(0) = y_0 = \nabla\varphi^*(x_0)$. Then the following holds for every $t > 0$:*

1) *If f is convex, then*
$$f(y(t)) - f(\bar{y}) \leq \frac{1}{t} D_\varphi(\bar{y}, y_0). \tag{13}$$

2) *If f is μ-strongly convex w.r.t. φ then*
$$D_\varphi(\bar{y}, y(t)) \leq D_\varphi(\bar{y}, y_0) e^{-\mu t}, \tag{14}$$
and in that case the system (4) is exponentially stable.

Proof. Let $u(t) = -\nabla f(\nabla\varphi^*(x(t)))$ be the state feedback law that achieves $V(x_0)$. Then
$$\begin{aligned} q(x(t), u(t)) &= -\dot{V}(x(t), u(t)) \\ &= \langle \nabla f(y(t)), y(t) - \bar{y}\rangle \\ &= f(y(t)) - f(\bar{y}) + D_f(\bar{y}, y(t)). \end{aligned}$$
where the first equality is by Lemma 1 and the last equality follows by rearranging and using the definition
$$D_f(y, y') = f(y) - f(y') - \langle f(y'), y - y'\rangle.$$
Therefore, using Theorem 1 and the fact that $D_f(\cdot,\cdot) \geq 0$, we have
$$\begin{aligned} V(x_0) &\geq \int_0^t q(x(s), u(s))\,\mathrm{d}s \\ &= \int_0^t \{f(y(s)) - f(\bar{y})\}\,\mathrm{d}s \\ &\geq t\big(f(y(t)) - f(\bar{y})\big), \end{aligned}$$
where the last inequality follows from the fact that the value of the objective f decreases along the output trajectory $y(t)$:
$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f(y(t)) &= \langle \nabla f(y(t)), \dot{y}(t)\rangle \\ &= \langle \nabla f(y(t)), \nabla^2\varphi^*(x(t))\dot{x}(t)\rangle \\ &= -\langle \nabla f(y(t)), \nabla^2\varphi^*(x(t))\nabla f(y(t))\rangle \leq 0 \end{aligned}$$
— since φ^* is C^2 and strictly convex, its Hessian $\nabla^2\varphi^*(x)$ is positive definite for all $x \in \mathbb{R}^n$. Dividing by t and using the fact that $V(x_0) = D_{\varphi^*}(x_0, \bar{x}) = D_\varphi(\bar{y}, y_0)$, we get (13).

When f is μ-strongly convex, we have
$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} V(x(t)) &= \dot{V}(x(t), u(t)) \\ &= \langle \nabla f(y(t)), \bar{y} - y(t)\rangle \\ &\leq f(\bar{y}) - f(y(t)) - \mu D_\varphi(\bar{y}, y(t)) \\ &= f(\bar{y}) - f(y(t)) - \mu D_{\varphi^*}(x(t), \bar{x}) \\ &= f(\bar{y}) - f(y(t)) - \mu V(x(t)). \end{aligned}$$

Integrating gives the estimate

$$V(x(t)) \le e^{-\mu t}V(x_0) + \int_0^t e^{-\mu(t-s)}\{f(\bar{y}) - f(y(s))\}\,\mathrm{d}s$$

which yields (14) since $f(\bar{y}) \le f(y)$ for all y. $\square$

E. A simple example

As a simple illustration, consider the quadratic objective $f(x) = \frac{1}{2}|Ax-b|^2$ with $A \in \mathbb{R}^{p\times n}$ and $b \in \mathbb{R}^p$ and the quadratic potential $\varphi(x) = \frac{1}{2}|x|^2$. Assume that $A^{\mathsf{T}}A$ is nonsingular. Then the cost $q(x,u)$ in (10) takes the form

$$q(x,u) = \frac{1}{2}|Ax-b|^2 + \frac{1}{2}\langle u, (A^{\mathsf{T}}A)^{-1}u\rangle - \frac{1}{2}|A\bar{y}-b|^2,$$

where $\bar{y} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}b$ is the unique minimizer of f. Thus, the control-theoretic interpretation of the gradient flow for this problem naturally leads to infinite-horizon optimal stabilization of a linear system with a quadratic cost.

In the general case, the cost $q(x,u)$ can be expressed as

$$q(x,u) = \big(f(\nabla\varphi^*(x)) - f(\bar{y})\big) + \big(f(\bar{y}) + f^*(-u) + \langle u, \bar{y}\rangle\big),$$

where the first term is the optimality gap at $y = \nabla\varphi^*(x)$, while the second term is nonnegative by the Fenchel–Young inequality and equals zero iff $u = -\nabla f(\bar{y}) \equiv 0$.

III. The stochastic problem

We now consider a stochastic version of continuous-time mirror descent, dubbed *mirror Langevin dynamics*, or MLD [5]–[7]. The MLD generates a pair of random trajectories $(X_t, Y_t)_{t\ge 0}$ according to (6). The words "Langevin dynamics" allude to the fact that, with the quadratic potential $\varphi(x) = \frac{1}{2}|x|^2$, (6) reduces to the usual Langevin dynamics

$$\mathrm{d}X_t = -\nabla f(X_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t.$$

The use of MLD is mainly in sampling, where one makes use of the fact that the steady-state probability density of Y_t is proportional to $e^{-f/\varepsilon}$. Since this limiting density concentrates on the set of global minimizers of f as $\varepsilon \downarrow 0$, the sampling problem is intimately related to the problem of minimizing f.

A. Some preliminaries

In addition to the conditions imposed on f and φ in Sec. II-A, we also assume the following:

- the gradient ∇f is Lipschitz-continuous;
- the potential function φ is C^2, strongly convex, cf. (8), and has the *modified self-concordance property* [7], i.e., there exists some constant $c > 0$, such that

$$\|\sqrt{\nabla^2\varphi(x)} - \sqrt{\nabla^2\varphi(x')}\|_2 \le c|x-x'|$$

  for all $x, x' \in \mathbb{R}^n$, where $\|\cdot\|_2$ is the 2-Schatten (or Hilbert–Schmidt) norm.

In particular, the above assumption on φ implies that φ^* has a Lipschitz-continuous gradient [13, Thm. 4.2.1] and that the map $x \mapsto \sqrt{(\nabla^2\varphi^*(x))^{-1}}$ is Lipschitz-continuous [7].

B. A finite-horizon optimal control problem

We work in the usual setting of controlled diffusion processes [14, §VI.3-4]. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbf{P})$ be a probability space with a complete and right-continuous filtration, and let $(W_t)_{t\ge0}$ be a standard n-dimensional $(\mathcal{F}_t)$-Brownian motion. Let a finite horizon $0 < T < \infty$ be given. An *admissible control* (of state feedback type) is any measurable function $u : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$, such that the Itô SDE

$$\mathrm{d}X_t = u(X_t,t)\,\mathrm{d}t + \sqrt{2\varepsilon(\nabla^2\varphi^*(X_t))^{-1}}\,\mathrm{d}W_t \tag{15}$$

has a unique strong solution for all $t \in [0,T]$ and for any deterministic initial condition $X_0 = x_0$ (cf. [15, §5.2] for details). For each $t \in [0,T]$ define the *expected cost-to-go*

$$J(x,t;u(\cdot)) := \mathbf{E}\left[\int_t^T q(X_s,u_s)\,\mathrm{d}s + D_{\varphi^*}(X_T,\bar{x})\,\middle|\,X_t = x\right], \tag{16}$$

with the instantaneous cost q the same as in (10), where u_t is shorthand for $u(X_t,t)$, and let

$$V(x,t) := \inf_{u(\cdot)\ \text{admissible}} J(x,t;u(\cdot)) \tag{17}$$

be the value function. We say that an admissible control $u(\cdot)$ is optimal if $J(x,t;u(\cdot)) = V(x,t)$ for all $x \in \mathbb{R}^n$ and all $t \in [0,T]$. Observe that, in contrast with the deterministic infinite-horizon problem posed in Sec. II-B, here we are dealing with a finite-horizon stochastic problem, and there is, in addition to the instantaneous cost q, also a terminal cost $D_{\varphi^*}(\cdot,\bar{x})$. The form of the performance criterion in (16) is reminiscent of the Brezis–Ekeland action functional (3).

C. The main result

Theorem 3. *The value function in (17) is equal to*

$$V(x,t) = D_{\varphi^*}(x,\bar{x}) + \varepsilon n(T-t), \tag{18}$$

and the feedback control $u(x,t) = -\nabla f(\nabla\varphi(x))$ is optimal.

Proof. We use the verification theorem from the theory of controlled diffusions [14, §VI.4]. We associate to the process (15) a family of infinitesimal generators $(\mathcal{L}^u : u \in \mathbb{R}^n)$, where $\mathcal{L}^u$ is the second-order linear differential operator

$$\mathcal{L}^u := \sum_{i=1}^n u_i \frac{\partial}{\partial x_i} + \varepsilon \sum_{i,j=1}^n \big(\nabla^2\varphi^*(x)\big)^{-1}_{ij} \frac{\partial^2}{\partial x_i \partial x_j}.$$

Then it is readily verified that the function V defined in (18) is a solution of the *Hamilton–Jacobi–Bellman equation*

$$\frac{\partial}{\partial t}V(x,t) + \min_{u\in\mathbb{R}^d}\big\{\mathcal{L}^u V(x,t) + q(x,u)\big\} = 0 \tag{19}$$

on $\mathbb{R}^n \times [0,T]$ with the terminal condition $V(x,T) = D_{\varphi^*}(x,\bar{x})$. Indeed, since

$$\frac{\partial}{\partial t}V(x,t) = -\varepsilon n,$$
$$\nabla V(x,t) = \nabla\varphi^*(x) - \nabla\varphi^*(\bar{x}),$$
$$\nabla^2 V(x,t) = \nabla^2\varphi^*(x)$$

we can follow the same argument as in the proof of Lemma 1 to show that, for any $u \in \mathbb{R}^n$, we have

$$\begin{aligned}&\frac{\partial}{\partial t}V(x,t)+\mathcal{L}^u V(x,t)+q(x,u)\\&=-\varepsilon n+\langle u,\nabla V(x,t)\rangle+\varepsilon\,\mathrm{tr}\big\{(\nabla^2\varphi^*(x))^{-1}\nabla^2 V(x,t)\big\}\\&\qquad+f(\nabla\varphi^*(x))+f^*(-u)+\langle u,\nabla\varphi^*(\bar{x})\rangle\\&=f(\nabla\varphi^*(x))+f^*(-u)+\langle u,\nabla\varphi^*(x)\rangle\geq 0,\end{aligned}$$

with equality iff $u = -\nabla f(\nabla\varphi^*(x))$. Thus, $V(x,t)$ is a solution of the HJB equation (19), and clearly $V(x,T) = D_{\varphi^*}(x,\bar{x})$. Then, by Theorem 4.1 in [14, §VI.4], $V(x,t)$ is the value function in (17), and the control given by

$$\begin{aligned}u(x,t)&=\operatorname*{arg\,min}_{u\in\mathbb{R}^n}\big\{\mathcal{L}^u V(x,t)+q(x,u)\big\}\\&=\operatorname*{arg\,min}_{u\in\mathbb{R}^n}\big\{f(\nabla\varphi^*(x))+f^*(-u)+\langle u,\nabla\varphi^*(x)\big\}\\&=-\nabla f(\nabla\varphi^*(x))\end{aligned}$$

is optimal (note that it is also time-invariant). This control is admissible since, by our assumptions on f and φ, the maps $x \mapsto -\nabla f(\nabla\varphi^*(x))$ and $x \mapsto \sqrt{2\varepsilon(\nabla^2\varphi^*(x))^{-1}}$ are Lipschitz-continuous and have at most linear growth. Consequently, with the choice of $u(x,t) = b(x)$, the SDE (15) has a unique strong solution [15, §5.2, Thm. 2.5], so $u(\cdot)$ is indeed admissible. ☐

D. Quantitative estimates

Theorem 4. *Let (X_t, Y_t), $t \geq 0$, be the random state and output trajectories of the mirror Langevin dynamics (6) with deterministic initial condition $X_0 = x_0$ and $Y_0 = y_0 = \nabla\varphi^*(x_0)$. Then the following holds for every $T > 0$:*

1) If f is convex, then

$$\frac{1}{T}\mathbf{E}\left[\int_0^T\{f(Y_t)-f(\bar{y})\}\,\mathrm{d}t\,\middle|\,Y_0=y_0\right]\leq\frac{1}{T}D_\varphi(\bar{y},y_0)+\varepsilon n. \quad (20)$$

2) If f is μ-strongly convex w.r.t. φ, for $\mu > 0$, then

$$\mathbf{E}[D_\varphi(\bar{y},Y_t)|Y_0=y_0]\leq D_\varphi(\bar{y},y_0)e^{-\mu t}+\frac{\varepsilon n}{\mu}(1-e^{-\mu T}). \quad (21)$$

Proof. Let $u_t := -\nabla f(\nabla\varphi^*(X_t))$. Then, proceeding just like in the proof of Theorem 2, we can write

$$\begin{aligned}q(X_t,u_t)&=f(Y_t)+f^*(-u_t)+\langle u_t,\bar{y}\rangle\\&=f(Y_t)-f(\bar{y})+D_f(\bar{y},Y_t)\\&\geq f(Y_t)-f(\bar{y}).\end{aligned}$$

Using this together with Theorem 3 gives

$$\begin{aligned}&D_\varphi(\bar{y},y_0)+\varepsilon nT=D_\varphi(x_0,\bar{x})+\varepsilon nT\\&=\mathbf{E}\left[\int_0^T q(X_t,u_t)\,\mathrm{d}t+D_\varphi(X_T,\bar{x})\,\middle|\,X_0=x_0\right]\\&\geq\mathbf{E}\left[\int_0^T\{f(Y_t)-f(\bar{y})\}\,\mathrm{d}t\,\middle|\,X_0=x_0\right].\end{aligned}$$

Dividing both sides by $T > 0$ and using the fact the σ-algebras $\sigma(X_t : t \in [0,T])$ and $\sigma(Y_t : t \in [0,T])$ coincide since $\nabla\varphi^*$ is a bijection, we obtain (20).

When f is μ-strongly convex, we have

$$f(\bar{y})-f(Y_t)\geq\langle\nabla f(Y_t),\bar{y}-Y_t\rangle+\mu D_\varphi(\bar{y},Y_t). \quad (22)$$

On the other hand, by Itô's lemma and by (19),

$$V(X_t,t)=V(X_0,0)+\int_0^t\langle\nabla f(Y_s),\bar{y}-Y_s\rangle\,\mathrm{d}s+M_t, \quad (23)$$

where M_t is a zero-mean $(\mathcal{F}_t)$-martingale. Since

$$\begin{aligned}V(X_s,s)&=D_{\varphi*}(X_s,\bar{x})+\varepsilon n(T-s)\\&=D_\varphi(\bar{y},Y_s)+\varepsilon n(T-s),\end{aligned}$$

combining (22) and (23) and then taking expectations given $Y_0 = y_0$ yields

$$\begin{aligned}&\mathbf{E}[D_\varphi(\bar{y},Y_t)|Y_0=y_0]\\&\leq D_\varphi(\bar{y},y_0)+\varepsilon nt-\mu\int_0^t\mathbf{E}[D_\varphi(\bar{y},Y_s)|Y_0=y_0]\,\mathrm{d}s\end{aligned}$$

for all $t \in [0,T]$. Grönwall's inequality gives (21). ☐

Note that, in contrast with the deterministic setting (cf. Theorem 2), when the objective function f is not strongly convex, we only have guarantees on the expected average objective $\mathbf{E}[\frac{1}{T}\int_0^T f(Y_t)\,\mathrm{d}t|Y_0=y_0]$, which, owing to the convexity of f, translates into an optimization error estimate for the time average of the trajectory, $\tilde{Y}_T := \frac{1}{T}\int_0^T Y_t\,\mathrm{d}t$:

$$\begin{aligned}&\mathbf{E}[f(\tilde{Y}_T)-f(\bar{y})|Y_0=y_0]\\&\leq\frac{1}{T}\mathbf{E}\left[\int_0^T\{f(Y_t)-f(\bar{y})\}\,\mathrm{d}y\,\middle|\,Y_0=y_0\right]\\&\leq\frac{1}{T}D_\varphi(\bar{y},y_0)+\varepsilon n.\end{aligned}$$

However, as the following result shows, in the low-noise regime (i.e., for all sufficiently small ε), with high probability, the MLD output trajectory $(Y_t)_{0\leq t\leq T}$ closely tracks the deterministic mirror-descent output trajectory $(y(t))_{0\leq t\leq T}$ with the same initial condition $Y_0 = y(0) = y_0$:

Theorem 5. *There exist positive time-independent constants C_i, $i = 1,2,3$, such that, for every $0 < \varepsilon \leq \frac{1}{C_1T^3}e^{-C_2T}$, the following estimate holds with probability at least $1-\delta$:*

$$\sup_{0\leq t\leq T}|f(Y_t)-f(y(t))|\leq\frac{C_3}{T}\sqrt{n\log\frac{n}{\delta}}. \quad (24)$$

Proof. Let $\Delta_t := |Y_t - y(t)|$. The following estimate holds by the Lipschitz continuity of ∇f:

$$f(Y_t)-f(y(t))\leq\langle\nabla f(y(t)),Y_t-y(t)\rangle+\frac{L_f}{2}|Y_t-y(t)|^2,$$

where L_f is the Lipschitz constant of ∇f. Moreover, the gradient norms $|\nabla f(y(t))|$ are uniformly bounded since

$$\begin{aligned}|\nabla f(y(t))| &\le |\nabla f(y(t)) - \nabla f(y(0))| + |\nabla f(y(0))| \\ &\le L_f |y(t) - y(0)| + |\nabla f(y(0))| \\ &\le L_f |y(t) - \bar{y}| + L_f |y(0) - \bar{y}| + |\nabla f(y(0))| \\ &\le 2L_f \sqrt{\frac{2}{\alpha} D_\varphi(\bar{y}, y(0))} + |\nabla f(y(0))| =: K_0,\end{aligned}$$

which in turn implies that

$$\sup_{0\le t\le T} |f(y(t)) - f(y(0))| \le K_0 \sup_{0\le t\le T} \Delta_t + \frac{L_f}{2} \sup_{0\le t\le T} \Delta_t^2. \tag{25}$$

Define the matrix-valued process $(\xi_t)_{0\le t\le T}$ by $\xi_t := \sqrt{(\nabla^2\varphi^*(X_t))^{-1}}$. For each $t \in [0,T]$, we have

$$\begin{aligned}&|X_t - x(t)| \\ &\le L_f \int_0^t |Y_s - y(s)| \, \mathrm{d}s + \sqrt{2\varepsilon} \sup_{0\le t\le T} \left| \int_0^t \xi_s \, \mathrm{d}W_s \right|.\end{aligned}$$

By our assumptions on φ, there exist positive constants $\kappa_2 \ge \kappa_1 > 0$, such that the eigenvalues of $\nabla^2\varphi^*(x)$ lie in the interval $[\kappa_1, \kappa_2]$. Hence, the process ξ_t is uniformly bounded, so the quadratic variations of the matrix entries $[\xi^{ij}]_t$, $1 \le i, j \le n$, are uniformly bounded by a positive multiple of t. Hence, by the time-change theorem for martingales [15, §3.4, Thm. 4.6], there exist a constant $\kappa > 0$ and a standard n-dimensional Brownian motion $(B_t)_{t\ge 0}$, such that

$$\sup_{0\le t\le T} \left| \int_0^t \xi_s \, \mathrm{d}W_s \right| \le \sup_{0\le t\le \kappa T} |B_t|.$$

Since $\nabla\varphi^*$ is Lipschitz-continuous, we have

$$\begin{aligned}\Delta_t &\le L_{\varphi^*} |X_t - x(t)| \\ &\le L_{\varphi^*} L_f \int_0^t \Delta_s \, \mathrm{d}s + \sqrt{2\varepsilon} L_{\varphi^*} \sup_{0\le t\le \kappa T} |B_t|.\end{aligned}$$

Grönwall's inequality therefore gives

$$\sup_{0\le t\le T} \Delta_t \le \sqrt{2\varepsilon} L_{\varphi^*} \sup_{0\le t\le \kappa T} |B_t| e^{L_{\varphi^*} L_f T}. \tag{26}$$

If $\varepsilon \le \frac{1}{L_{\varphi^*}^2 T^3} e^{-2L_{\varphi^*} L_f T}$, then, using (26) in (25), we obtain

$$\begin{aligned}&\sup_{0\le t\le T} |f(Y_t) - f(y(t))| \\ &\le \frac{\sqrt{2} K_0}{T^{3/2}} \sup_{0\le t\le \kappa T} |B_t| + \frac{L_f}{T^3} \sup_{0\le t\le \kappa T} |B_t|^2.\end{aligned}$$

By the reflection principle for the Brownian motion [15, p. 96],

$$\mathbf{P}\left\{ \sup_{0\le t\le \kappa T} |B_t| \ge r \right\} \le 2\mathbf{P}\left\{ |B_{\kappa T}| \ge r \right\} \le 4n e^{-r^2/2n\kappa T},$$

for every $r > 0$, and therefore

$$\sup_{0\le t\le T} |f(Y_t) - f(y(t))| \le \frac{\tilde{C}}{T} \sqrt{n \log \frac{n}{\delta}}$$

with probability at least $1 - \delta$, where $\tilde{C}$ is a constant that depends on K_0, L_f, κ. □

IV. Conclusion and Future Directions

In this paper, we have phrased continuous-time mirror descent methods in the framework of "inverse optimal control" [16]—that is, given an autonomous (i.e., control-free) dynamical system, identify a controlled dynamical system and a cost criterion, such that the autonomous dynamics can be viewed as the closed-loop system corresponding to an optimal control. An intriguing direction for future research is to interpret other optimization methods, such as the heavy-ball method [17], through the inverse optimal control lens. It would also be of interest to provide an optimal control perspective on the discrete-time variant of the Brezis–Ekeland principle [18].

Acknowledgments

The authors would like to thank Jelena Diakonikolas for pointing them to Ref. [2], which was the initial inspiration for this work, and also Anatoli Juditsky, Philippe Rigollet, and Matus Telgarsky for insightful comments and suggestions.

References

[1] H. Brezis and I. Ekeland, "Un principe variatonnel associé a certaines equations paraboliques. Le cas independant du temps." *C.R. Acad. Sci. Paris Sér. A*, vol. 282, pp. 971–974, 1976.

[2] N. Ghoussoub and L. Tzou, "A variational principle for gradient flows," *Mathematische Annalen*, vol. 330, pp. 519–549, 2004.

[3] E. D. Sontag, *Mathematical Control Theory: Deterministic Finite-Dimensional Systems*, 2nd ed. Springer, 1998.

[4] A. S. Nemirovski and D. B. Yudin, *Problem Complexity and Method Efficiency in Optimization*. New York: Wiley, 1983.

[5] Y.-P. Hsieh, A. Kavis, P. Rolland, and V. Cevher, "Mirrored Langevin dynamics," in *Neural Information Processing Systems*, vol. 32, 2018.

[6] S. Chewi, T. Le Gouic, C. Lu, T. Maunu, P. Rigollet, and A. Stromme, "Exponential ergodicity of mirror-Langevin diffusions," in *Neural Information Processing Systems*, vol. 34, 2020.

[7] R. Li, M. Tao, S. S. Vempala, and A. Wibisono, "The mirror Langevin algorithm converges with vanishing bias," in *Conference on Learning Theory*, 2022.

[8] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[9] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[10] W. Krichene, A. M. Bayen, and P. L. Bartlett, "Accelerated mirror descent in continuous and discrete time," in *Neural Information Processing Systems*, 2015.

[11] M. Telgarsky, "Stochastic linear optimization never overfits with quadratically-bounded losses on general data," in *Conference on Learning Theory*, 2022.

[12] P. L. Bartlett, E. Hazan, and A. Rakhlin, "Online adaptive gradient descent," in *Neural Information Processing Systems*, 2007.

[13] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis*. Springer, 2001.

[14] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*. Springer, 1975.

[15] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, 1998.

[16] J. Casti, "On the general inverse problem of optimal control theory," *Journal of Optimization Theory and Applications*, vol. 32, no. 4, pp. 491–497, 1980.

[17] B. T. Polyak, "Some methods of speeding up the convergence of iteration methods," *USSR Computational Mathematics and Mathematical Physics*, vol. 4, no. 5, pp. 1–17, 1964.

[18] U. Stefanelli, "the discrete Brezis-Ekeland principle," *Journal of Convex Analysis*, vol. 16, no. 1, pp. 71–87, 2009.