Algorithmic Optimality Guarantees for Nonsmooth H_∞ Output-Feedback Policy Search

Ashkan Soleymani and Patrick Jaillet

***Abstract*— We study continuous-time full-order dynamic output-feedback H_∞ policy search, a nonconvex and nonsmooth problem. Direct policy search is a central paradigm in reinforcement learning and continuous control, but rigorous guarantees remain scarce in robust output-feedback settings. The H_∞ problem is a canonical benchmark because it captures disturbance attenuation and robustness while exposing the hard nonsmooth geometry of policy-space optimization. We prove that on the exact identity-gauge slice of the extended convex lift, ε-stationarity yields $O(\varepsilon)$-suboptimality on compact exact slices, which in turn yields convergence-rate guarantees for nonsmooth policy-search methods. This result addresses the finite-time optimality-gap question raised by Guo and Hu [1] in the more general dynamic output-feedback H_∞ policy-search setting. We further use the established value equivalence supplied by extended convex lifting to formulate a nonstrict-feasibility bisection method with one final strict-feasibility recovery step, yielding an explicit ε-optimal stabilizing controller. These results provide a quantitative and algorithmic strengthening of prior qualitative optimality theory for nonsmooth H_∞ policy search.**

I. Introduction

Direct policy search, or policy optimization, refers to optimizing a parameterized controller directly against the closed-loop performance objective. It is a core paradigm in reinforcement learning (RL) and continuous control. The policy-gradient theorem of Sutton et al. [2] made this viewpoint algorithmic, and later deterministic, trust-region, and proximal variants became standard tools for continuous-action RL [3–5]. These methods have driven influential successes in simulated robotics and continuous control, including locomotion and end-to-end learning from high-dimensional observations [6, 7]. From a control viewpoint, direct policy search is attractive because it is end-to-end, accommodates richly parameterized controllers, and is compatible with model-based, model-free, and data-driven implementations [8, 9].

This viewpoint has recently reshaped parts of systems and control. A growing literature studies the geometry and algorithmics of policy optimization for the linear-quadratic regulator (LQR), linear-quadratic-Gaussian (LQG) control, output estimation, and robust control, revealing benign landscapes in some settings and geometric obstructions in others [8–11]. The partially observed case is especially delicate. Dynamic output-feedback controllers carry their own internal state, and controllers related by similarity transformations represent the same input-output law. This redundancy creates degenerate directions, complicates stationarity, and can obstruct naive Euclidean descent [12, 13]. Recent work on output estimation and LQG makes this difficulty explicit and shows that suitable reparameterizations or quotient-geometric viewpoints can restore meaningful first-order analysis [10, 13].

Among robust synthesis problems, H_∞ control is a natural stress test for direct policy search. It is classical and practically central, but in policy space the objective is nonconvex and nonsmooth. In the state-feedback setting, Guo and Hu [1] proved that every Clarke-stationary stabilizing gain is globally optimal and showed that Goldstein-type methods can recover the global solution. For dynamic output feedback, Tang and Zheng [14] established global optimality of Clarke-stationary controllers under a nondegenerate exact-certificate condition. More broadly, recent extended convex lifting (ECL) results reveal hidden convexity behind optimal and robust control problems and show that nondegenerate first-order stationary points are globally optimal across several benchmark settings, including output-feedback H_∞ synthesis [11, 15]. Recent work established weak convexity and deterministic subgradient rates for discrete-time robust control, although the weak Polyak–Łojasiewicz conclusion remains confined to the state-feedback case [16].

The remaining gap is quantitative and algorithmic. The theorem of Tang and Zheng [14] is qualitative: it certifies exact Clarke-stationary points, but it does not convert an approximate stationarity residual into a value-gap bound and, consequently, does not provide algorithms with convergence rates for finding an optimal policy. Nor does it separate the two roles played by the lifted variables. One role is global: the full nonstrict lift captures the closure of strict stabilizing upper-bound certificates and therefore should be used to compute the optimal value. The other is local: after exactification and gauge fixing, an exact identity slice provides the appropriate chart for first-order certification and, hence, a foundation for the analysis of nonsmooth optimization algorithms for policy search. Keeping these roles distinct leads to sharper geometric insight and a more usable algorithmic interpretation.

Our main message is that the ECL has complementary local and global roles; see Figure 1. The stable exact

Department of Electrical Engineering and Computer Science and the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. E-mail addresses: `ashkanso@mit.edu` and `jaillet@mit.edu`.

identity-gauge slice is the appropriate chart for local first-order certification and, hence, for analyzing the convergence rates of nonsmooth optimization algorithms for policy search, while the full nonstrict convex lifted set computes the global infimal value through closure and strict lifted points support stabilizing controller recovery. Building on this separation, our contributions are fourfold.

1) We prove an exact *reconditioning* lemma showing every exact nondegenerate certificate is realization-equivalent to one with $P_{12} = I$.
2) On that identity-gauge exact slice, we establish a pairwise first-order error bound, implying that on every compact exact slice, ε-stationarity yields $O(\varepsilon)$-suboptimality.
3) Building on this suboptimality bound, we study the convergence rates of off-the-shelf nonsmooth optimization algorithms for finding Goldstein stationary points to the optimal policy in the dynamic output-feedback H_∞ policy search problem.
4) Using the established value-equivalence theorem for extended convex lifting and its dynamic output-feedback H_∞ specialization [11, Theorem 2.1 and Corollary 4.1], we formulate a nonstrict-feasibility bisection procedure in the present generalized-plant coordinates. One final strict-feasibility recovery step returns an explicit ε-optimal stabilizing controller.

Relation to classical H_∞ synthesis: Classical Riccati and linear matrix inequality (LMI) methods already solve full-order H_∞ synthesis in model-based settings [17, 18]. Our focus is instead the nonsmooth policy-optimization viewpoint, where the extended convex lift enables quantitative first-order guarantees and algorithmic controller recovery. This perspective is particularly relevant to optimization-based, certificate-driven and potentially data-driven methods.

II. Problem Definition

A. Dynamic output-feedback H_∞ policy search

We consider the continuous-time linear time-invariant plant

$$\dot{x}(t) = Ax(t) + B_1 w(t) + B_2 u(t), \tag{1}$$
$$z(t) = C_1 x(t) + D_{11} w(t) + D_{12} u(t), \tag{2}$$
$$y(t) = C_2 x(t) + D_{21} w(t), \tag{3}$$

where $x(t) \in \mathbb{R}^{n_x}$ is the plant state, $w(t)$ is the exogenous input, $u(t)$ is the control input, $z(t)$ is the performance output, and $y(t)$ is the measured output. As in Tang and Zheng [14], we assume that (A, B_1) and (A, B_2) are controllable and that (C_1, A) and (C_2, A) are observable. We restrict attention to full-order dynamic controllers of the form

$$\dot{\xi}(t) = A_K \xi(t) + B_K y(t), \tag{4}$$
$$u(t) = C_K \xi(t) + D_K y(t), \tag{5}$$

where $\xi(t) \in \mathbb{R}^{n_x}$ is the controller state. We identify the controller with the block matrix

$$K = \begin{bmatrix} D_K & C_K \\ B_K & A_K \end{bmatrix}.$$

Throughout, controller space is viewed as a finite-dimensional Euclidean space equipped with the Frobenius inner product and its induced norm.

The closed-loop transfer matrix from w to z is

$$T_{zw}(s; K) = C_{\rm cl}(K)(sI - A_{\rm cl}(K))^{-1} B_{\rm cl}(K) + D_{\rm cl}(K),$$

where the closed-loop matrices are the standard ones induced by (1)–(5); explicitly,

$$\begin{aligned} A_{\rm cl}(K) &= \begin{bmatrix} A + B_2 D_K C_2 & B_2 C_K \\ B_K C_2 & A_K \end{bmatrix}, \\ B_{\rm cl}(K) &= \begin{bmatrix} B_1 + B_2 D_K D_{21} \\ B_K D_{21} \end{bmatrix}, \\ C_{\rm cl}(K) &= \begin{bmatrix} C_1 + D_{12} D_K C_2 & D_{12} C_K \end{bmatrix}, \\ D_{\rm cl}(K) &= D_{11} + D_{12} D_K D_{21}. \end{aligned}$$

The stabilizing set is

$$\mathcal{C} := \{K : A_{\rm cl}(K) \text{ is Hurwitz}\},$$

and the H_∞ objective is

$$J(K) := \|T_{zw}(\cdot; K)\| = \|T_{zw}(\cdot; K)\|_\infty, \qquad K \in \mathcal{C}.$$

Hence the policy search problem is

$$J^* := \inf_{K \in \mathcal{C}} J(K). \tag{6}$$

A key fact from Tang and Zheng [14, Proposition 1] is that J is locally Lipschitz and subdifferentially regular on $\mathcal{C}$.

Proposition 1. *For problem (6), the mapping $J : \mathcal{C} \to \mathbb{R}$ is locally Lipschitz and subdifferentially regular on $\mathcal{C}$. In particular, for every $K \in \mathcal{C}$ and every direction V, the directional derivative $J'(K; V)$ exists and satisfies*

$$J'(K; V) = J^\circ(K; V) = \max_{G \in \partial J(K)} \langle G, V \rangle,$$

where $\partial J(K)$ denotes the Clarke subdifferential.

For locally Lipschitz regular functions, the ordinary directional derivative coincides with the Clarke directional derivative, which in turn equals the support function of the Clarke subdifferential; see, for example, [19, Section 2.3] or [20, Chapter 8]. We conclude this section with the definition of the Clarke and Goldstein subdifferential and stationary points for completeness [19, 21].

Definition 2. *Given a locally Lipschitz function $f : C \to \mathbb{R}$, its Clarke subdifferential is defined as*

$$\partial f(x) = \operatorname{conv}\left\{ \lim_{x_k \to x} \nabla f(x_k) \;\middle|\; \nabla f(x_k) \text{ exists}, \; x_k \in \mathcal{C} \right\},$$

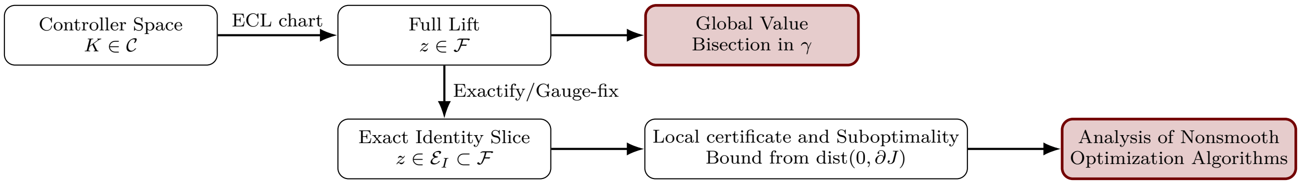

Fig. 1: Two roles of the lift. The full set $\mathcal{F}$ computes the global value, while the exact slice $\mathcal{E}_I$ yields local first-order certificates and supports the analysis of convergence rates for nonsmooth optimization algorithms.

and its Goldstein subdifferential is defined as

$$\partial_\delta f(x) = \mathrm{conv}\left\{ \bigcup_{y \in B_2(x,\delta)} \partial f(y) \right\}.$$

A point x is called an ε-Clarke stationary point if $\mathrm{dist}(0, \partial f(x)) \leq \varepsilon$, *and an (δ, ε)-Goldstein stationary point if* $\mathrm{dist}(0, \partial_\delta f(x)) \leq \varepsilon$, *where* dist *denotes the Euclidean distance.*

The Clarke stationary point serves as a primary solution concept for nonsmooth optimization; however, it suffers from significant computational intractability (e.g., there is no algorithm that can find Clarke stationary points in finite time [22, Theorem 5]). However, recent advances show that Goldstein stationarity, as a more relaxed solution concept, offers algorithmic tractability [22, 23].

B. Nondegenerate bounded-real certificates and Extended Convex Lifting

Define the matrix $N(K, P, \gamma) :=$

$$\begin{bmatrix} A_{\mathrm{cl}}(K)^\top P + P A_{\mathrm{cl}}(K) & P B_{\mathrm{cl}}(K) & C_{\mathrm{cl}}(K)^\top \\ B_{\mathrm{cl}}(K)^\top P & -\gamma I & D_{\mathrm{cl}}(K)^\top \\ C_{\mathrm{cl}}(K) & D_{\mathrm{cl}}(K) & -\gamma I \end{bmatrix}. \quad (8)$$

Following Tang and Zheng [14], let

$$\mathcal{S}_{\mathrm{nd}} := \Big\{ (K, P, \gamma) : P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^\top & P_{22} \end{bmatrix} \in \mathbb{S}^{2n_x}_{++}, \\ P_{12} \in GL_{n_x},\ N(K, P, \gamma) \preceq 0 \Big\}. \quad (9)$$

The nonstrict inequality in (9) does not by itself imply internal stability. If $(K, P, \gamma) \in \mathcal{S}_{\mathrm{nd}}$ and $K \in \mathcal{C}$, then the nonstrict bounded-real lemma yields $J(K) \leq \gamma$ [14]. We therefore distinguish the algebraic certificate set $\mathcal{S}_{\mathrm{nd}}$ from its stabilizing part

$$\mathcal{S}_{\mathrm{nd}}^{\mathrm{st}} := \{ (K, P, \gamma) \in \mathcal{S}_{\mathrm{nd}} : K \in \mathcal{C} \}.$$

The associated exact nondegenerate controller set is

$$\mathcal{C}_{\mathrm{nd}} := \{ K \in \mathcal{C} : \exists P \succ 0 \text{ such that } (K, P, J(K)) \in \mathcal{S}_{\mathrm{nd}} \}. \quad (10)$$

Tang and Zheng [14] prove that every *exact* Clarke-stationary controller $x \in \mathcal{C}_{\mathrm{nd}}$ $(0 \in \partial f(x))$ is globally optimal. Their analysis rests on an extended convex lifting. Introduce the lifted variable

$$z = (X, Y, M, H, F, L, \gamma),$$

where $X, Y \in \mathbb{S}^{n_x}$, $M \in \mathbb{R}^{n_x \times n_x}$, $H \in \mathbb{R}^{n_x \times n_y}$, $F \in \mathbb{R}^{n_u \times n_x}$, $L \in \mathbb{R}^{n_u \times n_y}$, and $\gamma \in \mathbb{R}$. Define the affine matrix map $\mathcal{M}(z)$ as in (7), and the convex lifted feasible set

$$\mathcal{F} := \left\{ z : \begin{bmatrix} X & I \\ I & Y \end{bmatrix} \succ 0,\ \mathcal{M}(z) \preceq 0 \right\}. \quad (11)$$

Tang and Zheng [14] construct a real-analytic diffeomorphism

$$\Phi : \mathcal{S}_{\mathrm{nd}} \to GL_{n_x} \times \mathcal{F}, \quad (12)$$

whose first coordinate is the certificate block P_{12}, and write $\Psi = \Phi^{-1}$ for its inverse [14, Proposition 2]. We will use only the following consequence.

Proposition 3 (ECL diffeomorphism)**.** *There exists a real-analytic diffeomorphism $\Phi : \mathcal{S}_{\mathrm{nd}} \to GL_{n_x} \times \mathcal{F}$ with inverse Ψ, and the first coordinate of $\Phi(K, P, \gamma)$ equals P_{12}. Equivalently, if*

$$\Psi(\Xi, z) = (K, P, \gamma),$$

then the $(1, 2)$ block of P equals Ξ. In particular, $\Psi(\Xi, z) \in \mathcal{S}_{\mathrm{nd}}$ for all $(\Xi, z) \in GL_{n_x} \times \mathcal{F}$. The congruence relation used to construct Φ and Ψ preserves strict definiteness; thus, if $\Psi(\Xi, z) = (K, P, \gamma)$, then $\mathcal{M}(z) \prec 0$ if and only if $N(K, P, \gamma) \prec 0$.

The result of [14] on the globality of exact Clarke stationary points is qualitative rather than quantitative, since exact Clarke stationary points lack algorithmic tractability, and prior to our work, no connection between the subdifferentials and the suboptimality gap had been established. In this work, our goal is to design an algorithmic framework to find optimal solutions for the dynamic output-feedback H_∞ policy search problem.

III. Exact reconditioning and the identity gauge slice

The first step is to show that the nondegenerate exact-certificate geometry can be normalized exactly by controller-state similarity.

Lemma 4 (Similarity invariance)**.** *For $S \in GL_{n_x}$ define*

$$\mathcal{T}_S(K) := \begin{bmatrix} D_K & C_K S^{-1} \\ S B_K & S A_K S^{-1} \end{bmatrix}. \quad (13)$$

Then $\mathcal{T}_S(K)$ is realization-equivalent to K. Consequently,

$$\mathcal{T}_S(K) \in \mathcal{C} \iff K \in \mathcal{C}, \qquad J\big(\mathcal{T}_S(K)\big) = J(K).$$

$$\mathcal{M}(z) := \begin{bmatrix} AX + B_2F + (AX + B_2F)^\mathsf{T} & M^\mathsf{T} + A + B_2LC_2 & B_1 + B_2LD_{21} & (C_1X + D_{12}F)^\mathsf{T} \\ M + (A + B_2LC_2)^\mathsf{T} & YA + HC_2 + (YA + HC_2)^\mathsf{T} & YB_1 + HD_{21} & (C_1 + D_{12}LC_2)^\mathsf{T} \\ (B_1 + B_2LD_{21})^\mathsf{T} & (YB_1 + HD_{21})^\mathsf{T} & -\gamma I & (D_{11} + D_{12}LD_{21})^\mathsf{T} \\ C_1X + D_{12}F & C_1 + D_{12}LC_2 & D_{11} + D_{12}LD_{21} & -\gamma I \end{bmatrix} \quad (7)$$

Proof. Let $\widehat{\xi} = S\xi$. Under this change of controller coordinates, (4)–(5) become

$$\dot{\widehat{\xi}} = SA_KS^{-1}\widehat{\xi} + SB_Ky, \qquad u = C_KS^{-1}\widehat{\xi} + D_Ky,$$

which is precisely the controller $\mathcal{T}_S(K)$. Therefore K and $\mathcal{T}_S(K)$ realize the same controller transfer matrix,

$$D_K + C_K(sI - A_K)^{-1}B_K.$$

Since the plant is unchanged, the closed-loop transfer matrix $T_{zw}(\cdot; K)$ is identical to $T_{zw}(\cdot; \mathcal{T}_S(K))$. Internal stability is preserved by state similarity, and so is the H_∞ norm of the closed-loop transfer matrix. Hence $K \in \mathcal{C}$ if and only if $\mathcal{T}_S(K) \in \mathcal{C}$, and $J\big(\mathcal{T}_S(K)\big) = J(K)$. ☐

Lemma 5 (Exact gauge fixing)**.** *Let $(K, P, \gamma) \in \mathcal{S}_{\rm nd}$, with*

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^\mathsf{T} & P_{22} \end{bmatrix}, \qquad P_{12} \in GL_{n_x}.$$

Set $S := P_{12}$ and define

$$\widehat{K} := \mathcal{T}_S(K), \qquad \widehat{P} := T^{-\mathsf{T}}PT^{-1}, \qquad T := \operatorname{diag}(I_{n_x}, S).$$

Then $(\widehat{K}, \widehat{P}, \gamma) \in \mathcal{S}_{\rm nd}$ and the off-diagonal block of $\widehat{P}$ equals the identity, i.e., $\widehat{P}_{12} = I_{n_x}$. Moreover, if $\gamma = J(K)$ then $\gamma = J(\widehat{K})$ as well.

Proof. By construction, T is invertible. Because $P \succ 0$, we have $\widehat{P} = T^{-\mathsf{T}}PT^{-1} \succ 0$. A direct block computation gives

$$\widehat{P} = \begin{bmatrix} P_{11} & P_{12}S^{-1} \\ S^{-\mathsf{T}}P_{12}^\mathsf{T} & S^{-\mathsf{T}}P_{22}S^{-1} \end{bmatrix} = \begin{bmatrix} P_{11} & I \\ I & S^{-\mathsf{T}}P_{22}S^{-1} \end{bmatrix},$$

so indeed $\widehat{P}_{12} = I$. Next, since $\widehat{K} = \mathcal{T}_S(K)$, Lemma 4 implies that the closed-loop matrices transform by controller-state similarity: $A_{\rm cl}(\widehat{K}) = TA_{\rm cl}(K)T^{-1}$, $B_{\rm cl}(\widehat{K}) = TB_{\rm cl}(K)$, $C_{\rm cl}(\widehat{K}) = C_{\rm cl}(K)T^{-1}$, and $D_{\rm cl}(\widehat{K}) = D_{\rm cl}(K)$. Therefore

$$N(\widehat{K}, \widehat{P}, \gamma) = \operatorname{diag}(T^{-\mathsf{T}}, I, I)\, N(K, P, \gamma)\, \operatorname{diag}(T^{-1}, I, I).$$

Since $N(K, P, \gamma) \preceq 0$ and congruence by an invertible matrix preserves semidefiniteness, we obtain $N(\widehat{K}, \widehat{P}, \gamma) \preceq 0$. Thus $(\widehat{K}, \widehat{P}, \gamma) \in \mathcal{S}_{\rm nd}$. Finally, Lemma 4 yields $J(\widehat{K}) = J(K)$. Hence if $\gamma = J(K)$, then also $\gamma = J(\widehat{K})$. ☐

Lemma 5 shows that every exact nondegenerate controller admits a realization-equivalent representative whose certificate satisfies $P_{12} = I$. This motivates the following slice as a representative of the set $\mathcal{C}_{nd}$.

Definition 6 (Identity gauge slice)**.** *Define*

$$\mathcal{S}_I := \{(K, P, \gamma) \in \mathcal{S}_{\rm nd} : P_{12} = I_{n_x}\}.$$

Let $\Psi_I : \mathcal{F} \to \mathcal{S}_I$ defined as $\Psi_I(z) := \Psi(I_{n_x}, z)$, and write $\Psi_I(z) = \big(K_I(z), P_I(z), \gamma(z)\big)$, where $\gamma(z)$ denotes the last coordinate of z.

Proposition 7 (Slice chart)**.** *The map Ψ_I is a real-analytic diffeomorphism from $\mathcal{F}$ onto $\mathcal{S}_I$. Define*

$$\mathcal{F}_{\rm st} := \{z \in \mathcal{F} : K_I(z) \in \mathcal{C}\}, \qquad \mathcal{S}_I^{\rm st} := \mathcal{S}_I \cap \mathcal{S}_{\rm nd}^{\rm st}.$$

Then $\mathcal{F}_{\rm st}$ is relatively open in $\mathcal{F}$, and Ψ_I restricts to a real-analytic bijection, with real-analytic inverse, between $\mathcal{F}_{\rm st}$ and $\mathcal{S}_I^{\rm st}$. Moreover, for every exact nondegenerate controller $K \in \mathcal{C}_{\rm nd}$ there exists $z \in \mathcal{F}_{\rm st}$ such that $K_I(z)$ is realization-equivalent to K and

$$\Psi_I(z) = \big(K_I(z), P_I(z), J(K)\big) \in \mathcal{S}_I^{\rm st}.$$

Proof. Because Ψ is a real-analytic diffeomorphism on $GL_{n_x} \times \mathcal{F}$, its restriction to the embedded submanifold $\{I_{n_x}\} \times \mathcal{F}$ is a real-analytic diffeomorphism onto its image. By Proposition 3, the first coordinate of $\Phi(K, P, \gamma)$ equals P_{12}. Hence $\Psi(I_{n_x}, z)$ has $P_{12} = I_{n_x}$ for every $z \in \mathcal{F}$, so the image is contained in $\mathcal{S}_I$.

Conversely, if $(K, P, \gamma) \in \mathcal{S}_I$, then $\Phi(K, P, \gamma) = (I_{n_x}, z)$ for a unique $z \in \mathcal{F}$, again because the first coordinate of Φ is P_{12}. Therefore $(K, P, \gamma) = \Psi(I_{n_x}, z) = \Psi_I(z)$, proving surjectivity onto $\mathcal{S}_I$. Since Ψ_I is the restriction of a diffeomorphism, it is itself a real-analytic diffeomorphism. Since the Hurwitz set $\mathcal{C}$ is open and K_I is continuous, $\mathcal{F}_{\rm st} = K_I^{-1}(\mathcal{C})$ is relatively open in $\mathcal{F}$. The stated restricted bijection and its analytic inverse follow directly from the definitions.

Now let $K \in \mathcal{C}_{\rm nd}$. By definition there exists $P \succ 0$ with $(K, P, J(K)) \in \mathcal{S}_{\rm nd}$. Applying Lemma 5, we obtain a realization-equivalent controller $\widehat{K}$ and a matrix $\widehat{P}$ such that $(\widehat{K}, \widehat{P}, J(K)) \in \mathcal{S}_I$. Since Ψ_I parameterizes $\mathcal{S}_I$, there exists $z \in \mathcal{F}$ with $\Psi_I(z) = \big(\widehat{K}, \widehat{P}, J(K)\big)$. Because $\widehat{K}$ is realization-equivalent to K, Lemma 4 gives $\widehat{K} \in \mathcal{C}$ and $J(\widehat{K}) = J(K)$. Hence $z \in \mathcal{F}_{\rm st}$ and the displayed certificate belongs to $\mathcal{S}_I^{\rm st}$. ☐

The identity gauge slice contains both exact and nonexact certificates. We isolate the exact part.

Definition 8 (Exact slice set)**.** *Define*

$$\mathcal{E}_I := \{z \in \mathcal{F}_{\rm st} : \gamma(z) = J(K_I(z))\}.$$

For a subset $\mathcal{Z} \subset \mathcal{F}_{\rm st}$, write $\mathcal{E}_I(\mathcal{Z}) := \mathcal{E}_I \cap \mathcal{Z}$.

Lemma 9. *The set $\mathcal{E}_I$ is relatively closed in $\mathcal{F}_{\rm st}$. Consequently, if $\mathcal{Z} \subset \mathcal{F}_{\rm st}$ is compact, then $\mathcal{E}_I(\mathcal{Z})$ is compact.*

Proof. By Proposition 7, K_I is continuous on $\mathcal{F}_{\rm st}$. Since J is locally Lipschitz on $\mathcal{C}$, it is continuous on $\mathcal{C}$. The map $z \mapsto \gamma(z)$ is continuous because it is just the last coordinate projection. Therefore

$$h(z) := \gamma(z) - J(K_I(z))$$

is continuous on $\mathcal{F}_{\rm st}$, and $\mathcal{E}_I = h^{-1}(\{0\})$ is relatively closed there. If $\mathcal{Z} \subset \mathcal{F}_{\rm st}$ is compact, then $\mathcal{E}_I(\mathcal{Z})$ is closed in $\mathcal{Z}$ and hence compact. □

IV. Quantitative first-order error bounds on the exact slice

Now, we present our main result, which connects the suboptimality of an exact slice member to its set of Clarke subdifferentials, thereby providing a backbone for the analysis of nonsmooth optimization algorithms.

Theorem 10 (Suboptimality Gap)**.** *Let $z, z^\star \in \mathcal{E}_I$, and*

$$\Psi_I(z) = \big(K, P, J(K)\big), \qquad \Psi_I(z^\star) = \big(K^\star, P^\star, J(K^\star)\big).$$

Then

$$J(K) - J(K^\star) \leq \operatorname{dist}\big(0, \partial J(K)\big) \cdot \big\| DK_I(z)[z^\star - z]\big\|. \quad (14)$$

In particular,

$$J(K) - J(K^\star) \leq \|DK_I(z)\|_{\rm op}\, \|z^\star - z\|\, \operatorname{dist}\big(0, \partial J(K)\big).$$

Proof. Fix $z, z^\star \in \mathcal{E}_I$ and define the segment

$$z_t := z + t(z^\star - z), \qquad t \in [0,1].$$

Since $\mathcal{F}$ is convex, $z_t \in \mathcal{F}$ for every $t \in [0,1]$. Let $\Psi_I(z_t) = \big(K_t, P_t, \gamma_t\big)$. Since $z \in \mathcal{E}_I \subset \mathcal{F}_{\rm st}$ and $\mathcal{F}_{\rm st}$ is relatively open in $\mathcal{F}$, there exists $t_0 \in (0,1]$ such that $z_t \in \mathcal{F}_{\rm st}$ for every $t \in [0, t_0]$. For such t, $K_t \in \mathcal{C}$ and $\Psi_I(z_t) \in \mathcal{S}_I \subset \mathcal{S}_{\rm nd}$, so the nonstrict bounded-real lemma yields

$$J(K_t) \leq \gamma_t, \qquad \forall t \in [0, t_0]. \quad (15)$$

Moreover, the γ-coordinate is affine along the segment:

$$\gamma_t = (1-t)\gamma(z) + t\gamma(z^\star).$$

Since $z, z^\star \in \mathcal{E}_I$, exactness gives $\gamma(z) = J(K), \gamma(z^\star) = J(K^\star)$. Combining this with (15), we obtain

$$\begin{aligned} \limsup_{t\downarrow 0} \frac{J(K_t) - J(K)}{t} &\leq \lim_{t\downarrow 0} \frac{\gamma_t - \gamma(z)}{t} \\ &= \gamma(z^\star) - \gamma(z) \\ &= J(K^\star) - J(K). \end{aligned} \quad (16)$$

Next, since K_I is real analytic on $\mathcal{F}$, it is differentiable at z and $K_t = K + tV + r(t)$, with

$$V := DK_I(z)[z^\star - z], \qquad \frac{\|r(t)\|}{t} \to 0 \quad (t \downarrow 0).$$

Because J is locally Lipschitz at K, there exist $L > 0$ and a neighborhood of K on which

$$|J(K_1) - J(K_2)| \leq L\|K_1 - K_2\|.$$

For all sufficiently small $t > 0$ we therefore have

$$\begin{aligned} \frac{J(K+tV) - J(K)}{t} &\leq \frac{J(K_t) - J(K)}{t} + \frac{|J(K+tV) - J(K_t)|}{t} \\ &\leq \frac{J(K_t) - J(K)}{t} + L\frac{\|K + tV - K_t\|}{t} \\ &= \frac{J(K_t) - J(K)}{t} + L\frac{\|r(t)\|}{t}. \end{aligned}$$

Taking the limit superior as $t \downarrow 0$ and using (16) yields

$$J'(K;V) = \lim_{t\downarrow 0} \frac{J(K+tV) - J(K)}{t} \leq J(K^\star) - J(K), \quad (17)$$

where the existence of the directional derivative follows from Proposition 1. Again, by Proposition 1,

$$J'(K;V) = \max_{G \in \partial J(K)} \langle G, V\rangle.$$

The Clarke subdifferential of a locally Lipschitz function is nonempty, compact, and convex, so there exists $G_{\min} \in \partial J(K)$ with $\|G_{\min}\| = \operatorname{dist}\big(0, \partial J(K)\big)$. Hence

$$\begin{aligned} J'(K;V) &\geq \langle G_{\min}, V\rangle \\ &\geq -\|G_{\min}\|\, \|V\| \\ &= -\operatorname{dist}\big(0, \partial J(K)\big)\, \|V\|. \end{aligned} \quad (18)$$

Combining (17) and (18),

$$J(K) - J(K^\star) \leq \operatorname{dist}\big(0, \partial J(K)\big)\, \|V\|,$$

which is (14). Finally,

$$\|V\| = \big\|DK_I(z)[z^\star - z]\big\| \leq \|DK_I(z)\|_{\rm op}\, \|z^\star - z\|,$$

which completes the proof. □

Theorem 10 is the quantitative refinement of the landscape argument in Tang and Zheng [14] on the reconditioned exact slice. It becomes especially useful on compact convex slices.

Corollary 11. *Let $\mathcal{Z} \subset \mathcal{F}_{\rm st}$ be compact and convex, and assume that $\mathcal{E}_I(\mathcal{Z}) \neq \varnothing$. Define*

$$J^{\rm ex}_{\mathcal{Z}} := \min_{z \in \mathcal{E}_I(\mathcal{Z})} J\big(K_I(z)\big),$$

which is well defined by Lemma 9. Also define

$$L_{\mathcal{Z}} := \sup_{z \in \mathcal{Z}} \|DK_I(z)\|_{\rm op}, \; D_{\mathcal{Z}} := \operatorname{diam}(\mathcal{Z}), \; C_{\mathcal{Z}} := L_{\mathcal{Z}} D_{\mathcal{Z}}.$$

Then for every $z \in \mathcal{E}_I(\mathcal{Z})$,

$$J\big(K_I(z)\big) - J^{\rm ex}_{\mathcal{Z}} \leq C_{\mathcal{Z}} \operatorname{dist}\big(0, \partial J(K_I(z))\big). \quad (19)$$

Proof. By Lemma 9, the set $\mathcal{E}_I(\mathcal{Z})$ is compact. Since $z \mapsto J(K_I(z))$ is continuous on $\mathcal{E}_I(\mathcal{Z})$, there exists $z^\star \in \mathcal{E}_I(\mathcal{Z})$ such that

$$J\big(K_I(z^\star)\big) = J^{\rm ex}_{\mathcal{Z}}.$$

Applying Theorem 10 to z and $z^\star$ gives

$$J\big(K_I(z)\big) - J^{\rm ex}_{\mathcal{Z}} \leq \|DK_I(z)\|_{\rm op}\, \|z - z^\star\|\, \operatorname{dist}\big(0, \partial J(K_I(z))\big).$$

Since $z, z^\star \in \mathcal{Z}$, we have $\|z - z^\star\| \leq D_{\mathcal{Z}}$ and $\|DK_I(z)\|_{\text{op}} \leq L_{\mathcal{Z}}$. Substituting these bounds yields (19) and proof is concluded. $\square$

This result can serve as a core for the analysis of nonsmooth optimization algorithms for output-feedback H_∞ policy search. If $\mathcal{E}_I(\mathcal{Z})$ contains a globally optimal controller for (6), then $J_{\mathcal{Z}}^{\text{ex}} = J^* = \inf_{K \in \mathcal{C}} J(K)$, and every Clarke-stationary $z \in \mathcal{E}_I(\mathcal{Z})$ is globally optimal. In addition, as we will show in the next section, the convex lift $\gamma(z)$ attains values arbitrarily close to J^*. As a result, we can establish a similar connection between global suboptimality and the Clarke subdifferential. We do not include this result in the interest of space.

In turn, we discuss algorithmic consequence of the value gap bounds.

Corollary 12. *Let $\mathcal{Z} \subset \mathcal{F}_{\text{st}}$ be compact and convex with $\mathcal{E}_I(\mathcal{Z}) \neq \varnothing$, and let $\{z_k\}_{k \geq 0} \subset \mathcal{E}_I(\mathcal{Z})$ be any sequence. Define the stationarity residual*

$$\varepsilon_k := \operatorname{dist}\big(0, \partial J(K_I(z_k))\big).$$

Then

$$J\big(K_I(z_k)\big) - J_{\mathcal{Z}}^{\text{ex}} \leq C_{\mathcal{Z}}\, \varepsilon_k, \qquad \forall k \geq 0. \tag{20}$$

In particular, if $\varepsilon_k \to 0$, then

$$J\big(K_I(z_k)\big) \to J_{\mathcal{Z}}^{\text{ex}}.$$

If $\mathcal{E}_I(\mathcal{Z})$ contains a globally optimal controller, then (20) is a global optimality-gap estimate.

Proof. This is an immediate specialization of Corollary 11 to the points z_k. $\square$

Corollary 12 is deliberately abstract. It does not prove that a particular Goldstein, gradient-sampling, or bundle method generates exact iterates in a compact slice, nor does it establish a complexity bound for driving ε_k to zero. These are separate algorithmic questions. What the corollary does provide is the missing bridge from approximate stationarity to value suboptimality *once* a reconditioned method is analyzed on the exact slice. This is the dynamic output-feedback counterpart (and hence more general) of the value-gap open problem explicitly raised in the state-feedback case by Guo and Hu [1], now considered on the nondegenerate ECL manifold introduced by Tang and Zheng [14].

It is widely known that even ε-Clarke stationarity faces computational intractability [1, 22]. Therefore, in order to obtain concrete convergence rates, we turn to more relaxed solution concepts such as (δ, ε)-Goldstein stationarity. While in general $\operatorname{dist}\big(0, \partial_\delta J(K)\big) \leq \operatorname{dist}\big(0, \partial J(K)\big)$, under an outer continuity assumption on the Clarke subdifferential map $K \mapsto \partial J(J)$, we can recover a reverse inequality up to an additive term. This, in turn, allows us to connect off-the-shelf convergence guarantees of nonsmooth optimization algorithms such as Zhang et al. [22], Davis et al. [23] for finding Goldstein stationary points to convergence of the objective value for the H_∞ policy search problem. We formalize this result in the following.

Assumption 13. *The Clarke subdifferential map $K \mapsto \partial J(K)$ is one-sided Hausdorff continuous; that is, for each $K \in \mathcal{C}_{nd}$, there exists a nondecreasing modulus $\psi_K : [0, \infty) \mapsto [0, \infty)$, with $\omega_K(r) \to 0$ as $r \downarrow 0$, such that*

$$\partial J(K') \subset \partial J(K) + \psi_K(\|K' - K\|)\mathbb{B}_2(0, 1),$$

for all K' near K.

A commonly used special case of Assumption 13 is the linear one, where $\psi_K(\|K' - K\|) = a\|K' - K\|$ for a positive constant a, which Kong and Lewis [24] refer to as upper Lipschitzness of the subdifferential.

In turn, we relate Assumption 13 to a relationship between $\operatorname{dist}\big(0, \partial_\delta J(K)\big)$ and $\operatorname{dist}\big(0, \partial J(K)\big)$.

Proposition 14. *Under Assumption 13, we have*

$$\operatorname{dist}\big(0, \partial J(K)\big) \leq \operatorname{dist}\big(0, \partial_\delta J(K)\big) + \psi_K(\delta).$$

Proof. Because the set $\partial J(K) + \psi_K(\|K' - K\|)\mathbb{B}_2(0, 1)$ is convex and contains every $\partial J(K')$ for all $K' \in \mathbb{B}_2(K, \delta)$, we have

$$\partial J_\delta(K) \subset \partial J(K) + \psi_K(\delta)\mathbb{B}_2(0, 1).$$

Therefore,

$$\operatorname{dist}\big(0, \partial J(K)\big) - \psi_K(\delta) \leq \operatorname{dist}\big(0, \partial_\delta J(K)\big),$$

and the proof is complete. $\square$

Theorem 15 (Convergence rate)**.** *Under Assumption 13, the Interpolated Normalized Gradient Descent (INGD) algorithm [22] takes $\widetilde{\mathcal{O}}(\frac{1}{\delta\varepsilon^3})$ gradient and function valuations to find a $z \in \mathcal{E}_I$ such that $K_I(z)$ is an (ε, δ)-Goldstein stationary point and more importantly $K_I(z)$ is $C_{\mathcal{Z}}\,(\varepsilon + \psi_K(\delta))$-suboptimal on the set $\mathcal{Z}$,*

$$J\big(K_I(z)\big) - J_{\mathcal{Z}}^{\text{ex}} \leq C_{\mathcal{Z}}\,(\varepsilon + \psi_K(\delta)).$$

Proof. The proof follows immediately by combining Proposition 14, Corollary 12, and the analysis of INGD from Davis et al. [23, Theorem 3.2]. $\square$

V. Global value computation and strict controller recovery in the convex lift

The exact-slice argument of Section IV is intrinsically local as it converts first-order information at an exact point into a value gap relative to an exact comparison point. The value equivalence used in this section is an established consequence of extended convex lifting (ECL). In particular, Zheng et al. [11, Theorem 2.1] show that a policy objective equipped with an ECL has the same infimum as its convex lifted problem, and Zheng et al. [11, Corollary 4.1] state this conclusion explicitly for dynamic output-feedback H_∞ control. We use this identity to distinguish nonstrict value computation from strict stabilizing-controller recovery and to formulate the resulting feasibility-bisection procedure. Although a nonstrict point of $\mathcal{F}$ need not decode to an internally stabilizing controller, strict lifted points do support stabilizing controller recovery.

A. Strict lifted points

Define $\mathcal{F}_{\mathrm{str}} := \{z \in \mathcal{F} : \mathcal{M}(z) \prec 0\}$.

Lemma 16. *For every $z \in \mathcal{F}_{\mathrm{str}}$,*

$$\Psi_I(z) = \big(K_I(z), P_I(z), \gamma(z)\big) \in \mathcal{S}_I \subset \mathcal{S}_{\mathrm{nd}},$$

and $K_I(z) \in \mathcal{C}$ with

$$J\big(K_I(z)\big) < \gamma(z).$$

Proof. By Proposition 7, $\Psi_I(z) \in \mathcal{S}_I \subset \mathcal{S}_{\mathrm{nd}}$. The strict-definiteness correspondence in Proposition 3 gives

$$N\big(K_I(z), P_I(z), \gamma(z)\big) \prec 0.$$

Its leading principal block therefore satisfies

$$A_{\mathrm{cl}}(K_I(z))^\top P_I(z) + P_I(z) A_{\mathrm{cl}}(K_I(z)) \prec 0.$$

Since $P_I(z) \succ 0$, the Lyapunov theorem implies $K_I(z) \in \mathcal{C}$. The strict bounded-real lemma then gives $J(K_I(z)) < \gamma(z)$. □

B. Strict nondegenerate approximation of stabilizing controllers

Lemma 17. *Let $K \in \mathcal{C}$ and let $\bar{\gamma} > J(K)$. Then there exists $P \in \mathbb{S}^{2n_x}_{++}$ such that*

$$N(K, P, \bar{\gamma}) \prec 0.$$

Proof. This is the standard strict bounded-real lemma in continuous time; see also the lemma invoked in the proof of Proposition 4 in Tang and Zheng [14]. □

Lemma 18 (Strict nondegenerate approximation)**.** *Let $K \in \mathcal{C}$ and let $\eta > 0$. Then there exists $z \in \mathcal{F}_{\mathrm{str}}$ such that $K_I(z)$ is realization-equivalent to K and*

$$\gamma(z) \le J(K) + \eta.$$

In fact, one may arrange $\gamma(z) = J(K) + \eta$.

Proof. Set $\bar{\gamma} := J(K) + \eta$. By Lemma 17, there exists $P \succ 0$ such that

$$N(K, P, \bar{\gamma}) \prec 0.$$

Because positive definiteness and strict negative definiteness are open conditions, there exists $\delta > 0$ such that whenever $E \in \mathbb{R}^{n_x \times n_x}$ satisfies $\|E\| < \delta$, the symmetric perturbation

$$\widetilde{P} := P + \begin{bmatrix} 0 & E \\ E^\top & 0 \end{bmatrix}$$

still obeys

$$\widetilde{P} \succ 0, \qquad N(K, \widetilde{P}, \bar{\gamma}) \prec 0.$$

Since GL_{n_x} is open and dense in $\mathbb{R}^{n_x \times n_x}$, we may choose such an E with $\widetilde{P}_{12} = P_{12} + E \in GL_{n_x}$. Hence

$$(K, \widetilde{P}, \bar{\gamma}) \in \mathcal{S}_{\mathrm{nd}}.$$

Applying Lemma 5, we obtain a controller $\widehat{K}$ realization-equivalent to K and a matrix $\widehat{P}$ such that

$$(\widehat{K}, \widehat{P}, \bar{\gamma}) \in \mathcal{S}_I, \qquad N(\widehat{K}, \widehat{P}, \bar{\gamma}) \prec 0,$$

where strictness is preserved by the congruence in the proof of Lemma 5. Now Proposition 7 yields a unique $z \in \mathcal{F}$ satisfying

$$\Psi_I(z) = \big(\widehat{K}, \widehat{P}, \bar{\gamma}\big).$$

By Proposition 3, $\mathcal{M}(z) \prec 0$, so $z \in \mathcal{F}_{\mathrm{str}}$. Thus $K_I(z) = \widehat{K}$ is realization-equivalent to K, and

$$\gamma(z) = \bar{\gamma} = J(K) + \eta,$$

which proves the claim. □

C. Value identity

Now, we reiterate the value identity of ECL which indicates that $\inf_{z \in \mathcal{F}} \gamma(z) = \inf_{K \in \mathcal{C}} J(K)$.

Proposition 19 (Lifted-value identity)**.** *Assume $\mathcal{C} \neq \varnothing$. Then the global optimal value of (6) satisfies*

$$J^\star := \inf_{K \in \mathcal{C}} J(K) = \inf_{z \in \mathcal{F}} \gamma(z) =: J_{\mathrm{lift}}. \tag{21}$$

No attainment assumption on a prescribed exact slice is required.

Proof. This is an application of the general extended-convex-lifting value theorem of Zheng et al. [11, Theorem 2.1]; its dynamic output-feedback H_∞ specialization is stated in Zheng et al. [11, Corollary 4.1]. For the generalized-plant coordinates used here, Proposition 3 supplies the lifting diffeomorphism and preserves the γ coordinate. Lemma 18 supplies strict lifted certificates arbitrarily close to every stabilizing-controller cost, while convex interpolation between any point of $\mathcal{F}$ and a strict point of $\mathcal{F}_{\mathrm{str}}$ supplies the required closure inclusion for the nonstrict lift. Thus the hypotheses of the cited value theorem hold and yield (21). □

Corollary 20. *Suppose there exists $z^\star \in \mathcal{F}_{\mathrm{st}}$ such that $\gamma(z^\star) = \min_{z \in \mathcal{F}} \gamma(z) = J_{\mathrm{lift}}$. Then $K_I(z^\star)$ is globally optimal and $J\big(K_I(z^\star)\big) = \gamma(z^\star) = J^\star$.*

Proof. Since $z^\star \in \mathcal{F}_{\mathrm{st}}$, $K_I(z^\star) \in \mathcal{C}$. Also $\Psi_I(z^\star) \in \mathcal{S}_I \subset \mathcal{S}_{\mathrm{nd}}$, so the nonstrict bounded-real lemma gives

$$J\big(K_I(z^\star)\big) \le \gamma(z^\star) = J_{\mathrm{lift}} = J^\star.$$

Since $J^\star$ is the infimum of J over $\mathcal{C}$, the reverse inequality is automatic. Hence equality holds throughout. □

Remark 21. *Proposition 19 records a statement about the optimal value, not about pointwise stable decoding of the full nonstrict lift. A minimizer in $\mathcal{F}$ need not lie in $\mathcal{F}_{\mathrm{st}}$. If such a boundary minimizer $z^\star$ exists, then convex* *combinations with any $z^\circ \in \mathcal{F}_{\rm str}$ yield strict lifted points whose decoded stabilizing-controller costs converge to $J^\star$. Degeneracy may also prevent the infimum from being attained by a point of $\mathcal{F}$, but neither phenomenon affects the value identity $J^\star = J_{\rm lift}$. Nondegeneracy and exactness re-enter when one wants local first-order certificates on the identity slice.*

D. Bisection over the convex lift

For a fixed scalar $\bar\gamma \in \mathbb{R}$, consider the feasibility problems

$$\mathrm{Feas}(\bar\gamma): \quad \exists (X,Y,M,H,F,L) \text{ such that}$$

$$\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \succ 0, \quad \text{and} \quad \mathcal{M}(X,Y,M,H,F,L,\bar\gamma) \preceq 0,$$

and

$$\mathrm{StrictFeas}(\bar\gamma): \quad \exists (X,Y,M,H,F,L) \text{ such that}$$

$$\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \succ 0, \quad \text{and} \quad \mathcal{M}(X,Y,M,H,F,L,\bar\gamma) \prec 0.$$

These are convex feasibility problems because the constraints are affine in the decision variables (X,Y,M,H,F,L).

Proposition 22 (Monotone feasibility threshold)**.** *The feasibility predicate* $\mathrm{Feas}(\bar\gamma)$ *is monotone nondecreasing in* $\bar\gamma$*: if* $\mathrm{Feas}(\bar\gamma)$ *holds and* $\bar\gamma' \geq \bar\gamma$*, then* $\mathrm{Feas}(\bar\gamma')$ *also holds. Moreover,*

$$\begin{aligned} &\mathrm{Feas}(\bar\gamma) \implies \bar\gamma \geq J^\star, \\ &\bar\gamma > J^\star \implies \mathrm{StrictFeas}(\bar\gamma) \implies \mathrm{Feas}(\bar\gamma). \end{aligned}$$

Proof. If $\mathrm{Feas}(\bar\gamma)$ holds, let (X,Y,M,H,F,L) be a feasible point. Replacing $\bar\gamma$ by a larger value $\bar\gamma' \geq \bar\gamma$ only subtracts the positive semidefinite block $(\bar\gamma' - \bar\gamma)\,\mathrm{diag}(0,0,I,I)$ from $\mathcal{M}$, so feasibility is preserved. This proves monotonicity.

If $\mathrm{Feas}(\bar\gamma)$ holds, then some $z \in \mathcal{F}$ has $\gamma(z) = \bar\gamma$. Proposition 19 therefore gives

$$\bar\gamma \geq \inf_{\tilde z \in \mathcal{F}} \gamma(\tilde z) = J^\star.$$

Conversely, suppose $\bar\gamma > J^\star$. By the definition of the infimum, there exists $K \in \mathcal{C}$ with $J(K) < \bar\gamma$. Applying Lemma 18 with $\eta = \bar\gamma - J(K)$ produces $z \in \mathcal{F}_{\rm str}$ with $\gamma(z) = \bar\gamma$. Hence $\mathrm{StrictFeas}(\bar\gamma)$ holds, and strict feasibility trivially implies nonstrict feasibility. ☐

Theorem 23 (Convex-lift bisection algorithm)**.** *Assume $\mathcal{C} \neq \varnothing$ and fix $\varepsilon > 0$. Consider the following procedure.*

1) *Starting from $m = 0$, test* $\mathrm{Feas}(2^m)$ *until the first feasible index $m = m_0$ is found.*
2) *Set $L_0 := 0$ and $U_0 := 2^{m_0}$.*
3) *Given $[L_j, U_j]$, form $G_j := (L_j + U_j)/2$.*
   - *If* $\mathrm{Feas}(G_j)$ *holds, set $L_{j+1} := L_j$ and $U_{j+1} := G_j$.*
   - *If* $\mathrm{Feas}(G_j)$ *fails, set $L_{j+1} := G_j$ and $U_{j+1} := U_j$.*
4) *Stop when $U_j - L_j \leq \varepsilon/2$, set*

$$\widehat\gamma := U_j + \frac{\varepsilon}{2},$$

solve $\mathrm{StrictFeas}(\widehat\gamma)$ *to obtain $\widehat z \in \mathcal{F}_{\rm str}$ with $\gamma(\widehat z) = \widehat\gamma$, and output*

$$K_\varepsilon := K_I(\widehat z).$$

Then the procedure uses at most $m_0 + 1 + \max\left\{0, \left\lceil \log_2\left(\frac{2U_0}{\varepsilon}\right)\right\rceil\right\}$ nonstrict feasibility tests and one strict-feasibility test. In particular, the total number of feasibility-oracle calls is $\mathcal{O}\big(1 + \log\max\{1, J^\star\} + \log\max\{1, \varepsilon^{-1}\}\big)$. On termination,

$$J^\star \leq U_j \leq J^\star + \frac{\varepsilon}{2}, \qquad K_\varepsilon \in \mathcal{C}, \qquad J(K_\varepsilon) < \widehat\gamma \leq J^\star + \varepsilon.$$

Hence K_ε is an ε-optimal stabilizing controller.

Proof. Since $\mathcal{C} \neq \varnothing$, Proposition 19 gives $J^\star < \infty$. Therefore $2^m > J^\star$ for all sufficiently large m, and Proposition 22 guarantees that the doubling phase terminates after $m_0 + 1$ feasibility tests.

Every interval $[L_j, U_j]$ constructed by the algorithm satisfies $L_j \leq J^\star \leq U_j$. The base case follows from $L_0 = 0 \leq J^\star$ and feasibility of U_0. If $\mathrm{Feas}(G_j)$ holds, Proposition 22 gives $J^\star \leq G_j = U_{j+1}$; if it fails, the implication $\bar\gamma > J^\star \Rightarrow \mathrm{Feas}(\bar\gamma)$ gives $G_j \leq J^\star$, so $L_{j+1} = G_j$. Thus the invariant is preserved.

After n bisection steps, the interval length is $U_0/2^n$. Hence the stopping condition is reached after at most $\max\left\{0, \left\lceil \log_2\left(\frac{2U_0}{\varepsilon}\right)\right\rceil\right\}$ bisection steps. The algorithm then performs one final strict-feasibility test.

At termination, $J^\star \leq U_j \leq L_j + \frac{\varepsilon}{2} \leq J^\star + \frac{\varepsilon}{2}$. Consequently, $\widehat\gamma = U_j + \frac{\varepsilon}{2} > J^\star$ and $\widehat\gamma \leq J^\star + \varepsilon$. Proposition 22 guarantees $\mathrm{StrictFeas}(\widehat\gamma)$. Lemma 16 applied to the recovered strict point gives $K_\varepsilon \in \mathcal{C}$ and $J(K_\varepsilon) < \widehat\gamma \leq J^\star + \varepsilon$. Thus K_ε is an ε-optimal stabilizing controller. ☐

VI. Conclusion

We gave a quantitative and algorithmic treatment of nonsmooth dynamic output-feedback H_∞ policy search via the extended convex lift. On the stable exact identity-gauge slice, we proved that approximate stationarity controls suboptimality on compact exact slices. For global computation, we used the established lifted-value identity to combine nonstrict-feasibility bisection with one strict-feasibility recovery step, producing an explicit ε-optimal stabilizing controller. Future work includes weakening the outer-continuity assumption on the Clarke subdifferential, analyzing concrete algorithms that operate near the exact slice without requiring exact iterates, and extending the framework to structured or reduced-order controllers and data-driven robust control.

ACKNOWLEDGMENT

AS and PJ were partially supported by the ONR grant N00014-24-1-2470. Part of this research was performed while the AS was visiting the Institute for Mathematical and Statistical Innovation (IMSI), which is supported by the National Science Foundation (Grant No. DMS-2425650).

References

[1] X. Guo and B. Hu, "Global convergence of direct policy search for state-feedback $\mathcal{H}_\infty$ robust control: A revisit of nonsmooth synthesis with goldstein subdifferential," *Advances in Neural Information Processing Systems*, vol. 35, pp. 32 801–32 815, 2022.

[2] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," *Advances in neural information processing systems*, vol. 12, 1999.

[3] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *International conference on machine learning*. Pmlr, 2014, pp. 387–395.

[4] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *International conference on machine learning*. PMLR, 2015, pp. 1889–1897.

[5] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[6] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[7] B. Recht, "A tour of reinforcement learning: The view from continuous control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, no. 1, pp. 253–279, 2019.

[8] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," in *International conference on machine learning*. PMLR, 2018, pp. 1467–1476.

[9] K. Zhang, B. Hu, and T. Basar, "Policy optimization for H_2 linear control with H_∞ robustness guarantee: Implicit regularization and global convergence," *SIAM Journal on Control and Optimization*, vol. 59, no. 6, pp. 4081–4109, 2021.

[10] J. Umenberger, M. Simchowitz, J. C. Perdomo, K. Zhang, and R. Tedrake, "Globally convergent policy search over dynamic filters for output estimation," *arXiv preprint arXiv:2202.11659*, 2022.

[11] Y. Zheng, C.-F. Pai, and Y. Tang, "Extended convex lifting for policy optimization of optimal and robust control," *Proceedings of Machine Learning Research vol*, vol. 283, pp. 1–13, 2025.

[12] B. Hu and Y. Zheng, "Connectivity of the feasible and sublevel sets of dynamic output feedback control with robustness constraints," *IEEE Control Systems Letters*, vol. 7, pp. 442–447, 2022.

[13] S. Kraisler and M. Mesbahi, "Output-feedback synthesis orbit geometry: Quotient manifolds and lqg direct policy optimization," *IEEE Control Systems Letters*, vol. 8, pp. 1577–1582, 2024.

[14] Y. Tang and Y. Zheng, "On the global optimality of direct policy search for nonsmooth $\mathcal{H}_\infty$ output-feedback control," in *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2023, pp. 6148–6153.

[15] Y. Zheng, C.-F. R. Pai, and Y. Tang, "Benign nonconvex landscapes in optimal and robust control, part i: Global optimality," *IEEE Transactions on Automatic Control*, 2026.

[16] Y. Watanabe, F.-Y. Liao, and Y. Zheng, "Policy optimization in robust control: Weak convexity and subgradient methods," *arXiv preprint arXiv:2509.25633*, 2025.

[17] J. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard H_2 and H_∞ control problems," in *1988 American control conference*. IEEE, 1988, pp. 1691–1696.

[18] I. Masubuchi, A. Ohara, and N. Suda, "Lmi-based controller synthesis: a unified formulation and solution," *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, vol. 8, no. 8, pp. 669–686, 1998.

[19] F. H. Clarke, *Optimization and nonsmooth analysis*. SIAM, 1990.

[20] R. T. Rockafellar and R. J. Wets, *Variational analysis*. Springer, 1998.

[21] A. A. Goldstein, "Optimization of lipschitz continuous functions," *Mathematical Programming*, vol. 13, no. 1, pp. 14–22, 1977.

[22] J. Zhang, H. Lin, S. Jegelka, S. Sra, and A. Jadbabaie, "Complexity of finding stationary points of nonconvex nonsmooth functions," in *International Conference on Machine Learning*. PMLR, 2020, pp. 11 173–11 182.

[23] D. Davis, D. Drusvyatskiy, Y. T. Lee, S. Padmanabhan, and G. Ye, "A gradient sampling method with complexity guarantees for lipschitz functions in high and low dimensions," *Advances in neural information processing systems*, vol. 35, pp. 6692–6703, 2022.

[24] S. Kong and A. S. Lewis, "Lipschitz minimization and the goldstein modulus: S. kong, as lewis," *Mathematical Programming*, pp. 1–30, 2025.