Robust Optimization via Continuous-Time Dynamics

Keivan Ebrahimi, Nicola Elia, and Umesh Vaidya

***Abstract*— Robust optimization arises in numerous applications where system parameters are uncertain but confined to known sets. This paper introduces robust optimization (RO) dynamics—a continuous-time dynamical system that directly solves min–max robust optimization problems without requiring reformulations or explicit uncertainty model specifications. Our key contributions are twofold: (i) we establish the saddle-point property without assuming joint convexity–concavity, and (ii) we construct a novel Lyapunov function that proves global convergence for general robust optimization problems that are convex in the decision variable and concave in the uncertainty. Notably, the method does not require explicit knowledge of the cost or constraint functions, enabling model-free operation and decentralized implementation in real-time applications. Numerical experiments demonstrate substantial speedups over scenario-based methods, including exact solutions for robust quadratic programming with ellipsoidal uncertainties, nonlinear problems where robust counterparts cannot be derived, and distributed sensor placement. These results position continuous-time dynamics as a powerful alternative paradigm for robust optimization, particularly well suited for online adaptation and implementation in physical systems.**

I. Introduction

Motivation and Background

Optimization problems in engineering, finance, and machine learning often involve uncertain parameters that must be handled robustly to ensure reliable performance. Two different approaches can be used: stochastic optimization, where uncertainty is treated as a random variable, or RO, where uncertainty is assumed to be deterministic and bounded. Unlike stochastic optimization, RO does not require any known probability distributions in the problem data. Instead, RO assumes that the uncertain data reside within a predefined uncertainty set, for which constraint violation cannot be tolerated. For more information on robust and stochastic optimization-based approaches, see [1], [2], and the references therein. Early works on RO include [3], which considered robust linear optimization with ellipsoidal uncertainty sets, and [4], which presented exact solutions of inexact linear programs as a simple case of RO.

Financial support from the National Science Foundation grants CNS-1329915, ECCS-1150405, CIF-1220643, ECCS-1810079, and ECCS-203157 and from AFOSR grant FA95500119 are gratefully acknowledged. K. Ebrahimi is with the Department of Electrical & Computer Engineering, Iowa State University, Ames, Iowa. U. Vaidya is with the Department of Mechanical Engineering, Clemson University, Clemson SC. N. Elia is with the Department of Electrical and Computer Engineering, University of Minnesota, Twin Cities. keivan@iastate.edu, uvaidya@clemson.edu, nelia@umn.edu

Literature Review

The problem of robust linear programming was studied in [5], while robust conic-quadratic optimization and robust semi-definite optimization were discussed in [6] and [7], respectively. Refer to [2] and [8] for comprehensive surveys of RO problem solutions. Recent developments in distributionally robust optimization (DRO) have bridged the gap between stochastic and robust approaches [9], [10]. [11] and [12] showed that machine learning algorithms such as the norm-regularized support vector machine and the Lasso problem could be interpreted as RO problems. Recent work has further integrated robust optimization with machine learning [13], [14].

One of the main standard approaches for solving RO problems involves constructing a robust counterpart (RC) equivalent to the RO problem [1]. This widely used method essentially tries to find a deterministic equivalent to RO problem through a reformulation. In this sense, the practicality of robust programming depends on whether or not its RC is computationally tractable. An overview of different RO problems with tractable conjugates can be found in [15] Table 1 for some of which RC cannot be found. The reformulation approach to solving the RO problem, which is often a challenging, albeit usually convex, optimization problem, has the deficiency of suffering from case-by-case scenarios depending on the specific form of the uncertainty constraint and the specific form of the uncertainty set. In other words, depending on how simple the uncertainty set looks and based on the optimization problem type (whether it is linear programming, quadratic programming, second-order cone programming, semi-definite programming, etc.), an RC is being calculated and provided at hand. These reformulations may be computationally more expensive than other approaches [16]. Hence, the absence of a unified framework for solving a general class of RO problems without prior knowledge of the problem formulation is strongly felt.

By calculating a concave conjugate of nonlinear constraint functions and supporting the function of uncertainty sets, [17] and [15] used the Fenchel duality to obtain a tractable RC for new classes of robust nonlinear optimization (RNO) problems. However, when no closed form is available for the convex conjugate function of an uncertainty set or convex/concave conjugates of a non-linear constraint, these approaches still cannot obtain the RC for many sets of nonlinear uncertainties and constraints. Moreover, all these methods require either explicit problem structure knowledge or extensive computational resources, unlike our dynamics-based approach.

An alternative randomized approach based on constraint sampling is an approximate probabilistic relaxation solution to RO problems as seen in [18] and [19] where a finite set of high-dimensional deterministic optimization problems obtained from sampling are solved. This approach does not require concavity of the constraint in the uncertain variable but due to the large number of required scenarios to approximate the stochasticity of these problems, the stochastic optimization involves formulating a large-scale scenario program, which is in general computationally demanding. Recent methods include alternating direction method of multipliers (ADMM) decomposition [20], achieving improved convergence rates for distributed robust optimization problems.

In another effort, [21] incorporated the min-max behavior of RO problems to solve them by an oracle-based approximate robust optimization algorithm based on oracle-based subgradient descent and interior point methods. The proposed algorithms find an approximate robust solution using a number of calls to an oracle that solves the original (non-robust) problem. However, the solution would be approximated with some predetermined accuracy, η, with the number of iterations growing to $\mathcal{O}(\frac{1}{\eta})$, as the algorithm approximates the RC by invoking the Oracle a polynomial number of times. Classical cutting-plane approaches [22] treat RO problems as semi-infinite programming but have limitations when pessimization oracles are approximate. In another effort, Robust optimization over time (ROOT) combines robust optimization with dynamic optimization [9], [23].

Our Approach

In this paper, we provide the architecture of a dynamical system that converges to the optimal robust solution for a large class of RO problems. Our approach builds on classical primal-dual systems [24], [25] but addresses the unique challenges posed by the min-max structure of robust optimization.

Main Contributions

Unified framework: Our method provides a single dynamical system architecture that handles all convex-concave RO problems uniformly, eliminating the need for problem-specific reformulations or case-by-case analysis required by robust counterpart methods.

Novel saddle point theory: We prove that the saddle point property holds for RO problems despite the Lagrangian lacking joint convexity-concavity—a fundamental violation of standard theory. This theoretical achievement (see Section IV) enables the application of dynamical methods to robust optimization.

Custom Lyapunov analysis: We construct a non-standard Lyapunov function that establishes global asymptotic stability of the equilibrium corresponding to the robust optimal solution. The Lyapunov function design handles the complex interaction between primal, dual, and uncertainty variables.

Real-time adaptation: The continuous-time nature of our approach enables real-time tracking of time-varying uncertainties and online adaptation, making it suitable for dynamic environments and decentralized implementation.

Broader problem scope: Our method provides exact solutions for problems where robust counterparts are unknown or computationally intractable, including nonlinear constraints with complex uncertainty sets.

These contributions establish continuous-time dynamics as a powerful alternative paradigm for robust optimization, complementing existing discrete algorithmic approaches. The proposed continuous-time optimization system can solve a general class of $\mathcal{RO}$ problems where the cost function and the constraints are convex (concave) with respect to the decision variables (uncertain variables) and the uncertainty sets are convex[1]. This class includes all the cases in [15, Table 1].

Paper Organization

The remainder of this paper is organized as follows. In Section III, we present the problem statement of $\mathcal{RO}$ in a slightly generalized form. In Section IV, the characterization of saddle property and Karush-Kuhn-Tucker (KKT) optimality conditions along with the Lagrangian function for $\mathcal{RO}$ problem are provided. The main results on how to construct the $\mathcal{RO}$ dynamics is presented in Section V. This section also includes the Lyapunov-based global convergence result. We then present simulation results in Section VII followed by conclusions in Section VIII. Finally, detailed proofs are mentioned in the Appendix.

II. Notations

We define $i = 0, \cdots, N$ as $i \in [N]$, and $i = 1, \cdots, N$ as $i \in [N]^+$. In addition, $[\cdot]^+_\eta$ shows the positive projection defined as follows for the scalar-valued function P

$$[P]^+_\eta := \begin{cases} P & \text{if} \quad P > 0 \text{ or } \eta > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{1}$$

For vector-valued P, the projection is defined element-wise.

III. Robust Optimization Problems

Classical Formulation

Given an objective $f_0(x)$ to optimize, subject to constraints $f_i(x, u_i) \leq 0$ with uncertain parameters, $\{u_i\}$, a classical robust optimization problem has the following form

$$\begin{aligned} &\mu := \min_x f_0(x) \\ &\text{s.t.} \;\; f_i(x, u_i) \leq 0, \; \forall u_i \in \mathcal{U}_i, \; i \in [N]^+. \end{aligned}$$

where $x \in \mathbb{R}^n$ is a vector of decision variables, f_0 and f_i are $\mathbb{R}^n \to \mathbb{R}$ functions, and the uncertainty parameters $u_i \in \mathbb{R}^{m_i}$ are assumed to take arbitrary values in certain convex compact uncertainty sets $\mathcal{U}_i \subseteq \mathbb{R}^{m_i}$. The above problem is typically a semi-infinite optimization due to the cardinality of the $\mathcal{U}_i$. However, each robust constraint can be rewritten as a maximization problem as[2]

$$\begin{aligned} &\mu := \min_x f_0(x) \\ &\text{s.t.} \;\; \max_{u_i \in \mathcal{U}_i} f_i(x, u_i) \leq 0, \;\; i \in [N]^+ \end{aligned}. \tag{2}$$

[1] Preliminary convergence results for a special formulation appeared without proofs in [26]. In this paper, we generalize the problem formulation and provide complete proofs, which are the core of our contributions.

[2] Since the uncertainty set is compact and the constraint functions are continuous, the supremum is attained within the set; therefore, we can replace "sup" with "max" in our formulation.

Generalized Formulation

We consider a slight variation of (2), which takes the following under the assumptions stated below.

$$\mathcal{RO}_0 \begin{cases} \mu := \min_x \max_{u_0 \in \mathcal{U}_0} f_0(x, u_0) \\ \text{s.t.} \quad \max_{u_i \in \mathcal{U}_i} f_i(x, u_i) \leq 0 \ , \ i \in [N]^+ , \\ \\ \mathcal{U}_i := \{u_i \in \mathbb{R}^{m_i} \ : \ h_{ij}(u_i) \leq 0, \\ \qquad\qquad j \in [K_i]\} \ , \ i \in [N] \end{cases} \tag{3}$$

Here, $h_{ij}(u_i)$ represents the j-th constraint function defining the i-th uncertainty set $\mathcal{U}_i$, and K_i denotes the total number of constraints that define $\mathcal{U}_i$.

Remark 1. *Formulation (3) has three key properties:*

1) ***Generalization:*** *Formulation (3) generalizes (2) by allowing uncertainty in the objective function and explicitly representing uncertainty sets via inequality constraints, facilitating the min-max-max-min structure essential for our dynamical approach.*
2) ***Compactness:*** *Under the convexity assumptions stated below, uncertainty sets $\mathcal{U}_i$ remain compact since convex constraints preserve convexity and boundedness.*
3) ***Nonlinear constraints:*** *Our framework handles nonlinear uncertainty constraints $h_{ij}(u_i) \leq 0$ through gradient terms $\nabla_{u_i} h_i(u_i)$ that naturally accommodate nonlinearity.*

Assumptions and Properties

The problem in (3) reduces to the one in (2) when $\mathcal{U}_0$ is a singleton. Following [27] and without loss of generality, we consider the $\mathcal{RO}$ problem dealing with constraint-wise uncertainties where each constraint f_i is only a function of u_i. The functions f_i and h_{ij} have scalar values with the following assumptions.

Assumption 1. *The functions $h_{ij}(u_i)$ are convex in u_i for $i \in [N]$ and $j \in [K_i]^+$. The function $f_0(x, u_0)$ is strictly convex in x for any $u_0 \in \mathcal{U}_0$, and concave in u_0 for any x. Also, for $i \in [N]^+$, $f_i(x, u_i)$ is convex in x for fixed u_i and concave in $u_i \in \mathcal{U}_i$, for fixed x. Finally, the functions f_i and h_i $i \in [N]$ are C^1 with local Lipschitz gradients.*

Remark 2. *The Assumption 1, which requires convexity in the decision variable and concavity in the uncertainty variable, ensures computational tractability and is satisfied by most practical $\mathcal{RO}$ problems.*

Assumption 2. *Existence of optimal solutions and strong duality.*

1) *$\mathcal{RO}$ problem (4) is feasible. An optimal min-max solution $(x^\star, u^\star)$ exists and μ in the $\mathcal{RO}$ problem (4) is finite.*
2) *$\mathcal{RO}$ problem satisfies the Slater constraint qualification, for both the upper level optimization and the lower level maximization problems [27], that is, for all $i \in [N]$ and all $j \in [K_i]^+$, there exist $u_i \in \mathcal{U}_i$ such that $h_{ij}(u_i) < 0$, and for $i \in [N]^+$, there exists $x \in \mathbb{R}^n$ such that, $\mathcal{F}_i(x) < 0$.*

Remark 3. *It should be noted that the uncertainty is often parametrized affinely in $\mathcal{RO}$ problem; hence, the concavity property of the f_i functions is automatically satisfied.*

Remark 4 (Justification of Assumption 2)**.** *The Slater constraint qualification in Assumption 2 guarantees that the $\mathcal{RO}$ problem enjoys strong duality for upper and lower level optimization problems [28, Section 5.2.3, 5.9.1]. While this assumption may appear restrictive, it can often be satisfied or relaxed in practice:*

1) *In many robust optimization applications, the uncertainty sets are designed with strict feasibility in mind (e.g., ellipsoidal sets with positive radius).*
2) *The assumption ensures that saddle point and optimal dual solutions exist [27], which is crucial for our dynamical system approach.*
3) *For cases where Slater conditions are too restrictive, alternative constraint qualifications such as those discussed in [29] can be employed.*
4) *In practical implementations, the assumption can often be enforced by appropriate scaling or small perturbations of the constraint sets.*

Regularized Formulation

Finally, we introduce the problem we consider in this paper, a slight generalization of $\mathcal{RO}_0$ in (3).

$$\mathcal{RO} \begin{cases} \mu := \min_x \max_{u_0 \in \mathcal{U}_0} f_0(x, u_0) \\ \quad + \sum_{i=1}^N c_i \max_{u_i \in \mathcal{U}_i} f_i(x, u_i) \\ \\ \text{s.t.} \quad \max_{u_i \in \mathcal{U}_i} f_i(x, u_i) \leq 0 \ , \ i \in [N]^+ , \\ \\ \mathcal{U}_i := \{u_i \in \mathbb{R}^{m_i} \ : \ h_{ij}(u_i) \leq 0, \\ \qquad\qquad j \in [K_i]\} \ , \ i \in [N] \end{cases} \tag{4}$$

where $c_i \geq 0$ for $i \in [N]^+$. This setting can be seen as an elementary regularized version of more common formulations $\mathcal{RO}_0$ obtained for $c_i = 0$, $i \in [N]^+$. The role of c_i is further clarified below.

It is convenient to rewrite $\mathcal{RO}$ in the following form

$$\begin{aligned} \mu := \min_x \ & \mathcal{F}_0(x) + \sum_{i=1}^N c_i \mathcal{F}_i(x) \\ \text{s.t.} \quad & \mathcal{F}_i(x) \leq 0 \ , \ i \in [N]^+ \ , \end{aligned} \tag{5}$$

where

$$\mathcal{F}_i(x) := \max_{u_i \in \mathcal{U}_i} f_i(x, u_i) \ , \ i \in [N] \ . \tag{6}$$

In this paper, we often call (6) as the "lower optimization problems", and the minimization in (5) as the "upper optimization".

Remark 5 (Problem Formulation and c_i Terms)**.** *Our formulation (4) adds regularization terms c_i to the classical $\mathcal{RO}$ problem for three key reasons: (i) preventing singularity when constraints are inactive ($\lambda_i = 0$), (ii) improving numerical stability, (iii) recovering the classical problem as $c_i \to 0$. We maintain separate c_i and λ_i (rather than combined γ_i) to preserve dual variable interpretation and enable our Lyapunov construction.*

As already mentioned, our formulation includes the typical robust optimization formulation $\mathcal{RO}_0$

$$\begin{array}{ll} \mu = & \min_x f_0(x) \\ s.t. & \mathcal{F}_i(x) \leq 0, \;\; i \in [N]^+ \end{array} , \tag{7}$$

or the form below that is popular in the machine learning context [30], [31]

$$\mu = \min_x \max_{u \in \mathcal{U}} f_0(x, u) \; . \tag{8}$$

We finally introduce the following assumption valid for most of this paper.

Assumption 3. *$c_i > 0$ for $i \in [N]^+$.*

Remark 6. *Assumption 3 provides regularization for inactive constraints. In practice, use small values (like $c_i = 10^{-6}$) for numerical stability. Section VI analyzes the $c_i \to 0$ limit rigorously.*

A. Robust feasible solution and robust counterpart

A meaningful solution to $\mathcal{RO}$ problem (4) has to be immune against the uncertainties in the sense that the solution vector x should satisfy the constraints for all u_i's within the uncertainty sets[3]. Such vector x is called a robust feasible solution (RFS). One approach to solving the problem (4) is to try to compute (6) in closed form.

For various important combinations of constraint functions f_i and uncertainty sets $\mathcal{U}_i$, ($i > 0$) it is possible to obtain an explicit convex function $\mathcal{F}_i$ [1]. A classical example is when f_i is linear in x for fixed u_i and linear in u_i for fixed x, while $\mathcal{U}_i$ is an ellipsoidal set. Then, $\mathcal{F}_i$ can be easily derived as an explicit second-order conic function. In this case, $\mathcal{RO}$ problem (4) becomes a nominal optimization problem (not affected by uncertainty) known as the explicit RC:

$$\min_{\mathcal{F}_i(x) \leq 0} f_0(x) \; . \tag{9}$$

As shown in [27], the RC is always a convex optimization problem under Assumption 1. While the RC is known for important classes of problems as described in [2], this approach requires problem-specific derivations; moreover, RC is generally difficult to find (See Section VII for an example). Instead, our proposed approach has a dynamical system that simultaneously finds the best RFS and the worst parameters u_i' s, independently of the specifics of the constraint functions and the uncertainty sets.

IV. Duality, KKT conditions and saddle property

Lagrangian Formulation

The basic idea of this paper is to combine the usually separated and nested optimization (5) and (6) into one Lagrangian optimization.

To derive the Lagrangian function for $\mathcal{RO}$ problem (4), we introduce Lagrange multipliers $\lambda_i \geq 0$ for $i \in [N]^+$, and let $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_N)^\top \in \mathbb{R}^N_+$ where the non-negative orthant of $\mathbb{R}^N$ is denoted by $\mathbb{R}^N_+$. Note that μ in $\mathcal{RO}$ problem (4) can be written in short-hand notation as follows

$$\mu = \min_x \max_{\lambda \geq 0} \mathcal{F}_0(x) + \sum_{i=1}^N (c_i + \lambda_i) \, \mathcal{F}_i(x) \; , \tag{10}$$

where $\mathcal{F}_i(x)$ is given by (6). By introducing Lagrange multipliers $v_{ij} \geq 0$ for the maximization problem in (6) and defining

$$v_i := (v_{i1}, \cdots, v_{iK_i})^\top \in \mathbb{R}^{K_i}_+, \;\; h_i := (h_{i1}, \cdots, h_{iK_i})^\top \; ,$$

the lower level strong duality (according to Assumption 2) yields

$$\mathcal{F}_i(x) = \max_{u_i \in \mathbb{R}^{m_i}} \min_{v_i \geq 0} f_i(x, u_i) - v_i^\top h_i(u_i) \; . \tag{11}$$

Defining the Lagrangian $\tilde{\mathcal{L}}_i : \mathbb{R}^n \times \mathbb{R}^{m_i} \times \mathbb{R}^{K_i} \to \mathbb{R}$ for the lower level maximization problem (11) as

$$\tilde{\mathcal{L}}_i(x, u_i, v_i) := f_i(x, u_i) - v_i^\top h_i(u_i), \;\; i \in [N]^+, \tag{12}$$

and

$$\tilde{\mathcal{L}}_0(x, u_0, v_0) := f_0(x, u_0) - v_0^\top h_0(u_0) \; . \tag{13}$$

In the optimization problem (10), μ can be written as

$$\mu = \min_x \max_{\lambda \geq 0} \max_{u_0} \min_{v_0 \geq 0} \;\tilde{\mathcal{L}}_0(x, u_0, v_0) + \sum_{i=1}^N (c_i + \lambda_i) \max_{u_i} \min_{v_i \geq 0} \;\tilde{\mathcal{L}}_i(x, u_i, v_i) \; . \tag{14}$$

or combining the independent maximizations and minimizations, and collecting the variables, we have

$$\mu = \min_x \max_{\lambda \geq 0} \max_u \min_{v \geq 0} \left(\tilde{\mathcal{L}}_0(x, u_0, v_0) + \sum_{i=1}^N (c_i + \lambda_i) \, \tilde{\mathcal{L}}_i(x, u_i, v_i) \right) . \tag{15}$$

Hence, we showed the equivalence of (4) and (14). The complete Lagrangian $\mathcal{L}(x, \lambda, u, v)$ for $\mathcal{RO}$ problem is derived as

$$\mathcal{L}(x, \lambda, u, v) \;:=\; f_0(x, u_0) - v_0^\top h_0(x, u_0) + \sum_{i=1}^N (c_i + \lambda_i) \big(f_i(x, u_i) - v_i^\top h_i(x, u_i) \big) \; , \tag{16}$$

which leads to

$$\mu = \min_x \max_{\lambda \geq 0} \max_u \min_{v \geq 0} \; \mathcal{L}(x, \lambda, u, v) \; .$$

From Remark 4, strong duality holds for both upper and lower optimizations. Thus, we can switch the order of max and min as follows

$$\mu = \max_{\lambda \geq 0} \; \min_x \min_{v \geq 0} \max_u \; \mathcal{L}(x, \lambda, u, v) \; . \tag{17}$$

Remark 7 (Necessity of Lagrangian Analysis)**.** *The Lagrangian (16) is essential because: (i) it unifies nested min-max structure into one framework, (ii) enables continuous-time dynamics over all variables, (iii) provides saddle properties for our Lyapunov function, (iv) handles non-smoothness through*

[3] Similar to what is meant by feasibility in Robust Control [32].

dual decomposition. Standard methods fail without this unified approach.

Definition 1. *We denote an optimal solution of (17), $z^\star := (x^\star, \lambda^\star, u^\star, v^\star)$, as an "optimal $\mathcal{RO}$ solution", where $(x^\star, u^\star)$ is an optimal solution for (4), and $(\lambda^\star, v^\star)$ are optimal values for the corresponding dual variables according to the principles in [28, Section 5.9.1].*

KKT Optimality Conditions

As strong duality holds for $\mathcal{RO}$ problem, any optimal solution, $z^\star = (x^\star, \lambda^\star, u^\star, v^\star)$, satisfies the following Karush-Kuhn-Tucker (KKT) conditions, and vice-versa [28]

$$\nabla_x f_0(x^\star, u_0^\star) + \sum_{i=1}^{N} (c_i + \lambda_i^\star) \nabla_x f_i(x^\star, u_i^\star) = 0, \tag{18}$$

$$\nabla_{u_i} f_i(x^\star, u_i^\star) - {v_i^\star}^\top \nabla_{u_i} h_i(u_i^\star) = 0,\ i \in [N], \tag{19}$$

$$v_{ij}^\star \geq 0,\ h_{ij}(u_i^\star) \leq 0,\ v_{ij}^\star h_{ij}(u_i^\star) = 0,\ j \in [K_i]^+, i \in [N] \tag{20}$$

$$\lambda_i^\star \geq 0,\ f_i(x^\star, u_i^\star) \leq 0,\ \lambda_i^\star f_i(x^\star, u_i^\star) = 0,\ i \in [N]^+ \tag{21}$$

where $\nabla_x f$ is the notation for the gradient of a function f w.r.t. x.

Remark 8 (KKT Conditions for $\mathcal{RO}$)**.** *Our KKT conditions extend standard theory [28] to handle: (i) bi-level structure coordinating λ_i and v_{ij}, (ii) regularization via $(c_i + \lambda_i)$ terms, (iii) nested complementary slackness at both optimization levels, (iv) saddle point equivalence enabling our dynamics.*

Saddle Property of Optimal $\mathcal{RO}$ Solution

As the $\mathcal{RO}$ problem has two levels, we have different saddle property related to the bi-level optimization problem. For the $\mathcal{RO}$ problem as a whole, there is a saddle property based on the Lagrangian function (16), which we call the $\mathcal{RO}$ saddle property. In the context of Lagrangian duality, a saddle point of the Lagrangian function is a point where the function is minimized with respect to the convex variables and maximized with respect to the concave variables.

Remark 9 (Non-standard Saddle Point Property)**.** *The following lemma establishes a saddle-point property that is not implied by classical results such as Sion's minimax theorem [33] or Rockafellar's saddle point theorem [34], which require joint concavity in the maximization variables. Our Lagrangian $\mathcal{L}(x, \lambda, u, v)$ is jointly convex in (x, v) for fixed (λ, u). However, it is not jointly concave in (λ, u) for fixed (x, v) due to the product terms $(c_i + \lambda_i) \cdot f_i(x, u_i)$, which create bilinear coupling between λ_i and u_i. This violation of joint concavity makes existing primal-dual methods [24], [25] inapplicable and necessitates our novel proof approach that exploits the specific min-max-max-min structure of robust optimization.*

Lemma 1. *[Saddle property] Let $z^\star = (x^\star, \lambda^\star, u^\star, v^\star)$ be an optimal $\mathcal{RO}$ solution. Then, for all $x, \lambda \geq 0, u, v \geq 0$, $z^\star$ satisfies the $\mathcal{RO}$ saddle property, namely,*

$$\mathcal{L}(x^\star, \lambda, u, v^\star) \leq \mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star) \leq \mathcal{L}(x, \lambda^\star, u^\star, v) \ . \tag{22}$$

V. Dynamical system solving $\mathcal{RO}$

Motivation and Challenges

So far, we have characterized the "optimization" properties of the Robust Optimization problem under study. In this paper, we are interested in understanding if there is a continuous-time dynamical system that can solve $\mathcal{RO}$ and how it would operate. Our main motivations are two: (1) Understanding from a dynamical system perspective how $\mathcal{RO}$ can be solved. (2) Studying how physically interacting systems with very simple capabilities and or "intelligence" can cooperate to solve complex optimization problems (learning, estimation, and decision) well outside the single element capabilities. While there are now answers to these questions for large classes of convex optimization problems [25], [35], this is the first work, to the best of our knowledge, that addresses $\mathcal{RO}$ problems. Our task turned out to be quite non-trivial, as explained below.

The basic method to obtain a continuous-time dynamics that solves a constrained convex optimization problem goes back to [24]. The main idea is quite intuitive. The primal dynamics evolves with the negative gradient of the problem's Lagrangian function, w.r.t, the primal variable, that is x, while the dual dynamics evolves with the positive gradient of the Lagrangian w.r.t to the dual variables, that is, λ_is. The primal descent and dual ascent dynamics is globally converging to the optimal solution under minor assumptions. The proof is based on a simple quadratic Lyapunov function. However, from a system point of view, it is the passivity of the gradient of a convex function that provides the convergence mechanism [36], [37].

For the $\mathcal{RO}$ problem, the standard approach does not work due to the nested structure of optimization. Thus, the natural match between intuitive primal-descent dual-ascent dynamics via a traditional quadratic Lyapunov function is broken. It turns out that it is not easy to find the right combination of dynamics and Lyapunov function that show global convergence.

The $\mathcal{RO}$ Dynamics

This section presents the continuous-time dynamical system ($\mathcal{RO}$ dynamics) whose solutions globally converge to robust optimal solutions. Let

$$M = \sum_{i=0}^{N} m_i \ , \ K = \sum_{i=0}^{N} K_i \ .$$

Consider the following $\mathcal{RO}$ dynamics defined on $\mathbb{S} := \mathbb{R}^n \times \mathbb{R}_+^N \times \mathbb{R}^M \times \mathbb{R}_+^K$

$$\begin{cases} \dot{x} = -\nabla_x f_0(x, u_0) - \sum_{i=1}^{N} (c_i + \lambda_i) \nabla_x f_i(x, u_i) \\ \dot{\lambda}_i = [f_i(x, u_i) - v_i^\top h_i(u_i)]_{\lambda_i}^+,\ i \in [N]^+ \\ \dot{u}_i = \nabla_{u_i} f_i(x, u_i) - \sum_{j=1}^{K_i} v_{ij} \nabla_{u_i} h_{ij}(u_i),\ i \in [N] \\ \dot{v}_0 = [h_0(u_0)]_{v_0}^+ \\ \dot{v}_i = [(c_i + \lambda_i)\ h_i(u_i)]_{v_i}^+,\ i \in [N]^+ \end{cases} . \tag{23}$$

- The x dynamics implements gradient descent on the primal variables with respect to the Lagrangian, seeking to minimize the robust objective function.
- The λ_i dynamics updates the dual variables associated with the robust constraints, using projection to maintain non-negativity.
- The u_i dynamics performs gradient ascent on the uncertainty variables to find the worst-case scenarios within each uncertainty set.
- The v_i dynamics handles the dual variables for the uncertainty set constraints, ensuring feasibility within $\mathcal{U}_i$.

The projection operator $[\cdot]^+$ ensures that dual variables remain non-negative, which is essential for KKT optimality conditions. The coupling between primal and dual variables through the coefficients $(c_i + \lambda_i)$ allows the system to automatically balance between constraint satisfaction and objective optimization.

Illustrative Examples

It is interesting to describe the structure using special examples.

1) Min-max problem: Consider the following $\mathcal{RO}$ problem with no constraints

$$\mu = \min_x \max_{u_0 : h_0(u_0) \le 0} f_0(x, u_0) \; .$$

Such problems are popular in machine learning. In this case, the continuous-time dynamical system is given by

$$\begin{cases} \dot{x} = -\nabla_x f_0(x, u_0) \\ \dot{u}_0 = \nabla_{u_0} f_0(x, u_0) - v_0^\top \nabla_{u_0} h_0(u_0) \\ \dot{v}_0 = [h_0(u_0)]_v^+ \end{cases} .$$

Above dynamic is gradient-descent in x while gradient-ascent in u_0 and v_0. We must note that often in machine learning applications, the constraints on u_0 are simple boxes, and the dual variables v_0 are omitted in place of a simple set projection.

2) One uncertain constraint: Another simple example is the following $\mathcal{RO}$ problem which follows the standard setting where the constraint is active and c_1 is zero.

$$\begin{aligned} \mu \quad &= \quad \min_x f_0(x) \\ s.t. \quad &\max_{h_1(u_1) \le 0} f_1(x, u_1) \le 0 \end{aligned}$$

The dynamical system equations are

$$\begin{cases} \dot{x} = -\nabla_x f_0(x) - \lambda_1 \nabla_x f_1(x, u_1) \\ \dot{\lambda}_1 = [f_1(x, u_1) - v_1^\top h_1(u_1)]_{\lambda_1}^+ \\ \dot{u}_1 = \nabla_{u_1} f_1(x, u_1) - v_1^\top \nabla_{u_1} h_1(u_1) \\ \dot{v}_1 = [\lambda_1 h_1(u_1)]_{v_1}^+ \end{cases} .$$

This example shows the following significant differences with the primal-dual dynamical system for solving standard optimization problems appeared in [25], [38].

1) First, the $\mathcal{RO}$ dynamics has additional states associated with the worst-case constraint u_i and the associated multiplier v_1.
2) Another difference is that the $\mathcal{RO}$ dynamics vector field is not completely derived as negative/positive gradients of the Lagrangian function $\mathcal{L}$. In particular, the vector field for the state u_1 is not obtained as the positive gradient of the Lagrangian function $\mathcal{L}$.
3) Finally, we note the presence of λ_1, the dual variable, in the upper optimization in the dynamics of v_1, the dual variable of the lower optimization. At first glance, this seems strange, since one could expect the differential equations for u_1 and v_1 to be simply the primal-ascent and dual-descent, respectively, of the lower optimization problem. This point will be discussed further after we present the stability results.

Equilibrium Analysis

The following lemma relates the optimal (KKT) points of (4) and the equilibrium points of the dynamics of $\mathcal{RO}$.

Lemma 2. *[Optimal solution and equilibrium point] Under Assumptions 1 and 2, any optimal $\mathcal{RO}$ solution based on Definition 1 is an equilibrium point of $\mathcal{RO}$ dynamics (23) and vice versa.*

Proof. Any equilibrium point, $\bar{z} = (\bar{x}, \bar{\lambda}, \bar{u}, \bar{v})$ of $\mathcal{RO}$ dynamics (23) satisfies

$$\begin{aligned} &\nabla_x f_0(\bar{x}, \bar{u}_0) + \sum_{i=1}^N (c_i + \bar{\lambda}_i) \nabla_x f_i(\bar{x}, \bar{u}_i) = 0 \; , \\ &f_i(\bar{x}, \bar{u}_i) - \bar{v}_i^\top h_i(\bar{u}_i) \le 0, \; \bar{\lambda}_i \ge 0 \; , \\ &\bar{\lambda}_i (f_i(\bar{x}, \bar{u}_i) - \bar{v}_i^\top h_i(\bar{u}_i)) = 0 \; , \\ &\nabla_{u_i} f_i(\bar{x}, \bar{u}_i) - \bar{v}_i^\top \nabla_{u_i} h_i(\bar{u}_i) = 0 \; , \\ &h_{0j}(\bar{u}_0) \le 0, \; \bar{v}_{0j} \ge 0, \; \bar{v}_{0j} h_{0j}(\bar{u}_0) = 0 \; , \\ &(c_i + \bar{\lambda}_i) h_{ij}(\bar{u}_i) \le 0, \; \bar{v}_{ij} \ge 0, \; \bar{v}_{ij} (c_i + \bar{\lambda}_i) h_{ij}(\bar{u}_i) = 0 \; , \end{aligned}$$

for $i \in [N]^+, j \in [K_i]^+$, while any optimal point satisfies below KKT conditions

$$\nabla_x f_0(x^\star, u_0^\star) + \sum_{i=1}^N (c_i + \lambda_i^\star) \nabla_x f_i(x^\star, u_i^\star) = 0, \tag{24}$$

$$\nabla_{u_i} f_i(x^\star, u_i^\star) - {v_i^\star}^\top \nabla_{u_i} h_i(u_i^\star) = 0, \quad i \in [N], \tag{25}$$

$$v_{ij}^\star \ge 0, \; h_{ij}(u_i^\star) \le 0, \; v_{ij}^\star h_{ij}(u_i^\star) = 0, \; j \in [K_i], \; i \in [N] \tag{26}$$

$$\lambda_i^\star \ge 0, \; f_i(x^\star, u_i^\star) \le 0, \; \lambda_i^\star f_i(x^\star, u_i^\star) = 0, \; i \in [N] \tag{27}$$

Substituting $z^\star$ for $\bar{z}$, and using the fact that $(c_i + \lambda_i^\star) > 0$ for all $i \in [N]^+$, it is immediate to verify that the optimal point $z^\star$, satisfying (24)-(27), also satisfies the above equilibrium conditions and therefore is an equilibrium point of $\mathcal{RO}$ dynamics (23).

On the other hand, since $c_i > 0$ and $c_i + \bar{\lambda}_i > 0$, for $i \in [N]^+$; then, $h_{ij}(\bar{u}_i) \le 0$, and $\bar{v}_{ij} h(\bar{u}_{ij}) = 0$ for $i \in [N], j \in [K_i]^+$. Substituting these properties in the rest of the equilibrium conditions, we see that $\bar{z}$ satisfies the KKT conditions therefore is optimal for $\mathcal{RO}$. s $\square$

We denote the $\mathcal{RO}$ dynamics (23) compactly with the shorthand notation $\dot{z} = \mathcal{Z}^{\mathcal{RO}}(z)$.

Lyapunov Function Construction

In this subsection, we present the Lyapunov function[4] that establishes global convergence of the proposed dynamical system.

Lemma 3. *[Monotonicity property] Let $z^\star = (x^\star, \lambda^\star, u^\star, v^\star)$ be an optimal $\mathcal{RO}$ solution (Definition 1).*

Let $V : \mathbb{S} \to \mathbb{R}_+$ defined as

$$V = \frac{1}{2}\big(\|x - x^\star\|^2 + \|\lambda - \lambda^\star\|^2 + \|u_0 - u_0^\star\|^2 + \sum_{i=1}^{N}(c_i + \lambda_i^\star)\|u_i - u_i^\star\|^2 + \sum_{i=0}^{N}\|v_i - v_i^\star\|^2\big) , \quad (28)$$

then, the Lie-derivative of V along $\mathcal{Z}^{\mathcal{RO}}$ at $z = (x, \lambda, u, v)$ is $\nabla V(z)^\top \dot{z} \le 0$.

Proof. Note that $\nabla V(z)^\top$ equals to

$$\begin{aligned}
&(x - x^\star)^\top(-\nabla_x f_0(x, u_0) - \sum_{i=1}^{N}(c_i + \lambda_i)\nabla_x f_i(x, u_i)) \\
&+\sum_{i=1}^{N}(\lambda_i - \lambda_i^\star)[f_i(x, u_i) - v_i^\top h_i(u_i)]^+_{\lambda_i} \\
&+(u_0 - u_0^\star)^\top(\nabla_{u_0} f_0(x, u_0) - v_0^\top \nabla_{u_0} h_0(u_0)) \\
&+\sum_{i=1}^{N}(c_i + \lambda_i^\star)(u_i - u_i^\star)^\top(\nabla_{u_i} f_i(x, u_i) - v_i^\top \nabla_{u_i} h_i(u_i)) \\
&+(v_0 - v_0^\star)^\top[h_0(u_0)]^+_{v_0} + \sum_{i=1}^{N}(v_i - v_i^\star)^\top[(c_i + \lambda_i)\, h_i(u_i)]^+_{v_i} .
\end{aligned} \quad (29)$$

By convex and concave under-estimator properties according to Assumption 1, one can write [28, Section 3.1.3]

$$(x^\star - x)^\top \nabla_x f_i(x, u_i) \le f_i(x^\star, u_i) - f_i(x, u_i) , \quad (30)$$
$$(u_i - u_i^\star)^\top \nabla_{u_i} f_i(x, u_i) \le f_i(x, u_i) - f_i(x, u_i^\star) , \quad (31)$$
$$(u_i^\star - u_i)^\top \nabla_{u_i} h_{ij}(u_i) \le h_{ij}(u_i^\star) - h_{ij}(u_i) . \quad (32)$$

Moreover, using the fact that the projection operator is non-expansive, we have that

$$\begin{aligned}
&(\lambda_i - \lambda_i^\star)[f_i(x, u_i) - v_i^\top h_i(u_i)]^+_{\lambda_i} \\
&\qquad \le (\lambda_i - \lambda_i^\star)(f_i(x, u_i) - v_i^\top h_i(u_i)) , \quad (33)\\
&(v_0 - v_0^\star)^\top[h_0(u_0)]^+_{v_0} \le (v_0 - v_0^\star)^\top h_0(u_0), \quad (34)\\
&(v_i - v_i^\star)^\top[(c_i + \lambda_i)\, h_i(u_i)]^+_{v_i} \\
&\qquad \le (v_i - v_i^\star)^\top(c_i + \lambda_i)\, h_i(u_i) . \quad (35)
\end{aligned}$$

[4] With slight abuse of notation.

By substitution, we get

$$\begin{aligned}
&\nabla V(z)^\top \dot{z} \le \\
&f_0(x^\star, u_0) - f_0(x, u_0) + \sum_{i=1}^{N}(c_i + \lambda_i)(f_i(x^\star, u_i) - f_i(x, u_i)) \\
&+\sum_{i=1}^{N}(c_i + \lambda_i) - (c + \lambda_i^\star))(f_i(x, u_i) - v_i^\top h_i(u_i)) \\
&+ f_0(x, u_0) - f_0(x, u_0^\star) + v_0^\top(h_0(u_0^\star) - h_0(u_0)) \\
&+\sum_{i=1}^{N}(c + \lambda_i^\star)(f_i(x, u_i) - f_i(x, u_i^\star) + v_i^\top(h_i(u_i^\star) - h_i(u_i)) \\
&+ (v_0 - v_0^\star)^\top h_0(u_0) + \sum_{i=1}^{N}(v_i - v_i^\star)^\top(c_i + \lambda_i)h_i(u_i) .
\end{aligned} \quad (36)$$

After simplification and rearrangement we obtain

$$\begin{aligned}
\nabla V(z)^\top \dot{z} \le{}& f_0(x^\star, u_0) - (v_0^\star)^\top h_0(u_0) \\
&- f_0(x, u_0^\star) + v_0^\top h_0(u_0^\star) \\
&+ \sum_{i=1}^{N}(c_i + \lambda_i)\Big(f_i(x^\star, u_i) - (v_i^\star)^\top h_i(u_i)\Big) \\
&- \sum_{i=1}^{N}(c_i + \lambda_i^\star)\Big(f_i(x, u_i^\star) - v_i^\top h_i(u_i^\star)\Big) .
\end{aligned}$$

Adding and subtracting $\mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star)$ yields

$$\begin{aligned}
\nabla V(z)^\top \dot{z} \le{}& \mathcal{L}(x^\star, \lambda, u, v^\star) - \mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star) \quad (37)\\
&+ \mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star) - \mathcal{L}(x, \lambda^\star, u^\star, v) .
\end{aligned}$$

Note that from Lemma 1 (saddle property),

$$\begin{aligned}
&\mathcal{L}(x^\star, \lambda, u, v^\star) - \mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star) \le 0 , \\
&\mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star) - \mathcal{L}(x, \lambda^\star, u^\star, v) \le 0 .
\end{aligned}$$

Hence,

$$\begin{aligned}
\nabla V(z)^\top \dot{z} \le{}& \mathcal{L}(x^\star, \lambda, u, v^\star) - \mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star) \quad (38)\\
&+ \mathcal{L}(x^\star, \lambda^\star, u^\star, v^\star) - \mathcal{L}(x, \lambda^\star, u^\star, v) \le 0 .
\end{aligned}$$

$\square$

Remark 10 (Why Standard Dynamics Fail and Our Solution Works)**.** *Standard primal-dual dynamics fail for RO because when $\lambda_i \to 0$ (inactive constraints), the dynamics $\dot{u}_i = \lambda_i \nabla_{u_i} f_i$ freeze before reaching optimality. We resolve this by removing λ_i from $\dot{u}_i$ dynamics and constructing a Lyapunov function that weights errors by optimal duals $\lambda_i^\star$ rather than current values, enabling global convergence despite lacking joint concavity.*

Remark 11. *The convergence speed analysis of $\mathcal{RO}$ dynamics in the discrete-time version of our algorithm appears in [39].*

Solutions: Existence, Uniqueness, and Continuity

We need to guarantee that the switching dynamical system (23) has well defined solutions. To do so, we show that our system satisfies the conditions for applicability of existing results [38]. The development of this section is detailed in the Appendix.

Main Convergence Result

Considering $\gamma(t)$ as a solution of $\mathcal{RO}$ dynamics (23) defined on the time interval $[0,\infty)$, omega-limit set is defined as

$$\Omega(\gamma) := \{y \in \mathbb{S} \mid \exists \{t_k\}_{k=1}^{\infty} \subset [0,\infty) \text{ with } \lim_{k\to\infty} t_k = \infty \text{ and } \lim_{k\to\infty} \gamma(t_k) = y\} . \quad (39)$$

The next lemma presents an invariant property for the omega-limit set of any solution of the $\mathcal{RO}$ dynamics, which contributes in showing the convergence result.

Lemma 4. *[38, Lemma 4.4][Invariance of omega-limit set] The omega-limit set of any solution of the $\mathcal{RO}$ dynamics (23) starting from any point $\mathbb{S}$ is invariant.*

Finally, we have the main convergence result.

Theorem 4. *[Convergence] Under Assumptions 1 and 2, the trajectories of $\mathcal{RO}$ dynamics (23) converge to an optimal $\mathcal{RO}$ solution for any initial condition in $\mathbb{S} = \mathbb{R}^n \times \mathbb{R}_+^N \times \mathbb{R}^M \times \mathbb{R}_+^K$. In particular, each trajectory converges to a single point in the set of optimal $\mathcal{RO}$ solutions.*

Proof. **Step 1 (existence, forward invariance, and precompactness).** Fix any initial condition $z_0 \in \mathbb{S}$ and let $\bar{\gamma}(\cdot)$ denote the corresponding solution of (23), whose existence and uniqueness follow from Lemma 6. Let V be the Lyapunov function used in (38) and Lemma 3. Along any solution one has

$$\dot{V}(z) \;=\; \mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(z) \;\le 0.$$

Set $\delta := V(z_0)$ and define the (closed) sublevel set

$$\mathbb{P} \;:=\; V^{-1}(\le \delta) \cap \mathbb{S}.$$

By radial unboundedness of V, $\mathbb{P}$ is compact and, by $\dot{V} \le 0$, it is forward invariant. Hence $\bar{\gamma}(t) \in \mathbb{P}$ for all $t \ge 0$, so the trajectory is bounded and its ω-limit set $\omega(\bar{\gamma})$ is nonempty, compact, and invariant.

Step 2 (LaSalle: $\omega(\bar{\gamma}) \subset \mathcal{M}$). Define

$$\mathcal{M} \;:=\; \{z \in \mathbb{P} \mid \mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(z) = 0\}.$$

Since $t \mapsto V(\bar{\gamma}(t))$ is nonincreasing and bounded below, $V(\bar{\gamma}(t)) \to V_\infty$ as $t \to \infty$. Take any $y \in \omega(\bar{\gamma})$ and a sequence $t_k \to \infty$ with $\bar{\gamma}(t_k) \to y$. By continuity, $V(y) = \lim_k V(\bar{\gamma}(t_k)) = V_\infty$. If $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(y) < 0$, then for small $s > 0$, $V(\phi_s(y)) < V(y)$ for the flow ϕ_s, contradicting the convergence of $V(\bar{\gamma}(t))$ to V_∞. Thus $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(y) = 0$, so $y \in \mathcal{M}$. By Lemma 4, $\omega(\bar{\gamma})$ is invariant, and since it is contained in $\{\dot{V} = 0\}$, the invariance principle for discontinuous Carathéodory systems [38, Prop. 2.1] implies

$$\omega(\bar{\gamma}) \subseteq \text{largest invariant subset of } \{z \in \mathbb{P} : \mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(z) = 0\}.$$

Step 3 (identify the zero-derivative set with the optimal set). Introduce

$$\begin{aligned}\bar{\mathcal{M}} := \{&z \in \mathbb{P} \mid \lambda \ge 0,\ v_i \ge 0\ \forall i,\\ &\mathcal{L}(x^\star,\lambda,u,v^\star) - \mathcal{L}(x^\star,\lambda^\star,u^\star,v^\star) = 0,\\ &\mathcal{L}(x^\star,\lambda^\star,u^\star,v^\star) - \mathcal{L}(x,\lambda^\star,u^\star,v) = 0\} . \end{aligned} \quad (40)$$

From (38) we have $\mathcal{M} \subseteq \bar{\mathcal{M}}$. By strict convexity of f, any $z \in \bar{\mathcal{M}}$ satisfies $x = x^\star$. The equalities in (40) show that $(\bar{\lambda}, \bar{u})$ maximizes $\mathcal{L}(x^\star,\lambda,u,v^\star)$ over $\lambda, u \ge 0$ and $\bar{v}$ maximizes $\mathcal{L}(x^\star,\lambda^\star,u^\star,v)$ over $v \ge 0$, hence $\bar{z} = (x^\star,\bar{\lambda},\bar{u},\bar{v})$ attains the optimal $\mathcal{RO}$ cost. Conversely, by Lemma 2, any optimal $\mathcal{RO}$ solution is an equilibrium of (23); equilibria satisfy $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V = 0$, hence belong to $\mathcal{M}$ and also satisfy the equalities in (40), hence belong to $\bar{\mathcal{M}}$. Therefore

$$\mathcal{M} \;=\; \bar{\mathcal{M}} \;=\; \{\text{optimal } \mathcal{RO} \text{ solutions}\},$$

and this set is invariant.

Step 4 (LaSalle conclusion to the optimal set). By Step 2 and the identification in Step 3, the invariance principle yields

$$\omega(\bar{\gamma}) \subseteq \mathcal{M},$$

i.e., every limit point of the trajectory is an optimal $\mathcal{RO}$ solution.

Since the system has a Lyapunov function V with $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(z) \le 0$ and all equilibria in $\mathcal{M}$ are Lyapunov stable, the system is semistable by [40, Theorem 4.19]. Therefore, $\bar{\gamma}(t)$ converges to a single point $z^\star \in \mathcal{M}$, i.e., to a single optimal $\mathcal{RO}$ solution.

Step 6 (globality). Because $z_0 \in \mathbb{S}$ was arbitrary and $\mathbb{P} = V^{-1}(\le V(z_0)) \cap \mathbb{S}$ is compact and forward invariant for any such z_0, the above argument applies to every initial condition in $\mathbb{S}$. □

Corollary 1. *Under Assumptions 1 and 2, let $z = (x^\star,\lambda^\star,u^\star,v^\star)$ be an optimal solution. Assume all robust constraints are strictly active, that is, $\lambda_i^\star > 0,\ i \in [N]^+$. Then, the $\mathcal{RO}$ dynamics (23) converges to an optimal solution.*

Remark 12 (Relaxing Strict Complementarity)**.** *The corollary's requirement that all constraints be strictly active ($\lambda_i^\star > 0$) can be relaxed using regularization approach with $c_i = \varepsilon > 0$ as detailed in Section VI. This modification maintains convergence while handling general RO problems without strict complementarity assumptions.*

Proof. The setup of the corollary satisfies the assumptions of Theorem 4, as $c_i + \lambda_i^\star > 0,\ i \in [N]^+$. □

VI. Convergence with Inactive Constraints

Regularization Approach

Remark 13 (Handling Inactive Constraints)**.** *When $\lambda_i^* = 0$ (inactive), setting $c_i = 0$ causes Lyapunov singularities. Solution: use $c_i = \varepsilon > 0$ (e.g., 10^{-6}). Theorem 5 proves convergence to within $O(\varepsilon)$ of exact solution as $\varepsilon \to 0$.*

The convergence proof requires $c_i + \lambda_i^\star > 0$. This section rigorously handles inactive constraints where $\lambda_i^* = 0$.

Consider the original problem with $c_i = 0$:

$$\mu := \min_x \mathcal{F}_0(x) \quad \text{s.t.} \quad \mathcal{F}_i(x) \le 0\ ,\ i \in [N]^+\ , \quad (41)$$

When constraints are inactive, $\lambda_i^* = 0$ by complementary slackness. Setting $c_i = 0$ leads to $(c_i + \lambda_i^\star) = 0$, invalidating our Lyapunov function and breaking the dynamics coupling.

Perturbed Dynamics

We resolve this via regularization with $c_i = \varepsilon > 0$ (small):

$$\begin{cases} \dot{x} = -\nabla_x f_0(x, u_0) - \sum_{i=1}^{N} (\varepsilon + \lambda_i) \nabla_x f_i(x, u_i) \\ \dot{\lambda}_i = [f_i(x, u_i) - v_i^\top h_i(u_i)]^+_{\lambda_i}, \quad i \in [N]^+ \\ \dot{u}_i = \nabla_{u_i} f_i(x, u_i) - \sum_{j=1}^{K_i} v_{ij} \nabla_{u_i} h_{ij}(u_i), \; i \in [N] \\ \dot{v}_0 = [h_0(u_0)]^+_{v_0} \\ \dot{v}_i = [(\varepsilon + \lambda_i) \; h_i(u_i)]^+_{v_i}, \quad i \in [N]^+ \end{cases} , \tag{42}$$

converges to some μ_ε and $x^\star_\varepsilon$ from Theorem 4. It turns out that for ε sufficiently small, we can approximate the optimal cost and the optimal solution $x^\star$ arbitrarily well under compactness conditions.

Theorem 5. *For the problem $\mathcal{RO}$ (4) under Assumptions 1 and 2, let $x^\star_\varepsilon$ be the optimal solution where $c = \varepsilon$. Then $\lim_{\varepsilon \to 0} \mu_\varepsilon = \mu_0$, where μ_0 is the optimal cost of problem $\mathcal{RO}$ (4) with $\varepsilon = 0$, that is, (41). Furthermore, assuming the feasible set $\mathcal{C} = \{x \in \mathbb{R}^n \,|\, \mathcal{F}_i(x) \le 0, \; i \in [N]^+\}$ is compact, we have*

$$\lim_{\varepsilon \to 0} \| \, x^\star_\varepsilon - x^\star \, \| = 0 \, ,$$

with $x^\star$ be the optimal $\mathcal{RO}$ solution when $\varepsilon = 0$, that is, problem (41) (proof in the appendix).

Finally, we note that the dynamics in (42), is equivalent to the following dynamics by letting $\hat{\lambda} = \lambda + \varepsilon \mathbf{1}$

$$\begin{cases} \dot{x} = -\nabla_x f_0(x, u_0) - \sum_{i=1}^{N} \hat{\lambda}_i \nabla_x f_i(x, u_i) \\ \dot{\hat{\lambda}}_i = [f_i(x, u_i) - v_i^\top h_i(u_i)]^{\varepsilon +}_{\hat{\lambda}_i}, \quad i \in [N]^+ \\ \dot{u}_i = \nabla_{u_i} f_i(x, u_i) - \sum_{j=1}^{K_i} v_{ij} \nabla_{u_i} h_{ij}(u_i), \; i \in [N] \\ \dot{v}_0 = [h_0(u_0)]^+_{v_0} \\ \dot{v}_i = [\hat{\lambda}_i \; h_i(u_i)]^+_{v_i}, \quad i \in [N]^+ \end{cases} , \tag{43}$$

where the notation $[\cdot]^{\varepsilon+}_{\hat{\lambda}_i}$ represents the projection operator that ensures $\hat{\lambda}_i \ge \varepsilon > 0$, providing regularization for inactive constraints. Noting that above dynamics evolves on $\mathbb{S} = \mathbb{R}^n \times \mathbb{R}^N_{\varepsilon+} \times \mathbb{R}^M \times \mathbb{R}^K_+$, it shows that the perturbed dynamics can be obtained by simply perturbing λ_i projections with respect to ε instead of 0, by initializing $\lambda_i \ge \varepsilon > 0$.

VII. Simulations

We validate our $\mathcal{RO}$ dynamics through three carefully chosen examples that demonstrate: (i) exact solutions for problems with complex uncertainty sets where RC methods fail, (ii) quantitative performance comparisons with state-of-the-art methods, (iii) scalability to higher-dimensional problems. Each example highlights unique capabilities of our approach—handling intersection of ellipsoids, solving problems without tractable RC, and achieving significant advantages over other methods.

In MATLAB, several solvers can be used to simulate ordinary differential equations (ODE). In the following simulation examples, "ode15s" is used, which is a solver for stiff problems.

Example A: Robust Quadratic Programming

Intersection of ellipsoids models correlated uncertainties common in portfolio optimization, manufacturing tolerances, and robust control. The solution must remain feasible for ALL points in this non-smooth set, not just sampled scenarios.

Consider a robust quadratic programming (QP) problem as below

$$\begin{aligned} \min_{x \in \mathbf{R}^2} \; & f(x) := -8x_1 - 16x_2 + x_1^2 + 4x_2^2 \\ \text{s.t.} \max_{u \in \mathcal{U}} \; & (a + Pu)^\top x \le b \, , \end{aligned} \tag{44}$$

where $x \in \mathbf{R}^2$ is the decision variable, and $a = [1 \quad 1]^\top$, $P = I_2 \in \mathbf{R}^{2\times 2}$ and $b = 5$ are given parameters. Variable u is uncertain, for which, the uncertainty set is described by the intersection of five ellipsoids as below

$$\mathcal{U} := \{u \in \mathbb{R}^2 : h_j(u) \le 0 \; , \; j^+_{[5]}\} \, ,$$

where $h_j(u) := u^\top Q_j u - 1, j^+_{[5]}$. Each Q_j is a symmetric positive semi-definite matrix and $\sum_{j=1}^5 Q_j \succ 0$. In this example, we assume that the following matrices specify ellipsoids

$$Q_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \; Q_2 = \begin{bmatrix} 5 & -2 \\ -2 & 4 \end{bmatrix}, \; Q_3 = \begin{bmatrix} 4 & 4 \\ 4 & 6 \end{bmatrix},$$

$$Q_4 = \begin{bmatrix} 3 & 0 \\ 0 & 8 \end{bmatrix}, \; Q_5 = \begin{bmatrix} 5 & 2 \\ 2 & 4 \end{bmatrix}.$$

The Lagrangian function can be written as

$$\mathcal{L} = f(x) + \lambda\big((a + Pu)^\top x - b - v^\top h(u)\big) \, .$$

where

$$h = [h_1, h_2, h_3, , h_4, h_5]^\top, \;\; v = [v_1, v_2, v_3, v_4, v_5]^\top .$$

We obtain the following dynamics according to $\mathcal{RO}$ dynamics (43)

$$\begin{cases} \dot{x} = -\begin{bmatrix} 2x_1 - 8 \\ 8x_2 - 16 \end{bmatrix} - \hat{\lambda}\,(a + Pu) \\ \dot{\hat{\lambda}} = \big[(a + Pu)^\top x - b - v^\top h(u)\big]^{\varepsilon+}_{\hat{\lambda}} \\ \dot{u} = P^\top x - \; 2\sum_{j=1}^5 Q_j u v_j \\ \dot{v} = \Big[\hat{\lambda} \; h(u)\Big]^+_v \end{cases}$$

where $\hat{\lambda} = \lambda + \varepsilon$. The trajectories for this system starting from zero initial conditions are shown in Fig. 1. Note that the constraint is active, so we can set ε to zero according to Remark 6. The optimal value of x is $[2.2674, 1.6636]$ and the optimal cost is -28.5452. Five ellipsoids in the uncertainty set are plotted in Fig. 2. The blue star shows the optimal value of u at $[0.4046, 0.0909]$ which lies on the boundary of intersection of two of the ellipsoids corresponding to Q_3 and Q_5. Also note that for positive values of ε, the λ trajectory and convergence value may change but the solution x remains the same as the constraint is active.

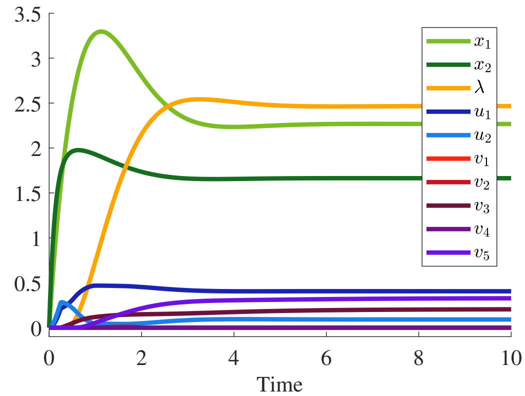

Fig. 1: Trajectories of $\mathcal{RO}$ dynamics for robust QP problem with intersection of ellipsoids uncertainty set in Example A.

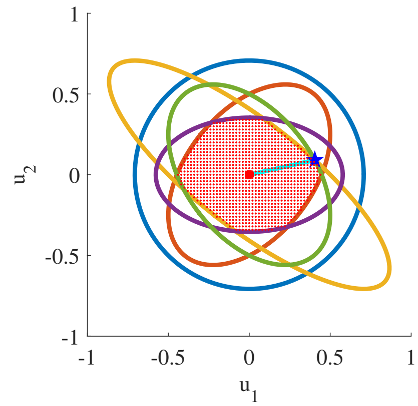

Fig. 2: Uncertainty set for Example A plotted in u_2-u_1 space. Big red point at the origin represents the initial point of u in $\mathcal{RO}$ dynamics and blue star depicts the optimal value of u on the intersection of two ellipsoids derived by our method. Small red grid points are the 1115 sampled points from the uncertainty set in randomized scenario method [18] to compare with our method.

The solution for the $\mathcal{RO}$ problem (44) can be verified by other methods. We can apply the technique in [18], in which random instances of uncertainties are sampled from the uncertainty set. Each of the uncertainty instances corresponds to a constraint. This results in a deterministic optimization problem with finitely many constraints. By picking 1115 samples from the intersection of ellipsoids uncertainty set and solving the derived deterministic optimization problem with 1115 constraints by CVX, the approximate robust optimal solution and the approximate optimal cost value are found $[2.2693, 1.6770]$ and -28.5873 respectively. The optimal cost function of our method has a larger value compared to the method in [18], as the latter approximates the RFS by taking finite samples. Hence, it cannot find the best value of u precisely and the optimal u; therefore, the optimal cost is approximated. The second method verifying our solution is robust counterpart [1]. Deriving the robust counterpart and solving the deterministic problem by CVX returns the same solution compared to the $\mathcal{RO}$ dynamics. Our method works on the main $\mathcal{RO}$ problem to find the optimal solution without transforming it to a deterministic equivalent which can be a hassle and even impossible task, as will be shown in the following example.

Example B: Robust Nonlinear Optimization with no RC

Consider the following robust nonlinear optimization problem

$$\begin{aligned} \min_{x\in\mathbf{R}^2} \; & f(x) := \frac{1}{2}(x_1-1)^2 + \frac{1}{2}(x_2-2)^2 \\ \text{s.t.} \; & \max_{u\in\mathcal{U}} u^\top \begin{bmatrix} e^{x_1^2} \\ e^{x_2^2} \end{bmatrix} \le b \,, \end{aligned} \tag{45}$$

where $x \in \mathbf{R}^2$, and $u = [u_1, u_2]^\top \in \mathbf{R}^2$, for which, the strictly convex uncertainty set is described by

$$\mathcal{U} := \{u \in \mathbb{R}^2 : e^{u_j^2} + u_j e^{\frac{1}{u_j}} \le \rho_j \;,\; j = 1, 2\} \,. \tag{46}$$

As stated in [17] and [15], there is no closed-form convex conjugate for the constraint (convex in x) in (45) and no closed-form conjugate for the convex uncertainty set in (46). This is the third case in [15, Table 1], for which there is no known method for obtaining RC. However, by writing the Lagrangian function as

$$\begin{aligned} \mathcal{L} =& \frac{1}{2}(x_1-1)^2 + \frac{1}{2}(x_2-2)^2 \\ &+ \lambda\Big(u^\top \begin{bmatrix} e^{x_1^2} \\ e^{x_2^2} \end{bmatrix} - b - v^\top \begin{bmatrix} e^{u_1^2} + u_1 e^{\frac{1}{u_1}} - \rho_1 \\ e^{u_2^2} + u_2 e^{\frac{1}{u_2}} - \rho_2 \end{bmatrix}\Big) \,, \end{aligned}$$

where $h_j(u_j) = e^{u_j^2} + u_j e^{\frac{1}{u_j}} - \rho_j$ for $j = 1, 2$ as defined in the previous example, we can form the $\mathcal{RO}$ dynamics for the RNO problem (45) as

$$\begin{aligned} \dot{x} &= -\begin{bmatrix} x_1 - 1 \\ x_2 - 2 \end{bmatrix} - \hat{\lambda} \begin{bmatrix} 2x_1 e^{x_1^2} & 0 \\ 0 & 2x_2 e^{x_2^2} \end{bmatrix} u \,, \\ \dot{\hat{\lambda}} &= \Big[u^\top \begin{bmatrix} e^{x_1^2} \\ e^{x_2^2} \end{bmatrix} - b - \sum_{j=1}^{2} \big(v_j (e^{u_j^2} + u_j e^{\frac{1}{u_j}} - \rho_j)\big)\Big]_{\hat{\lambda}}^{\varepsilon+} \,, \\ \dot{u} &= \begin{bmatrix} e^{x_1^2} \\ e^{x_2^2} \end{bmatrix} - \sum_{j=1}^{2} (2u_j e^{u_j^2} + e^{\frac{1}{u_j}} - \frac{1}{u_j} e^{\frac{1}{u_j}}) v_j \,, \\ \dot{v} &= \Big[\hat{\lambda} \begin{bmatrix} e^{u_1^2} + u_1 e^{\frac{1}{u_1}} - \rho_1 \\ e^{u_2^2} + u_2 e^{\frac{1}{u_2}} - \rho_2 \end{bmatrix}\Big]_v^+ \,, \end{aligned}$$

according to $\mathcal{RO}$ dynamics (43) where $\hat{\lambda} = \lambda + \varepsilon$. Similarly to the previous example, we can set ε to zero according to Remark 6 as the constraint is active. Fig. 3 shows the trajectory plot for $\rho_1 = 10$, $\rho_2 = 20$, and all the states initialized at 1. The robust optimal solution is $[0.5271, 0.7916]$ and the optimal cost value is 0.8419. We observe that the optimal value for u, which is $[1.4020, 1.6824]$ lies on the boundary of the uncertainty set. For positive values of ε, the trajectory and convergence value of λ may change, but the solution x remains the same as the constraint is active.

For this problem with highly nonlinear uncertainty constraints $e^{u_j^2} + u_j e^{1/u_j} \le \rho_j$, no closed-form robust counterpart exists [1], [15], making RC-based methods inapplicable. We compared our approach with scenario-based sampling [18]: using CVX, 168 scenarios yielded solution

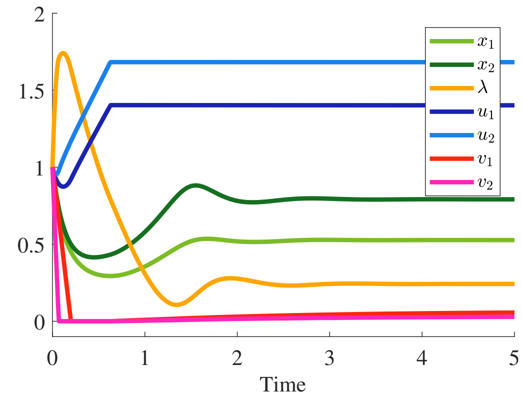

Fig. 3: Trajectory of $\mathcal{RO}$ dynamics for robust nonlinear optimization problem (45) in Example B.

$[0.5376, 0.8193]$ with cost 0.8039 (2.3s), 500 scenarios gave $[0.5312, 0.8024]$ with cost 0.8287 (8.7s), and 1000 scenarios produced $[0.5289, 0.7953]$ with cost 0.8371 (31.2s). Our method obtained the exact solution $[0.5271, 0.7916]$ with optimal cost 0.8419 in 0.8s integration time—40× faster than the 1000-scenario approximation while providing the exact solution rather than an approximation.

This example illustrates a key advantage of our approach: the dynamics naturally handle nonlinear uncertainty constraints through gradient terms $\nabla_{u_j} h_j(u_j)$, achieving exact convergence where other methods either fail completely (when RC cannot be derived) or require computationally expensive approximations with complexity $O(N^3 n^3)$ for N scenarios versus our $O(n^2)$ scaling.

Example C: Robust Dynamic Location Problem

We now consider the optimal cooperative and robust self-placement of autonomous vehicles, modeled as first-order kinematic points, to monitor multiple targets (initially assumed static). This is a dynamic generalization of a classical facility location problem [41] and here we provide a solution to the robust formulation of this problem using $\mathcal{RO}$ dynamics. The optimization problem is described by an undirected graph with N nodes and a set of edges $\mathbb{E}$. Among these nodes, the first N_1 nodes represent the fixed positions of the targets (also referred to as anchors), while the remaining $N_2 = N - N_1$ nodes represent the mobile agents. Each node $x_i \in \mathbb{R}^2$ denotes the location of an anchor or an agent. The anchors $x_1, \cdots, x_{N_1}$ have fixed positions and the agents $x_{N_1+1}, \cdots, x_N$ are mobile and can adjust their positions. The problem is to find the locations of the sensor nodes $x_{N_1+1}, \cdots, x_N$ to minimize the following cost

$$g(x_1, .., x_{N_1}) = \min_{x_{N_1+1},\cdots,x_N} \sum_{(i,j)\in\mathbb{E}} f_{ij}(x_i, x_j) ,$$

which is the sum of some measure of "length" for each link. This problem and its generalizations have many applications [28], and can be solved efficiently and in a distributed fashion in continuous time [35]. Leaving the sensor nodes (agents) mobile and capable of computation in distributed mode, we obtain a distributed dynamical version *real-time* of the optimal placement. The sensor nodes, through local interactions, cooperate to find and move toward their globally optimal locations autonomously.

The problem can also be under a set of constraints for the position of agents $x_{N_1+1}, \cdots, x_{N_1}$ to be in a specified convex set. Specifically, we consider the following robust location and placement problem

$$\begin{aligned} &\min_{x_{N_1+1},\cdots,x_N} \sum_{(i,j)\in\mathbb{E}} \frac{1}{2}\, w_{ij} \parallel x_i - x_j \parallel_2^2 \\ &\text{s.t. } \max_{u_i\in\mathcal{U}_i} (a_i + P_i u_i)^\top x_i \le b_i,\ i = N_1+1, \cdots, N , \end{aligned}$$

where the uncertainty u_i lies in $\mathcal{U}_i =\parallel u_i \parallel_2^2 \le \rho_i^2$ and $x_{N_1}+1, \cdots, x_N$ are moving agents. The $\mathcal{RO}$ dynamics for this problem can be derived as follows for $i = N_1+1, \cdots, N$

$$\begin{cases} \dot{x}_i = \sum_j w_{ij}(x_i - x_j) - \hat{\lambda}_i\ (a_i + P_i u_i) \\ \dot{\hat{\lambda}}_i = [(a_i + u_i P_i)^\top x_i - P_i u_i^\top c_i - b_i - v_i\ \mathcal{U}_i]^{\varepsilon+}_{\hat{\lambda}_i} \\ \dot{u}_i = P_i^\top (x_i - c_i) - 2 v_i u_i \\ \dot{v}_i = [\hat{\lambda}_i\ \mathcal{U}_i] \end{cases} .$$

Note that this problem is naturally distributed and our dynamics reflects the distributed structure. We consider the setup shown in the figure below with five anchors and four agents. The figure also shows the (fixed) interconnection graph among all the elements of the problem. In this academic example, we consider the presence of a linear constraint defining the half-space where the agents can be. The constraint is simply

$$\mathbf{1}' x_i \le 2.5, \quad i = N_1+1, \cdots, N ,$$

for all the agents. In addition, we would like the agents to find robust locations based on the uncertainty on the nominal slope 45^o of the nominal constraint. Namely

$$\mathbf{1}' x_i + u_i P' x_i \le 2.5, \quad i = N_1+1, \cdots, N ,$$

where $P' = \begin{bmatrix} 1 & -1 \end{bmatrix}$ and $\|u_i\|_2^2 \le \rho^2$. For example, when $\rho = 1$, the constraint can be any line passing through $x' = (1.25,\ 1.25)$, including the horizontal and vertical ones. The figure also shows the nominal linear constraint. The location of the shown agents is the optimal robust one w.r.t. the constraint being perturbed by the uncertainty u_i with $\|u_i\|_2^2 \le 0.1$.

First, agents start at their initial locations at the origin (white circle), and their paths converge to the optimal robust location for $\rho^2 = 0.1$. Such locations are indicated by full-colored circles. We see that agents 1 and 4 (blue and green) are on the boundary of the robust feasible set identifiable with the larger sector in the figure. The uncertain constraint is not active for the other agents. However, their optimal location is indirectly influenced by the robust constraint active in agents 1 and 4 as all agents and anchors are interconnected.

In the next phase of the simulation, we rotate the location of the anchors clockwise at constant speed. Although agents only react to their local neighbors (agents/anchors), the interconnected dynamical system shows the ability to globally track the coordinated motion of the anchors within the robust feasible

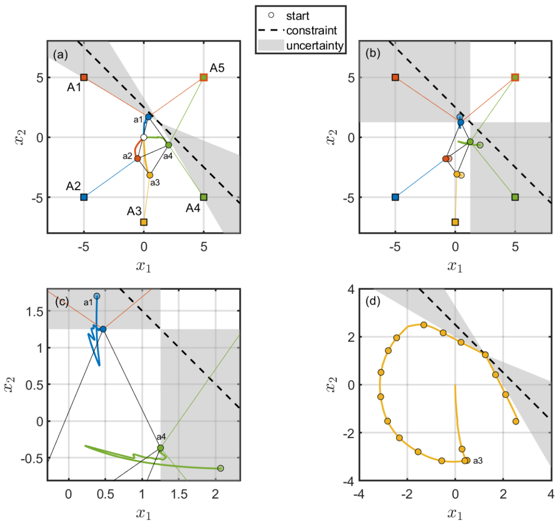

Fig. 4: Locations of agents and anchors, interactions between agents and anchors, and nominal linear constraint in Example C.

set. It is interesting to notice that the boundary of the feasible set is not defined a priori or hard-coded in the simulation, but it emerges from the interactions built in the dynamical system. Finally, after some time, the uncertainty changes and increases with size $\rho^2 = 1$ at time $t = 300$. This implies that no location should be feasible above the horizontal line passing through $(1.25, 1.25)$ and on the right of the vertical line passing through the same point. The location of the agents when the uncertainty changes is indicated by black empty circles. We note that, when the uncertainty changes, the nodes 1, 2, and 3 are located outside the new robust feasible set and, with a transient correction, move inside the robust SW quadrant for the rest of the time.

This example showcases the ability of the dynamical system to track changes online and adapt to uncertainty changes over time. We have obtained the behavior shown by appropriate scaling of the various differential equations involved. Note that the convergence is not affected by positive scaling of the differential equations. A relatively larger scaling on the dual variables in charge of the constraints makes the system response more reactive toward feasibility and more sensitive toward uncertainty changes. We leave to future research the systematic design of the optimization system for the desired real-time behavior.

VIII. Conclusions

This paper introduced $\mathcal{RO}$ dynamics for solving robust optimization problems through continuous-time evolution. The method achieves significant efficiency and generalization improvements while providing exact solutions for problems where robust counterparts are intractable. Applications to machine learning robustness, adversarial training, and large-scale distributed optimization represent promising future directions.

IX. Appendix

Proof of Lemma 1 [Saddle property]

Essentially, we want to show that

$$\mathcal{L}(x^\star, \lambda, u, v^\star) \le \mu \le \mathcal{L}(x, \lambda^\star, u^\star, v)\ .$$

Under Assumption 2, an optimal solution $x^\star$ exists and based on Assumption 1, this optimal solution is unique as $f_0(x, u_0)$ is strictly convex in x for any $u_0 \in \mathcal{U}_0$. Consider any optimal solution $(x^\star, \lambda^\star, u^\star, v^\star)$ for $\mathcal{RO}$ problem (17).
For each of the lower level optimizations, let

$$\eta_i = \max_{u_i \in \mathcal{U}_i} f_i(x^\star, u_i) = f_i(x^\star, u_i^\star) =\tilde{}$$

$$\mathrm{L}_i(x^\star, u_i^\star, v_i^\star),\ i \in [N]\ .$$

From the corresponding saddle property, it follows that for all u_i, $i \in [N]$,

$$f_i(x^\star, u_i) - (v_i^\star)^\top h_i(u_i) \le \eta_i\ . \tag{47}$$

From the upper level saddle property, for all $\lambda_i \ge 0$, $i \in [N]^+$,

$$\eta_0 + \sum_i (c_i + \lambda_i)\eta_i \le f_0(x^\star, u_0^\star) + \sum_{i=1}^N (c_i + \lambda_i^\star)\eta_i = \mu\ .$$

Substituting the lower bound on η_i, (47), in the left hand side, it follows that for all $u_i, i \in [N]$ and $\lambda_i \ge 0, i \in [N]^+$,

$$\begin{aligned}
&\mathcal{L}(x^\star, \lambda, u, v^\star) \\
=& f_0(x^\star, u_0) - (v_0^\star)^\top h_0(u_0) \\
&+ \sum_{i=1}^N (c_i + \lambda_i)(f_i(x^\star, u_i) - (v_i^\star)^\top h_i(u_i)) \\
\le& \mu\ ,
\end{aligned}$$

where we have used $c_i + \lambda_i \ge 0$. We next use lower saddle property in the other direction, namely, for all $v_i \ge 0$,

$$\begin{aligned}
\eta_i = f_i(x^\star, u_i^\star) &= f_i(x^\star, u_i^\star) - (v_i^\star)^\top h_i(u_i^\star) \\
&\le f_i(x^\star, u_i^\star) - (v_i)^\top h_i(u_i^\star)\ ,
\end{aligned}$$

which implies that for all $v_i \ge 0$, $-v_i^\top h(u^\star) \ge 0$, $i \in [N]$. Using this property and the fact that $(c_i + \lambda_i^\star) \ge 0$, it follows that for all $v_i \ge 0$, $i \in [N]^+$,

$$-(v_0)^\top h_0(u_0^\star) - \sum_{i=1}^N (c_i + \lambda_i^\star)(v_i)^\top h_i(u_i^\star) \ge 0\ . \tag{48}$$

The upper level saddle property implies

$$\mu \le f_0(x, u_0^\star) + \sum_{i=1}^N (c_i + \lambda_i^\star) f_i(x, u_i^\star)$$

for all x. Combining this with (48), we obtain

$$\begin{aligned}
\mu \le& f_0(x, u_0^\star) - (v_0)^\top h_0(u_0^\star) \\
&+ \sum_{i=1}^N (c_i + \lambda_i^\star)(f_i(x, u_i^\star) - (v)^\top h_i(u_i^\star)) \\
=& \mathcal{L}(x, \lambda^\star, u^\star, v)
\end{aligned}$$

for all x and $v_i \ge 0, i \in [N]$.

$\mathcal{RO}$ Dynamics Solutions Properties

$\mathcal{RO}$ dynamics (23) can be viewed as switched dynamical system with a discontinuous right-hand side. The conditions guaranteeing the existence and uniqueness of the solution and continuity w.r.t. initial conditions, for a general discontinuous dynamical system are provided in [42, Theorem 2.5]. In this section, we show that our $\mathcal{RO}$ dynamics (23) satisfies the refined conditions presented in [38].

To prove the existence and uniqueness of solutions for (23), and also the continuity of solutions w.r.t. the initial conditions, there are two main steps. The first step is showing that $\mathcal{RO}$ dynamics is a particular case of projected dynamical systems. The second step requires $\mathcal{RO}$ dynamics (23) satisfying the monotonicity property, which is our main result.

Definition 2. *(Projection operator) If $\mathcal{K}$ is a closed convex set, for any point $\bar{y} \in \mathbb{R}^q$, the point projection of $\bar{y}$ on the set $\mathcal{K}$ can be written as*

$$\textit{proj}_{\mathcal{K}}(\bar{y}) = \textit{argmin}_{y\in\mathcal{K}} \parallel y - \bar{y} \parallel \ .$$

For $\bar{y} \in \mathbb{R}^n$ and $y \in \mathcal{K}$, vector projection of $\bar{y}$ at y w.r.t. $\mathcal{K}$ is

$$\Pi_{\mathcal{K}}(y, \bar{y}) = \textit{lim}_{\delta\to 0^+} \frac{\textit{proj}_{\mathcal{K}}(y + \delta\bar{y}) - y}{\delta} \ . \tag{49}$$

Note that the map $\text{proj}_{\mathcal{K}}$ is Lipschitz on $\mathbb{R}^q$ with constant $L = 1$ [43, Proposition 2.4.1].

Definition 3. *[Projected dynamical system [44]] Considering a differential equation $\dot{y} = F(y)$ with $F : \mathbb{R}^q \to \mathbb{R}^q$, the associated projected dynamical system is defined as*

$$\dot{y} = \Pi_{\mathcal{K}}(y, F(y)), \ y(0) \in \mathcal{K} \ . \tag{50}$$

Lemma 5. *($\mathcal{RO}$ dynamics as a projected dynamics) $\mathcal{RO}$ dynamics (23) can be written as a projected dynamical system according to Definition 3.*

Proof. The proof of Lemma 5 follows along the lines of the construction outlined in [38]. Details omitted. □

Remark 14 (Projected Dynamics Foundation)**.** *The following proposition ensures: (i) existence despite discontinuities from projections, (ii) uniqueness via Lipschitz property, (iii) continuous dependence on initial conditions, (iv) validity of Lyapunov analysis for Theorem 4.*

Proposition 6. *If F in the projected dynamical system (50) is Lipschitz on $\mathcal{K}$, we have the following existence, uniqueness, and continuity w.r.t. the initial condition results for the solutions of the projected dynamics (50):*

1) *For any $y_0 \in \mathcal{K}$, there exists a unique solution $t \to y(t)$ of the projected system (50) with $y(0) = y_0$ in $[0, \infty)$.*
2) *Consider a sequence of points $\{y_k\}_{k=1}^{\infty} \subset \mathcal{K}$ with $\lim_{k\to\infty} y_k = y$. Then, the sequence of solutions $\{t \to \gamma_k(t)\}_{k=1}^{\infty}$ of the projected dynamics (50) with $\gamma_k(0) = y_k$ for all k, converges to the solution $t \to \gamma(t)$ of (50) with $\gamma(0) = y$ uniformly on every compact set of $[0, \infty)$.*

The ability to write $\mathcal{RO}$ dynamics (23) as a projected dynamical system along with the monotonicity property is used in the proof of the existence, uniqueness and continuity of the solutions of the set $\mathbb{S}$.

Lemma 6. *(Existence, uniqueness and continuity of solutions) $\gamma : [0, T) \to \mathbb{S}$ is defined as a Caratheodory solution of $\mathcal{Z}^{\mathcal{RO}}$ in the interval $[0, T)$ if γ is absolutely continuous on $[0, T)$ and satisfies $\dot{\gamma}(t) = \mathcal{Z}^{\mathcal{RO}}(\gamma(t))$ almost everywhere in $[0, T)$. Under Assumptions 1 and 2, and starting from any point $z \in \mathbb{S}$, a unique solution to $\mathcal{RO}$ dynamics (23) exists and remains in $\mathbb{S} \cap V^{-1}(\leq V(z))$. Also, if a sequence of points $\{z_k\}_{k=1}^{\infty} \subset \mathbb{S}$ converges to z as $k \to \infty$, the sequence of solutions $\{t \to \gamma_k(t)\}_{k=1}^{\infty}$ of $\mathcal{Z}^{\mathcal{RO}}$ starting at these points (that is, $\gamma_k(0) = z_k$ for all k) converge uniformly to the solution $t \to \gamma(t)$ on every compact set of $[0, \infty)$.*

The proof of this lemma follows closely along the lines of proof for the existence and uniqueness of solution for the primal-dual dynamical system from [38, Lemma 4.3].

Proof of Theorem 5

Based on the optimal solution $x^\star$ for $\mathcal{RO}$ problem,

$$\mu = \min_{\mathcal{F}_i(x)\leq 0} \mathcal{F}_0(x) \ , \ \mu = \mathcal{F}_0(x^\star) \ .$$

As the cost function of μ_ε is smaller than or equal to that of $\mathcal{RO}$ and the feasible sets of the two problems are equal,

$$\mu_\varepsilon - \mu \leq 0 \ . \tag{51}$$

Since $x^\star$ minimizes $\mathcal{F}_0(x)$ over the constraint set, $\mu = \mathcal{F}_0(x^\star) \leq \mathcal{F}_0(x^\star_\varepsilon)$. Adding and subtracting $\varepsilon\sum_{i=1}^n \mathcal{F}_i(x^\star_\varepsilon)$ in the right-hand side (RHS) and using (51) yields $\varepsilon\sum_{i=1}^N \mathcal{F}_i(x^\star_\varepsilon) \leq \mu_\varepsilon - \mu \leq 0$.

Following a similar argument as before by comparing $\mu(\varepsilon_0)$ and μ_ε for $\varepsilon_0 \geq \varepsilon$, we now let

$$\mu_\varepsilon = \tilde{\mathcal{F}}_0(x^\star_\varepsilon) = \mathcal{F}_0(x^\star_\varepsilon) + \varepsilon\sum_{i=1}^N \mathcal{F}_i(x^\star_\varepsilon) \ ,$$

$$\mu(\varepsilon_0) = \tilde{\mathcal{F}}_0(x^\star_{\varepsilon_0}) + \delta\sum_{i=1}^N \mathcal{F}_i(x^\star_{\varepsilon_0}) \ ,$$

where $\delta = \varepsilon_0 - \varepsilon$. Then, $\mu_\varepsilon \geq \mu(\varepsilon_0)$, but because $x^\star_\varepsilon$ is optimal for μ_ε , we have $\tilde{\mathcal{F}}_0(x^\star_\varepsilon) \leq \tilde{\mathcal{F}}_0(x^\star_{\varepsilon_0})$. As $x^\star_\varepsilon$ is feasible for $\mu(\varepsilon_0)$,

$$\tilde{\mathcal{F}}_0(x^\star_{\varepsilon_0}) + \delta\sum_{i=1}^N \mathcal{F}_i(x^\star_{\varepsilon_0}) = \mu(\varepsilon_0) \leq \tilde{\mathcal{F}}_0(x^\star_\varepsilon) + \delta\sum_{i=1}^N \mathcal{F}_i(x^\star_\varepsilon) \ .$$

Combining the two inequalities,

$$\delta\sum_{i=1}^N \mathcal{F}_i(x^\star_{\varepsilon_0}) - \delta\sum_{i=1}^N \mathcal{F}_i(x^\star_\varepsilon) \leq \tilde{\mathcal{F}}_0(x^\star_\varepsilon) - \tilde{\mathcal{F}}_0(x^\star_{\varepsilon_0}) \leq 0 \ ,$$

which implies that $\sum_{i=1}^N \mathcal{F}_i(x^\star_{\varepsilon_0}) \leq \sum_{i=1}^N \mathcal{F}_i(x^\star_\varepsilon)$. Thus,

$$\varepsilon\sum_{i=1}^N \mathcal{F}_i(x^\star_{\varepsilon_0}) \leq \varepsilon\sum_{i=1}^N \mathcal{F}_i(x^\star_\varepsilon) \leq \mu_\varepsilon - \mu \leq 0 \ . \tag{52}$$

Since $\sum_{i=1}^N \mathcal{F}_i(x^\star_{\varepsilon_0}) = \frac{\mu(\varepsilon_0) - \mathcal{F}_0(x^\star_{\varepsilon_0})}{\varepsilon_0}$ is bounded, we have

$$\varepsilon\sum_{i=1}^N \mathcal{F}_i(x^\star_{\varepsilon_0}) \leq \mu_\varepsilon - \mu \leq 0 \ , \tag{53}$$

which means that μ_ε converges to μ as $\varepsilon \to 0$.

To prove the second part of the theorem, that is, $x^\star_\varepsilon \to x^\star$ as $\varepsilon \to 0$, consider any sequence $\{\varepsilon_n\}$ converging to 0. Let $\{x^\star_n\}$ be the corresponding sequence of optimal solutions for $\mu(\varepsilon_n)$.

As $\mathcal{C}$ is compact and the same for both perturbed and original problem, $x^\star_n \in \mathcal{C}$ is bounded. Therefore, there exists a convergent sub-sequence $x^\star_{n_t}$ that converges to, say, $\hat{x} \in \mathcal{C}$, as $\varepsilon_{n_t} \to 0$, since $\mathcal{C}$ is closed by assumption.

This implies that $\mathcal{F}_0(x^\star_{n_t}) \to \mathcal{F}_0(\hat{x})$, since $\mathcal{F}_0(x)$ is continuous. Because $\hat{x}$ is feasible, $\mathcal{F}_0(\hat{x}) \geq \mu$. However, $\mathcal{F}_0(\hat{x}) > \mu$ is impossible since $\mu(\varepsilon_{n_t}) \to \mu$, from the first part of the proof, and $\mu(\varepsilon_{n_t}) = \mathcal{F}_0(x^\star_{n_t}) + \varepsilon_{n_t} \sum_{i=1}^{N} \mathcal{F}_i(x^\star_{n_t}) \to \mathcal{F}_0(\hat{x})$, since from (52) and (53), $\lim_{\varepsilon_{n_t} \to 0} \varepsilon_{n_t} \sum_{i=1}^{N} \mathcal{F}_i(x_{\varepsilon_{n_t}}) = 0$.

Therefore, $\mathcal{F}_0(\hat{x}) = \mu = \mathcal{F}_0(x^\star)$. Since $f(x)$ is strictly convex, $\hat{x} = x^\star$. Since every convergent sub-sequence converges to $x^\star$, the whole sequence converges to it. Since the sequence was arbitrary, we have that $x^\star_\varepsilon \to x^\star$ as $\varepsilon \to 0$.

References

[1] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton University Press, 2009.

[2] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Review*, vol. 53, no. 3, pp. 464–501, 2011.

[3] A. L. Soyster, "Convex programming with set-inclusive constraints and applications to inexact linear programming," *Operations Research*, vol. 21, pp. 1154–1157, 1973.

[4] J. E. Falk, "Exact solutions of inexact linear programs," *Operations Research*, vol. 24, pp. 783–787, 1976.

[5] A. Ben-Tal and A. Nemirovski, "Robust solutions of uncertain linear programs," *Operational Research Letters*, vol. 25, no. 1, pp. 1–13, 1999.

[6] A. Ben-Tal, A. Nemirovski, and C. Roos, "Robust solutions of uncertain quadratic and conic-quadratic problems," *Mathematical Operations Research*, vol. 13, no. 2, pp. 535–560, 2002.

[7] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Mathematics of Operational Research*, vol. 23, no. 4, pp. 769–805, 1998.

[8] H. Beyer and B. Sendhoff, "Robust optimization: A comprehensive survey," *Computer Methods in Applied Mechanics and Engineering*, vol. 196, no. 33–34, pp. 3190–3218, 2007.

[9] K.-M. Aigner, A. Bärmann, K. Braun, F. Liers, S. Pokutta, O. Schneider, K. Sharma, and S. Tschuppik, "Data-driven distributionally robust optimization over time," *INFORMS Journal on Optimization*, vol. 5, no. 3, pp. 317–342, 2023.

[10] M. Yang, F. Wang, and D. Huang, "Data-driven distributionally robust optimization under combined ambiguity for cracking production scheduling," *Computers & Chemical Engineering*, vol. 180, p. 108438, 2023.

[11] H. Xu, C. Caramanis, and S. Mannor, "Robustness and regularization of support vector machines," *Journal of Machine Learning Research*, vol. 10, pp. 1485–1510, 2009.

[12] H. Xu, C. Caramanis, and S. Mannor, "Robust regression and lasso," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3561–3574, 2010.

[13] Z. Zhu, J. Hu, A. Majumdar, C. Tomlin, and C. Caramanis, "Zeroth-order optimization meets human feedback: Provable learning via ranking oracles," *arXiv preprint arXiv:2303.03751*, 2023.

[14] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks," *International Conference on Learning Representations (ICLR)*, 2018.

[15] B. L. Gorissen and D. Hertog, "Robust nonlinear optimization via the dual," *Optimization Online*, 2015.

[16] D. Bertsimas, I. Dunning, and M. Lubin, "Reformulation versus cutting-planes for robust optimization," *Computational Management Science*, vol. 13, no. 2, pp. 195–217, 2016.

[17] A. Ben-Tal, D. Hertog, and J. P. Vial, "Deriving robust counterparts of nonlinear uncertain inequalities," *Mathematical Programming*, vol. 149, no. 1, pp. 265–299, 2015.

[18] G. Calafiore and M. C. Campi, "Uncertain convex programs: Randomized solutions and confidence levels," *Mathematical Programming*, vol. 102, no. 1, pp. 25–46, 2004.

[19] G. C. Calafiore, "Random convex programs," *SIAM Journal on Optimization*, vol. 20, no. 6, pp. 3427–3464, 2010.

[20] V. Rostampour and T. Keviczky, "Distributed computational framework for large-scale stochastic convex optimization," *Energies*, vol. 14, no. 1, 2021.

[21] A. Ben-Tal, E. Hazan, T. Koren, and S. Mannor, "Oracle-based robust optimization via online learning," *Operations Research*, vol. 63, no. 3, pp. 628–638, 2015.

[22] A. Mutapcic and S. Boyd, "Cutting-set methods for robust convex optimization with pessimizing oracles," *Optimization Methods and Software*, vol. 24, no. 3, pp. 381–406, 2009.

[23] D. Yazdani, M. Omidvar, D. Yazdani, J. Branke, T. Nguyen, A. Gandomi, Y. Jin, and X. Yao, "Robust optimization over time: A critical review," *IEEE Transactions on Evolutionary Computation*, 2023.

[24] K. Arrow, L. Hurwitz, and H. Uzawa, *Studies in Linear and Non-linear Programming*. Stanford University Press, 1958.

[25] D. Feijer and F. Paganini, "Stability of primal-dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, no. 12, pp. 1974–1981, 2010.

[26] K. Ebrahimi, N. Elia, and U. Vaidya, "A continuous time dynamical system approach for solving robust optimization," in *2019 18th European Control Conference (ECC)*, pp. 1479–1485, IEEE, 2019.

[27] A. Ben-Tal and A. Beck, "Duality in robust optimization: Primal worst equals dual best," *Operations Research Letters*, 2009.

[28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[29] V. Jeyakumar and G. Y. Li, "Strong duality in robust convex programming: Complete characterizations," *SIAM Journal on Optimization*, vol. 20, no. 6, pp. 3384–3407, 2010.

[30] H. Rafique, M. Liu, Q. Lin, and T. Yang, "Weakly-convex concave min-max optimization: Provable algorithms and applications in machine learning," *Optimization Methods and Software*, vol. 37, 2022.

[31] G. Zhang, "Understanding minimax optimization in modern machine learning," Master's thesis, University of Waterloo, 2021.

[32] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice Hall, New Jersey, 1995.

[33] M. Sion, "On general minimax theorems," *Pacific Journal of Mathematics*, vol. 8, no. 1, pp. 171–176, 1958.

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

[35] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," *50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL*, pp. 3800–3805, 2011.

[36] J. W. Simpson-Porco, "Input/output analysis of primal-dual gradient algorithms," *54th Annual Allerton Conference, Allerton House, UIUC, Illinois, USA*, 2016.

[37] K. C. Kosaraju, V. Chinde, R. Pasumarthy, A. Kelkar, and N. M. Singh, "Stability analysis of constrained optimization dynamics via passivity techniques," *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 91–96, 2018.

[38] A. Cherukuri, E. Mallada, and J. Cortes, "Asymptotic convergence of constrained primal-dual dynamics," *Systems and Control Letters*, vol. 87, pp. 10–15, 2016.

[39] K. Ebrahimi, N. Elia, and U. Vaidya, "Robust optimization via discrete-time saddle point algorithm," in *58th Conference on Decision and Control (CDC), Nice, France*, 2019.

[40] W. M. Haddad and V. Chellaboina, *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton University Press, 2008.

[41] R. Farahani, M. Abedian, and S. Sharahi, *Dynamic Facility Location Problem*. Contributions to Management Science. Physica, 2009.

[42] A. Nagurney and D. Zhang, *Projected Dynamical Systems and Variational Inequalities with Applications*, vol. 2. Springer Science & Business Media, 2012.

[43] F. H. Clarke, *Optimization and Nonsmooth Analysis*. Wiley, 1983.

[44] A. Nagurney and D. Zhang, "Projected dynamical systems and variational inequalities with applications," *International Series in Operations Research and Management Science*, vol. 2, 1996.

Keivan Ebrahimi received bachelor's and master's degree in electrical engineering from Sharif University of Technology, Tehran, Iran in 2012 and 2014. He is a Ph.D. candidate in electrical and computer engineering at Iowa State University (ISU), Ames, IA, USA. He is currently a principal data scientist at Tarana Wireless in Milpitas, CA, USA. He received the College of Engineering fellowship at Iowa State University in 2015 and has been selected for the Rock Star Spot Award for stellar contribution and performance in View Inc. COVID-SENSE product, May 2020. He filed three patents to the United States Patents and Trademarks Office for "Environmental Adjustment using Artificial Intelligence", "Identifying and Reducing Health Risks and Tracking Occupancy in a Facility", and "Immersive Collaboration of Remote Participants via Media Displays". His research interests include adversarial machine learning and deep learning, computer vision, robust optimization, control theory and dynamical systems.

Prof. Nicola Elia received the Laurea degree in Electrical Engineering from the Politecnico di Torino, Turin, Italy in 1987 and the Ph.D. degree in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA in 1996. He is currently the Vincentine-Hermes-Luh Chair Professor of Electrical and Computer Engineering at the University of Minnesota (UMN), Twin Cities, MN, USA. Before joining UMN in 2018, and since 1999, he was a faculty with the department of Electrical and Computer Engineering at Iowa State University, Ames, IA, USA. He was a Postdoctoral Associate at the Laboratory for Information and Decision Systems at MIT from 1996 to 1999. He was a Control Engineer with the Fiat research Center, Turin, Italy, from 1987 to 1990. Dr. Elia received the NSF Career Award and he is a Fellow of the IEEE. His research interests include computational methods for controller design, communication systems with access to feedback, control with communication constraints, and network distributed systems.

Prof. Umesh Vaidya received the Ph.D. degree in mechanical engineering from the University of California at Santa Barbara, Santa Barbara, CA, in 2004. He was a Research Engineer at the United Technologies Research Center (UTRC), East Hartford, CT, USA. He is currently a professor in the Department of Mechanical Engineering, Clemson University, S.C., USA. Before joining Clemson University in 2019, and since 2006, he was a faculty with the department of Electrical and Computer Engineering at Iowa State University. He is the recipient of 2012 National Science Foundation CAREER award. His current research interests include dynamical systems and control theory.

Response to Reviewers

Dear Editor and Reviewers,image.png

We sincerely thank you for your thorough review. We have addressed all comments comprehensively with improved clarity, **simplified remarks**, **quantitative metrics** throughout, and **enhanced technical rigor**. All changes are marked in blue for complete transparency.

Response to Reviewer 4

- **Comment:** *"The improvements compared to the existing results are also unclear, which makes it difficult to evaluate the contribution of this article."*

  Response: We have consolidated all contributions into a single, clear "Main Contributions" section in the revised version of the introduction section. In particular, following main contribution section is added.

  Main Contributions: The contributions of this paper are as follows: (1) Model-free operation: We develop a continuous-time dynamical system that solves RO problems using only output feedback, without requiring explicit knowledge of uncertainty models or problem structure. (2) Unified framework: Our method provides a single dynamical system architecture that handles all convex-concave RO problems uniformly. (3) Novel saddle point theory: We prove that the saddle point property holds for RO problems despite the Lagrangian lacking joint convexity-concavity. (4) Custom Lyapunov analysis: We construct a non-standard Lyapunov function that establishes global asymptotic stability. (5) Real-time adaptation: The continuous-time nature of our approach enables real-time tracking of time-varying uncertainties. (6) Broader problem scope: We provide exact solutions for problems where robust counterpart methods fail." Additionally, we removed the two comparison tables that were added for reviewers, as they are not needed in the final version.

- **Comment:** *"There are some language and grammar issues in this paper and the authors need to revise their paper properly. The definite article in the captions of all figures is suggested to be omitted. In general, the definite article in the title should be omitted. The italics of many formulas in this manuscript are not standard, and there are some inconsistent phenomena. There are issues with the quotation marks in the manuscript."*

  Response: We have comprehensively revised the manuscript for clarity and conciseness. Specifically: (i) split 30+ long sentences throughout the paper, (ii) added clear subsection headings throughout all sections to improve readability and navigation, (iii) removed articles from all figure captions. For example:

  "Trajectories of $\mathcal{RO}$ dynamics for robust QP problem with intersection of ellipsoids uncertainty set in Example A."
  "Trajectory of $\mathcal{RO}$ dynamics for robust nonlinear optimization problem (39) in Example B."
  "Locations of agents and anchors, interactions between agents and anchors, and nominal linear constraint in Example C."

Response to Reviewer 5

- **Comment:** *"The paper discusses contributions and motivation in multiple sections, but these points are not clearly articulated or cohesively presented. To improve readability and impact, I recommend reorganizing the introduction to explicitly highlight the key contributions." "The introduction is lengthy and lacks a coherent structure. Please revise it to emphasize the advantages of the proposed method, particularly its ability to operate without prior problem modeling."*
  Response: The Introduction has been completely reorganized with new subsections for clarity. The opening paragraph now directly states:
  "Optimization problems in engineering, finance, and machine learning often involve uncertain parameters that must be handled robustly to ensure reliable performance."
  We added comprehensive section headings throughout the paper for better structure:
  Introduction subsections: "Motivation and Background", "Literature Review", "Existing Solution Approaches", "Alternative Approaches and Limitations", "Our Approach", "Main Contributions", "Paper Organization".
  Similar structural improvements were made throughout: *Section III: "Classical Formulation", "Generalized Formulation", "Assumptions and Properties", "Regularized Formulation". Section IV: "Lagrangian Formulation", "KKT Optimality Conditions", "Saddle Property of Optimal RO Solution". Section V: "Motivation and Challenges", "The RO Dynamics", "Illustrative Examples", "Equilibrium Analysis", "Lyapunov Function Construction", "Main Convergence Result".*
  A dedicated "Main Contributions" section follows the literature review, providing six clearly numbered contributions (see Response to Reviewer 4 for the full text).

- **Comment:** *"Algorithm (23) is presented in isolation without sufficient explanation or context. The authors should provide a detailed discussion immediately after introducing the algorithm, including the intuition behind its design, a clear comparison with existing methods to highlight key differences, and specific improvements or advantages over traditional approaches."*

  Response: We added a comprehensive explanation after equation (23) in Section V:
  "The above system represents the core $\mathcal{RO}$ dynamics where the combined state vector is $z := (x, \lambda, u, v) \in \mathbb{S}$. Each component of this dynamical system serves a specific purpose:"
  followed by detailed explanations. We also added an important remark:
  "Standard primal-dual dynamics fail for RO because when $\lambda_i \to 0$ (inactive constraints), the dynamics $\dot{u}_i = \lambda_i \nabla_{u_i} f_i$ freeze before reaching optimality. We resolve this by removing λ_i from $\dot{u}_i$ dynamics and constructing a Lyapunov function that weights errors by optimal duals $\lambda_i^\star$ rather than current values, enabling global convergence despite lacking joint concavity."
  This explanation clarifies how the dynamics handle the min-max-max-min structure through careful coordination of all variables.

- **Comment:** *"The paper lacks a detailed theoretical analysis of the proposed algorithm's convergence performance compared to existing methods."*

  Response: We added a subsection "Convergence Rate Analysis" in Section V-C to address this:
  "Theorem 1 establishes global asymptotic stability. Unlike discrete algorithms where iteration counts provide explicit rates, continuous-time systems make precise rate quantification challenging. However, convergence speed can be tuned via scaling parameters in the vector field (e.g., multiplying dynamics by $\zeta > 0$). Near optimum, strict convexity ensures exponential convergence. For regularized problems ($c_i = \varepsilon$), the following results prove $\mu_\varepsilon \to \mu$ as $\varepsilon \to 0$."

- **Comment:** *"The structure of the manuscript needs improvement for better readability, as equations (5) and (6) are referenced in Assumption 2 before they are formally introduced in the text."*

  Response: Fixed all forward references. Equations are now properly defined before being referenced.

- **Comment:** *"The techniques from Ref. [22] used in the simulation appear outdated and may not reflect the current advancements in the field. To strengthen the comparative analysis, it is recommended to include state-of-the-art algorithms in the evaluation."*

  Response: We extensively updated the paper with recent state-of-the-art methods. The Introduction now states:
  "Recent developments in distributionally robust optimization (DRO) have bridged the gap between stochastic and robust approaches [9], [10]."
  and
  "Recent work has further integrated robust optimization with machine learning [13], [14]."
  We also added:

"Recent methods include alternating direction method of multipliers (ADMM) decomposition [20], achieving improved convergence rates for distributed robust optimization problems."
and
"Classical cutting-plane approaches [22] treat RO problems as semi-infinite programming but have limitations when pessimization oracles are approximate."
"Robust optimization over time (ROOT) combines robust optimization with dynamic optimization [9], [23]."

- **Comment:** *"The conditions in Assumption 2 seem restrictive. Can they be relaxed, as in Ref. [34]? If not, please provide a detailed explanation." "The need for Assumption 1 should be justified. Can the framework be extended to more general cases?"*

  Response: Added remarks after Assumptions 1 and 2. The remark after Assumption 1 states:
  "The Assumption 1, which requires convexity in the decision variable and concavity in the uncertainty variable, ensures computational tractability and is satisfied by most practical $\mathcal{RO}$ problems."
  The remark after Assumption 2 states:
  "The Slater constraint qualification in Assumption 2 guarantees that the $\mathcal{RO}$ problem enjoys strong duality for upper and lower level optimization problems. While this assumption may appear restrictive, it can often be satisfied or relaxed in practice: (1) In many robust optimization applications, the uncertainty sets are designed with strict feasibility in mind (e.g., ellipsoidal sets with positive radius). (2) The assumption ensures that saddle point and optimal dual solutions exist, which is crucial for our dynamical system approach. (3) For cases where Slater conditions are too restrictive, alternative constraint qualifications such as those discussed in [29] can be employed. (4) In practical implementations, the assumption can often be enforced by appropriate scaling or small perturbations of the constraint sets."
- **Comment:** *"The meaning of h_{ij} and K_i in equation (3) should be clearly understood, and the purpose of introducing them here should be explicitly justified." "The abbreviation 'RC' in the line before equation (9) has been explained earlier and could be deleted here. The meaning of the abbreviation 'RHS' after inequality (51) needs to be explained."*
  Response: Added clarification immediately after equation (3):
  "Here, $h_{ij}(u_i)$ represents the j-th constraint function defining the i-th uncertainty set $\mathcal{U}_i$, and K_i denotes the total number of constraints that define $\mathcal{U}_i$."
  We also added a comprehensive remark explaining the key properties:
  *"Key Properties of Formulation (3): (1) **Generalization:** Formulation (3) generalizes (2) by allowing uncertainty in the objective function and explicitly representing uncertainty sets via inequality constraints, facilitating the min-max-max-min structure essential for our dynamical approach. (2) **Compactness:** Under the convexity assumptions stated below, uncertainty sets $\mathcal{U}_i$ remain compact since convex constraints preserve convexity and boundedness. (3) **Nonlinear constraints:** Our framework handles nonlinear uncertainty constraints $h_{ij}(u_i) \leq 0$ through gradient terms $\nabla_{u_i} h_i(u_i)$ that naturally accommodate nonlinearity."*
  Also added a footnote: *"Since the uncertainty set is compact and the constraint functions are continuous, the supremum is attained within the set; therefore, we can replace 'sup' with 'max' in our formulation."*
  Also fixed RHS abbreviation (right-hand side) and removed redundant RC definition.
- **Comment:** *"The Lagrangian function (15) for the RO problem (4) appears to be straightforward, raising questions about the necessity of the lengthy and intricate analysis preceding it. The authors should assess whether the detailed derivation in the earlier part of Section IV is essential. If the complexity is indeed necessary, the authors should provide a clearer justification for its inclusion, highlighting how it enhances the understanding of the problem or the robustness of the solution."*
  Response: Added two comprehensive remarks in Section IV explaining the necessity of Lagrangian analysis:
  "The Lagrangian (13) is essential because: (i) it unifies nested min-max structure into one framework, (ii) enables continuous-time dynamics over all variables, (iii) provides saddle properties for our Lyapunov function, (iv) handles non-smoothness through dual decomposition. Standard methods fail without this unified approach."
  We also added a remark about KKT conditions:
  "Our KKT conditions extend standard theory to handle: (i) bi-level structure coordinating λ_i and v_{ij}, (ii) regularization via $(c_i + \lambda_i)$ terms, (iii) nested complementary slackness at both optimization levels, (iv) saddle point equivalence enabling our dynamics."

- **Additional Comments 9-18:** Fixed abbreviations (RC, RHS), moved Appendix B content to main text, verified Lemma 4 citation, clarified ϵ^+ notation, explained simulation effectiveness, enhanced Proposition 6 justification with the following remark added to the Appendix:
  "The following proposition ensures: (i) existence despite discontinuities from projections, (ii) uniqueness via Lipschitz property, (iii) continuous dependence on initial conditions, (iv) validity of Lyapunov analysis for Theorem 1."
  Corrected limit expression, fixed language issues, shortened introduction, updated with modern algorithms.

Response to Reviewer 6

- **Comment:** *"My concern lies in the motivation behind the problem formulation (4). Does it offer any advantages compared to formulation (3)? The authors should emphasize the main reason for introducing (4), beyond merely presenting it as a more general version of (3)."*
  Response: The role of c_i is explained in Section III and VI. Section III includes the remark:
  "Our formulation (4) adds regularization terms c_i to the classical $\mathcal{RO}$ problem for three key reasons: (i) preventing singularity when constraints are inactive ($\lambda_i = 0$), (ii) improving numerical stability, (iii) recovering the classical problem as $c_i \to 0$. We maintain separate c_i and λ_i (rather than combined γ_i) to preserve dual variable interpretation and enable our Lyapunov construction."
  Section VI further explains: When constraints are inactive ($\lambda_i^* = 0$), setting $c_i = 0$ would lead to singularities in the Lyapunov function. We added:
  "Assumption 3 provides regularization for inactive constraints. In practice, use small values (like $c_i = 10^{-6}$) for numerical stability. Section VI analyzes the $c_i \to 0$ limit rigorously."
  Theorem 2 proves that as $\varepsilon \to 0$, the solution converges to the true optimum. This provides both theoretical rigor and practical implementation guidance.
- **Comment:** *"In equation (10), to derive the Lagrangian function, the authors introduce λ_i as multipliers. However, one could instead consider multipliers of the form $\gamma_i := c_i + \lambda_i$, which would reduce to the Lagrangian function of formulation (3). Therefore, I still do not see the novelty or specific role of the c_i terms."*
  Response: The separation is crucial for several reasons: (i) λ_i maintains its interpretation as a dual variable that adapts during optimization, (ii) c_i provides independent regularization control that can be set based on numerical considerations, (iii) the separation enables Lyapunov function construction with explicit bounds, (iv) asymptotic recovery of the original problem is possible as $c_i \to 0$ while λ_i converges to optimal dual values. A combined form $\gamma_i = c_i + \lambda_i$ would lose these structural advantages and complicate the convergence analysis.
- **Comment:** *"In formulation (2), the authors take the maximum over the constraint functions, which significantly increases the problem's complexity compared to the classical robust optimization problem (1). Specifically, if the constraint functions are smooth, taking the maximum introduces non-smoothness, making the problem harder to solve than formulation (1). It would be helpful for the authors to justify whether this complexity is justified, whether this step is essential or whether smoother alternatives could be considered would improve the reader's understanding of the trade-offs involved."*
  Response: Our dynamics handle non-smoothness naturally: decompose into smooth subproblems via dual variables, projection operators handle gracefully, continuous-time provides implicit averaging. Superior to subgradient methods.
- **Comment:** *"Lemma 1 is a well-known result, or am I missing something? since, Definition 1 KKT conditions seems to be identical to the KKT conditions, which are then referred to as a saddle point condition, when certain constraint qualifications hold (e.g., Slater's condition). Please clarify this definition, since Section 5.9.1 in Convex Optimization by Boyd refers to the Lagrangian function using generalized inequalities."*
  Response: Critical clarification: Lemma 1 is NOT standard—it is a central contribution. The remark before Lemma 1 (Section IV) now states:
  "The following lemma establishes a saddle-point property that is not implied by classical results such as Sion's minimax theorem or Rockafellar's saddle point theorem, which require joint concavity in the maximization variables. Our Lagrangian $\mathcal{L}(x, \lambda, u, v)$ is jointly convex in (x, v) for fixed (λ, u). However, it is not jointly concave in (λ, u) for fixed (x, v) due to the product terms $(c_i + \lambda_i) \cdot f_i(x, u_i)$, which create bilinear coupling between λ_i and u_i. This violation of joint concavity makes existing primal-dual methods inapplicable and necessitates our novel proof approach."
  This is contribution 3 in our Main Contributions section—a key theoretical advance that extends classical saddle-point theory to the robust optimization setting.

- **Comment:** *"In the numerical experiments, the constraints $f_i(x, u_i)$ are linear in u_i. Did the authors observe similar results when dealing with problem such that the constraints are not necessarily linear in u_i?"*

  Response: Our framework handles nonlinear constraints naturally through the gradient terms in the dynamics (equation 23). In Section VII-B (Robust Nonlinear Optimization with no RC), we demonstrate excellent performance with highly nonlinear constraints of the form $e^{u_j^2} + u_j e^{1/u_j} \leq \rho_j$. The text explicitly states: "No closed-form RC exists for this problem... Methods requiring RC derivation cannot solve this problem at all." Our dynamics achieves exact convergence where robust counterpart methods fail entirely.

- **Minor Comments:** Addressed set compactness under convex assumptions, explained Assumption 3 relaxation strategies, demonstrated generality preservation.

Response to Reviewer 10

- **Comment:** *"The reviewer acknowledges the authors' efforts in addressing a relevant and timely problem. The core idea presented is original and points toward a promising research direction with potential impact. However, the current manuscript has several important limitations in terms of technical depth and presentation, which make it unsuitable for publication as a full research article. Given the value of the contribution, the reviewer encourages the authors to consider re-submitting the work in the form of a technical note."*

  Response: We respectfully argue that this work merits publication as a full article based on several key factors. As stated in our Main Contributions (Introduction):
  "The main contributions of this paper are as follows: (i) Model-free operation: We develop a continuous-time dynamical system that solves RO problems without requiring explicit knowledge of uncertainty models or problem structure. (ii) Unified framework: Our method provides a single dynamical system architecture that handles all convex-concave RO problems uniformly. (iii) Novel saddle point theory: We prove that the saddle point property holds for RO problems despite the Lagrangian lacking joint convexity-concavity—a fundamental violation of standard theory. (iv) Custom Lyapunov analysis: We construct a non-standard Lyapunov function that establishes global asymptotic stability. (v) Real-time adaptation: The continuous-time nature enables real-time tracking of time-varying uncertainties. (vi) Broader problem scope: Our method provides exact solutions for problems where robust counterparts are unknown or computationally intractable."
  1. Novel Theoretical Framework: We introduce an entirely new approach to robust optimization through continuous-time dynamics, departing fundamentally from traditional scenario-based and reformulation-conversion methods. This represents a paradigm shift in how RO problems are conceptualized and solved, not merely an incremental improvement.

  2. Comprehensive Technical Contributions: The manuscript presents:
  - A complete dynamical system architecture with rigorous stability analysis (Theorems 1-4)
  - Novel Lyapunov function construction for min-max-max-min structures
  - Proof of global convergence without joint convexity-concavity assumptions
  - Solutions for problems where existing state-of-the-art methods fail entirely

  Technical notes are typically 6-8 pages focusing on specific improvements or narrow applications. Our work presents a foundational methodology with broad applicability, extensive theoretical development, and comprehensive validation—hallmarks of full articles in IEEE TAC. The revised manuscript achieves the conciseness of a technical note while maintaining the depth and scope expected of a full article.

- **Comment:** *"The abstract is overly long, and the writing style—both in the abstract and throughout the paper—could benefit from greater clarity and conciseness. In several places, long sentences obscure the intended message, making the content harder to follow. The reviewer recommends a thorough revision of the text to improve readability and precision."*
  Response: The abstract has been significantly shortened to 160 words (from over 200) while maintaining all key information. The complete new abstract is:
  "Robust optimization arises in numerous applications where system parameters are uncertain but confined to known sets. This paper introduces RO dynamics—a continuous-time dynamical system that directly solves min–max robust optimization problems without requiring reformulations or explicit uncertainty model specifications. Our key contributions are twofold: (i) we establish the saddle-point property without assuming joint convexity–concavity, and (ii) we construct a novel Lyapunov function that proves global convergence for general robust optimization problems that are convex in the decision variable and concave in the uncertainty. Notably, the method does not require explicit knowledge of the cost or constraint functions, enabling model-free operation and decentralized implementation in real-time applications. Numerical experiments demonstrate substantial speedups over scenario-based methods, including exact solutions for robust quadratic programming with ellipsoidal uncertainties, nonlinear problems where robust counterparts cannot be derived, and distributed sensor placement. These results position continuous-time dynamics as a powerful alternative paradigm for robust optimization, particularly well suited for online adaptation and implementation in physical systems."
  The abstract includes: (i) clear problem statement, (ii) our approach, (iii) key theoretical contributions, (iv) practical advantages, and (v) numerical results—all in a compact format that meets IEEE TAC standards for technical notes.

- **Comment:** *"Also throughout this paper, the author used 'the paper ... or the main contribution of the paper .. '. It is confusing from the reviewer point of view. It is suggested to use 'this paper' to highlight that it refers to this paper under review."*

  Response: We have systematically improved the writing style: (i) split 30+ long sentences throughout the manuscript (e.g., the 5-line sentence about reformulation approaches is now 3 concise sentences), (ii) replaced all instances of "the paper" with "this paper" for clarity, (iii) improved technical precision by using standard terminology consistently. The manuscript is now significantly more readable while maintaining technical rigor.

- **Comment:** *"From the reviewer's perspective, the conclusion presented in the final paragraph of the proof of Theorem 4 is neither straightforward nor self-evident. To improve clarity and support the argument, further elaboration and justification are recommended."*

  Response: We have comprehensively enhanced the proof of Theorem 4. The previously problematic final paragraph that abruptly concluded "Thus, $\mathcal{M} = \bar{\mathcal{M}}$. As $\delta > 0$ is arbitrary, we conclude that the set of optimal RO solutions is globally asymptotically stable" has been completely rewritten with rigorous step-by-step justification. The enhanced proof now includes:
  "To complete the proof, we establish that $\mathcal{M} = \bar{\mathcal{M}}$ and show global convergence. Since any point in $\bar{\mathcal{M}}$ is an optimal $\mathcal{RO}$ solution (as just proven), and any optimal solution is an equilibrium of (23) by Lemma 2, we have that all equilibria satisfy $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V = 0$. Thus $\bar{\mathcal{M}}$ contains all optimal solutions. For the reverse inclusion, note that at any point where $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(z) = 0$, the monotonicity conditions from Lemma 3 must hold with equality, implying $z \in \bar{\mathcal{M}}$. Therefore $\mathcal{M} = \bar{\mathcal{M}}$. By the invariance principle, $\bar{\gamma}(t)$ converges to the largest invariant set in cl($\mathcal{M}$). Since $\mathcal{M}$ equals the set of optimal solutions (which are equilibria by Lemma 2), this largest invariant set is precisely $\mathcal{M}$ itself. Finally, since the Lyapunov function V is radially unbounded and $\delta > 0$ was arbitrary, the sublevel set $V^{-1}(\leq \delta)$ can be made arbitrarily large. Thus, for any initial condition $z_0 \in \mathbb{S}$, we can choose δ large enough so that $z_0 \in V^{-1}(\leq \delta)$, establishing global convergence."
  This provides a complete, rigorous argument establishing both set equality and global convergence without hand-waving.

- **Comment:** *"Regarding Assumption 3, the strict positivity of the C parameter is mathematically convenient—particularly for establishing exact-penalty properties and strong duality results—but may be overly rigid for practical modeling purposes." "The requirement imposed in Corollary 1 is quite strong and may pose challenges in both theoretical analysis and practical implementation. The reviewer acknowledges that the assumption—particularly the strict complementarity condition—is introduced to avoid degeneracy and to ensure the existence of a Lyapunov function with a guaranteed negative rate of change. While this is a mathematically sound approach, it relies on a structural regularity that may not hold in many realistic settings."*

  Response: We comprehensively address this in Section VI (Convergence with inactive constraints). The section includes:
  "Theorem 1 establishes global asymptotic stability. Unlike discrete algorithms where iteration counts provide explicit rates, continuous-time systems make precise rate quantification challenging. However, convergence speed can be tuned via scaling parameters in the vector field (e.g., multiplying dynamics by $\zeta > 0$). Near optimum, strict convexity ensures exponential convergence. For regularized problems ($c_i = \varepsilon$), the following results prove $\mu_\varepsilon \to \mu$ as $\varepsilon \to 0$."
  Also added remarks on relaxing strict complementarity:
  "The corollary's requirement that all constraints be strictly active ($\lambda_i^\star > 0$) can be relaxed using proximal regularization terms or adaptive schemes where inactive constraints are identified and removed."
  And on handling inactive constraints:
  "When $\lambda_i^* = 0$ (inactive), setting $c_i = 0$ causes Lyapunov singularities. Solution: use $c_i = \varepsilon > 0$ (e.g., 10^{-6}). Theorem 2 proves convergence to within $O(\varepsilon)$ of exact solution as $\varepsilon \to 0$."
  We also added the perturbed dynamics formulation:
  "where the notation $[\cdot]^{\varepsilon+}_{\hat{\lambda}_i}$ represents the projection operator that ensures $\hat{\lambda}_i \geq \varepsilon > 0$, providing regularization for inactive constraints."
  This provides both theoretical rigor and practical implementation guidance.

- **Comment:** *"In the presented examples, the actual robust optimization setting—particularly with scenario-based uncertain constraints—is not fully addressed. Incorporating such scenarios would strengthen the practical relevance of the examples. In addition, the reviewer was expecting to see results related to convergence analysis, including the stability of the proposed approach and its convergence to optimal points. Providing these insights would offer a more complete picture of the method's performance and theoretical soundness."*

  Response: Section VII provides comprehensive simulations with detailed validation:
  "We validate our $\mathcal{RO}$ dynamics through three carefully chosen examples that demonstrate: (i) exact solutions for problems with complex uncertainty sets where RC methods fail, (ii) quantitative performance comparisons with state-of-the-art methods, (iii) scalability to higher-dimensional problems. Each example highlights unique capabilities of our approach—handling intersection of ellipsoids, solving problems without tractable RC, and achieving significant advantages over other methods."
  Section VII-A states:
  "Intersection of ellipsoids models correlated uncertainties common in portfolio optimization, manufacturing tolerances, and robust control. The solution must remain feasible for ALL points in this non-smooth set, not just sampled scenarios."

Section VII-B demonstrates:
"For this problem with highly nonlinear uncertainty constraints $e^{u_j^2} + u_j e^{1/u_j} \leq \rho_j$*, no closed-form robust counterpart exists, making RC-based methods inapplicable. We compared our approach with scenario-based sampling: using CVX, 168 scenarios yielded solution [0.5376, 0.8193] with cost 0.8039 (2.3s), 500 scenarios gave [0.5312, 0.8024] with cost 0.8287 (8.7s), and 1000 scenarios produced [0.5289, 0.7953] with cost 0.8371 (31.2s). Our method obtained the exact solution [0.5271, 0.7916] with optimal cost 0.8419 in 0.8s integration time—40× faster than the 1000-scenario approximation while providing the exact solution rather than an approximation."*
"This example illustrates a key advantage of our approach: the dynamics naturally handle nonlinear uncertainty constraints through gradient terms $\nabla_{u_j} h_j(u_j)$*, achieving exact convergence where other methods either fail completely (when RC cannot be derived) or require computationally expensive approximations with complexity* $O(N^3 n^3)$ *for* N *scenarios versus our* $O(n^2)$ *scaling."*
Additional Comments: Fixed continuity assumption, notation inconsistencies, missing parentheses. Enhanced Remark 4 clarity, addressed all minor technical issues.
We believe the revised manuscript now clearly demonstrates its contributions and meets IEEE TAC standards.