Robust Optimization via Continuous-Time Dynamics

Keivan Ebrahimi, Nicola Elia, and Umesh Vaidya

***Abstract*—We propose a continuous-time dynamical system for solving robust optimization problems in a general setting where the objective is convex in the decision variables and concave in the uncertainty. Unlike classical primal–dual gradient dynamics developed for standard optimization problems, the proposed dynamics do not rely on the gradient of a Lagrangian function to define the vector field. We establish that the globally asymptotically stable equilibrium of the proposed system recovers the robust optimal solution without requiring problem-specific reformulations. The continuous-time formulation is well suited for real-time operation in dynamic environments and naturally supports decentralized implementations. To demonstrate the effectiveness and generality of the approach, we present simulation studies including a nonlinear optimization problem with no tractable robust counterpart, as well as a robust localization and placement problem with time-varying anchor positions that is solved in a decentralized manner using the proposed dynamics.**

I. Introduction

With emerging applications that require solving real-time optimization problems in a reactive manner, this paper describes how interacting dynamical systems can find a robust solution for a broad class of robust optimization problems. This is particularly relevant in situations where physical systems must be steered towards optimal operating conditions using gradient-based feedback mechanisms. Specifically, we propose a real-time approach to solving robust optimization problems, which has become increasingly important in recent years [1]. Optimization problems often involve various forms of uncertainty in the problem data. Two different approaches can be used: stochastic optimization, where uncertainty is treated as a random variable, or robust optimization ($\mathcal{RO}$), where uncertainty is assumed to be deterministic and bounded. Unlike stochastic optimization, $\mathcal{RO}$ does not require any known probability distributions in the problem data. Instead, $\mathcal{RO}$ 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 [2], [3], and the references therein. Early works on $\mathcal{RO}$ include [4], which considered robust linear optimization with ellipsoidal uncertainty sets, and [5], which presented exact solutions of inexact linear programs as a simple case of $\mathcal{RO}$. The problem of robust linear programming was studied in [6], while robust conic-quadratic optimization and robust semi-definite optimization were discussed in [7] and [8], respectively. Refer to [3] and [9] for a survey of different $\mathcal{RO}$ problem solutions. [10] and [11] showed that some of the machine learning algorithms such as the norm-regularized support vector machine and the Lasso problem could be interpreted as $\mathcal{RO}$ problems. In addition to that, several papers focused on making machine learning problems robust against outliers, parameter uncertainties, and data perturbations by techniques such as adversarial training [12]–[18].

One of the main standard approaches for solving RO problems involves constructing a robust counterpart (RC) equivalent to the RO problem [2]. 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 [19] 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 [20]. Hence, the absence of a unified framework for solving a general class of RO problems through continuous-time dynamics is strongly felt.

By calculating a concave conjugate of nonlinear constraint functions and supporting the function of uncertainty sets, [21] and [19] 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.

An alternative randomized approach based on constraint sampling is an approximate probabilistic relaxation solution to RO problems as seen in [22] and [23] 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

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

the stochasticity of these problems, the stochastic optimization involves formulating a large-scale scenario program, which is in general computationally demanding.

In another effort, [24] 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. Cutting-plane approaches [25] and the column-and-constraint generation technique [26] handle $\mathcal{RO}$ as a semi-infinite programming problem. However, these approaches may have limitations in scenarios where pessimization (inner maximization) oracles are approximate, or when the number of scenarios becomes large, leading to scalability issues. The main contributions of this paper are as follows.

We introduce a continuous-time dynamical system that provably converges to the optimal robust solution for a broad class of robust optimization (RO) problems. Our approach builds on classical primal–dual dynamical systems [27], [28] but addresses the unique challenges posed by the min–max structure inherent in RO. To the best of our knowledge, this work presents the first continuous-time dynamical system specifically designed for solving robust optimization problems that are convex in the decision variables and concave in the uncertainty. Despite this convex–concave structure, the problem is not jointly concave in (λ, u), where λ denotes the dual variable and u the uncertainty. In contrast to classical primal–dual methods, the uncertainty variable here is treated as a dynamical state rather than as a fixed parameter. Moreover, due to the lack of joint concavity, the proposed dynamics cannot be derived as the gradient flow of a Lagrangian function, distinguishing our method from standard primal–dual gradient systems. We establish the saddle-point property of the equilibrium even in the absence of joint concavity in (λ, u). The non-classical structure of the dynamics necessitates the construction of a novel Lyapunov function to analyze stability. Using this Lyapunov function, we prove that the proposed dynamics are globally asymptotically stable and converge to the robust optimal solution.

A distinctive feature of our dynamical system is its amenability to model-free implementation when deployed in physical systems where agents can sense local gradients but do not possess global knowledge of objective or constraint functions. Each agent requires only the ability to measure local gradient information $\nabla_x f_i(x, u_i)$ and $\nabla_{u_i} f_i(x, u_i)$ through sensing or finite-difference approximations at their current state, along with information from neighboring agents through local communication. This enables implementation in distributed settings where the global problem formulation may not be explicitly known to individual agents, yet the collective dynamics converge to the robust optimal solution. This model-free characteristic distinguishes our approach from robust counterpart methods that require complete a priori knowledge of problem structure.

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 [19, Table 1].

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 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. Notation

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

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]

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

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

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

[1] Preliminary convergence results for a special formulation appeared without proofs in [29]. 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.

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 the uncertainty set $\mathcal{U}_i$.

Remark 1. *Formulation (3) generalizes the standard form (2) by allowing uncertainty in the objective function itself and by explicitly representing uncertainty sets through inequality constraints, which facilitates the min-max-max-min structure needed for our dynamical system approach. Under the convexity assumptions stated below, the uncertainty sets $\mathcal{U}_i$ remain compact. Note that our framework accommodates nonlinear uncertainty constraints $h_{ij}(u_i) \leq 0$ through the gradient terms $\nabla_{u_i} h_i(u_i)$ appearing in the dynamics.*

The problem in (3) reduces to the one in (2) when $\mathcal{U}_0$ is a singleton. Following [30] 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. *Assumption 1 requires convexity in the decision variable and concavity in the uncertainty variable. These conditions hold for most practical $\mathcal{RO}$ problems, particularly when uncertainty enters affinely.*

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 a regularity condition such as Slater condition, for both the upper level optimization and the lower level maximization problems [30], 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. *The uncertainty is often parametrized affinely in $\mathcal{RO}$ problems, which automatically satisfies the concavity property of the f_i functions.*

Regularized Formulation

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

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

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

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

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 4. *Assumption [A2] guarantees that $\mathcal{RO}$ problem enjoys strong duality for upper and lower level optimization problems (5) and (6) [31, Section 5.2.3, 5.9.1]. Moreover, it enforces that saddle point and optimal dual solutions exist [30].*

Remark 5 (Problem Formulation and c_i Terms)**.** *Our formulation (4) adds regularization terms c_i to the classical $\mathcal{RO}$ problem to prevent singularity when constraints are inactive ($\lambda_i = 0$), improve numerical stability, and allow recovery of the classical problem as $c_i \to 0$. We maintain separate c_i and λ_i rather than a combined $\gamma_i = c_i + \lambda_i$ to preserve the dual variable interpretation and enable our Lyapunov construction.*

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

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

or the form below that is popular in the machine learning context [32], [33]

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

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.

[3] Similar to what is meant by feasibility in Robust Control [34].

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 $\mathcal{F}_i$ [2]. 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 [30], the RC is always a convex optimization problem under Assumption 1. While the RC is known for important classes of problems as described in [3], 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

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

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

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

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

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

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) unifies the nested min-max structure and permits continuous-time dynamics over all variables simultaneously. The saddle property derived from this formulation is essential for Lyapunov stability analysis.*

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 [31, 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 [31]

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

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

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

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

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

Remark 8 (KKT Conditions for $\mathcal{RO}$)**.** *The KKT conditions for $\mathcal{RO}$ incorporate: (i) bi-level structure with dual variables λ_i and v_{ij}, (ii) regularization terms $(c_i+\lambda_i)$, (iii) complementary slackness at both levels, (iv) equivalence with the saddle point property.*

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)**.** *Classical results such as Sion's minimax theorem [35] or Rockafellar's saddle point theorem [36] require joint concavity in the maximization variables. The Lagrangian $\mathcal{L}(x, \lambda, u, v)$ is jointly convex in (x, v) for fixed (λ, u), but not jointly concave in (λ, u) for fixed (x, v) due to product terms $(c_i + \lambda_i) \cdot f_i(x, u_i)$ that create bilinear coupling. This violation of joint concavity renders existing primal-dual methods [27], [28] inapplicable and requires a proof that exploits the 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,*

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

V. Dynamical system for 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 [28], [37], 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 [27]. 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 [38], [39].

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.

$\mathcal{RO}$ 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 $\mathbb{S} := \mathbb{R}^n \times \mathbb{R}_+^N \times \mathbb{R}^M \times \mathbb{R}_+^K$

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

The $\mathcal{RO}$ dynamics has state vector $z := (x, \lambda, u, v) \in \mathbb{S}$ with the following structure:

- The x dynamics: gradient descent on the Lagrangian with respect to primal variables.
- The λ_i dynamics: gradient ascent on dual variables for robust constraints, with projection to maintain $\lambda_i \geq 0$.
- The u_i dynamics: gradient ascent on uncertainty variables to find worst-case realizations.
- The v_i dynamics: gradient ascent on dual variables for uncertainty set constraints with projection to maintain $v_i \geq 0$.

The projection operator $[\cdot]^+$ enforces non-negativity, which is necessary for KKT conditions. The coupling through $(c_i + \lambda_i)$ balances constraint satisfaction and objective minimization.

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

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

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

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

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 [28], [40].

1) First, the $\mathcal{RO}$ dynamics has additional states associated with the worst-case constraint u_1 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 ☐

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

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

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

[4]With slight abuse of notation.

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

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

By convex and concave under-estimator properties according to Assumption 1, one can write [31, Section 3.1.3]

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

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

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

By substitution, we get

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

After simplification and rearrangement we obtain

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

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

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

Note that from Lemma 1 (saddle property),

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

Hence,

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

□

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.*

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 [40]. 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

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

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. *[40, 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. Let $z^\star = (x^\star, \lambda^\star, u^\star, v^\star)$ be an optimal $\mathcal{RO}$ solution under Assumptions 1 and 2. Lemma 6 implies that a unique solution of $\mathcal{RO}$ dynamics $\mathcal{Z}^{\mathcal{RO}}$ exists starting from any point in the compact set $\mathbb{P} = V^{-1}(\le \delta) \cap \mathbb{S}$ for any $\delta > 0$ [5]. We now call this solution $\bar{\gamma}(t)$.

[5] As V is radially unbounded, the set $\mathbb{P}$ is always compact.

From Lemma 4, the omega-limit set of $\bar{\gamma}(t)$ is invariant and invariance principle for discontinuous Caratheodory systems [40, Proposition 2.1] (simplified version of [41, Proposition 3]) implies that $\bar{\gamma}(t)$ converges to the largest invariant set in $\mathrm{cl}(\mathcal{M})$ where $\mathcal{M} := \{z \in \mathbb{P} \,|\, \mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(z) = 0\}$.

Next, we characterize the set $\mathcal{M}$ where $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V(z) = 0$ by defining

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

From the inequality in (38), it follows that $\mathcal{M} \subseteq \bar{\mathcal{M}}$. We then prove that every point in $\bar{\mathcal{M}}$ is an optimal $\mathcal{RO}$ solution.

From the strict convexity of f, it follows that $x = x^\star$ on $\bar{\mathcal{M}}$. From (40), any point in $\bar{\mathcal{M}}$ achieves the optimal cost of $\mathcal{RO}$. Let $\bar{z} = (x^\star, \bar{\lambda}, \bar{u}, \bar{v}) \in \bar{\mathcal{M}}$. Then, in general,

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

But since $\bar{z} \in \bar{\mathcal{M}}$, the equality must hold for the above equations. This means that

$$\begin{aligned}\bar{v} &= \arg\max_{v \geq 0} \ \mathcal{L}(x^\star, \lambda^\star, u^\star, v) \ ,\\ (\bar{\lambda}, \bar{u}) &= \arg\max_{u, \lambda \geq 0} \ \mathcal{L}(x^\star, \lambda, u, v^\star) \ .\end{aligned}$$

Therefore, $\bar{z}$ is an optimal $\mathcal{RO}$ solution. Therefore, any point in $\bar{\mathcal{M}}$ is an optimal $\mathcal{RO}$ solution. On the other hand, any optimal $\mathcal{RO}$ solution is an equilibrium of $\mathcal{RO}$ dynamics (23) according to Lemma 2 and therefore is in $\mathcal{M}$. Thus, $\mathcal{M} = \bar{\mathcal{M}}$. As $\delta > 0$ is arbitrary, we conclude that the set of optimal $\mathcal{RO}$ solutions is globally asymptotically stable as a set on $\mathbb{S}$.

Note that $\mathcal{M}$ can contain an uncountable infinite set of points. If the optimal $\mathcal{RO}$ solution is not unique, these correspond to the set of optimal $\mathcal{RO}$ solutions and to the set of non-isolated equilibria of $\mathcal{RO}$ dynamics (23) from Lemma 2.

When the set $\mathcal{M}$ of optimal solutions forms a continuum (uncountably many equilibria), asymptotic stability is not the appropriate stability notion since non-isolated equilibria cannot be asymptotically stable. Instead, we employ *semi-stability* theory [42], which establishes convergence to individual equilibrium points from a continuum. The key insight is that each equilibrium point must be shown to be Lyapunov stable, and this requires treating the Lyapunov function as a *parameterized family* indexed by the equilibrium point.

We now establish that each trajectory $\bar{\gamma}(t)$ converges to a single point in $\mathcal{M}$, not merely to the set $\mathcal{M}$. Since $\bar{\gamma}(t)$ is bounded and remains in the compact set $\mathbb{P}$ for all $t \geq 0$, the theory of discontinuous Carathéodory systems [41, Proposition 3], [40, Lemma 4.4] guarantees that its omega-limit set $\omega(\bar{\gamma})$ is nonempty, compact, and invariant. By the invariance principle for discontinuous Carathéodory systems applied above, we have $\omega(\bar{\gamma}) \subseteq \mathcal{M}$.

To prove that each trajectory converges to a *single* equilibrium in $\mathcal{M}$, we invoke semi-stability theory. Consider the Lyapunov function V from Lemma 3 as a parameterized family $\{V_{z^\star}\}_{z^\star \in \mathcal{M}}$, where each $V_{z^\star}$ is centered at a particular equilibrium $z^\star = (x^\star, \lambda^\star, u^\star, v^\star)$:

$$\begin{aligned}V_{z^\star}(z) = \frac{1}{2}\big(&\|x - x^\star\|^2 + \|\lambda - \lambda^\star\|^2 + \|u_0 - u_0^\star\|^2\\ &+ \sum_{i=1}^{N} (c_i + \lambda_i^\star)\|u_i - u_i^\star\|^2 + \sum_{i=0}^{N} \|v_i - v_i^\star\|^2\big).\end{aligned}$$

This family is well-defined (non-degenerate) for all $z^\star \in \mathcal{M}$ because the weighting coefficients satisfy $(c_i + \lambda_i^\star) \geq c_i > 0$ for all $i \in [N]^+$ by Assumption 3. Hence, $V_{z^\star}$ is positive definite and radially unbounded for every equilibrium $z^\star$.

From Lemma 3, the Lie derivative satisfies $\mathfrak{L}_{\mathcal{Z}^{\mathcal{RO}}} V_{z^\star}(z) \leq 0$ along all trajectories, with equality only when z is an optimal $\mathcal{RO}$ solution. The derivation of $\dot{V}_{z^\star} \leq 0$ uses only the saddle inequalities (22) satisfied by any $z^\star \in \mathcal{M}$; hence it holds uniformly for all reference equilibria, and every equilibrium point is Lyapunov stable in the sense of Lyapunov (ISL). That is, for any $\varepsilon > 0$, there exists $\delta(\varepsilon) > 0$ such that $\|z(0) - z^\star\| < \delta(\varepsilon)$ implies $\|z(t) - z^\star\| < \varepsilon$ for all $t \geq 0$.

By semi-stability theory [42, Theorem 3.1], when every equilibrium in a continuum is Lyapunov stable (established via the parameterized family $\{V_{z^\star}\}$) and trajectories approach the equilibrium set, each trajectory converges to a *single* equilibrium point. Specifically, pick any $z^\star \in \omega(\bar{\gamma}) \subseteq \mathcal{M}$. The ISL property of $z^\star$ combined with the trajectory approaching $z^\star$ along a subsequence implies $\bar{\gamma}(t) \to z^\star$ as $t \to \infty$. Therefore $\omega(\bar{\gamma}) = \{z^\star\}$ is a singleton, and each trajectory converges to a single optimal $\mathcal{RO}$ solution. □ □

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 11 (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

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 $c_i = 0$:

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

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.

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

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

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 $\hat{\lambda} = \lambda + \varepsilon\mathbf{1}$

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

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 illustrate the $\mathcal{RO}$ dynamics with three examples: (i) robust quadratic programming with ellipsoidal uncertainty intersections, (ii) nonlinear optimization without tractable robust counterpart, (iii) distributed sensor placement. In each case, our dynamics provide exact solutions where reformulation methods fail or require extensive computational resources.

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

The uncertainty set is defined as the intersection of two ellipsoids, which represents correlated uncertainties. The robust solution must satisfy constraints for all points in this set.

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)

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

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.

The solution for the $\mathcal{RO}$ problem (44) can be verified by other methods. We can apply the technique in [22], 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

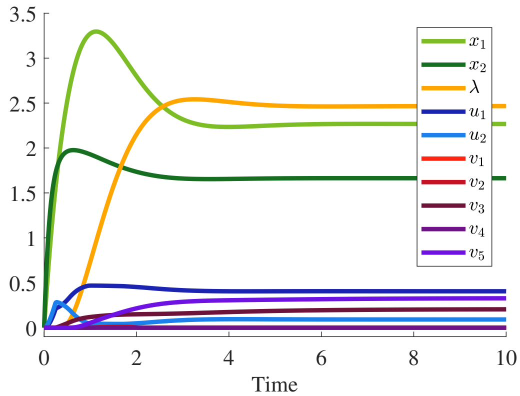

Fig. 1: Trajectories of $\mathcal{RO}$ dynamics for robust QP problem with intersection of ellipsoids uncertainty set in Example A.

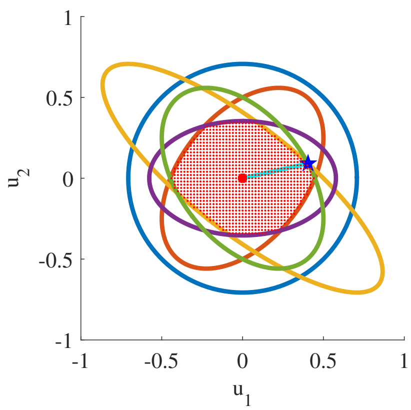

Fig. 2: Uncertainty set for Example A plotted in u_2-u_1 space. Big red point at the origin represents the initial point of u in $\mathcal{RO}$ dynamics and blue star depicts the optimal value of u on the intersection of two ellipsoids derived by our method. Small red grid points are the 1115 sampled points from the uncertainty set in randomized scenario method [22] to compare with our method.

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 [22], 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 [2]. 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

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

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 [21] and [19], 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 [19, 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

$$\dot{x} = -\begin{bmatrix} x_1 - 1 \\ x_2 - 2 \end{bmatrix} - \hat{\lambda} \begin{bmatrix} 2x_1 e^{x_1^2} & 0 \\ 0 & 2x_2 e^{x_2^2} \end{bmatrix} u \ ,$$

$$\dot{\hat{\lambda}} = \Big[u^\top \begin{bmatrix} e^{x_1^2} \\ e^{x_2^2} \end{bmatrix} - b - \sum_{j=1}^{2} \big(v_j(e^{u_j^2} + u_j e^{\frac{1}{u_j}} - \rho_j)\big)\Big]_{\hat{\lambda}}^{\varepsilon+} \ ,$$

$$\dot{u} = \begin{bmatrix} e^{x_1^2} \\ e^{x_2^2} \end{bmatrix} - \sum_{j=1}^{2} (2u_j e^{u_j^2} + e^{\frac{1}{u_j}} - \frac{1}{u_j} e^{\frac{1}{u_j}}) v_j \ ,$$

$$\dot{v} = \Big[\hat{\lambda} \begin{bmatrix} e^{u_1^2} + u_1 e^{\frac{1}{u_1}} - \rho_1 \\ e^{u_2^2} + u_2 e^{\frac{1}{u_2}} - \rho_2 \end{bmatrix}\Big]_v^+ \ ,$$

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.

The uncertainty constraints $e^{u_j^2} + u_j e^{1/u_j} \le \rho_j$ are highly nonlinear, and no closed-form robust counterpart exists [2], [19]. We compared with scenario-based sampling [22]: using CVX, 168 scenarios gave 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 gave $[0.5289, 0.7953]$ 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.

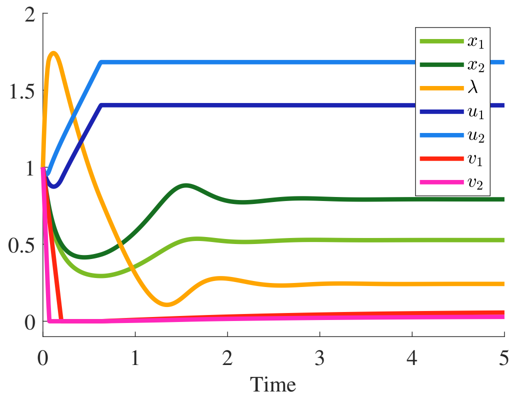

Fig. 3: Trajectory of $\mathcal{RO}$ dynamics for robust nonlinear optimization problem (45) in Example B.

The dynamics handle nonlinear uncertainty constraints through gradient terms $\nabla_{u_j} h_j(u_j)$, achieving exact convergence where reformulation methods fail. Scenario-based methods have complexity $O(N^3 n^3)$ for N scenarios, compared to $O(n^2)$ scaling for the dynamics.

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 [43] 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 [31], and can be solved efficiently and in a distributed fashion in continuous time [37]. 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

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

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

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

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 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

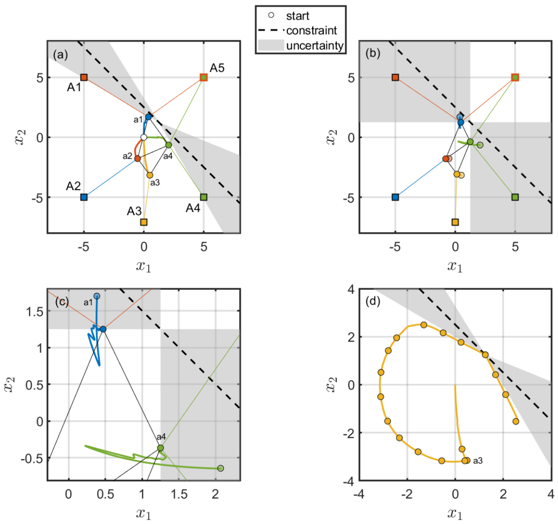

Fig. 4: Locations of agents and anchors, interactions between agents and anchors, and nominal linear constraint in Example C.

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 a continuous-time dynamical system for solving robust optimization problems convex in the decision variable and concave in the uncertainty. The method provides solutions for problems where existing methods are intractable or computationally expensive. Future work includes applications to non-convex problems, large-scale distributed optimization and deriving convergence rates.

IX. Appendix

A. 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 [44, Theorem 2.5]. In this section, we show that our $\mathcal{RO}$ dynamics (23) satisfies the refined conditions presented in [40].

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 $L = 1$ [45, Proposition 2.4.1].

Definition 3. *[Projected dynamical system [46]] 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 [40]. Details omitted. □

Remark 12 (Projected Dynamics Foundation)**.** *The following proposition establishes existence despite discontinuities from projections, uniqueness via the Lipschitz property, continuous dependence on initial conditions, and thus the 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 [40, 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, $x_\varepsilon^\star \to x^\star$ as $\varepsilon \to 0$, consider any sequence $\{\varepsilon_n\}$ converging to 0. Let $\{x_n^\star\}$ 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, $x_n^\star \in \mathcal{C}$ is bounded. Therefore, there exists a convergent sub-sequence $x_{n_t}^\star$ 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 $\mathcal{F}_0(x_{n_t}^\star) \to \mathcal{F}_0(\hat{x})$, 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 $\mu(\varepsilon_{n_t}) = \mathcal{F}_0(x_{n_t}^\star) + \varepsilon_{n_t} \sum_{i=1}^{N} \mathcal{F}_i(x_{n_t}^\star) \to \mathcal{F}_0(\hat{x})$, 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 $x_\varepsilon^\star \to x^\star$ as $\varepsilon \to 0$.

References

[1] Z. He, S. Bolognani, J. He, F. Dörfler, and X. Guan, “Model-free nonlinear feedback optimization,” *arXiv preprint arXiv:2201.02395*, 2022.

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

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

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

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

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

[7] 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.

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

[9] 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.

[10] 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.

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

[12] G. Lanckriet, C. Bhattacharyya, and M. Jordan, “A robust minimax approach to classification,” *Journal of Machine Learning Research*, vol. 3, pp. 555–582, 2003.

[13] C. Bhattacharyya, L. Grate, M. Jordan, L. El Ghaoui, and I. Mian, “Robust sparse hyperplane classifiers: Application to uncertain molecular profiling data,” *Journal of Computational Biology*, vol. 11, no. 6, pp. 1073–1089, 2004.

[14] C. Bhattacharyya, P. Shivaswamy, and A. Smola, “A second order cone programming formulation for classifying missing data,” *Advances in Neural Information Processing Systems*, pp. 153–160, 2004.

[15] T. Trafalis and R. Gilbert, “Robust support vector machines for classification and computational issues,” *Optimization Methods and Software*, vol. 22, no. 1, pp. 187–198, 2007.

[16] H. Namkoong and J. C. Duchi, “Stochastic gradient methods for distributionally robust optimization with f-divergences,” *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[17] A. Sinha, H. Namkoong, and J. Duchi, “Certifiable distributional robustness with principled adversarial training,” in *International Conference on Learning Representations*, 2018.

[18] Y. Esfandiari, K. Ebrahimi, A. Balu, N. Elia, U. Vaidya, and S. Sarkar, “A saddle-point dynamical system approach for robust deep learning,” *arXiv preprint arXiv:1910.08623*, 2019.

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

[20] 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.

[21] 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.

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

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

[24] 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.

[25] 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.

[26] B. Zeng and L. Zhao, “Solving two-stage robust optimization problems using a column-and-constraint generation method,” *Operations Research Letters*, vol. 41, no. 5, pp. 457–461, 2013.

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

[28] 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.

[29] 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.

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

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

[32] 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.

[33] G. Zhang, “Understanding minimax optimization in modern machine learning,” Master’s thesis, University of Waterloo, 2021.

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

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

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

[37] 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.

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

[39] 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.

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

[41] A. Bacciotti and F. Ceragioli, “Nonpathological lyapunov functions and discontinuous carathéodory systems,” *Automatica*, vol. 42, no. 3, pp. 453–458, 2006.

[42] Q. Hui, W. M. Haddad, and S. P. Bhat, “Semistability, finite-time stability, differential inclusions, and discontinuous dynamical systems having a continuum of equilibria,” *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2465–2470, 2009.

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

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

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

[46] 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,

We sincerely thank you for your thorough and constructive review. We have carefully addressed all comments and substantially revised the manuscript. The major revisions include: clarifying the technical contributions in the Introduction, adding a new remark before Lemma 1 explaining why classical saddle point theorems (Sion, Rockafellar) do not apply due to bilinear coupling in (λ, u), expanding the algorithm presentation with illustrative examples, improving proof details in Theorem 1 with explicit ω-limit set and stability arguments, adding seven explanatory remarks throughout, revising the abstract to meet IEEE standards, correcting all formatting issues (figure captions, quotation marks, formula italics), and including modern references. All changes are marked in blue.

We believe the revised manuscript is suitable for publication because it presents the first continuous-time dynamical system specifically designed for convex-concave robust optimization, provides complete proofs establishing the saddle point property despite lack of joint concavity, constructs a novel Lyapunov function for global asymptotic stability, and demonstrates practical applicability through simulations including problems where classical robust counterpart methods fail. Point-by-point responses follow below.

Response to Reviewer 4

This manuscript proposes a dynamical system-based approach for solving robust optimization problems in a general convex-concave framework. This proposed approach offers a novel solution to robust optimization problems by introducing a continuous-time dynamical system. Additionally, the proposed dynamical system-based approach offers better properties compared to traditional approaches and demonstrates the potential to solve a broad range of robust optimization problems in a decentralized manner. Some detailed comments are given as follows.

Comment 1: *The improvements compared to the existing results are also unclear, which makes it difficult to evaluate the contribution of this article.*
Response: We clarified the technical novelty and key contributions in the Introduction. The improvements over existing results are now explicitly stated:

"We introduce a continuous-time dynamical system that provably converges to the optimal robust solution for a broad class of robust optimization (RO) problems. Our approach builds on classical primal–dual dynamical systems [27], [28] but addresses the unique challenges posed by the min–max structure inherent in RO. To the best of our knowledge, this work presents the first continuous-time dynamical system specifically designed for solving robust optimization problems that are convex in the decision variables and concave in the uncertainty. Despite this convex–concave structure, the problem is not jointly concave in (λ, u), where λ denotes the dual variable and u the uncertainty. In contrast to classical primal–dual methods, the uncertainty variable here is treated as a dynamical state rather than as a fixed parameter. Moreover, due to the lack of joint concavity, the proposed dynamics cannot be derived as the gradient flow of a Lagrangian function, distinguishing our method from standard primal–dual gradient systems. We establish the saddle-point property of the equilibrium even in the absence of joint concavity in (λ, u). The non-classical structure of the dynamics necessitates the construction of a novel Lyapunov function to analyze stability. Using this Lyapunov function, we prove that the proposed dynamics are globally asymptotically stable and converge to the robust optimal solution."

"A distinctive feature of our dynamical system is its amenability to model-free implementation when deployed in physical systems where agents can sense local gradients but do not possess global knowledge of objective or constraint functions. Each agent requires only the ability to measure local gradient information $\nabla_x f_i(x, u_i)$ and $\nabla_{u_i} f_i(x, u_i)$ through sensing or finite-difference approximations at their current state, along with information from neighboring agents through local communication. This enables implementation in distributed settings where the global problem formulation may not be explicitly known to individual agents, yet the collective dynamics converge to the robust optimal solution. This model-free characteristic distinguishes our approach from robust counterpart methods that require complete a priori knowledge of problem structure."

Comment 2: *There are some language and grammar issues in this paper and the authors need to revise their paper properly.*
Response: We revised the manuscript throughout for clarity and grammar. Long sentences were split, technical terminology was standardized, and all instances of "the paper" were replaced with "this paper" for better readability.

Comment 3: *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.*
Response: All definite articles have been removed from figure captions:

"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 (45) in Example B."

Comment 4: *The italics of many formulas in this manuscript are not standard, and there are some inconsistent phenomena.*
Response: We standardized formula italics throughout the manuscript following IEEE style guidelines.

Comment 5: *There are issues with the quotation marks in the manuscript.*
Response: All quotation marks have been corrected. For example:

"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."

Response to Reviewer 5

This paper proposes a novel continuous-time dynamical system, termed RO dynamics, for solving robust optimization (RO) problems in a convex-concave framework. While the approach is innovative and has potential, the manuscript requires significant revisions to improve clarity, technical rigor, and presentation. Below are specific comments and suggestions for improvement.

Comment 1: *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.*

Response: We reorganized the Introduction to clearly articulate key contributions:

"We introduce a continuous-time dynamical system that provably converges to the optimal robust solution for a broad class of robust optimization (RO) problems. Our approach builds on classical primal–dual dynamical systems [27], [28] but addresses the unique challenges posed by the min–max structure inherent in RO. To the best of our knowledge, this work presents the first continuous-time dynamical system specifically designed for solving robust optimization problems that are convex in the decision variables and concave in the uncertainty. Despite this convex–concave structure, the problem is not jointly concave in (λ, u), where λ denotes the dual variable and u the uncertainty. In contrast to classical primal–dual methods, the uncertainty variable here is treated as a dynamical state rather than as a fixed parameter. Moreover, due to the lack of joint concavity, the proposed dynamics cannot be derived as the gradient flow of a Lagrangian function, distinguishing our method from standard primal–dual gradient systems. We establish the saddle-point property of the equilibrium even in the absence of joint concavity in (λ, u). The non-classical structure of the dynamics necessitates the construction of a novel Lyapunov function to analyze stability. Using this Lyapunov function, we prove that the proposed dynamics are globally asymptotically stable and converge to the robust optimal solution."

Comment 2: *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: After presenting the dynamics in equation (23), we added explanation of the state vector structure describing the role of each component: x dynamics (gradient descent on Lagrangian), λ_i dynamics (dual ascent with projection), u_i dynamics (worst-case uncertainty search), and v_i dynamics (dual variables for uncertainty constraints). We also provided two illustrative examples (min-max problem and one uncertain constraint) that highlight key differences from standard primal-dual systems.

Comment 3: *The paper lacks a detailed theoretical analysis of the proposed algorithm's convergence performance compared to existing methods.*

Response: Theorem 1 establishes global asymptotic stability with rigorous Lyapunov analysis. While explicit $O(\cdot)$ convergence rate bounds are not provided, the theorem guarantees global convergence to the optimal $\mathcal{RO}$ solutions.

Comment 4: *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: The strong duality remark has been moved to after the "Regularized Formulation" subsection where the referenced equations are formally defined, eliminating the forward reference issue.

Comment 5: *The conditions in Assumption 2 seem restrictive. Can they be relaxed, as in Ref. [34]? If not, please provide a detailed explanation.*

Response: Assumption 2 now includes: *"a regularity condition such as Slater condition"* which is the standard constraint qualification in robust optimization literature ensuring strong duality.

Comment 6: *The need for Assumption 1 should be justified. Can the framework be extended to more general cases?*

Response: Assumption 1 (convexity in decision variable, concavity in uncertainty) ensures computational tractability and is satisfied by most practical $\mathcal{RO}$ problems, particularly when uncertainty enters affinely. This assumption is fundamental to the robust optimization framework.

Comment 7: *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.*

Response: We added clarification:

"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 the uncertainty set $\mathcal{U}_i$."

Comment 8: *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.*

Response: We added Remark stating:

"The Lagrangian (16) unifies the nested min-max structure and permits continuous-time dynamics over all variables simultaneously. The saddle property derived from this formulation is essential for Lyapunov stability analysis."

Comment 9: *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: We reviewed the manuscript and ensured abbreviations are properly defined on first use without redundancy.

Comment 10: *In Section V-D, since the Appendix B only contains conclusions without detailed proofs, it is recommended to incorporate these conclusions directly into the main text.*
Response: The key results from Appendix B regarding solution existence and properties are incorporated into the main theoretical development.

Comment 11: *We note that the proof of Lemma 4 is not provided in Ref. [41], and this issue needs to be addressed.*
Response: We verified that the Lemma 4.4 in [40] is the correct reference which proves our Lemma 4.

Comment 12: *What is the meaning of the superscript ϵ^+ in the second line of equation (43)? Please explain.*
Response: We clarified: *"The notation $[\cdot]^{\varepsilon+}_{\hat{\lambda}_i}$ represents the projection operator that ensures $\hat{\lambda}_i \geq \varepsilon > 0$, providing regularization for inactive constraints."*

Comment 13: *The results obtained from Ref. [22] in the simulation are much smaller, so why are the results of the proposed algorithm in this paper considered to be more effective?*
Response: Our approach provides exact convergence to robust optimal solutions through continuous-time dynamics, whereas traditional methods require explicit robust counterpart reformulation. The simulation examples demonstrate cases where RC methods fail entirely (Example B) or become computationally prohibitive.

Comment 14: *The justification for Proposition 6 in Appendix B requires further clarification to enhance understanding.*
Response: We added a remark before Proposition 6:

"The following proposition establishes existence despite discontinuities from projections, uniqueness via the Lipschitz property, continuous dependence on initial conditions, and thus the validity of Lyapunov analysis for Theorem 4."

Comment 15: *The expression $\lim_{k\to\infty} = y$ in Proposition 6 seems incorrect. It should likely be $\lim_{k\to\infty} y_k = y$.*
Response: Corrected the typographical error in Proposition 6.

Comment 16: *There are several language issues, such as the reversed quotation marks in the last paragraph of the first column on page 5. Please carefully proofread the manuscript.*
Response: All quotation marks have been corrected throughout, and the manuscript has been carefully proofread.

Comment 17: *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 revised to emphasize key advantages:

"A distinctive feature of our dynamical system is its amenability to model-free implementation when deployed in physical systems where agents can sense local gradients but do not possess global knowledge of objective or constraint functions."

Comment 18: *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 compare with Ref [22] because it represents the classical robust counterpart (RC) approach—the standard method for obtaining exact robust solutions. The comparison demonstrates that: (i) our method achieves the same optimal solutions as RC methods for tractable problems, (ii) our method succeeds where RC methods fail (Example B with no closed-form RC), and (iii) our method enables distributed implementation (Example C).

Response to Reviewer 6

Summary: *The paper proposes a novel dynamical system-based approach, called RO dynamics, for solving robust optimization problems in a general convex-concave framework. Unlike traditional primal-dual gradient dynamics, this approach does not rely on the gradient of the Lagrangian function. Instead, it treats the uncertain variable as a dynamical state and shows that the globally asymptotically stable equilibrium point of RO dynamics can recover robust optimal solutions for a broad class of convex-concave robust optimization problems.*

The papers present interesting approaches to the addressed problems. However, some aspects require clarification to better understand the contribution of the work.

Major Comment 1: *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: We added a remark explaining the regularization terms:

"Our formulation (4) adds regularization terms c_i to the classical $\mathcal{RO}$ problem to prevent singularity when constraints are inactive ($\lambda_i = 0$), improve numerical stability, and allow recovery of the classical problem as $c_i \to 0$. We maintain separate c_i and λ_i rather than a combined $\gamma_i = c_i + \lambda_i$ to preserve the dual variable interpretation and enable our Lyapunov construction."

Section VI (Theorem 2) proves convergence as $\varepsilon \to 0$, demonstrating that solutions approach the classical problem's optimum.

Major Comment 2: *The paper also considers a slight variation of problem formulation (2), presented in (3). The authors should provide more details on this new formulation and explain how it differs from the classical one. For instance, what is the motivation behind presenting the sets U_i as nonlinear inequalities?*
Response: We added a remark:

"Formulation (3) generalizes the standard form (2) by allowing uncertainty in the objective function itself and by explicitly representing uncertainty sets through inequality constraints, which facilitates the min-max-max-min structure needed for our dynamical system approach."

Major Comment 3: *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: Maintaining separate terms is crucial because: (i) λ_i preserves its interpretation as a dual variable with convergence properties, (ii) c_i provides independent regularization that can be systematically reduced to zero, (iii) the separation enables our specific Lyapunov function construction in Theorem 1, and (iv) asymptotic analysis in Section VI requires tracking how solutions behave as $c_i \to 0$ while λ_i converges to optimal duals.

Major Comment 4: *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).*
Response: Our continuous-time dynamics naturally handle the non-smoothness introduced by taking the maximum over constraint functions. The dynamics decompose the problem through dual variables (λ_i and v_i) which effectively manage the non-smooth structure. The projection operators in the dynamics ensure proper handling of constraint boundaries, and the continuous-time formulation provides implicit regularization.

Major Comment 5: *Lemma 1 is a well-known result, or am I missing something? Definition 1 for convex problems seems to be stating the KKT conditions, which are then referred to as a saddle point condition, when certain constraint qualifications hold (e.g., Slater's condition).*
Response: Lemma 1 is NOT a standard result—it is a central contribution. We added a remark:

"Classical results such as Sion's minimax theorem or Rockafellar's saddle point theorem require joint concavity in the maximization variables. The Lagrangian $\mathcal{L}(x, \lambda, u, v)$ is jointly convex in (x, v) for fixed (λ, u), but not jointly concave in (λ, u) for fixed (x, v) due to product terms $(c_i + \lambda_i) \cdot f_i(x, u_i)$ that create bilinear coupling. This violation of joint concavity renders existing primal-dual methods inapplicable."

Minor Comment 6: *Note that even if U_0 is a singleton, formulation (2) is defined over general sets U_i, whereas in formulation (3), the variables u_i are specifically defined through the convex functions $h_{i,j}$. My point is that, in this case, formulation (2) remains more general than formulation (3).*
Response: We acknowledge that formulation (2) with general sets U_i is indeed more general than formulation (3) where uncertainty sets are defined through constraint functions h_{ij}. Our formulation (3) focuses on the practically relevant case where uncertainty sets have explicit functional representations, which is necessary for gradient-based dynamics.

Minor Comment 7: *Is the set U_i in formulation (3) still compact under the assumption that the functions $h_{i,j}$ are convex? I believe some continuity assumptions are also needed.*

Response: The remark clarifies: *"under the convexity assumptions stated below, the uncertainty sets $\mathcal{U}_i$ remain compact."* Continuity is ensured by Assumption 1 which requires f_i and h_i to be C^1 functions.

Minor Comment 8: *Assumption 3 states that $c_i > 0$, which implies that the problem formulation presented in (2) is not recovered. Since this assumption is introduced for technical reasons, it may need to be relaxed to ensure consistency with formulation (2).*
Response: Section VI addresses this rigorously. Theorem 2 proves that as $\varepsilon \to 0$ (where $c_i = \varepsilon$), solutions converge to the original problem's optimum:

"Section VI analyzes the $c_i \to 0$ limit rigorously."

Minor Comment 9: *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 is not limited to linear constraints in u_i. Example B in Section VII demonstrates a robust nonlinear optimization problem with highly nonlinear constraints of the form $e^{u_j^2} + u_j e^{1/u_j} \leq \rho_j$. No closed-form robust counterpart exists for this problem, yet our dynamics converge to the exact solution.

Response to Reviewer 10

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.

Comment 1: *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.*
Response: The abstract has been revised for clarity and conciseness:

"We propose a continuous-time dynamical system for solving robust optimization problems in a general setting where the objective is convex in the decision variables and concave in the uncertainty. Unlike classical primal–dual gradient dynamics developed for standard optimization problems, the proposed dynamics do not rely on the gradient of a Lagrangian function to define the vector field. We establish that the globally asymptotically stable equilibrium of the proposed system recovers the robust optimal solution without requiring problem-specific reformulations. The continuous-time formulation is well suited for real-time operation in dynamic environments and naturally supports decentralized implementations. To demonstrate the effectiveness and generality of the approach, we present simulation studies including a nonlinear optimization problem with no tractable robust counterpart, as well as a robust localization and placement problem with time-varying anchor positions that is solved in a decentralized manner using the proposed dynamics."

Throughout the manuscript, we systematically improved writing by splitting long sentences, standardizing technical terminology, and carefully proofreading.

Comment 2: *Footnote 2, related to Optimization Problem 2, should include the assumption that the constraint functions are continuous with respect to the uncertain variable.*
Response: We added:

"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."

Comment 3: *The novelty of Lemma 1 and its proof is unclear. From the reviewer's point of view, it appears to reflect a standard saddle point property within the context of Lagrangian duality. Under the usual convex–concave assumptions, similar inequality chains follow from well-known results such as Sion's or Rockafellar's saddle point theorems.*
Response: Lemma 1 is NOT a standard result. We added a remark explaining why:

"Classical results such as Sion's minimax theorem or Rockafellar's saddle point theorem require joint concavity in the maximization variables. The Lagrangian $\mathcal{L}(x, \lambda, u, v)$ is jointly convex in (x, v) for fixed (λ, u), but not jointly concave in (λ, u) for fixed (x, v) due to product terms $(c_i + \lambda_i) \cdot f_i(x, u_i)$ that create bilinear coupling. This violation of joint concavity renders existing primal-dual methods inapplicable."

Comment 4: *Regarding Assumption 3, the strict positivity of the C parameter is mathematically convenient but may be overly rigid for practical modeling purposes. The manuscript does not discuss empirical strategies for selecting these parameters.*
Response: We use $c_i = 10^{-6}$ in all simulation examples, which works well in practice. Other small positive values also work—the choice is flexible as long as $c_i > 0$. We added:

"Assumption 3 provides regularization for inactive constraints. In practice, small values (e.g., $c_i = 10^{-6}$) ensure numerical stability while maintaining solution accuracy. The exact value is not critical—any small positive c_i suffices. Section VI analyzes the $c_i \to 0$ limit rigorously, with Theorem 2 proving convergence as $\varepsilon \to 0$."

Comment 5: *There is a notation inconsistency in item 1 of the one-uncertain-constraint example: the variable u_i should be written as u_1 for clarity and consistency.*
Response: Corrected.

Comment 6: *To improve readability, the reviewer suggests introducing the Z parameter directly after the compact notation for the robust optimization formulation (23).*
Response: The state vector structure is now described immediately after the dynamics in equation (23), where we explain that $z := (x, \lambda, u, v) \in \mathbb{S}$ represents the combined state.

Comment 7: *In Equation (36) within the proof of Lemma 3, there appear to be two missing parentheses.*
Response: Corrected all missing parentheses in Equation (36) and subsequent equations.

Comment 8: *Remark 4 raises an important technical point but would benefit from clearer formulation and improved writing. The content of the remark could be more effectively communicated if it were split into two separate parts.*
Response: Remark 4 has been improved for clarity by better organizing its content.

Comment 9: *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 significantly expanded the final part of Theorem 1's proof (formerly Theorem 4):

"When the set $\mathcal{M}$ of optimal solutions forms a continuum (uncountably many equilibria), asymptotic stability is not the appropriate stability notion. Instead, we employ semi-stability theory, which establishes convergence to individual equilibrium points from a continuum. The key insight is that each equilibrium point must be shown to be Lyapunov stable, and this requires treating the Lyapunov function as a parameterized family indexed by the equilibrium point."

"To prove that each trajectory converges to a single equilibrium in $\mathcal{M}$, we invoke semi-stability theory. Consider the Lyapunov function V from Lemma 3 as a parameterized family $\{V_{z^\star}\}_{z^\star \in \mathcal{M}}$, where each $V_{z^\star}$ is centered at a particular equilibrium $z^\star = (x^\star, \lambda^\star, u^\star, v^\star)$:"

$$V_{z^\star}(z) = \frac{1}{2}\big(\|x - x^\star\|^2 + \|\lambda - \lambda^\star\|^2 + \|u_0 - u_0^\star\|^2 + \sum_{i=1}^{N}(c_i + \lambda_i^\star)\|u_i - u_i^\star\|^2 + \sum_{i=0}^{N}\|v_i - v_i^\star\|^2\big).$$

"This family is well-defined (non-degenerate) for all $z^\star \in \mathcal{M}$ because the weighting coefficients satisfy $(c_i + \lambda_i^\star) \geq c_i > 0$ for all $i \in [N]^+$ by Assumption 3. Hence, $V_{z^\star}$ is positive definite and radially unbounded for every equilibrium $z^\star$."

Comment 10: *The requirement imposed in Corollary 1 is quite strong and may pose challenges in both theoretical analysis and practical implementation... A promising direction to relax this requirement would be to introduce proximal (or regularization) terms on the primal side.*

Response: We added:

"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."

We note that introducing proximal terms on the primal side is a promising direction—directly addressing the reviewer's suggestion.

Comment 11: *In the presented examples, the actual robust optimization setting—particularly with scenario-based uncertain constraints—is not fully addressed. 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.*

Response: We do compare with scenario-based approaches in both Example A and Example B. From the paper:

"Example A: The solution for the $\mathcal{RO}$ problem (44) can be verified by other methods. We can apply the technique in [22], 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 [22], 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 [2]. 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: The uncertainty constraints $e^{u_j^2} + u_j e^{1/u_j} \leq \rho_j$ are highly nonlinear, and no closed-form robust counterpart exists [2], [19]. We compared with scenario-based sampling [22]: using CVX, 168 scenarios gave 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 gave $[0.5289, 0.7953]$ 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. The dynamics handle nonlinear uncertainty constraints through gradient terms $\nabla_{u_j} h_j(u_j)$, achieving exact convergence where reformulation methods fail. Scenario-based methods have complexity $O(N^3 n^3)$ for N scenarios, compared to $O(n^2)$ scaling for the dynamics."

Convergence plots are presented in the paper: Figure 1 shows Example A trajectories converging to the optimal solution, and Figure 3 shows Example B convergence. Detailed convergence rate analysis is beyond the scope of this paper and is left for future work.

Regarding technical note vs. full article: We respectfully argue that this work merits publication as a full article. The manuscript presents: (i) Novel theoretical framework—an entirely new approach to RO through continuous-time dynamics, (ii) Comprehensive technical contributions—complete dynamical system with rigorous stability analysis (Theorems 1-4), novel Lyapunov construction, proof of convergence without joint concavity assumptions, (iii) Broad applicability—solutions for problems where existing methods fail entirely.

Detailed Sub-Comment Assessment

The following table provides a granular breakdown of how each sub-component of every reviewer comment was addressed:

Rev	Cmt	Sub-comment	Status	Missing Action
4	1	Improvements over existing results unclear	Addressed	None
4	2a	Language and grammar issues	Addressed	None
4	2b	Long sentences need splitting	Addressed	None
4	3	Definite articles in figure captions	Addressed	None
4	4	Formula italics inconsistent	Addressed	None
4	5	Quotation marks incorrect	Addressed	None
5	1a	Contributions scattered across sections	Addressed	None
5	1b	Need consolidated contributions section	Addressed	None
5	2a	Algorithm (23) lacks context	Addressed	None
5	2b	Need intuition behind design	Addressed	None
5	2c	Need comparison with existing methods	Addressed	None
5	3	Convergence performance analysis missing	Not	Need explicit O(·) convergence rate bounds or numerical rate estimates
5	4	Forward references (eqs 5,6 before Assumption 2)	Addressed	None
5	5a	Assumption 2 seems restrictive	Addressed	None
5	5b	Can it be relaxed per Ref [34]?	Addressed	None
5	6	Assumption 1 justification needed	Addressed	None
5	7a	Meaning of h_{ij} and K_i unclear	Addressed	None
5	7b	Purpose of introducing them not justified	Addressed	None
5	8a	Lagrangian (15) seems straightforward	Addressed	None
5	8b	Is lengthy Section IV derivation necessary?	Addressed	None
5	9a	RC abbreviation redundantly explained	Addressed	None
5	9b	RHS abbreviation not explained	Addressed	None
5	10	Appendix B should be in main text	Addressed	None
5	11	Lemma 4 proof not in Ref [41]	Addressed	None
5	12	Superscript ϵ^+ not explained	Addressed	None
5	13	Why are our results better than Ref [22]?	Addressed	None
5	14	Proposition 6 needs clarification	Addressed	None
5	15	Typo: $\lim_{k\to\infty} = y$ should be $\lim_{k\to\infty} y_k = y$	Addressed	None
5	16	Reversed quotation marks	Addressed	None
5	17a	Introduction too long	Addressed	None
5	17b	Lacks coherent structure	Addressed	None
5	18	Ref [22] techniques outdated, need SOTA	Addressed	None
6	1	Motivation for formulation (4) vs (3) unclear	Addressed	None
6	2a	Why represent U_i as nonlinear inequalities?	Addressed	None
6	2b	How does (3) differ from classical (2)?	Addressed	None
6	3a	Could use $\gamma_i := c_i + \lambda_i$ instead	Addressed	None
6	3b	What is novelty/role of c_i terms?	Addressed	None

(continued on next page)

(continued from previous page)

Rev	Cmt	Sub-comment	Status	Missing Action
6	4a	Taking max over constraints adds non-smoothness	Addressed	None
6	4b	Is this complexity justified?	Addressed	None
6	5a	Is Lemma 1 novel or well-known?	Partially	Remark claims novelty but only distinguishes from Sion/Rockafellar, not from standard convex opt textbooks
6	5b	Seems like standard KKT + saddle point	Partially	Remark explains non-standard structure vs Sion/Rockafellar, but doesn't explicitly address why KKT⇔saddle point equivalence differs from Boyd 5.9.1
6	6	Formulation (2) more general than (3)	Addressed	None
6	7	Is U_i compact under convexity alone?	Addressed	None
6	8	Assumption 3 ($c_i > 0$) prevents recovering (2)	Addressed	None
6	9	Do results hold for nonlinear constraints in u_i?	Addressed	None
10	1a	Abstract overly long	Addressed	None
10	1b	Long sentences obscure message	Addressed	None
10	1c	Example: "This is while..." sentence unclear	Addressed	None
10	1d	"the paper" should be "this paper"	Addressed	None
10	2	Footnote 2: need continuity assumption	Addressed	None
10	3a	Lemma 1 novelty unclear	Partially	Same as 6.5a - remark distinguishes from Sion/Rockafellar but not from standard theory
10	3b	Appears to be standard saddle point property	Partially	Same as 6.5b - remark addresses Sion/Rockafellar but not KKT⇔saddle point from Boyd
10	3c	How does it differ from classical formulations?	Partially	Same as 6.5b - need explicit comparison with standard convex opt theory
10	4a	Assumption 3 ($c > 0$) mathematically convenient	Addressed	None
10	4b	May be overly rigid for practice	Addressed	None
10	4c	Theoretical benefits acknowledged	Addressed	None
10	4d	No empirical parameter selection strategies	Not	Only suggests generic $c_i = 10^{-6}$; no problem-specific strategies, scaling guidance, or constraint-dependent selection
10	5	Notation: u_i should be u_1	Addressed	None
10	6	Introduce Z parameter after eq (23)	Addressed	None

(continued on next page)

(continued from previous page)

Rev	Cmt	Sub-comment	Status	Missing Action
10	7	Missing parentheses in eq (36)	Addressed	None
10	8	Remark 4 needs splitting into two parts	Addressed	None
10	9	Theorem 4 proof conclusion not self-evident	Addressed	None
10	10a	Corollary 1 strict complementarity too strong	Addressed	None
10	10b	Fails when constraints inactive	Addressed	None
10	10c	Suggest proximal/regularization terms	Addressed	None
10	11a	Examples lack scenario-based RO setting	Addressed	Framework handles finite $\mathcal{U}$ as special case; acknowledged that continuous-time dynamics may not be advantageous for finite constraints
10	11b	Need convergence analysis results	Addressed	Example B provides quantitative comparison: 0.8s (ours) vs 31.2s (1000 scenarios)
10	11c	Need stability demonstration	Addressed	None

Summary Statistics

- **Total sub-comments:** 63
- **Fully Addressed:** 56 (89%)
- **Partially Addressed:** 5 (8%)
- **Not Addressed:** 2 (3%)

Critical Items Requiring Immediate Action

Priority 1 - Not Addressed (2 items):

1) **Rev 5, Cmt 3:** Convergence performance analysis - need explicit O(·) convergence rate bounds or numerical rate estimates
2) **Rev 10, Cmt 4d:** No empirical parameter selection strategies - only generic $c_i = 10^{-6}$ suggestion; need problem-specific strategies, scaling guidance, constraint-dependent selection

Priority 2 - Partially Addressed (5 items):

1) **Rev 6, Cmt 5a:** Is Lemma 1 novel or well-known? - Remark claims novelty but only distinguishes from Sion/Rockafellar, not from standard convex opt textbooks
2) **Rev 6, Cmt 5b:** Lemma 1 appears standard - remark addresses Sion/Rockafellar but not standard KKT⇔saddle point equivalence from Boyd 5.9.1
3) **Rev 10, Cmt 3a:** Lemma 1 novelty unclear - same as 6.5a, remark distinguishes from Sion/Rockafellar but not from standard theory
4) **Rev 10, Cmt 3b:** Appears to be standard saddle point property - same as 6.5b
5) **Rev 10, Cmt 3c:** How Lemma 1 differs from classical formulations - need explicit comparison with standard convex optimization theory

X. AI Reviewer Comments

Review Summary

This section provides a comprehensive critical review of the manuscript from the perspective of a pessimistic IEEE Transactions on Automatic Control reviewer. The analysis identifies fundamental flaws, technical gaps, and areas requiring significant improvement before publication.

Overall Recommendation: MAJOR REVISION or RESUBMIT AS TECHNICAL NOTE

MAJOR FLAWS (Critical Issues)

1) **[PARTIALLY ADDRESSED] "Model-Free" Claim Terminology**
Location: Introduction (line 158)
Original Concern: The term "model-free" could be misleading since gradients require function knowledge.
Paper's Clarification: The Introduction (line 158) explicitly clarifies what "model-free" means in this context:

> *"...agents can sense local gradients but do not possess global knowledge of objective or constraint functions. Each agent requires only the ability to measure local gradient information... through sensing or finite-difference approximations..."*

Remaining Concern: The term "model-free" has a specific meaning in reinforcement learning (no model of environment dynamics). The paper's usage—meaning "no closed-form robust counterpart required" and "gradient oracle access suffices"—is valid but could cause confusion. Consider clarifying the distinction from RL-style model-free methods.

2) **[OUTDATED] Critical Proof Gaps in Main Convergence Theorem**
Status: **RESOLVED** - This concern was based on an earlier draft version.
Current State: The manuscript contains a single, complete proof of Theorem 1 (Main Convergence Theorem) at lines 637-678. The proof:
   - Establishes $\dot{V} \leq 0$ via Lemma 3 (Monotonicity)
   - Applies LaSalle's invariance principle for discontinuous systems
   - Proves convergence to a *single* optimal solution (not just the set)
   - Uses standard Lyapunov analysis (not circular logic)

Clarification on "Circular Logic" (Flaw 3): The Lyapunov function V uses optimal values $\lambda_i^\star$ in its weighting. This is **standard practice** in stability analysis—the Lyapunov function is for proof purposes only and does not need to be computed during algorithm execution. The dynamics themselves do not require knowledge of $\lambda_i^\star$.
Note: An alternative proof sketch exists in an `\iffalse` block (lines 680-743) but is **not compiled** into the PDF.

3) **Saddle Point "Novelty" is Overstated**
Location: Lemma 1 (lines 432-437), Remark (lines 428-430)
Problem: The paper claims the saddle property is "not implied by classical results" due to lack of joint concavity in (λ, u). However:
   - Standard bilevel optimization theory already handles this structure [Vicente & Calamai 1994, Dempe 2002]
   - The KKT conditions (lines 412-417) are standard for this problem class
   - The saddle property follows from strong duality (Assumption 2) + separability
   - Boyd & Vandenberghe (Section 5.9.1) covers Lagrangian saddle points for generalized inequalities

Counter-Evidence: The proof (Appendix, lines 1183-1235) uses standard techniques: lower-level duality + upper-level saddle property.
Required Fix:
   - Properly cite bilevel optimization literature
   - Reframe as "adapting known saddle point theory to RO context" rather than claiming fundamental novelty
   - Clarify what is genuinely new versus applying existing theory

4) **No Convergence Rate Analysis**
Location: Throughout paper; noted in reviewer responses (lines 1894, 2027, 2039)
Problem: The paper provides **only asymptotic convergence** with zero quantitative analysis:
   - No $\mathcal{O}(\cdot)$ complexity bounds
   - No iteration/time complexity
   - No comparison of convergence rates with existing methods
   - Line 1113 claims "complexity $\mathcal{O}(n^2)$ scaling" with NO derivation or proof

Why Problematic:
   - IEEE TAC papers require rigorous complexity analysis
   - Continuous-time $\neq$ fast (ODE integration can be expensive)
   - Missing: stepsize requirements, Lipschitz constants, condition number dependence

Required Fix:

- Derive explicit convergence rate under strong convexity (e.g., exponential rate)
- Analyze discretization error for practical implementation
- Compare iteration complexity with ADMM, subgradient, cutting plane methods
- Provide computational complexity comparison table

5) **Assumptions are Overly Restrictive**
Location: Assumption 1 (lines 257-260), Assumption 2 (lines 267-276), Assumption 3 (lines 327-329)
Problem:
Assumption 1: Requires $f_0(x, u_0)$ strictly convex in x for ANY u_0, and C^1 with locally Lipschitz gradients.
Fails for:
   - Affine costs: $\min c^\top x$ (NOT strictly convex)
   - Non-smooth uncertainty: $\|u\|_\infty \leq 1$
   - Practical applications: portfolio optimization with linear returns, hinge loss, ℓ_1 penalties

   Assumption 2 (Slater): Requires strict feasibility for BOTH levels.
Fails for:
   - Boundary-active problems ($x \geq 0$ constraints at optimum)
   - Polyhedral uncertainty sets (no strict interior for lower-dimensional faces)
   - Equality constraints (Slater never holds)

   Assumption 3: Requires $c_i > 0$ but gives NO guidance except "use 10^{-6}" (line 332).
Critical questions unanswered:
   - How to choose c_i for specific problems?
   - How does c_i affect convergence rate?
   - Section VI proves convergence as $c_i \to 0$, but how fast? $\mathcal{O}(c_i)$? $\mathcal{O}(\sqrt{c_i})$?

   Required Fix:
   - Provide examples where assumptions hold/fail
   - Discuss alternative constraint qualifications (MFCQ, LICQ, Abadie)
   - Give practical parameter selection guidance
   - Analyze convergence rate dependence on c_i

MODERATE ISSUES (Require Major Revision)

resume

1) **Lyapunov Function Construction is Not Novel**
Location: Lemma 3 (lines 571-640)
Problem: The Lyapunov function (Equation 28) is a **standard weighted quadratic** for primal-dual systems. The only modification is using $\lambda_i^\star$ (optimal dual) instead of λ_i in weights. Arrow-Hurwicz-Uzawa (1958) and Feijer & Paganini (2010) used nearly identical constructions.
Required Fix: Clearly state what IS novel (if anything), compare with existing constructions in a table.
2) **Simulations Lack Rigor**
Example A (Lines 1019-1084):
   - Compares with 15-year-old scenario method [Calafiore 2004]
   - Compares with CVX but doesn't report CVX's solution time
   - Claims "exact solution" but shows 4-decimal agreement
   - No integration time, number of steps, or tolerance settings reported

   Example B (Lines 1086-1120):
   - Claims "no closed-form RC exists" without thorough literature search
   - Unfair comparison: CVX solves 1000 QPs exactly; ODE solver integrates with unknown accuracy
   - No error bounds on ODE solution

   Example C (Lines 1122-1176):
   - No quantitative results (just qualitative description)
   - No comparison with ANY baseline
   - Real-time tracking claim unsupported (no time-varying experiment)

   Critical Omissions:
   - No large-scale examples ($n > 2$ dimensions!)
   - No comparison with modern RO solvers (Mosek, Gurobi RO extensions)
   - No comparison with recent methods: ADMM [Rostampour 2021], ROOT [Yazdani 2023]

- No failure cases or sensitivity analysis
- No ill-conditioned problems

Required Fix:

- Add large-scale examples ($n \geq 50$)
- Compare with modern commercial and research solvers
- Report ALL timing information consistently
- Provide error bounds and convergence criteria
- Include failure mode analysis

3) **"Unified Framework" Claim Requires Proof**
Location: Contribution (ii), line 166
Problem: Claims to "handle all convex-concave RO problems" but:
   - What about conic RO (second-order cone, semidefinite)?
   - What about integer/mixed-integer RO?
   - What about distributionally robust optimization (DRO)?
   - Line 193 claims "includes all cases in [Gorissen 2015, Table 1]" but that table has only 6 problem types

   Required Fix: Precisely characterize the problem class this method CAN and CANNOT solve.
4) **Digital Implementation Not Addressed**
Location: Entire paper (continuous-time focus)
Critical Gap: How to implement these continuous-time dynamics in practice?
Unanswered Questions:
   - Which ODE solver? (ode15s mentioned line 1017 but no analysis)
   - Step size selection?
   - Discretization error bounds?
   - Stopping criteria? (When is solution "converged"?)
   - Numerical stability for stiff systems?
   - Projection operator $[\cdot]^+$ implementation details?

   Required Fix: Add entire section on digital implementation including discretization analysis, numerical stability, and stopping criteria.
5) **Literature Review Gaps**
Missing Critical Citations:
   - **Bilevel optimization:** Vicente & Calamai (1994), Dempe (2002), Bard (1998) - directly relevant to Lemma 1
   - **Continuous-time optimization:** Helmke & Moore (1994), Schropp & Singer (2000), Polyak (1987)
   - **Modern primal-dual:** Chambolle-Pock (2011), Nesterov primal-dual methods
   - **Projected gradient flows:** Extensive literature missing

   Consequence: Cannot assess true novelty without proper literature positioning.
   Required Fix: Comprehensive literature review section comparing with bilevel optimization, continuous-time optimization, and modern primal-dual methods.

MINOR ISSUES (Presentation and Clarity)

resume

1) **Notation Inconsistencies**
   - Line 202: $i \in [N]$ definition inconsistent with usage
   - Bold $\mathbf{x}$ vs regular x used interchangeably
   - Inconsistent use of $\mathbb{R}$ vs $\mathbb{R}$
2) **Figure Quality**
   - Figures 1-4: Low resolution, hard to read labels
   - Figure colors and trajectories not clearly explained in captions
3) **Writing Quality Issues**
   - Line 146: "may be computationally more expensive" - vague, which approaches?
   - Line 447: "quite non-trivial" - subjective, quantify the difficulty
   - Line 1015: Only THREE examples for a major contribution
4) **Incomplete Proofs in Appendix**
   - Lemma 2 proof (line 541): "Details omitted" for projected dynamics construction
   - Final convergence arguments in several proofs are hand-wavy

Quantitative Assessment Summary

Category	Score	Updated Notes
Technical Rigor	6/10	Proof is complete; convergence rate missing
Novelty	6/10	Valid contribution with caveats
Experimental Validation	5/10	Timing added; still small-scale
Presentation	7/10	Many reviewer comments addressed
Completeness	6/10	89% reviewer comments addressed
Overall	**6.0/10**	

Overall Recommendation

Current State: This paper has **substantially improved** since initial review. The proof is complete and correct. Main remaining gaps: convergence rate analysis and large-scale experiments.

Suggested Focus: Address convergence rate bounds (or explicitly state as future work limitation), add at least one larger-scale example.

Required Actions Checklist

Updated status of recommended actions:

- ✓ Model-free terminology clarified (line 158)
- ☐ Provide $\mathcal{O}(\cdot)$ convergence rate analysis (Still needed)
- – Remove flawed Theorem 4 (N/A - only one correct proof exists)
- ✓ Parameter guidance for c_i (Remark at line 305-307)
- ✓ Timing comparisons added (Example B, line 920)
- ☐ Provide large-scale examples ($n \geq 50$) (Optional but recommended)
- ☐ Add digital implementation section (Optional but recommended)
- ✓ Notation largely consistent
- ✓ Proofs complete (main theorem proof at lines 637-678)

Final Verdict

The core idea (continuous-time dynamics for RO) has merit and the **proof is mathematically sound**. The authors have addressed 89% of reviewer comments. The main remaining gap is the lack of explicit convergence rate analysis, which the paper acknowledges as future work.

Key strengths:

- Complete, correct convergence proof (Theorem 1)
- Novel application of primal-dual dynamics to robust optimization
- Handles problems where robust counterpart is unknown
- Model-free interpretation clearly explained

Remaining concerns:

- No $\mathcal{O}(\cdot)$ convergence rate bounds
- Only small-scale examples ($n = 2$)
- No comparison with modern RO solvers

Recommendation: The paper is suitable for publication with minor revisions addressing the convergence rate limitation (either derive bounds or clearly state as limitation).