Distributionally Robust Optimization Techniques for Stochastic Optimal Control

Chun Man Oswin So

In Partial Fulfillment of the
Requirements for the Degree of
Bachelor of Science in Computer Science
at the College of Computing

Georgia Institute of Technology
December 2021

Distributionally Robust Optimization Techniques for Stochastic Optimal Control

Approved: DocuSigned by: Evangelos Theodorou C5C98A0325B9474...
____
Dr. Evangelos Theodorou

Approved: DocuSigned by: Jim Rehg 633C5D0B85E442E...
____
Dr. James M. Rehg

Abstract

Distributionally robust optimal control is a relatively new field of robust control that tries to address the issue of safety by hedging against the worst-cast distributions. However, because probability distributions are infinite-dimensional, this problem is in general computationally intractable. This thesis provides an overview of applications of distributionally robust optimization for stochastic optimal control. In particular, we look at existing and potentially new computationally tractable methods for performing distributionally robust optimal control using the Wasserstein metric.

Contents

Chapter 1

Introduction

Safety is an important issue for controls and planning algorithms in situations where failure could be catastrophic. An important part of being able to plan safely is being able to take into disturbances from the environment and model uncertainty into account, and has been a central area of study in stochastic optimal control [1]. Most of the literature on this topic makes the assumption that the disturbances (whether in the form of control noise or parametric uncertainty) are Gaussian. However, it is often hard to accurately determine the exact distributions, and making wrong assumptions about the disturbance distribution can be worse than not modeling the disturbances at all (eg. [2]).

Distributionally Robust Optimization (DRO), also known as ambiguous stochastic optimization, is a relatively new field that addresses this problem by assuming that the underlying probability distribution lies in an *ambiguity set* of probability distributions [3], and attempts to hedge against this ambiguity by planning for the worst case scenario.

There are many choices for the describing the ambiguity set, though this can mainly be split into two groups: *moment-based*, and *discrepancy-based* approaches. Early DRO models defined ambiguity sets using information about the moments of the distribution [4], and these approaches have been applied to optimal control [5]. Discrepancy-based approaches define the ambiguity set as some sort of neighbourhood or ball around a *nominal* or *baseline* estimate of the underlying probability distribution, usually from data. A popular measure is the family of ϕ-divergences [6] , which contains divergences such as the popular Kullback-Leibler(KL) divergence and χ^2 divergence. Another measure that has come under much interest recently, especially in operations research and machine learning communities is the Wasserstein or Optimal Transport metric [7], applied to optimal control in [8][9][10][11]. In contrast to the information-theoretic nature of the ϕ-divergences, the Wasserstein metric is a proper metric on the space of probability distributions and takes into account the structure of underlying metric space. This advantage of the Wasserstein metric over the ϕ-divergences was shown in [12], where the Wasserstein metric was able to greatly improve the stability of training.

In this thesis, we will present an overview of various existing applications of distributionally robust control for stochastic optimal control. In particular, we will explore the Wasserstein metric in depth, and look at current existing and potential new *computationally tractable* methods of applying the Wasserstein metric.

Chapter 2

Related Works

There is a wealth of information on the fields of DRO, stochastic optimal control and optimal transport separately, but the only recently has the optimal control community applied DRO techniques to optimal control.

The earliest approaches to DRO focused on moment matching of the underlying disturbance distribution. In [13], the authors propose a general approach to deriving robust solutions for stochastic optimization programs based on mean-covariance information. In [14], the authors develop tractable semidefinite programming based approximations for distributionally robust chance constraints with mean and covariance matching and show that joint chance constraints can be approximated by worst-case Conditional Value-at-Risk (CVaR) constraints.

However, moment matching approaches assume that only the moments are known exactly and nothing else relevant about the underlying distribution is known. As it is often the case that one has some samples of observations of the disturbance, moment matching based approaches can yield overly conservative results. In [15], it is shown that in the newsvendor problem, by hedging against an unlikely worst-case distribution, the solution does not perform well in other more likely distributions.

To improve on this flaw of moment matching approaches, statistical-distance approaches accounts for distributions that are close (with respect to the statistical-distance) to some empirical distribution. In [16], an ambiguity set defined using the ϕ-divergences is used to formulate stochastic programs with distributionally robust chance constraint, and the structure of the corresponding worst-case distribution is derived. In [17], the ϕ-divergences are applied to distributionally robust two stage stochastic programs, and a decomposition-based solution algorithm is proposed.

Although ϕ-divergences are very popular in a large number of fields, there are a number of problems that arise when using them for distributionally robust optimization. For example, if the KL-Divergence is used, only distributions that have support on points where the nominal distribution also has support are considered. Another fundamental problem with the ϕ-divergences are that they do not consider the underlying topology of the space and hence cannot account for distances between the two distributions. The authors in [18] provide an example of this by showing that an ambiguity set defined using the KL-divergence around a histogram of an image perturbed by a low contrast transformation could potentially exclude the true image but not a pathological image that is visually very different from the perturbed image.

To solve this issue, other statistical-distances that can better incorporate a notion of the distance between two distributions have been considered. The *Wasserstein distance*, originally from the field of optimal transport, has risen in popularity the past few years as one such statistical-distance which does take into account the underlying metric space of distributions. The field of modern optimal transport, which essentially began with Kantorovich's work on using linear programming into tackle the original problem proposed by Monge, is quite old. However, only recently has other communities such as the stochastic optimization community and machine learning communities begun to incorporate this into their fields.

In [7], the Wasserstein-1 metric is used to define an ambiguity set for a DRO problem, and the infinite-dimensional problem is reformulated as a finite convex program under assumptions of the convexity of the underlying space. Similarly, in [19], the Wasserstein-1 metric is used to define a two-stage stochastic optimization problem, and the worst-case distribution is derived. In [18], the authors relax the assumptions on the order of the Wasserstein distance and various other assumptions, identify necessary and sufficient conditions for the existence of the worst-case distribution and show that data-driven DRO problem can be approximated by robust optimization problems.

The recent techniques in DRO have only recently been applied to problems in optimal control. In [20], Yang examines the problem of distributionally robust markov decision processes using the Wasserstein distance, and uses the Kantorovich duality to transform the problem into a tractable finite-dimensional problem. In [10], Yang provides a similar reformulation of the infinite dimensional problem and converts it into a semi-infinite problem which can be solved with algorithms such as discretization and sampling-based methods. In [9], Yang applies the above algorithm to the domain of wind power ramp management, where he is able to exploit the piecewise-linear structure of the cost function to convert the semi-infinite program into a linear program.

In a separate, but closely aligned direction, the authors in [11] propose a risk-aware motion control scheme with the constraints based on Conditional Value-at-Risk (CVaR) of an ambiguity set defined by a Wasserstein ball around empirical measurements. To resolve the infinite-dimensional issue, the authors combine the Kantorovich dual formulation from [18] with an expression for the upper bound on the CVaR to reduce the infinite-dimensional problem to a finite-dimensional nonlinear program.

Chapter 3

Methods

3.1 Problem Formulation and Motivation

Consider the problem of solving for the solution x in the following general stochastic program:

$$J^* := \inf_{u \in \mathbb{U}} \left\{ \int_{\Xi} h(u, \xi) \, \mathrm{d}\mathbb{P}(\xi) \right\} \tag{3.1}$$

with feasible set $\mathbb{U} \subset \mathbb{R}^n$, uncertainty set $\Xi \subset \mathbb{R}^m$ and objective function $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ for some unknown probability measure $\mathbb{P} \in P(\mathcal{X})$. Because the distribution $\mathbb{P}$ is not usually precisely known, it is not possible to solve (3.1) exactly. However, we are often able to observe a finite number of realizations $\{\hat{\xi}^{(1)}, \ldots, \hat{\xi}^{(N)}\}$ of $\mathbb{P}$. Thus, we can construct an empirical distribution *approximation* of $\mathbb{P}$:

$$\hat{\mathbb{P}}_N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\hat{\xi}^{(i)}} \tag{3.2}$$

where $\delta_{\hat{\xi}^{(i)}}$ is the Dirac delta measure concentrated at the sample $\hat{\xi}^{(i)}$.

Then, one can solve the original stochastic program (3.1) with the Sample-Average Approximation (SAA) problem:

$$\begin{aligned} J_{\mathrm{SAA}} &:= \inf_{u \in \mathbb{U}} \left\{ \int_{\mathcal{X}} h(u, \xi) \, \mathrm{d}\hat{\mathbb{P}}(\xi) \right\} \\ &:= \inf_{u \in \mathbb{U}} \left\{ \frac{1}{N} \sum_{i=1}^{N} h(u, \hat{\xi}^{(i)}) \right\} \end{aligned} \tag{3.3}$$

As the number of samples N tends to infinity, the empirical distribution $\hat{\mathbb{P}}$ will approach the true distribution $\mathbb{P}$, and the optimal value and optimal solution of the SAA problem (3.3) will converge almost surely to the original stochastic program (3.1) [21].

However, it is not always possible to obtain a large number of samples N in situations where data collection may be expensive or impossible, and the performance of SAA on out-of-sample situations due to a small N tends to be poor. Furthermore, the empirical distribution $\hat{\mathbb{P}}$ may be biased and not reflect the true distribution $\mathbb{P}$. To resolve theses issues, we can formulate (3.1) instead as a *distributionally robust* optimization which hedges against the worst case distribution in some ambiguity set

$\mathbb{B}$ designed to characterize the samples in empirical distribution $\hat{\mathbb{P}}$. This way, we can define the certificate $\hat{J}_N$ as the optimal value of the following problem which minimizes the *worst-case* expected cost:

$$\hat{J}_N := \inf_{u \in \mathbb{U}} \sup_{\mathbb{Q} \in \mathbb{B}} \left\{ \int_{\mathcal{X}} h(u, \xi) \, \mathrm{d}\hat{\mathbb{Q}}(\xi) \right\} \tag{3.4}$$

As mentioned in the introduction, there are many different ways in which the ambiguity set $\mathbb{B}$ can be defined. Due to the beneficial properties of the Wasserstein ambiguity set, we will focus on this method in the sections below.

3.2 Methods for DRO with the Wasserstein Ambiguity Set

3.2.1 Introduction and Definitions

Here, we give an overview of methods for where the ambiguity set is described using the Wasserstein metric.

Let $\mathcal{X}$ and $\mathcal{Y}$ be some Polish space (complete, separable metric spaces), and denote the space of Borel probability measures on $\mathcal{X}$ and $\mathcal{Y}$ as $P(\mathcal{X})$ and $P(\mathcal{Y})$ respectively. Let $\Pi(\mu, \nu)$ be the set of all joint probability measures on $\mathcal{X} \times \mathcal{Y}$ whose marginals on $\mathcal{X}$ and $\mathcal{Y}$ are μ and ν. Let $\mathrm{law}(X)$ denote the law of a random variable X defined on some probability space $(\Omega, \mathcal{F}, \mathcal{P})$, equivalent to $X_{\#}\mathcal{P}$. We will define the p-Wasserstein distance as follows:

Definition 1. Let $(\mathcal{X}, d)$ be a Polish space (i.e. a complete, separable metric space), and let $p \in [0, \infty)$. For any two probability measures μ, ν on $\mathcal{X}$, the p-Wasserstein distance between μ and ν is defined by

$$\begin{aligned} W_p(\mu, \nu) &= \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X}} d(x, y)^p \, \mathrm{d}\pi(x, y) \right)^{\frac{1}{p}} \\ &= \inf_{X,Y} \left\{ \mathbb{E}[d(X, Y))^p]^{\frac{1}{p}}, \quad \mathrm{law}(X) = \mu, \quad \mathrm{law}(Y) = \nu \right\} \end{aligned} \tag{3.5}$$

$\pi(x, y)$ can be viewed as a *transport plan* between the measures μ and ν. Thus, finding the optimal π can be viewed as finding the optimal transport plan to transport probability mass from the measures μ to ν, minimizing the cost $d(x, y)^p$ that is incurred from moving from x to y.

The dual problem of the above (sometimes called the *Monge* problem) is the Kantorovich-dual, which is an important property that we will use later on.

Theorem 2 (Kantorovich-Rubinstein distance)**.** *For any μ, ν in $P(\mathcal{X})$ such that $\int_{\mathcal{X}} d(x_0, x)\mu(\mathrm{d}x)$ is finite for any $x_0 \in \mathcal{X}$,*

$$W_1(\mu, \nu) = \sup_{\|\phi\|_{Lip} \leq 1} \left\{ \int_{\mathcal{X}} \phi(y) \, \mathrm{d}\mu(y) - \int_{\mathcal{X}} \phi(x) \, \mathrm{d}\nu(x) \right\} \tag{3.6}$$

For a detailed proof of Theorem 2, see Villani [22, p. 107]. The dual problem can be interpreted as two probability measures μ, ν being close to each other if and only if all 1-Lipschitz functions ϕ have similar values when integrated under μ and ν.

We will now use the 1-Wasserstein distance to define the ambiguity set $\mathbb{B}_{W_1}$ for some probability measure μ:

$$\mathbb{B}_{W_p}(\mu, \epsilon) := \{\nu \in P(X) : W_p(\mu, \nu) \leq \epsilon\} \tag{3.7}$$

where $\epsilon > 0$ defines the radius of the ball of the ambiguity set $\mathbb{B}_{W_1}$. In words, this is the set of all probability measures ν that are no more than ϵ away from ν under the 1-Wasserstein distance.

Under this definition of the ambiguity, the DRO problem (3.4) now has the form

$$\hat{J}_N := \inf_{u \in \mathbb{U}} \sup_{\mathbb{Q} \in \mathbb{B}(\hat{\mathbb{P}}_N, \epsilon)} \left\{ \int_{\Xi} h(u, \xi) \, \mathrm{d}\mathbb{Q}(\xi) \right\} \tag{3.8}$$

3.2.2 Solving the Wasserstein DRO problem

However, one of the biggest issues with (3.8) is that this is an infinite dimensional minimax optimization problem — the inner maximization problem is done over all probability measures $\mathbb{Q}$ inside the Wasserstein ball $B(\hat{\mathbb{P}}_N, \epsilon)$, which is an infinite dimensional space. Solving such an infinite dimensional problem is generally computationally intractable. In order to resolve this, we need to make a few assumptions on the problem structure (3.4):

Assumption 3 (Convexity of Ξ)**.** *The uncertainty set $\Xi \subset \mathbb{R}^m$ is convex and closed*

Assumption 4 (Decomposition of objective function h)**.** *The objective function $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ can be decomposed into the pointwise maximum of K elementary measurable functions $l_k : \mathbb{R}^m \to \mathbb{R}, \quad k \leq K$, i.e.,*

$$h(u, \xi) = \max_{k \leq K} l_k^u(\xi) \tag{3.9}$$

such that each of the negative elementary functions l_k are proper, convex and lower semicontinuous for all $k \leq K$. Furthermore, we assume that l_k is not identically $-\infty$ on Ξ for all $k \leq K$.

With these assumptions, we now have the following theorem for reformulating (3.8) as a finite-dimensional convex program [7, Theorem 4.2]:

Theorem 5 (Convex reduction)**.** *Suppose that the assumptions Assumption 3 and Assumption 4 hold. Then, for any $\epsilon \geq 0$, the DRO problem is equal to the optimal value of the following finite convex program:*

$$\inf_{\lambda, s_i, z_{ik}, \nu_{ik}} \quad \lambda\epsilon + \frac{1}{N}\sum_{i=1}^{N} s_i \tag{3.10}$$

$$s.t. \quad [-l_k]^*(z_{ik} - \nu_{ik}) + \sigma_{\Xi}(\nu_{ik}) - \langle z_{ik}, \hat{\xi}_i \rangle \leq s_i \tag{3.11}$$

$$\|z_{ik}\|_* \leq \lambda \tag{3.12}$$

where $[-l_k]^*(z_{ik} - \nu_{ik})$ denotes the conjugate of $-l_k$ evaluated at $z_{ik} - \nu_{ik}$, $\|z_{ik}\|_*$ denotes the dual norm of z_{ik}, χ_{Ξ} denotes the characteristic function of Ξ, and σ_{Ξ} is the conjugate of the characteristic function.

The proof of Theorem 5 can be found in [7], and where it transforms (3.8) by first using the Kantorovich dual formulation Theorem 2, formulating the Lagrangian and solving for the dual problem, then using the definitions of the dual norm and conjugacy. Strong duality is shown for the dual problem by making use of assumption 3 and an extended version of a strong duality result for moment problems [23].

3.3 Wasserstein Penalty Problem

A related problem to (3.8) is the *Wasserstein Penalty Problem* [10], where the hard constraint $\mathbb{Q} \in \mathbb{B}(\hat{\mathbb{P}}_N, \epsilon)$ is instead replaced with a soft constraint term $\lambda W_p(\mathbb{Q}, \hat{\mathbb{P}})^p$ in the objective function, where the term λ can be thought of as a Lagrange multiplier for the hard constraint on $\mathbb{Q}$. This unconstrained optimization problem is often easier to solve compared to the normal constrained problem and a solution can be solved for explicitly in the Linear Quadratic case [10, 24].

3.3.1 Infinite Horizon Wasserstein Penalty Problem

We now consider the (unconstrained) discrete time infinite horizon distributionally robust optimal control problem

$$\inf_{u \in \mathbb{U}} \sup_{\mathbb{Q} \in P(\Xi)} \int_{\Xi} h(u, \xi) \, \mathrm{d}\mathbb{Q}(\xi) \tag{3.13}$$

where the cost functional is now the infinite horizon cost functional

$$h(u, \xi) = \sum_{t=0} \alpha^t \left(c(x_t, u_t) - \lambda W_p(\mathbb{Q}, \hat{\mathbb{P}}_N)^p \right) \tag{3.14}$$

for some discount factor $\alpha \in (0, 1]$ and running cost $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ and Lagrange multiplier λ, subject to the dynamics constraints

$$x_{t+1} = f(x_t, u_t, \xi_t) \tag{3.15}$$

$$x_0 = \boldsymbol{x} \tag{3.16}$$

where $x_t \in \mathbb{X} \subset \mathbb{R}^{n_x}$ and $u_t \in \mathbb{U} \subset \mathbb{R}^{n_u}$ denote the system state and control input respectively.

In this setting, we need impose some additional assumptions for measurable selection in semicontinuous models to guarantee the existence of ϵ-optimal controls in the minimax setting [25] [26]:

Assumption 6. *The function c is lower semicontinuous on $\mathbb{X} \times \mathbb{U}$, and*

$$|c(x, u)| \leq b\zeta(x) \quad \forall (x, u) \in \mathbb{X} \times \mathbb{U} \tag{3.17}$$

for some constant $b \geq 0$ and continuous function $\zeta : \mathbb{X} \to [1, \infty)$ such that $\zeta'(x, u) := \int_{\Xi} \zeta(f(x, u, \xi)) \, \mathrm{d}\mathbb{P}(\xi)$ is continuous on $\mathbb{X} \times \mathbb{U}$ for any $\mathbb{P} \in P(\Xi)$.

In addition, there exists a constant $\beta \in [0, \frac{1}{\alpha})$ such that

$$\zeta'(x, u) \leq \beta\zeta(x) \quad \forall (x, u) \in \mathbb{X} \times \mathbb{U} \tag{3.18}$$

Under this assumption, we can define the Bellman operator T'_λ of the Wasserstein penalty problem (3.13) for $x \in \mathbb{X}$:

$$(T'_\lambda V)(x) := \inf_{u \in \mathbb{U}} \sup_{\mathbb{P} \in P(\Xi)} \int_\Xi c(x,u) - \lambda W_p(\mathbb{P}, \hat{\mathbb{P}}_N)^p + \alpha V(f(x,u,\xi)) \, \mathrm{d}\mathbb{P}(\xi) \tag{3.19}$$

By using the strong duality result from [18], we obtain the following result:

Proposition 7. *Suppose the function $\xi \mapsto V(f(x,u,\xi))$ lies in $L^1(\mathrm{d}\hat{\mathbb{P}}_N)$ for each $(x,u) \in \mathbb{X} \times \mathbb{U}$. Then, the Bellman operator T'_λ, for all $x \in \mathbb{X}$, can be expressed as*

$$(T'_\lambda)(x) := \inf_{u \in \mathbb{U}} \left[c(x,u) + \frac{1}{N} \sum_{i=1}^{N} \sup_{\xi' \in \Xi} \left\{ \alpha V(f(x,u,\xi')) - \lambda d(\hat{\xi}^{(i)}, \xi'^{(i)})^p \right\} \right] \tag{3.20}$$

Furthermore, we have

$$(Tv)(x) = \inf_{\lambda \geq 0} [(T'_\lambda v)(x) + \lambda \theta^p] \quad \forall x \in \mathbb{X} \tag{3.21}$$

where T is the Bellman operator for the original problem:

$$(Tv)(x) := \inf_{u \in \mathbb{U}} \sup_{\mathbb{P} \in \mathbb{B}_{W_p}(\hat{\mathbb{P}}_N, \epsilon)} \left\{ c(x,u) + \alpha \int_\Xi V(f(x,u,\xi)) \, \mathrm{d}\mathbb{P}(\xi) \right\} \tag{3.22}$$

Linear-Quadratic Problems

We now consider the case for when the the dynamics f are *linear*:

$$x_{t+1} = f(x_t, u_t, \xi_t) := Ax_t + Bu_t + C\xi_t \tag{3.23}$$

for matrices $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$ and $C \in \mathbb{R}^{n_x \times m}$, and the cost function c is quadratic in x and u:

$$c(x_t, u_t) := x_t^\top Q x_t + u_t^\top R u_t \tag{3.24}$$

for positive semidefinite $Q \in \mathbb{R}^{n_x \times n_x}$ and positive definite $R \in \mathbb{R}^{n_u \times n_u}$.

Additionally, we choose $p = 2$ for the definition of the ambiguity set $\mathbb{B}(\hat{\mathbb{P}}, \epsilon)$, i.e.,

$$\mathbb{B}_{W_2}(\mu, \epsilon) := \{\nu \in P(X) : W_2(\mu, \nu) \leq \epsilon\} \tag{3.25}$$

and assume that $\hat{\mathbb{P}}$ is zero mean, i.e.,

$$\int_\Xi \xi \, \mathrm{d}\hat{\mathbb{P}}(\xi) = \frac{1}{N} \sum_{i=1}^{N} \hat{\xi}^{(i)} = 0 \tag{3.26}$$

By using dynamic programming, we can obtain the following explicit solution of the LQ problem [9]:

Theorem 8 (Infinite Horizon LQ Wasserstein Penalty Problem)**.** *Suppose that there exists a symmetric positive semidefinite matrix $P \in \mathbb{R}^{n_x \times n_x}$ that solves the following Ricatti equation for a sufficiently large λ:*

$$P = Q + \alpha A^\top P A + \alpha^2 A^\top S(P) A \tag{3.27}$$

where $S(P)$ is defined to be

$$\begin{aligned} S(P) &:= PCD^{-1}C^\top P - [I + \alpha CDC^\top P]^\top PB \\ &\qquad \times [R + \alpha B^\top \{P + \alpha PCDC^\top P\} B]^{-1} \\ &\qquad\qquad \times B^\top P[I + \alpha CDC]^\top P] \end{aligned} \tag{3.28}$$

$$D := (\lambda I - \alpha C^\top PC)^{-1} \tag{3.29}$$

Then, there exists a positive constant $\bar{\lambda}$ such that for any $\lambda \geq \bar{\lambda}$, the optimal value function V^* and the unique optimal policy u^* of the Wasserstein penalty problem (3.13) are given by

$$V^*(x) = x^\top Px + z \tag{3.30}$$

$$u^*(x) = Kx \tag{3.31}$$

where

$$z := \frac{\lambda}{1-\alpha} \operatorname{tr}\Big[(\lambda D - I)\Sigma\Big] \tag{3.32}$$

$$K := -\Big[R + \alpha B^\top (P + \alpha PCDC^\top P)B\Big]^{-1} \alpha B^\top P^\top \Big[I + \alpha CDC^\top P\Big] A \tag{3.33}$$

$$\Sigma := \int_\Xi \xi\xi^\top \, \mathrm{d}\hat{\mathbb{P}}(\xi) = \sum_{i=1}^N \hat{\xi}^{(i)} (\hat{\xi}^{(i)})^\top \tag{3.34}$$

Furthermore, the deterministic stationary policy γ^* defined by

$$\gamma^*(x) = \frac{1}{N} \sum_{i=1}^N \delta_{\xi^{*(i)}} \quad \forall x \in \mathbb{R}^{n_x} \tag{3.35}$$

characterizes the worst case distribution for each $x \in \mathbb{R}^{n_x}$, where $\xi^{*(i)}$ is defined by

$$\xi^{*(i)}(x) := D[\alpha C^\top P(A + BK)x + \lambda \xi^{(i)}] \tag{3.36}$$

Proof. Define the function $v : \mathbb{R}^{n_x} \to \mathbb{R}$ by

$$v(x) := x^\top Px + z \quad \forall x \in \mathbb{R}^{n_x} \tag{3.37}$$

where $P \in \mathbb{R}^{n_x \times n_x}$ is some symmetric positive semidefinite matrix to be specified later, and $z \in \mathbb{R}$.

We now compute $T'_\lambda v$ by using Proposition 7. Denoting the inner supremum in (3.20) as ϕ, in this case we have that

$$\begin{aligned} \phi(u, \xi) &:= \sup_{\xi' \in \Xi} \{\alpha V(f(x, u, \xi')) - \lambda d(\xi, \xi')^p\} && (3.38) \\ &= \sup_{\xi' \in \Xi} \Big\{\alpha V(Ax + Bu + C\xi') - \lambda \|\xi - \xi'\|^2\Big\} && (3.39) \\ &= \sup_{\xi' \in \Xi} \Big\{\alpha(Ax + Bu + C\xi')^\top P(Ax + Bu + C\xi') + \alpha z - \lambda \|\xi - \xi'\|^2\Big\} && (3.40) \end{aligned}$$

Gathering the terms that are quadratic in ξ', we have that

$$\phi(u, \xi) = \sup_{\xi' \in \Xi} \Big\{\cdots + \xi'^\top (\alpha C^\top PC - \lambda I)\xi'\Big\} \tag{3.41}$$

Hence, there exists some constant $\bar{\lambda} > 0$ dependent on P such that for any $\lambda \geq \bar{\lambda}$, the objective function of the maximization problem above is strictly concave in ξ, i.e., $\alpha C^\top PC - \lambda I$ is strictly negative definite. Hence, the unique maximizer ξ^* is given by:

$$\xi^* := (\lambda I - \alpha C^\top PC)^{-1}[\alpha C^\top P(Ax + Bu) + \lambda\xi] \tag{3.42}$$
$$= D[\alpha C^\top P(Ax + Bu) + \lambda\xi] \tag{3.43}$$

where we have defined $D := (\lambda I - \alpha C^\top PC)^{-1}$ for convenience.

Plugging this back into (3.40), we obtain:

$$\begin{aligned}\phi(u, \xi) = \alpha(&x^\top A^\top PAx + u^\top B^\top PBu + 2x^\top A^\top PBu + z) \\ &+ (\alpha C^\top P(Ax + Bu) + \lambda\xi)^\top D(\alpha C^\top P(Ax + Bu) + \lambda\xi) - \lambda \|\xi\|^2\end{aligned} \tag{3.44}$$

Since $\hat{\mathbb{P}}_N$ is zero mean, and denoting the variance of $\mathbb{P}$ as $\Sigma := \mathbb{E}_{\xi\sim\hat{\mathbb{P}}_N}\left[\xi\xi^\top\right]$, taking the expectation of the above:

$$\begin{aligned}\mathbb{E}_{\xi\sim\hat{\mathbb{P}}_N}[\phi(u, \xi)] = u^\top &\left[\alpha B^\top PB + \alpha^2 B^\top PCDC^\top PB\right] u \\ &+ 2\alpha\left[x^\top A^\top + \alpha x^\top A^\top PCDC^\top\right] PBu \\ &+ x^\top\left[\alpha A^\top PA + \alpha^2 A^\top PCDC^\top PA\right] x \\ &+ \alpha z + \lambda^2\operatorname{tr}[D\Sigma] - \lambda\operatorname{tr}[\Sigma]\end{aligned} \tag{3.45}$$

Plugging the above into (3.20), we obtain:

$$(T'_\lambda v)(x) = \inf_{u\in\mathbb{U}}\left\{x^\top Qx + u^\top Ru + \mathbb{E}_{\xi\sim\hat{\mathbb{P}}_N}[\phi(u, \xi)]\right\} \tag{3.46}$$

Looking at the terms quadratic in u, we obtain

$$R + \alpha B^\top PB + \alpha^2 B^\top PCDC^\top PB \tag{3.47}$$

This is positive definite for $\lambda \geq \bar{\lambda}$, because D is positive definite for $\lambda \geq \bar{\lambda}$, and R is positive definite. Hence, we can compute the explicit minimizer u^* as

$$u^* := -[R + \alpha B^\top(P + \alpha PCDC^\top P)B]^{-1}\alpha B^\top P^\top[I + \alpha CDC^\top P]Ax \tag{3.48}$$
$$= Kx \tag{3.49}$$

where the feedback matrix K has the form

$$K = -[R + \alpha B^\top(P + \alpha PCDC^\top P)B]^{-1}\alpha B^\top P^\top[I + \alpha CDC^\top P]A \tag{3.50}$$

Plugging u^* into (3.46), we now have that

$$(T'_\lambda v)(x) = x^\top(Q + \alpha A^\top PA + \alpha^2 A^\top SA)x + \alpha z + \lambda\operatorname{tr}[(\lambda D - I)\Sigma] \tag{3.51}$$

where S is defined as in (3.28) Hence, if P and z satisfy

$$P = Q + \alpha A^\top PA + \alpha^2 A^\top SA \tag{3.52}$$
$$(1 - \alpha)z = \lambda\operatorname{tr}[(\lambda D - I)\Sigma] \tag{3.53}$$

then we will obtain

$$(T'_\lambda v) = v \tag{3.54}$$

and thus v corresponds to the optimal value function V of the Wasserstein penalty problem.

Now, to characterize the worst case distribution, plugging in $\xi = \hat{\xi}^{(i)}$ and $u = Kx$ into the definition of ξ^* (3.42), we obtain

$$\xi^{*(i)}(x) = D[\alpha C^\top P(Ax + BKx) + \lambda \hat{\xi}^{(i)}] \tag{3.55}$$

Now, define $\gamma^*(x) := \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi^*(x)^{(i)}}$. Then, using the Monge definition of W_2 and the fact that both $\gamma^*(x)$ and $\hat{\mathbb{P}}_n$ are empirical distribution, we have that

$$W_2(\gamma^*(x), \hat{\mathbb{P}}_N)^2 = \inf_{\kappa \in \Pi(\gamma^*(x), \hat{\mathbb{P}}_N)} \int_{\mathcal{X}} \|x - y\|^2 \, \mathrm{d}\kappa(x, y) \tag{3.56}$$

$$= \min \left\{ \sum_{i,j=1}^{N} \kappa_{i,j} \left\| \xi^{*(i)} - \hat{\xi}^{(j)} \right\| \;\middle|\; \sum_{j=1}^{N} \kappa_{i,j} = \frac{1}{N}, \sum_{i=1}^{N} \kappa_{i,j} = \frac{1}{N} \right\} \tag{3.57}$$

$$\leq \frac{1}{N} \sum_{i=1}^{N} \left\| \xi^{*(i)} - \hat{\xi}^{(i)} \right\| \tag{3.58}$$

Hence, examining the inside of the integral in the definition of $T'_\lambda v$, we have that

$$\mathbb{E}_{\xi \sim \gamma^*(x)} \left[c(x, u^*) - \lambda W_2(\gamma^*(x), \hat{\mathbb{P}}_N)^2 + \alpha V(f(x, u^*, \xi) \right] \tag{3.59}$$

$$\geq c(x, u^*) - \lambda \sum_{i=1}^{N} \left\| \xi^{*(i)} - \hat{\xi}^{(i)} \right\| + \frac{\alpha}{N} \sum_{i=1}^{N} V(f(x, u^*, \xi^{*(i)})) \tag{3.60}$$

$$= c(x, u^*) + \frac{1}{N} \sum_{i=1}^{N} \sup_{\xi \in \Xi} \left\{ \alpha V(f(x, u, \xi)) - \lambda \left\| \xi - \hat{\xi}^{(i)} \right\|^2 \right\} \tag{3.61}$$

On the other hand, by Proposition 7, we have

$$\begin{aligned} &\sup_{\mathbb{Q} \in P(\Xi)} \mathbb{E}_{\xi \sim \mathbb{Q}} \left[c(x, u^*) - \lambda W_2(\mathbb{Q}, \hat{\mathbb{P}}_N)^2 + \alpha V(f(x, u^*, \xi)) \right] \\ &= c(x, u^*) + \frac{1}{N} \sum_{i=1}^{N} \sup_{\xi \in \Xi} \left\{ \alpha V(f(x, u, \xi)) - \lambda \left\| \xi - \hat{\xi}^{(i)} \right\|^2 \right\} \end{aligned} \tag{3.62}$$

Hence, we must have that $\gamma^*(x)$ is a worst-cast distribution for the LQ Wasserstein penalty problem (3.13). □

3.3.2 Finite Horizon Wasserstein Penalty Differential Dynamic Programming

We now consider a finite-horizon version of the Wasserstein penalty problem. Additionally, in this subsection, we extend the derivation in [24] to consider *general nonlinear dynamics* by performing a 2nd order Taylor expansion at each step and solving the resulting linear quadratic Wasserstein penalty problem.

The (unconstrained) discrete-time finite-horizon Wasserstein penalty problem has the form

$$\inf_{u \in \mathbb{U}} \sup_{\mathbb{Q} \in P(\Xi)} \int_{\Xi} h(u, \xi) \, \mathrm{d}\mathbb{Q}(\xi) \tag{3.63}$$

where the cost functional is now a finite-sum over some horizon $T > 0$

$$h(u, \xi) = \Phi(x_T) + \sum_{t=0}^{T-1} l(x_t, u_t) - \lambda W_p(\mathbb{Q}, \hat{\mathbb{P}}_N)^2 \tag{3.64}$$

subject to the following noise-affine discrete dynamics constraints for convenience.

$$x_{t+1} = f(x_t, u_t) + f_\xi \xi_t \tag{3.65}$$
$$x_0 = \boldsymbol{x} \tag{3.66}$$

The derivation can be easily be extended to the case where linearization is done for f_ξ as well.

We now follow the approach in DDP and derive update laws via dynamic programming. Let $(\bar{x}, \bar{u})$ denote some nominal trajectory. We perform a quadratic approximation of the costs around $(\bar{x}, \bar{u})$:

$$l(\bar{x} + \delta x, \bar{u} + \delta u) \approx l(\bar{x}, \bar{u}) + \begin{bmatrix} l_x \\ l_u \end{bmatrix}^\top \begin{bmatrix} \delta x \\ \delta u \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \delta x \\ \delta u \end{bmatrix}^\top \begin{bmatrix} l_{xx} & l_{xu} \\ l_{ux} & l_{uu} \end{bmatrix} \begin{bmatrix} \delta x \\ \delta u \end{bmatrix}, \tag{3.67}$$

and a linear approximation of the dynamics around the same nominal trajectory:

$$f(\bar{x} + \delta x, \bar{u} + \delta u) \approx f(\bar{x}, \bar{u}) + f_x^\top \delta x + f_u^\top \delta u. \tag{3.68}$$

Define the value function $V : \mathcal{X} \to \mathbb{R}$ recursively as usual via Bellman's equation, where we drop time indices and denote the value function at the next timestep by V' (not to be confused by the Jacobian):

$$V(x) = \inf_{u \in \mathbb{U}} \sup_{\mathbb{Q} \in P(\Xi)} \left\{ l(x, u) - \lambda W_2(\mathbb{Q}, \hat{\mathbb{P}}_N)^2 + \int_\Xi V'(f(x, u)) \, \mathrm{d}\mathbb{Q}(\Xi) \right\} \tag{3.69}$$

As before, we reformulate the infinite-dimensional optimization over $P(\Xi)$ by applying Kantorovich duality theorem 2:

$$V(x) = \inf_{u \in \mathbb{U}} \sup_{\mathbb{Q} \in P(\Xi)} \left\{ l(x, u) + \frac{1}{N} \sum_{i=1}^{N} \sup_{\xi \in \Xi} \{ V'(f(x, u)) - \lambda \left\| \hat{\xi}^{(i)} - \xi \right\|^2 \right\} \tag{3.70}$$

We first prove the form of the optimal control update δu^* in the following lemma (generalization of Lemma 1 in [24] using notation from the DDP domain):

Lemma 9. *Suppose that V' is quadratic in x with the form*

$$V'(\bar{x} + \delta x) = \frac{1}{2} \delta x^\top V_x' x \delta x + {V_x'}^\top \delta x + \bar{V}' \tag{3.71}$$

where V_x' and $V_x' x$ denote the first and second derivatives of V' with respect to x and $\bar{V}' := V'(\bar{x})$. Furthermore, suppose that the penalty parameter λ satisfies $\lambda > \left\| f_\xi^\top V_{xx}' f_\xi \right\|$ and that $l_{uu} + f_u^\top V_{xx}' f_u$ is positive-definite. Then, the minimax problem in (3.70) admits a unique saddle point $(\delta u^*, \xi^*)$ characterized by the following coupled system of equations:

$$Q_{\xi\xi} \xi^{*,(i)} + Q_{\xi x} \delta x + Q_{\xi u} \delta u + Q_{\xi^i} = 0 \tag{3.72}$$

$$Q_{uu} \delta u^* + Q_{ux} \delta x + Q_{u\xi} \frac{1}{N} \sum_{i=1}^{N} \xi^{*,(i)} + Q_u = 0 \tag{3.73}$$

where

$$Q_{\xi\xi} = f_\xi^\top V'_{xx} f_\xi - \lambda I, \qquad Q_{\xi x} = f_\xi^\top V'_{xx} f_x, \qquad Q_{\xi u} = Q_{u\xi}^\top = f_\xi^\top V'_{xx} f_u, \tag{3.74}$$

$$Q_{uu} = l_u u + f_u^\top V'_{xx} f_u, \quad Q_{ux} = l_{ux} + f_u^\top V'_{xx} f_x, \tag{3.75}$$

$$Q_{\xi^i} = f_\xi^\top V'_x + \lambda \hat{\xi}^{(i)}, \qquad Q_u = l_u + f_u^\top V'_x. \tag{3.76}$$

The above coupled system of equations can be solved for δu^* to yield the closed-form solution

$$\delta u^* = -\big(Q_{uu} - Q_{u\xi} Q_{\xi\xi}^{-1} Q_{\xi u}\big)^{-1} \big(\tilde{Q}_u + \tilde{Q}_{ux}\delta x\big) = k + K\delta x \tag{3.77}$$

where $\tilde{Q}_u$ and $\tilde{Q}_{ux}$ have the form

$$\tilde{Q}_u = Q_u - Q_{u\xi} Q_{\xi\xi}^{-1} (f_\xi^\top V'_x - \lambda \bar{\xi}) \tag{3.78}$$

$$\tilde{Q}_{ux} = Q_{ux} - Q_{u\xi} Q_{\xi\xi}^{-1} Q_{\xi x} \tag{3.79}$$

Proof. The assumption that $\lambda > \left\|\Xi^\top V'_{xx} \Xi\right\|$ guarantees that the inner optimization problem (3.70) is strictly concave in ξ. Solving for the first-order optimality conditions for ξ^i yields

$$0 = f_\xi^\top V'_{xx}(f_x\delta x + f_u\delta u + f_\xi \xi^i) + f_\xi^\top V'_x + \lambda(\hat{\xi}^{(i)} - \xi^{(i)}) \tag{3.80}$$

$$= Q_{\xi\xi}\xi^i + Q_{\xi x}\delta x + Q_{\xi u}\delta u + Q_{\xi^i} \tag{3.81}$$

Doing the same for δu gives

$$0 = l_u + l_{uu}\delta u + l_{ux}\delta x + \frac{1}{N}\sum_{i=1}^{N} \{ f_u^\top V'_{xx}(f_x\delta x + f_u\delta u + f_\xi \xi^{(i)}) + f_u^\top V'_x \} \tag{3.82}$$

$$= Q_{uu}\delta u + Q_{ux}\delta x + Q_{u\xi}\frac{1}{N}\sum_{i=1}^{N} \xi^{(i)} + Q_u \tag{3.83}$$

Finally, to obtain a closed-form solution for δu^* (and thus $\xi^{*,i}$), we first eliminate $\xi^{*,i}$:

$$\xi^{*,i} = -Q_{\xi\xi}^{-1}\big(Q_{\xi x}\delta x + Q_{\xi u}\delta u + Q_{\xi^i}\big) \tag{3.84}$$

and then substitute it into δu^*, yielding

$$0 = Q_{uu}\delta u^* + Q_{ux}\delta x - Q_{u\xi}Q_{\xi\xi}^{-1}\big(Q_{\xi x}\delta x + Q_{\xi u}\delta u + Q_{\xi^i}\big) + Q_u \tag{3.85}$$

$$\implies \delta u^* = -(Q_{uu} - Q_{u\xi}Q_{\xi\xi}^{-1}Q_{\xi u})^{-1}\big(\tilde{Q}_u + \tilde{Q}_{ux}\delta x\big) \tag{3.86}$$

□

We next derive the backward pass in the following lemma.

Lemma 10. *Suppose that V' is quadratic in x with the form*

$$V'(\bar{x} + \delta x) = \frac{1}{2}\delta x^\top V'_{xx}\delta x + {V'_x}^\top \delta x + \bar{V}' \tag{3.87}$$

Then V is also quadratic and has the form

$$V(\bar{x} + \delta x) = \frac{1}{2}\delta x^\top V_{xx}\delta x + V_x^\top \delta x + \bar{V} \tag{3.88}$$

where V_{xx}, V_x and ΔV have the form

$$\begin{aligned} V_{xx} = Q_{xx} + Q_{xu}K + K^\top Q_{ux} + K^\top Q_{uu}K + Q_{x\xi}K_\xi + K_\xi^\top Q_{x\xi} \\ + K^\top Q_{u\xi}K_\xi + K_\xi^\top Q_{\xi u}K + K_\xi^\top Q_{\xi\xi}K_\xi \end{aligned} \tag{3.89}$$

$$\begin{aligned} V_x = Q_x + Q_u^\top K + k^\top Q_{ux} + k^\top Q_{uu}K + Q_{\xi^i}K_\xi + k^T Q_{u\xi}K_\xi \\ + \frac{1}{N}\sum_{i=1}^{N}\Big\{k_{\xi^i}Q_{x\xi} + k_{\xi^i}Q_{\xi u}K + k_{\xi^i}Q_{\xi\xi}K_\xi\Big\} \end{aligned} \tag{3.90}$$

$$\bar{V} = \bar{Q} + Q_u^\top k + \frac{1}{2}k^\top Q_{uu}k + \frac{1}{N}\sum_{i=1}^{N}\Big\{Q_{\xi^i}k_{\xi^i} + k^\top Q_{u\xi}k_{\xi^i} + \frac{1}{2}k_{\xi^i}^\top Q_{\xi\xi}k_{\xi^i}\Big\} \tag{3.91}$$

Proof. Plugging in the optimal controls $\delta u^* = k + K\delta x$ into $\xi^{*,i}$ gives

$$\xi^{*,i} = -Q_{\xi\xi}^{-1}\Big(Q_{\xi x}\delta x + Q_{\xi u}\delta u^* + Q_{\xi^i}\Big) \tag{3.92}$$

$$= -Q_{\xi\xi}^{-1}\Big(\big[Q_{\xi x} + Q_{\xi u}K\big]\delta x + \big[Q_{\xi^i} + Q_{\xi u}k\big]\Big) \tag{3.93}$$

$$= K_\xi\delta x + k_{\xi^i} \tag{3.94}$$

Substituting δu^* and $\xi^{*,i}$ back into value function (3.70) then yields

$$\begin{aligned} V(\bar{x} + \delta x) = Q(\bar{x}) + Q_x\delta x + \frac{1}{2}\delta x^\top Q_{xx}\delta x + \delta x^\top Q_{xu}(k + K\delta x) + \frac{1}{2}(k + K\delta x)^\top Q_{uu}(k + K\delta x) \\ + {Q_u}^\top(k + K\delta x) + \delta x^\top Q_{x\xi}(\bar{k}_\xi + K_\xi\delta x) + \delta u^\top Q_{u\xi}(\bar{k}_\xi + K_\xi\delta x) \\ + \frac{1}{N}\sum_{i=1}^{N}(k_{\xi^i} + K_\xi\delta x)^\top Q_{\xi\xi}(\bar{k}_\xi + K_\xi\delta x) + \frac{1}{N}\sum_{i=1}^{N}Q_{\xi^i}(k_{\xi^i} + K\delta x) \end{aligned} \tag{3.95}$$

$$= \frac{1}{2}\delta x^\top V_{xx}\delta x + V_x\delta x + \bar{V} \tag{3.96}$$

where we have defined $\bar{k}_\xi := \frac{1}{N}\sum_{i=1}^{N} k_{\xi^i}$ to denote the mean k_{ξ^i}. □

Chapter 4

Conclusion

In this work, we have presented an overview of methods for applying techniques from distributionally robust optimization into stochastic optimal control. By making use of the Kantorovich duality for the Wasserstein metric, we are able to transform the original computationally intractable infinite dimensional minimax problem into a tractable problem by taking advantage of the discrete nature of the empirical distribution $\hat{\mathbb{P}}_N$. We have shown an application of these of techniques to the case of Linear Quadratic control as presented in [9], where a relaxed version of the original problem can be solved to obtain an explicit solution. Additionally, we have generalized the finite-horizon case of the Wasserstein penalty problem to the nonlinear case via iterative linear and quadratic approximations of the problem.

4.1 Future Work

Although it is possible to solve the relaxed version of the original problem in the Linear Quadratic regime (and thus for all discrete time systems via the use of Differential Dynamic Programming), the question of how to solve the original constrained problem in a computationally efficient manner is still an open question. One method of doing so would be to utilize techniques from constrained optimization techniques such as the Augmented Lagrangian (AL) or Alternating Directions Method of Multipliers (ADMM) to iteratively solve the constrained problem via the relaxed version. Extending the methods in this paper with the above techniques is left for future work.

Bibliography

[1] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*, ser. Applications of mathematics. New York: Springer, 1999, no. 43.

[2] A. Nilim and L. El Ghaoui, “Robust Control of Markov Decision Processes with Uncertain Transition Matrices,” *Operations Research*, vol. 53, no. 5, pp. 780–798, Oct. 2005. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/opre.1050.0216

[3] H. Rahimian and S. Mehrotra, “Distributionally Robust Optimization: A Review,” *arXiv:1908.05659 [cs, math, stat]*, Aug. 2019, arXiv: 1908.05659. [Online]. Available: http://arxiv.org/abs/1908.05659

[4] I. Popescu, “A Semidefinite Programming Approach to Optimal-Moment Bounds for Convex Classes of Distributions,” *Mathematics of Operations Research*, vol. 30, no. 3, pp. 632–657, Aug. 2005. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/moor.1040.0137

[5] B. Van Parys, D. Kuhn, P. Goulart, and M. Morari, “Distributionally Robust Control of Constrained Stochastic Systems,” *IEEE Transactions on Automatic Control*, pp. 1–1, 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7122259/

[6] Z. Hu and L. J. Hong, “Kullback-Leibler Divergence Constrained Distributionally Robust Optimization,” p. 34, 2012.

[7] P. M. Esfahani and D. Kuhn, “Data-driven Distributionally Robust Optimization Using the Wasserstein Metric: Performance Guarantees and Tractable Reformulations,” *arXiv:1505.05116 [math, stat]*, Jun. 2017, arXiv: 1505.05116. [Online]. Available: http://arxiv.org/abs/1505.05116

[8] I. Yang, “Distributionally robust stochastic control with conic confidence sets,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. Melbourne, Australia: IEEE, Dec. 2017, pp. 4291–4296. [Online]. Available: http://ieeexplore.ieee.org/document/8264292/

[9] ——, “Data-Driven Distributionally Robust Stochastic Control of Energy Storage for Wind Power Ramp Management Using the Wasserstein Metric,” *Energies*, vol. 12, no. 23, p. 4577, Dec. 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/23/4577

[10] ——, “Wasserstein Distributionally Robust Stochastic Control: A Data-Driven Approach,” *arXiv:1812.09808 [cs, math]*, Nov. 2019, arXiv: 1812.09808. [Online]. Available: http://arxiv.org/abs/1812.09808

[11] A. Hakobyan and I. Yang, "Wasserstein Distributionally Robust Motion Control for Collision Avoidance Using Conditional Value-at-Risk," *arXiv:2001.04727 [cs, eess]*, Jan. 2020, arXiv: 2001.04727. [Online]. Available: http://arxiv.org/abs/2001.04727

[12] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," *arXiv:1701.07875 [cs, stat]*, Dec. 2017, arXiv: 1701.07875. [Online]. Available: http://arxiv.org/abs/1701.07875

[13] I. Popescu, "Robust Mean-Covariance Solutions for Stochastic Optimization," *Operations Research*, vol. 55, no. 1, pp. 98–112, Feb. 2007. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/opre.1060.0353

[14] S. Zymler, D. Kuhn, and B. Rustem, "Distributionally robust joint chance constraints with second-order moment information," *Mathematical Programming*, vol. 137, no. 1-2, pp. 167–198, Feb. 2013. [Online]. Available: http://link.springer.com/10.1007/s10107-011-0494-7

[15] Z. Wang, P. Glynn, and Y. Ye, "Likelihood Robust Optimization for Data-driven Problems," *arXiv:1307.6279 [math]*, Oct. 2014, arXiv: 1307.6279. [Online]. Available: http://arxiv.org/abs/1307.6279

[16] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Mathematical Programming*, vol. 158, no. 1-2, pp. 291–327, Jul. 2016. [Online]. Available: http://link.springer.com/10.1007/s10107-015-0929-7

[17] D. K. Love, "Phi-Divergence Constrained Ambiguous Stochastic Programs for Data-Driven Optimization," p. 32, Oct. 2015.

[18] R. Gao and A. J. Kleywegt, "Distributionally Robust Stochastic Optimization with Wasserstein Distance," *arXiv:1604.02199 [math]*, Jul. 2016, arXiv: 1604.02199. [Online]. Available: http://arxiv.org/abs/1604.02199

[19] C. Zhao and Y. Guan, "Data-driven risk-averse stochastic optimization with Wasserstein metric," *Operations Research Letters*, vol. 46, no. 2, pp. 262–267, Mar. 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0167637718300506

[20] I. Yang, "A Convex Optimization Approach to Distributionally Robust Markov Decision Processes With Wasserstein Distance," *IEEE Control Systems Letters*, vol. 1, no. 1, pp. 164–169, Jul. 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7938642/

[21] A. Shapiro and D. Dentcheva, *Lectures on Stochastic Programming: Modeling and Theory*. Society for Industrial and Applied Mathematics, 2014.

[22] C. Villani, *Optimal Transport: Old and New*, ser. Grundlehren der mathematischen Wissenschaften, 2009.

[23] A. Shapiro, "On duality theory of conic linear problems," in *Semi-infinite programming*. Springer, 2001, pp. 135–165.

[24] K. Kim and I. Yang, "Minimax control of ambiguous linear stochastic systems using the Wasserstein metric," *arXiv:2003.13258 [cs, eess, math]*, Mar. 2020, arXiv: 2003.13258. [Online]. Available: http://arxiv.org/abs/2003.13258

[25] J. I. González-Trejo, O. Hernández-Lerma, and L. F. Hoyos-Reyes, "Minimax Control of Discrete-Time Stochastic Systems," *SIAM Journal on Control and Optimization*, vol. 41, no. 5, pp. 1626–1659, Jan. 2002. [Online]. Available: http://epubs.siam.org/doi/10.1137/S0363012901383837

[26] A. S. Nowak, "Measurable selection theorems for minimax stochastic optimization problems," *SIAM Journal on Control and Optimization*, vol. 23, no. 3, pp. 466–476, 1985.