Bounding the Distance to Unsafe Sets
With Convex Optimization

Jared Miller, Member, IEEE, and Mario Sznaier, Fellow, IEEE

Abstract—This work proposes an algorithm to bind the minimum distance between points on trajectories of a dynamical system and points on an unsafe set. Prior work on certifying safety of trajectories includes barrier and density methods, which do not provide a margin of proximity to the unsafe set in terms of distance. The distance estimation problem is relaxed to a Monge–Kantorovich-type optimal transport problem based on existing occupation-measure methods of peak estimation. Specialized programs may be developed for polyhedral norm distances (e.g., L1 and Linfinity) and for scenarios where a shape is traveling along trajectories (e.g., rigid body motion). The distance estimation problem will be correlatively sparse when the distance objective is separable.

Index Terms—Linear matrix inequality (LMI), numerical optimization, peak estimation, safety, sum of squares (SOS).

I. INTRODUCTION

A TRAJECTORY is safe with respect to an unsafe set $X_u$ if no point along the trajectory contacts or enters $X_u$. Safety of trajectories may be quantified by the distance of closest approach to $X_u$, which will be positive for all safe trajectories and zero for all unsafe trajectories. The task of finding this distance of closest approach will also be referred to as “distance estimation.” In this setting, an agent with state $x$ is restricted to a state space $X \subseteq \mathbb{R}^n$ and starts in an initial set $X_0 \subset X$. The trajectory of an agent evolving according to locally Lipschitz dynamics $\dot{x} = f(t, x(t))$ starting at an initial condition $x_0 \in X_0$ is denoted by $x(t \mid x_0)$. The closest approach as measured by a distance function $c$ that any trajectory takes to the unsafe set $X_u$ in a time horizon of $t \in [0, T]$ can be found by solving

$$P^* = \inf_{t, x_0, y} c(x(t \mid x_0), y)$$

Solving (1) requires optimizing over all points $(t, x_0, y) \in [0, T] \times X_0 \times X_u$, which is generically a nonconvex and difficult task. Upper bounds to $P^*$ may be found by sampling points $(x_0, y)$ and evaluating $c(x(t \mid x_0), y)$ along these sampled trajectories. Lower bounds to $P^*$ are a universal property of all trajectories, and will satisfy $P^* > 0$ if all trajectories starting from $X_0$ in the time horizon $[0, T]$ are safe with respect to $X_u$.

This article proposes an occupation-measure-based method to compute lower bounds of $P^*$ through a converging hierarchy of convex semidefinite programs (SDPs) [1]. These SDPs arise from a finite truncation of infinite dimensional linear programs (LPs) in measures [2]. Occupation measures are Borel measures that contain information about the distribution of states evolving along trajectories of a dynamical system. The distance estimation LP formulation is based on measure LP arising from peak estimation of dynamical systems [3], [4], [5] because the state function to be minimized along trajectories is the point-set distance function between $x \in X$ and $X_u$. Inspired by optimal transport theory [6], [7], [8], the distance function $c(x, y)$ between points $x \in X$ on trajectories and $y \in X_u$ is relaxed to an expectation of the distance $c(x, y)$ with respect to probability distributions over $X$ and $X_u$.

Occupation measure LP for control problems were first formulated in [9], and their linear matrix inequality (LMI) relaxations were detailed in [10]. These occupation measure methods have also been applied to region of attraction estimation and backwards reachable set maximizing control [11], [12], [13].

Prior work on verifying safety of trajectories includes barrier functions [14], [15], density functions [16], and safety margins [17]. Barrier and density functions offer binary indications of safety/unsafety; if a barrier/density function exists, then all trajectories starting from $X_0$ are safe. Barrier/density functions may be nonunique, and the existence of such a function does not yield a measure of closeness to the unsafe set. Safety margins are a measure of constraint violation, and a negative safety margin verifies safety of trajectories. Safety margins can vary with constraint reparameterization, even in the same coordinate system (e.g., multiplying all defining constraints of $X_u$ by a positive constant scales the safety margin by that constant), and therefore yield a qualitative certificate of safety. The distance of closest approach $P^*$ is independent of constraint reparameterization, returning quantifiable and geometrically interpretable information about safety of trajectories.
The contributions of this article include the following.
1) A measure LP to lower bound the distance estimation task (1).
2) A proof of convergence to $P^*$ within arbitrary accuracy as the degree of LMI approximations approaches infinity.
3) A decomposition of the distance estimation LP using correlative sparsity when the cost $c(x, y)$ is separable.
4) Extensions, such as finding the distance of closest approach between a shape with evolving orientations and the unsafe set.

Parts of this article were presented at the 61st Conference on Decision and Control [18]. Contributions of this work over and above the conference version include the following.
1) A discussion of the scaling properties of safety margins.
2) An application of correlative sparsity in order to reduce the computational cost of finding distance estimates.
3) An extension to bounding the set–set distance between a moving shape and the unsafe set.
4) A presentation of a lifting framework for polyhedral norm distance functions.
5) A full proof of strong duality.

The rest of this article is organized as follows. Section II reviews preliminaries such as notation and measures for peak and safety estimation. Section III proposes an infinite-dimensional LP to bound the distance of closest approach between points along trajectories and points on the unsafe set. Section IV truncates the infinite-dimensional LP into SDP through the momentum-sum of squares (SOS) hierarchy, and studies numerical considerations associated with these SDP. Section V utilizes correlative sparsity to create SDP relaxations of distance estimation with smaller positive semidefinite (PSD) matrix constraints. Distance estimation problems for shapes traveling along trajectories are posed in Section VI. Examples of the distance estimation problem are presented in Section VII. Section VIII details extensions to the distance estimation problem, including uncertainty, polyhedral norm distances, and application of correlative sparsity. Finally, Section IX concludes this article. Appendix I offers a proof of strong duality between infinite-dimensional LPs for distance estimation. Appendix II summarizes the moment-SOS hierarchy.

II. PRELIMINARIES

A. Notation and Measure Theory

Let $\mathbb{R}$ be the set of real numbers and $\mathbb{R}^n$ be an $n$-dimensional real Euclidean space. Let $\mathbb{N}$ be the set of natural numbers and $\mathbb{N}^n$ be the set of $n$-dimensional multi-indices. The total degree of a multi-index $\alpha \in \mathbb{N}^n$ is $|\alpha| = \sum \alpha_i$. A monomial $\prod_{i=1}^{n} x_i^{\alpha_i}$ may be expressed in multi-index notation as $x^n$. The set of polynomials with real coefficients is $\mathbb{R}[x]$, and polynomials $p(x) \in \mathbb{R}[x]$ may be represented as the sum over a finite index set $J \subset \mathbb{N}^n$ of $p(x) = \sum_{\alpha \in J} p_{\alpha} x^\alpha$. The set of polynomials with monomials up to degree $|\alpha| = d$ is $\mathbb{R}[x]_{\leq d}$. A metric function $c(x, x) = c(y, y) > 0$ for $x \neq y$ satisfies the following properties [19]:

\[
c(x, x) = 0 \quad (2a)
\]
\[
c(x, y) \leq c(x, z) + c(z, y) \quad \forall z \in S. \quad (2b)
\]

The set of metrics are closed under addition and pointwise maximums. Every norm $|| \cdot ||$ inspires a metric $c_{|| \cdot ||}(x, y) = ||x - y||$. The point-set distance function $c(x; Y)$ between a point $x \in X$ and a closed set $Y \subset X$ is defined by

\[
c(x; Y) = \inf_{y \in Y} c(x, y). \quad (3)
\]

The set of continuous functions over a Banach space $S$ is denoted as $C(S)$, the set of finite signed Borel measures over $S$ is $M(S)$, and the set of nonnegative Borel measures over $S$ is $M_+(S)$. A duality pairing exists between all functions $f \in C(S)$ and measures $\mu \in M_+(S)$ by Lebesgue integration:

\[
\langle f, \mu \rangle = \int_S f(s) d\mu(s) \quad (2c)
\]

The point-set distance function $c(x; Y)$ between a point $x \in X$ and a closed set $Y \subset X$ is defined by

\[
c(x; Y) = \inf_{y \in Y} c(x, y). \quad (3)
\]

The set of continuous functions over a Banach space $S$ is denoted as $C(S)$, the set of finite signed Borel measures over $S$ is $M(S)$, and the set of nonnegative Borel measures over $S$ is $M_+(S)$. A duality pairing exists between all functions $f \in C(S)$ and measures $\mu \in M_+(S)$ by Lebesgue integration:

\[
\langle f, \mu \rangle = \int_S f(s) d\mu(s) \quad (2c)
\]

The set of continuous functions over a Banach space $S$ is denoted as $C(S)$, the set of finite signed Borel measures over $S$ is $M(S)$, and the set of nonnegative Borel measures over $S$ is $M_+(S)$. A duality pairing exists between all functions $f \in C(S)$ and measures $\mu \in M_+(S)$ by Lebesgue integration:

\[
\langle f, \mu \rangle = \int_S f(s) d\mu(s) \quad (2c)
\]

The set of continuous functions over a Banach space $S$ is denoted as $C(S)$, the set of finite signed Borel measures over $S$ is $M(S)$, and the set of nonnegative Borel measures over $S$ is $M_+(S)$. A duality pairing exists between all functions $f \in C(S)$ and measures $\mu \in M_+(S)$ by Lebesgue integration:

\[
\langle f, \mu \rangle = \int_S f(s) d\mu(s) \quad (2c)
\]
starting from \( X_0 \subset X \subset \mathbb{R}^n \) and remaining in \( X \)

\[
P^* = \sup_{t \in [0,T], x_0 \in X_0} p(x(t \mid x_0))
\]

\[
x(0) = x_0, \quad \dot{x}(t) = f(t, x(t)). \tag{4}
\]

Every optimal trajectory of (4) (if one exists) may be described by a tuple \((x_0^*, t_p^*)\) with \( x_0^+ = x(t_p^* \mid x_0^*) \) satisfying \( P^* = p(x_0^+) = p(x(t_p^* \mid x_0^*)). \) A persistent example throughout this article will be the flow system of [14]

\[
\dot{x} = \begin{bmatrix}
x_2 \\
-x_1 - x_2 + \frac{1}{2} x_1^3
\end{bmatrix}. \tag{5}
\]

Fig. 1 plots trajectories of the flow system in cyan for times \( t \in [0,5] \), starting from the initial set \( X_0 = \{ x \mid (x_1 - 1.5)^2 + x_2 \leq 0.4^2 \} \) in the black circle. The minimum value of \( x_2 \) along these trajectories is \( \min x_2 \approx -0.5734 \). The optimizing trajectory is shown in dark blue, starting at the blue circle \( x_0^* \approx (1.4889, -0.3998) \) and reaching optimality at \( x_0^* \approx (0.6767, -0.5734) \) in time \( t_p^* \approx 1.6627 \).

The work in [4] developed a measure LP to find an upper bound \( p^* \geq P^* \). This measure LP involves an initial measure \( \mu_0 \in \mathcal{M}_+(X_0) \), a peak measure \( \mu_p \in \mathcal{M}_+(\{0, T\} \times X) \), and an occupation measure \( \mu \in \mathcal{M}_+(\{0, T\} \times X) \) connecting together \( \mu_0 \) and \( \mu_p \). Given a distribution of initial conditions \( \mu_0 \in \mathcal{M}_+(X_0) \) and a stopping time \( 0 \leq t^* \leq T \), the occupation measure \( \mu \) of a set \( A \times B \) with \( A \subseteq [0, T] \), \( B \subseteq X \) is defined by

\[
\mu(A \times B) = \int_{[0,t^*] \times X_0} I_{A \times B}((t, x(t \mid x_0))) \, dt \, d\mu_0(x_0). \tag{6}
\]

The measure \( \mu(A \times B) \) is the \( \mu_0 \)-averaged amount of time a trajectory will dwell in the box \( A \times B \). With ODE dynamics \( \dot{x}(t) = f(t, x(t)) \), the Lie derivative \( \mathcal{L}_f \) along a test function \( v \in C^1([0, T] \times X) \) is

\[
\mathcal{L}_fv(t, x) = \partial_t v(t, x) + f(t, x) \cdot \nabla_x v(t, x). \tag{7}
\]

Liouville’s equation expresses the constraint that \( \mu_0 \) is connected to \( \mu_p \) by trajectories with dynamics \( f \) for all test functions \( v \in C^1([0, T] \times X) \)

\[
\langle v(t, x), \mu_p \rangle = \langle v(0, x), \mu_0 \rangle + \langle \mathcal{L}_fv(t, x), \mu \rangle \tag{8}
\]

\[
\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f \mu. \tag{9}
\]

Equation (9) is an equivalent short-hand expression to (8) for all \( v \). Substituting in the test functions \( v = 1, v = t \) to Liouville’s equation returns the relations \( \langle 1, \mu_0 \rangle = \langle 1, \mu_p \rangle \) and \( \langle 1, \mu \rangle = \langle 1, \mu_p \rangle \).

The measure LP corresponding to (4) with optimization variables \((\mu_0, \mu_p, \mu)\) is [4]

\[
p^* = \sup \langle p(x), \mu_p \rangle \tag{10a}
\]

\[
\mu_p = \delta_0 \otimes \mu_0 + \mathcal{L}_f \mu \tag{10b}
\]

\[
\langle 1, \mu_0 \rangle = 1 \tag{10c}
\]

\[
\mu, \mu_p \in \mathcal{M}_+([0, T] \times X) \tag{10d}
\]

\[
\mu_0 \in \mathcal{M}_+(X_0). \tag{10e}
\]

The initial measure \( \mu_0 \) is a probability measure by (10c). The relation \((1, \mu_p) = (1, \mu_0)\) from (10b) imposes that \( \mu_p \) is also a probability measure. A set of measures \((\mu_0, \mu_p, \mu)\) may be derived from a trajectory with initial condition \( x_0^* \in X_0 \) and a stopping time \( t_p^* \) in which \( x_0^+ = x(t_p^* \mid x_0^*) \), \( p(x_0^+) = P^* \), and \( x(t \mid x_0^*) \in X \forall t \in [0, t_p^*] \). The atomic measures are \( \mu_0 = \delta_{x=x_0^*}, \mu_p = \delta_{t=t_p^*} \otimes \delta_{x=x_0^*} \), and \( \mu \) is the occupation measure in times \([0, t_p^*]\) along \( t \mapsto (t, x(t \mid x_0^*)) \). These measures are solutions to constraints (10b)–(10e), which implies that \( p^* \geq P^* \). There is no relaxation gap \((p^* = P^*)\) if the set \([0, T] \times X\) is compact with \( X_0 \subset X \) (Section II-C of [5] and [9]). The moment-SOS hierarchy of SDP may be used to find a sequence of upper bounds to \( p^* \). The method in [5] approaches the moment-SOS hierarchy from the dual side, involving SOS constraints in terms of an auxiliary function \( v(t, x) \) [dual variable to constraint (10b)]. Near-optimal trajectory extraction can be attempted through SDP solution matrix factorization [17] (if a low-rank condition holds) and through sublevel set methods [5, 20].

C. Safety

This section reviews methods to verify that trajectories starting from \( X_0 \subset X \) do not enter an unsafe set \( X_u \subset X \). In Fig. 2, the unsafe set \( X_u = \{ x \mid x_1^2 + x_2^2 + (x_2 + 0.7)^2 \leq 0.5^2 \}, \sqrt{2}/2(x_1 + x_2 - 0.7) \leq 0 \} \) is the red half-circle to the bottom-left of trajectories.

Sufficient conditions certifying safety can be obtained using barrier functions [14, 15]. However, these conditions do not provide a quantitative measurement for the safety of trajectories. Safety margins, as introduced in [17], quantify the safety of trajectories through the use of maximin peak estimation. Assume that \( X_u \) is a basic semialgebraic set with description \( X_u = \{ x \mid p_i(x) \geq 0, i = 1, \ldots, N_u \} \). A point \( x \) is in \( X_u \) if all \( p_i(x) \geq 0 \). If at least one \( p_i(x) \) remains negative for all points along trajectories \( x(t \mid x_0), x_0 \in X_0 \), then no point starting from \( X_0 \)
Flow system is safe, \( p^* \leq -0.2831 \).

Safety margin scaling contours.

enters \( X_u \), and trajectories are therefore safe. The value \( p^* = \max_x \min_t x_p(t \mid x_0) \) is called the safety margin, and a negative safety margin \( p^* < 0 \) certifies safety. The moment-SOS hierarchy (Appendix II) can be used to find upper bounds \( p_d^* > p^* \) at degrees \( d \), and safety is assured if any upper bound is negative \( 0 > p_d^* > p^* \). Fig. 3 visualizes the safety margin for the flow system (5), where the bound of \( p^* \leq -0.2831 \) was found at the degree-4 relaxation.

The safety margin of trajectories will generally change if the unsafe set \( X_u \) is reparameterized even in the same coordinate system. Let \( q \leq 0 \) and \( s > 0 \) be violation and scaling parameters for the enlarged unsafe set \( (X_u^s)_q = \{ x \mid q \leq 0.5^2 - x_1^2 + (x_2 + 0.7)^2, \; q \leq -s(x_1 + x_2 - 0.7) \} \). The original unsafe set may be interpreted as \( X_u = (X_u^{s=\sqrt{2}/2})_{q=0} \). Fig. 4 visualizes contours of regions \( (X_u^s)_q \) as \( q \) decreases from 0 down to \(-2\) for sets with scaling parameters \( s = 5 \) and \( s = 1 \). The safety margins of trajectories with respect to \( X_u^s \) will vary as \( s \) changes, even as the same set \( X_u \) is represented in all cases. This is precisely the difficulty addressed in this article: developing scale invariant quantitative safety metrics.

### III. DISTANCE ESTIMATION PROGRAM

The goal of this article is to develop a computationally tractable framework to compute the worst case (over all initial conditions) distance of closest approach to an unsafe set. Specifically, we aim to solve the following problem.

**Problem 1 (Distance calculation):** Given semialgebraic initial condition \( (X_0) \) and unsafe \( (X_u) \) sets, solve the optimization problem (1).

In many practical situations, it is sufficient to obtain interpretable lower bounds on the minimum distance. Thus, the following problem is also of interest.

**Problem 2 (Distance estimation):** Given a semialgebraic initial condition set \( (X_0) \), a semialgebraic unsafe \( (X_u) \) set, and a positive integer \( d \) (degree), find a lower bound \( p_d^* \leq P^* \) to the solution of optimization (1).

As we will show in this article (and under mild compactness and regularity conditions), a convergent sequence of lower bounds \( \{p_d^*\} \) that rise to \( \lim_{d \to \infty} p_d^* = P^* \) may be constructed where each bound \( p_d^* \) is obtained by solving a finite dimensional SDP.

An optimizing trajectory of the distance program (1) may be described by a tuple \( T^* = (y^*, x_0^*, t^*_p) \) using Table I.

The relationship between these quantities for an optimal trajectory of (1) is

\[
P^* = c \left( x(t_p^* \mid x_0^*) \mid X_u \right) = c \left( x(t_p^* \mid x_0^*) \mid y^* \right)
\]  

Fig. 5 plots the trajectory of closest approach to \( X_u \) in dark blue. This minimal \( L_2 \) distance is 0.2831, and the red curve is the level set of all points with a point-set distance 0.2831 to \( X_u \). On the optimal trajectory, the blue circle is \( x^*_1 = (1.489, -0.3998) \), the blue star is \( x^*_p = x(t^* \mid x_0) \approx (0, -0.2997) \), and the blue square is \( y^* \approx (-0.2002, -0.4998) \). The closest approach of 0.2831 occurred at time \( t^* \approx 0.6180 \). Fig. 6 plots the distance and safety margin contours for the set \( X_u \). These distance contours for a given metric \( c \) are independent of the way that \( X_u \) is defined (within the same coordinate system).

### A. Assumptions

The following assumptions are made in Program (1).

#### TABLE I

| \( y^* \) | location on unsafe set of closest approach |
| \( x_0^* \) | initial condition to produce closest approach |
| \( t^*_p \) | time to reach closest approach from \( x_0^* \) |

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
A1: The sets \([0, T], X, X_u, X_0\) are all compact, \(X_0 \subset X\).

A2: The function \(f(t, x)\) is Lipschitz in each argument in the compact set \([0, T] \times X\).

A3: The cost \(c(x, y)\) is \(C^0\) in \(X \times X_u\).

A4: If \(x(t | x_0) \in \partial X\) for some \(t \in [0, T]\), \(x_0 \in X_0\), then \(x(t' | x_0) \notin X \forall t' \in (t, T]\).

A3 relaxes the requirement that \(c\) should be a metric, allowing for costs, such as \(\|x - y\|_2^2\) in addition to the metric \(\|x - y\|_2\).

The combination of A1 and A3 enforce that \(c\) is bounded, \(\delta_x \leq c\) allows for a measure \(\delta_x \otimes \delta_y\) inside \(X \times X_u\), by the Weierstrass extreme value theorem. Assumption A4 requires that trajectories leave \(X_0\) immediately after contacting the boundary \(\partial X\).

**Remark 1:** A strict \(\epsilon\)-superset \(X^*\) is a set \(X^* \supset X\) in which the boundaries of \(X^*\) and \(X\) have a positive distance. If trajectories starting in \(X_0\) remain in \(X\) at all times \(t \in [0, T]\), then any strict \(\epsilon\)-superset \(X^*\) satisfies A4. However, \(X\) may not satisfy A4, because there might exist a trajectory remaining in \(X\) that is tangent to \(\partial X\).

**B. Measure Program**

The problem of \(c^* = \min_{(x, y)} c(x, y)\) is identical to \(c^* = \min_{(x, y)} \langle c(x, y), \delta_x \otimes \delta_y \rangle\) for Dirac measures \(\delta_x\) and \(\delta_y\). The Dirac restriction may be relaxed to minimization over the set of probability measures \(c^* = \langle c(x, y), \eta \rangle, \eta \in \mathcal{M}_+(X \times X_u), (1, \eta) = 1\) with no change in the objective value \(c^*\). An infinite-dimensional convex LP in measures \((\mu_0, \mu_p, \mu, \eta)\) to bound from below the distance closest approach to \(X_u\) starting from \(X_0\) may be developed.

**Theorem 3.1:** Suppose that \(f \in C^0\) and A3 holds. Further impose that A4 holds if \(X_0 \subset X\) are both compact. Under these conditions, a lower bound for \(P^*\) is

\[
\begin{align*}
\text{Proof:} &\text{ Let } T = (y, x_0, t_p) \in X_u \times X_0 \times [0, T] \text{ be a tuple representing a trajectory with } x_p = x(t_p | x_0) \text{ achieving a distance } P = c(x_p, y). \text{ A set of measures (12e) and (12f) satisfying constraints (12b)–(12f) may be constructed from the tuple } T. \\
&\text{The initial measure } \mu_0 = \delta_x = \mu_0, \text{ the peak (free-terminal-time) measure } \mu_p = \delta_{x = x_p} \otimes \delta_{y = y_p}, \text{ with } x_p = x(t_p | x_0), \text{ and the joint measure } \eta = \delta_{x_p} \otimes \delta_{y_p}, \text{ are all rank-one atomic probability measures. The measure } \mu \text{ is the occupation measure of } t \mapsto (t, x(t | x_0)) \text{ in times } [0, t_p]. \text{ The distance objective (12a) for the tuple } T \text{ may be evaluated as}
\end{align*}
\]

\[
\langle c(x, y), \eta \rangle = \langle c(x, y), \delta_{x = x_p} \otimes \delta_{y = y_p} \rangle = c(x_p, y) = P.
\]

The feasible set of (12b)–(12f) contains all measures constructed from trajectories by the abovementioned process, which immediately implies that \(P^* \leq P^\dagger\).

**Remark 2:** As a reminder, the term \(\pi_{\#}^x\) from constraint (12b) is the operator performing \(x\)-marginalization. Constraint (12b) ensures that the \(x\)-marginals of \(\eta\) and \(\mu_p\) are equal: \(\forall \eta \in C(X) : \langle w(x), \eta(x, y) \rangle = \langle w(x), \mu_p(t, x) \rangle\).

We now prove that the measure program in (12) has the same objective value as the trajectory program in (1) under assumptions A1–A4. In order to accomplish this task, we require a pair of lemmas.

**Lemma 3.1:** Under assumptions A1–A4, the following measure LP has the same optimal value as (1):

\[
\begin{align*}
p^\circ &= \inf \langle c(x; X_u), \mu_p(t, x) \rangle & (14a) \\
\mu_p &= \delta_0 \otimes \mu_0 + \mathcal{L}_f^1 \mu & (14b) \\
\langle 1, \mu_0 \rangle &= 1 & (14c) \\
\mu_0 &\in \mathcal{M}_+(X_0), \mu_p, \mu \in \mathcal{M}_+([0, T] \times X). & (14d)
\end{align*}
\]

**Proof:** Problem (10) is a peak estimation instance of (1) with a continuous (A3) objective of \(p(x) = -c(x; X_u)\). Theorem 2.1 of [9] states that the peak estimation LP (14) will equal the true peak estimation problem (4) (distance estimation problem (1)). The measures in (14d) contain trajectories that stay within \(X\) and terminate on \(\partial X\) (agreeing with the nonreturn assumption A4).

**Lemma 3.2:** Under the assumptions that A1 and A3 hold and that \(\nu \in \mathcal{M}_+(X)\) is a probability measure, it follows that:

\[
\langle c(x; X_u), \nu(x) \rangle = \inf_{\eta \in \mathcal{M}_+(X \times X_u)} \langle c, \eta \rangle : \pi_{\#}^x \eta = \nu. 
\]

**Proof:** This follows by Theorem 2.2(a) of [21], given that \(X \times X_u\) is compact and \(c\) is continuous.

**Remark 3:** The parameterized method of [21] assumes that \(\nu\) has a positive density with respect to the Lebesgue measure on \(X\). However, this assumption of positive density is not required in the statement nor the proof of Theorem 2.2(a) used in [21] (and therefore in Lemma 3.2 in this article).

**Theorem 3.2:** Under assumptions A1–A4, \(p^* = P^\dagger\).

**Proof:** Lemma 3.1 states that \(p^\circ = P^*\) under assumptions A1–A4. For any solution \((\mu_0, \mu_p, \mu)\) to constraints (14a) and (14b), Lemma 3.2 allows for a measure \(\eta\) to be chosen under \(\nu = \pi_{\#}^x \mu_p\) with cost \(\langle c(x; X_u), \pi_{\#}^x \mu_p(x) \rangle = \langle c, \eta \rangle\). Furthermore, it is not possible to choose an \(\eta\) such that \(\langle c(x; X_u), \pi_{\#}^x \mu_p(x) \rangle \geq \langle c, \eta \rangle\). The infimal objectives \(p^\circ = P^\dagger\) are the same, which implies that \(p^* = P^\dagger\).

**C. Function Program**

Dual variables \(v(t, x) \in C^1([0, T] \times X), w(x) \in C(X), \gamma \in \mathbb{R}\) over constraints (12b)–(12d) must be introduced to derive the dual LP to (12). The Lagrangian \(\mathcal{L}\) of problem (12) is

\[
\mathcal{L} = \langle c(x, y), \eta \rangle + \langle v(t, x), \delta_0 \otimes \mu_0 + \mathcal{L}_f^1 \mu - \mu_p \rangle + \langle w(x), \pi_{\#}^x \mu_p - \pi_{\#}^x \eta \rangle + \gamma (1 - \langle 1, \mu_0 \rangle). 
\]
Recalling that $\forall \eta \in \mathcal{M}_+ (X \times Y)$, $w \in C(X)$ the relation that $(w(x), \eta(x,y)) = (w(x), \pi_y \eta(x))$ holds, the Lagrangian $\mathcal{L}$ in (16) may be reformulated as

$$
\mathcal{L} = c + \langle v(0,x) - c, \mu_0 \rangle + \langle c(x,y) - w(x), \eta \rangle
+ \langle w(x) - v(t,x), \mu_p \rangle + \langle \mathcal{L}_f v(t,x), \mu \rangle.
$$

(17)

The dual program (12) is provided by

$$
d^* = \sup_{\gamma, v, w} \inf_{\mu_0, \mu, \eta} \mathcal{L} = \sup_{\gamma} \inf_{\mu_0, \mu, \eta} \gamma = v(0,x) \geq \gamma \implies \forall x \in X_0
\text{ (18c)}
c(x,y) \geq w(x) \implies \forall (x,y) \in X \times X_u \text{ (18d)}
w(x) \geq v(t,x) \implies \forall (t,x) \in [0, T] \times X \text{ (18e)}
\mathcal{L}_f v(t,x) \geq 0 \implies \forall (t,x) \in [0, T] \times X \text{ (18f)}
w \in C(X) \text{ (18g)}
v \in C^1 ([0, T] \times X) \text{ (18h)}.

Theorem 3.3: Strong duality with $p^* = d^*$ and attainment of optima occurs under assumptions A1–A4.

Proof: See Appendix 1. \hfill \blacksquare

Remark 4: The continuous function $w(x)$ is a lower bound on the point set distance $c(x; X_u)$ by constraint (18d). The auxiliary function $v(t,x)$ is in turn a lower bound on $w(x)$ by constraint (18e). This establishes a chain of lower bounds $v(t,x) \leq w(x) \leq c(x; X_u)$ holding $\forall (t,x) \in [0, T] \times X$.

IV. FINITE-DIMENSIONAL PROGRAMS

This section presents finite-dimensional SDP truncations to the infinite-dimensional LP (12) and (18). We note that (14) may possess a nonpolynomial (but semialgebraic) cost $c(x; X_u)$, and therefore the lift to (12) is required to utilize the moment-SOS hierarchy.

A. Approximation Preliminaries

We introduce notation and concepts about moments and SOS polynomials that will be used in subsequent finite-dimensional programs. Refer to Appendix II for further detail (e.g., Archimedean structure, moment-SOS hierarchy, conditions of convergence). A basic semialgebraic set $\mathbb{K} = \{x | g_i(x) \geq 0, i = 1, \ldots, N_c\}$ is a set formed by a finite set of bounded-degree polynomial constraints. The $\alpha$-moment of a measure $\mu$ is $\mathbf{m}_\alpha = (c^{\alpha}, \mu)$. Assuming that each constraint polynomial $g_i(x)$ has a representation as $g_i(x) = \sum_{\sigma \in S}\gamma_{i,\sigma} x^\sigma$, then the matrix $\mathbf{M}_d(\mathbb{K}; \mathbf{m})$ formed by a moment sequence $\mathbf{m}$ is the block-diagonal matrix formed by diag($[\mathbf{m}_{\alpha,\beta}]$), with $\alpha, \beta \in S_+$. A polynomial $p(x)$ is SOS ($p(x) \in \Sigma[x]$) if there exists a finite integer $s$, a polynomial vector $v(x) \in \mathbb{R}[x]^s$, and a PSD matrix $Q \in \mathbb{S}_+^s$, such that $p(x) = v(x)^T Q v(x)$. SOS polynomials are nonnegative over $\mathbb{R}^n$. A polynomial is weighted sum of squares (WSOS) over a set $\mathbb{K}$ (expressed as $p(x) \in \Sigma[\mathbb{K}]$)

B. LMI Approximation

In the case where $c(x,y)$ and $f(t,x)$ are polynomial, (12) may be approximated with a converging hierarchy of SDPs. Assume that that $X_0$, $X$, and $X_u$ are Archimedean basic semialgebraic sets, each defined by a finite number of bounded-degree polynomial inequality constraints $X_0 = \{x | g_i^0(x) \geq 0\}$, $X = \{x | g_i^0(x) \geq 0\}$, and $X_u = \{x | g_i^u(x) \geq 0\}$.

The polynomial inequality constraints for $X_0$, $X$, and $X_u$ are degrees $d_k$, $d_k$, $d_k$, respectively. The Liouville equation in (12c) enforces a countably infinite set of linear constraints indexed by all possible $\alpha \in \mathbb{N}^n$, $\beta \in \mathbb{N}$.

$$
\langle x^\alpha, \mu_0 \rangle \leq \langle \mathcal{L}_f(x^\alpha t^{\beta}), \mu \rangle - \langle x^\alpha t^{\beta}, \mu_p \rangle = 0.
$$

(19)

The expression $\delta_{\beta} = 0$ is the Kronecker delta taking a value $\delta_{\beta} = 1$ when $\beta = 0$ and $\delta_{\beta} = 0$ when $\beta \neq 0$. Let ($m^0$, $m^p$, $m^\alpha$) be moment sequences for the measures ($\mu_0$, $\mu_p$, $\mu$). Define $\mathbf{m}_{\alpha} = (m^0, m^p, m^\alpha)$ as the linear relation induced by (19) at the test function $x^\alpha t^\beta$ in terms of moment sequences. The polynomial metric $c(x,y)$ may be expressed as $\sum_{\alpha, \gamma} c_{\alpha,\gamma} x^\alpha y^\gamma$ for multi-indices $\alpha, \gamma \in \mathbb{N}^n$. The complexity of dynamics $f$ induces a degree $d$ as $d = d + \lceil \deg(f)/2 \rceil - 1$. The degree-$d$ LMI relaxation of (12) with moment sequence variables ($m^0$, $m^p$, $m^\alpha$) is

$$
p_d = \min_{\alpha, \gamma} \sum_{\alpha, \gamma} c_{\alpha,\gamma} m^\alpha \gamma
$$

(20a)

$$
m^\alpha_0 = m^0_0 \quad \forall \alpha \in \mathbb{N}^n_d
$$

(20b)

$$
\text{Liou}_{\alpha, \beta}(m^0, m^p, m^\alpha) = 0 \quad \forall (\alpha, \beta) \in \mathbb{N}^{n+1}_d
$$

(20c)

$$
m^0_0 = 1
$$

(20d)

$$
\mathbf{m}_d(\mathbb{K}; m^0) \geq 0
$$

(20e)

$$
\mathbf{m}_d(\mathbb{K}; m^0) \geq 0
$$

(20f)

$$
\mathbf{m}_d(\mathbb{K}; m^0) \geq 0
$$

(20g)

$$
\mathbf{m}_d(\mathbb{K}; m^0) \geq 0
$$

(20h)

Constraints (20b)–(20d) are finite-dimensional versions of constraints (12b)–(12d) from the measure LP. In order to ensure convergence $\lim_{d \to \infty} p_d = p^*$, we must establish that all moments of measures are bounded.

Lemma 4.1: The masses of all measures in (12) are finite (uniformly bounded) if A1–A4 hold.

Proof: Constraint (12d) imposes that $\langle 1, \mu_0 \rangle = 1$, which further requires that $\langle 1, \mu_p \rangle = 1$ by constraint (12e) $(v(t,x) = 1)$ and $\langle 1, \mu_p \rangle = 1$ by constraint (12f) $(v(t,x) = 1)$. The occupation measure $\mu$ likewise has bounded mass with $\langle 1, \mu \rangle \leq T$ by constraint (12c) $(v(t,x) = t)$.

Lemma 4.2: The measures $\mu_0$, $\mu_p$, $\mu$ all have finite moments under Assumptions A1–A4.

Proof: A sufficient condition for a measure $\tau \in \mathcal{M}_+(X)$ with compact support to be bounded is to have finite mass $\langle 1, \tau \rangle$. In our case, the support of all measures $\mu_0$, $\mu_p$, $\mu$
are compact sets by A1. Further, under Assumptions A1–A4, all of these measures have bounded mass (Lemma 4.1).

This sufficiency is satisfied by all measures \((\mu_0, \mu_p, \mu, \eta)\).

**Theorem 4.1:** When \(T\) is finite and \(X_0, X, X_u\) are all Archimedean, the sequence of lower bounds \(p_d^{*} \leq p_{d+1}^{*} \leq p_{d+2}^{*} \ldots\) will approach \(p^*\) as \(d\) tends toward \(\infty\).

**Proof:** This convergence is assured by Corollary 8 of (22) under the Archimedean assumption and Lemma 4.2.

**Remark 5:** Nonpolynomial \(C^0\) cost functions \(c(x, y)\) may be approximated by polynomials \(\tilde{c}(x, y)\) through the Stone–Weierstrass theorem in the compact set \(X \times Y\). For every \(\epsilon > 0\), there exists a \(\tilde{c}(x, y) \in \mathbb{R}[x, y]\) such that \(\max_{x \in X, y \in Y} |c(x, y) - \tilde{c}(x, y)| \leq \epsilon\). Solving the peak estimation problem (12) with cost \(\tilde{c}(x, y)\) as \(\epsilon \to 0\) will yield convergent bounds to \(P^n\) with cost \(c(x, y)\). Section VIII-B offers an alternative peak estimation problem using polyhedral lifts for costs comprised by the maximum of a set of functions.

### C. Numerical Considerations

A moment matrix with \(n\) variables in degree \(d\) has dimension \((n+d)^2\). The sizes of moment matrices associated with a \(d\) relaxation of Program (20) with state \(x \in \mathbb{R}^n\), dynamics \(f(x, t)\), and induced dynamic degree \(\delta\), are listed in Table II.

The computational complexity of solving the SDP formulation of LMI (20) scales polynomially as the largest matrix size in Table II, usually \(M_d(m^n)\), except in cases where \(f(x, t)\) has a high polynomial degree.

**Remark 6:** The measures \(\mu_p\) and \(\eta\) may in principle be combined into a larger measure \(\tilde{\eta} \in M_{\gamma}(\mathbb{R})\) cost functions \(c(x, y)\) through the Stone–Weierstrass theorem in the compact set \(X \times Y\). For every \(\epsilon > 0\), there exists a \(\tilde{c}(x, y) \in \mathbb{R}[x, y]\) such that \(\max_{x \in X, y \in Y} |c(x, y) - \tilde{c}(x, y)| \leq \epsilon\). Solving the peak estimation problem (12) with cost \(\tilde{c}(x, y)\) as \(\epsilon \to 0\) will yield convergent bounds to \(P^n\) with cost \(c(x, y)\). Section VIII-B offers an alternative peak estimation problem using polyhedral lifts for costs comprised by the maximum of a set of functions.

### D. SOS Approximation

The degree-\(d\) WSOS truncation of program (18) is

\[
d_d^* = \max_{\gamma \in \mathbb{R}} \gamma \quad \text{subject to}\quad v(0, x) - \gamma \geq \sum_{x \in X} \mathds{1}_{x \in X} c(x, y) \in \Sigma[X, X] \leq d \quad \text{and} \quad w(x) - v(t, x) \in \Sigma[0, T] \times X \leq d \quad \text{and} \quad \mathcal{L}f v(t, x) \in \Sigma[0, T] \times X \leq d \quad \text{and} \quad w \in \mathbb{R}[x] \leq d \quad \text{and} \quad v \in \mathbb{R}[t, x] \leq d.
\]

**Theorem 4.2:** Strong duality holds with \(p_k^* = d_k^*\) for all \(k \in \mathbb{N}\) between (21) and (20) under assumptions A1–A5.

**Proof:** Refer to Corollary 8 of (22) (Archimedean condition and bounded masses), as well as to the proof of Theorem 4 and Lemma 4 in Appendix D of [11].

### V. Exploiting Correlative Sparsity

Many costs \(c(x, y)\) exhibit an additively separable structure, such that \(c\) can be decomposed into the sum of new terms \(c(x, y) = \sum_i c_i(x_i, y_i)\). Each term \(c_i\) in the sum is a function purely of \((x_i, y_i)\). Examples include the \(L_2\) family of distance functions, such as the squared \(L_2\) cost \(c(x, y) = \sum_i (x_i - y_i)^2\).

The theory of correlative sparsity in polynomial optimization, briefly reviewed in the following, can be used to substantially reduce the computational complexity entailed in solving the distance estimation SDP when \(c\) is additively separable [23].

This decomposition does not require prior structure on the set \(X \times X_u\). Other types of reducible structure (if applicable) include term sparsity [24], symmetry [25], and network dynamics [26]. These forms of structure may be combined if present, such as in correlative and term sparsity [27].

### A. Correlative Sparsity Background

Let \(K = \{x \mid g_k(x) \geq 0, k = 1, \ldots, N\}\) be an Archimedean basic semialgebraic set and \(\phi(x)\) be a polynomial. The correlative sparsity pattern (CSP) associated to \((\phi(x), g)\) is a graph \(G(V, E)\) with vertices \(V\) and edges \(E\). Each of the \(n\) vertices in \(V\) corresponds to a variable \(x_1, \ldots, x_n\). An edge \((x_i, x_j) \in E\) appears if variables \(x_i\) and \(x_j\) are multiplied together in a monomial in \(\phi(x)\), or if they appear together in at least one constraint \(g_k(x)\) [23].
The correlative sparsity pattern of \((\phi(x), g)\) may be characterized by sets \(I\) of variables and sets \(J\) of constraints. The \(p\) sets \(I\) should satisfy the following two properties.

1) (Coverage) \(\bigcup_{j=1}^{p} I_j = \mathcal{V}\)

2) (Running intersection property) for all \(k = 1, \ldots, p-1:\)
\[ I_{k+1} \cap \bigcup_{j=1}^{k} I_j \subseteq I_s \quad \text{for some} \ s \leq k \]

Equivalently, the sets \(I\) are the maximal cliques of a chordal extension of \(G(\mathcal{V}, \mathcal{E})\) [28]. The sets \(J = \{J_i\}_{i=1}^{n}\) are a partition over constraints \(g_k(x) \geq 0\). The number \(k\) is in \(J_i\) for \(k = 1, \ldots, N_X\) if all variables involved in the constraint polynomial \(g_k(x)\) are contained within the set \(I_i\). Let the notation \(x(I_i)\) denote the variables in \(x\) that are members of the set \(I_i\).

A sufficient sparse representation of positivity certificates may be developed for \((\phi(x), g)\) satisfying an admissible correlative sparsity pattern \((I, J)\) [29]

\[
\phi(x) = \sum_{i=1}^{p} \sigma_{a_0}(x(I_i)) + \sum_{k \in J_i} \sigma_k(x(I_i)) g_k(x) \\
\sigma_{a_0}(x) \in \Sigma[x(I_i)] \quad \sigma_k(x) \in \Sigma[x(I_i)] \quad \forall i = 1, \ldots, p. 
\]

(22)

The abovementioned equation is a sparse version of the Putinar certificate in (55). The sparse certificate (22) is additionally necessary for the \(G\)-sparse polynomial \(\phi(x)\) to be positive over \(\mathbb{K}\) if \((I, J)\) satisfies the running intersection property and a sparse Archimedean property: holds: that there exist finite constants \(R_i > 0\) for \(i = 1..n\) such that \(R_i^2 - |x(I_i)|^2\) is in the quadratic module (53) of constraints \(Q\{g_k\}_{k \in J_i}\) [29].

B. Correlative Sparsity for Distance Estimation

Constraint (18d) will exhibit correlative sparsity when \(c(x, y)\) is additively separable

\[
\sum_{i=1}^{n} c_i(x_i, y_i) - w(x) \geq 0 \quad \forall (x, y) \in X \times X_u. 
\]

(23)

The product-structure support set of (23) may be expressed as

\[
X \times X_u = \{(x, y) \mid g_1(x) \geq 0, \ldots, g_{N_X}(x) \geq 0, g_{N_X+1}(y) \geq 0, \ldots, g_{N_X+N_U}(y) \geq 0\}.
\]

(24)

The correlative sparsity graph of (23) is the graph Cartesian product of the complete graph \(K_n\) by \(K_2\), and is visualized at \(n = 4\) by the nodes and black lines in Fig. 7. Black lines imply that there is a link between variables. The black lines are drawn between each pair \((x_i, y_i)\) from the cost term \(c_i\). The polynomial \(w(x)\) involves mixed monominals of all variables \(x = (x_1, x_2, x_3, x_4)\). Prior knowledge on the constraints of \(X_u\) are not assumed in advance, so the variables in \((y) = (y_1, y_2, y_3, y_4)\) are joined together. A CSP \((I, J)\) associated with this system is

\[
I_1 = \{x_1, x_2, x_3, x_4, y_1\} \quad J_1 = \{1, \ldots, N_X\} \\
I_2 = \{x_2, x_3, x_4, y_1, y_2\} \quad J_2 = \emptyset \\
I_3 = \{x_3, y_4, y_2, y_3\} \quad J_3 = \emptyset 
\]

Fig. 7. CSP with 4-states and chordal extension.

\[
I_4 = \{x_4, y_1, y_2, y_3, y_4\} \quad J_4 = \{N_X + 1, \ldots, N_X + N_U\}
\]

A total of \((n-1)/2\) new edges are added in the chordal extension. Letting \(y_{1:i}\) be the collection of variables \((y_1, y_2, \ldots, y_i)\) for an index \(i \in 1..n\) (and with a similar definition for \(x_{1:n}\)), a correlative sparse certificate of positivity for constraint (18d) is

\[
\sum_{i=1}^{n} c_i(x_i, y_i) - w(x) = \sum_{i=1}^{n} \sigma_{a_0}(x_{1:n}, y_{1:i}) + \sum_{k=1}^{N_X} \sigma_k(x, y) g_k(x) + \sum_{k=N_X+1}^{N_X+N_U} \sigma_k(x, y) g_k(y)
\]

with SOS multipliers

\[
\sigma_{a_0}(x, y) \in \Sigma[x_{1:n}, y_{1:i}] \quad \forall i = 1, \ldots, p \\
\sigma_k(x, y) \in \Sigma[x_{1:n}] \quad \forall k = 1, \ldots, N_X \\
\sigma_k(x, y) \in \Sigma[x_{1:n}] \quad \forall k = N_X + 1, \ldots, N_X + N_U.
\]

(25)

The application of correlative sparsity to the distance problem replaces constraint (21c) by (26).

Remark 9: The CSP decomposition in (25) is nonunique. As an example, the following decompositions are all valid for \(n = 3\) (satisfy running intersection property):

\[
I_1 = \{x_1, x_2, x_3, y_1\} \quad I_1' = \{x_1, x_2, x_3, y_3\} \\
I_2 = \{x_2, x_3, y_1, y_3\} \quad I_2' = \{x_1, x_2, y_2, y_3\} \\
I_3 = \{x_2, y_1, y_2, y_3\} \quad I_3' = \{x_1, y_1, y_2, y_3\}.
\]
Angular Velocity = 0 rad/sec  Angular Velocity = 1 rad/sec

Fig. 8. Shape moving and rotating along flow (5) trajectories.

The original constraint (18d) is dual to the joint measure $\eta \in \mathcal{M}_+(X \times Y)$. Correlative sparsity may be applied to the measure program by splitting $\eta$ into new measures $\eta_i \in \mathcal{M}_+(X \times \mathbb{R})$, $\eta_n \in \mathcal{M}_+(\mathbb{R} \times X_u)$ and $\eta_l \in \mathcal{M}_+(\mathbb{R}^{n+1})$ for $i = 2, \ldots, n - 1$ following the procedure in [29]. These measures will align on overlaps with $\pi_l \eta_l \eta_i = \pi_l \eta_i \eta_l$, $\forall i = 1, \ldots, n - 1$. At a degree-$d$ relaxation, the moment matrix of $\eta$ in (20) has size $(2n^2 + d)$. Each of the $n$ moment matrices of $\{\eta_i\}_{i=1}^n$ has a size of $(n^2 + d + 1)$ for $n = 4$, $d = 4$ will have a moment matrix for $\eta$ of size $(4^2) = 49$ while the moment matrices for each of the $\eta_i$ are of size $(4_i) = 126$.

VI. SHAPE SAFETY

The distance estimation problem may be extended to sets or shapes travelling along trajectories, bounding the minimum distance between points on the shape and the unsafe set. An example application is in quantifying safety of rigid body dynamics, such as finding the closest distance between all points on an airplane and points on a mountain.

A. Shape Safety Background

Let $X \subset \mathbb{R}^n$ be a region of space with unsafe set $X_u$, and $c(x, y)$ be a distance function. The state $x \in \Omega$ (such as position and angular orientation) follows dynamics $\dot{x}(t) = f(x, t)$ between times $t \in [0, T]$. A trajectory is $x(t \mid \omega_0)$ for some initial state $\omega_0 \in \Omega_0 \subset \Omega$. The shape of the object is a set $S$. There exists a mapping $A(s; \omega) : S \times \Omega \to X$ that provides the transformation between local coordinates on the shape $(s)$ to global coordinates in $X$.

Examples of a shape traveling along trajectories are detailed in Fig. 8. The shape $S = [-0.1, 0.1]^2$ is the pink square. The left hand plot is a pure translation after a $5\pi/12$ radian rotation, and the right plot involves a rigid body transformation.

The distance estimation task with shapes is to bound

$$P^* = \inf_{t, \omega_0 \in \Omega_0, x \in X_u} \left\{ c \left( A(s; \omega(t \mid \omega_0)) , y \right) \right\}$$

$$\omega(t) = f(t, \omega), \quad \forall t \in [0, T].$$

(28)

For each trajectory in state $x(t \mid \omega_0)$, problem (28) ranges over all points in the shape $s \in S$ and points in the unsafe set $y \in X_u$ to find the closest approach. An optimal trajectory of the shape distance program may be expressed as $T^*_p = (y^*, s^*, \omega^*_p, t^*_p)$ with $\omega^*_p = \omega(t^*_p \mid \omega^*_0), x^*_p = A(s^*; \omega^*_p)$, and

$$P^* = c \left( A(s^*; \omega^*_p) , X_u \right) = c \left( A(s^*; \omega(t^*_p \mid \omega^*_0)) , y^* \right).$$

Remark 10: The objective in (28) can be expressed using

$$c_A(\omega; S, X_u) = \inf_{s \in S, y \in X_u} c(A(s; \omega), y)$$

as $c_A(\omega(t \mid \omega_0); S, X_u)$.

B. Assumptions

The following assumptions are made in the shape distance program (28).

A1’ : The sets $[0, T], \Omega, S, X, X_u$ are compact and $\Omega_0 \subset \Omega$.

A2’ : The function $f(t, \omega)$ is Lipschitz in each argument.

A3’ : The cost $c(x, y)$ is $C^0$.

A4’ : The coordinate transformation function $A(s; \omega) = C^0$.

A5’ : If $\omega(t \mid \omega_0) \in \partial \Omega$ for some $t \in [0, T]$, $\omega_0 \in \Omega_0$, then

$$\omega(t \mid \omega_0) \not\in \Omega \forall t \in (t, T].$$

A6’ : If $s \in S$ such that $A(s; \omega(t \mid \omega_0)) \not\in X$ or $A(s; \omega(t \mid \omega_0)) \in \partial X$ for some $t \in [0, T], \omega_0 \in \Omega_0$, then

$$A(s; \omega(t \mid \omega_0)) \not\in X \forall t \in (t, T].$$

An alternative assumption used instead of A5’–A6’ is that $\omega(t \mid \Omega_0)$ stays in $\Omega$ for all $\omega_0 \in \Omega_0$ and $A(s; \omega(t \mid \omega_0)) \in X$ for all $s \in S, t \in [0, T]$.

C. Shape Distance Measure Program

Program (28) involves a distance objective $c(x, y)$, where the point $x = A(s; \omega)$ is given by a coordinate transformation between body coordinates $s$ and the evolving orientation $\omega$. In order to formulate a measure program to (28), a shape measure $\mu_s \in \mathcal{M}_+(S \times \Omega)$ may be added to bridge the gap between the changing orientation $\omega$ and the comparison distance $x$. The shape measure contains information on the orientation $\omega$ and body coordinate $s$ that yields the closest point

$$\langle z(\omega), \mu_p(t, \omega) \rangle = \langle z(\omega), \mu_s(s, \omega) \rangle \quad \forall z \in C(\Omega)$$

(30a)

$$\langle w(x), \eta(x, y) \rangle = \langle w(A(s; \omega), \mu_s(s, \omega) \rangle \quad \forall w \in C(X).$$

(30b)

The shape measure $\mu_s$ chooses the worst-case body coordinate $s$ and orientation $\omega$ from $\mu_p$ (30a), such that the point $x = A(s; \omega)$ comes as close as possible to the unsafe set’s coordinate $y$ (30b). We retain the coordinate $x$ in order to decrease the computational complexity of the SDP, as elaborated upon further in Remark 6.

The infinite-dimensional measure program that lower bounds (28) is

$$p^* = \inf \langle c(x, y), \eta \rangle$$

(31a)

$$\mu_p = \delta_0 \otimes \mu + L_p^\mu$$

(31b)

$$\pi^\mu_p = \pi^\mu_s$$

(31c)

$$\pi^\mu_s = A(s; \omega)^\mu \otimes \mu_s$$

(31d)

$$\langle 1, \mu_0 \rangle = 1$$

(31e)

$$\mu_0 \in \mathcal{M}_+(\Omega_0), \quad \eta \in \mathcal{M}_+(X \times X_u)$$

(31f)

$$\mu_s \in \mathcal{M}_+(\Omega \times S)$$

(31g)
\begin{equation}
\mu_p, \mu \in \mathcal{M}_+ ([0, T] \times \Omega) .
\end{equation}

Constraint (12b) in the original distance formulation is now split between (31c) and (31d) [which are equivalent to (30b) and (30a)]. Problem (31) inherits all convergence and duality properties of the original (12) under the appropriately modified set of assumptions A1’–A6’.

Theorem 6.1: Under A3’–A4’ (and in addition A5’–A6’ when all sets in A1’ are compact possibly excluding [0, T]), the shape programs (28) and (31) are related by \( p^* \leq P^* \).

Proof: This proof will follow the same pattern as Theorem 3.1’s proof. A set of measures that are feasible solutions for the constraints of (31) may be constructed for any trajectory \( T = (y, s, \omega_0, t_p) \) with \( \omega_p = \omega(t_p \mid \omega_0) \), \( x_p = A(s; \omega_p) \). One choice of these measures are \( \mu_0 = \delta_{x=x_0} \), \( \mu_s = \delta_{x=s_0} \), \( \eta = \delta_{x=x_p} \delta_{y=y_p} \), \( \mu_\omega = \delta_{x=x_p} \delta_{y=\omega_0} \) and \( \mu \) as the occupation measure \( t \mapsto (t, \omega(t) \mid \omega_0) \) in times \([0, t_p^*] \). The feasible set of the constraints contains all trajectory-contracted measures, so \( p^* \leq P^* \).

Lemma 6.1: All measures in (31) have bounded mass under Assumption A1’.

Proof: This follows from the steps of Lemma 4.1. The conditions hold that 1 = \( (1, \mu_0) = (1, \mu_p) \) (31b), \( (1, \mu_p) = (1, \mu_s) \) (31c), \( (1, \mu_s) = (1, \eta) \) (31d), and \( (1, \mu) \leq T \) by (31b).

Lemma 6.2: The following peak estimation problem has the same optimal value as (28) under A1’–A6’:

\begin{align}
p^*_p &= \inf \{ c_A (\omega; S, X_u), \mu_p (t, \omega) \} \\
\mu_p &= \delta_{x=x_0} \otimes \mu_0 + L_f^\dagger \mu \\
(1, \mu_0) &= 1 \\
(1, \mu) &= 1 \\
(1, \mu) &= 1 \\
\mu_0 &\in \mathcal{M}_+ (\Omega_0), \eta \in \mathcal{M}_+ (X \times X_u) \\
\mu_s &\in \mathcal{M}_+ (X \times S) \\
\mu_p, \mu &\in \mathcal{M}_+ ([0, T] \times \Omega).
\end{align}

Proof: Refer to the proof of Lemma 3.1, with a shape-objective from (29).

Theorem 6.2: Under A1’–A6’, the optimal values of (31) and (28) are equal (\( P^* = p^* \)).

Proof: This proof repeats same process used in Theorem 3.2. Lemma 6.2 is used in place of Lemma 3.1. The reasoning of Lemma 3.2 is employed to construct infima-achieving measures \( \mu_s, \eta \) given a \( \mu_p \) from (32f) consistent with the marginal constraints (31c) and (31d).

D. Shape Distance Function Program

Defining a new dual function \( z(\omega) \) against constraint (31c) [also observed in (30a)], the Lagrangian of problem (31) is

\begin{align}
\mathcal{L} &= \langle c(x, y), \eta \rangle + \langle v(t, x), \delta_0 \otimes \mu_0 + L_f^\dagger \mu - \mu_p \rangle \\
&+ \langle z(\omega), \pi_\# (\mu_p - \mu_s) \rangle + \gamma (1 - \langle 1, \omega_0 \rangle) \\
&+ \langle v(x), A(s; \omega)_\# \mu_s - \pi_\# \eta \rangle.
\end{align}

The Lagrangian in (33) may be manipulated into

\begin{align}
\mathcal{L} &= \gamma + \langle c(x, y) - w(x), \eta \rangle + \langle v(0, \omega), - \gamma, \mu_0 \rangle \\
&+ \langle \mathcal{L}_f v(t, \omega), \mu \rangle + \langle z(\omega) - v(t, \omega), \mu_p \rangle \\
&+ \langle w(A; s; \omega) - z(\omega), \mu_s \rangle.
\end{align}

The dual program of (31) provided by minimizing the Lagrangian (34) with respect to \( (\eta, \mu_s, \mu_p, \mu_0) \) is

\begin{align}
d^* &= \sup_{\gamma \in \mathbb{R}} \gamma \\
v(0, \omega) &= \gamma \\
c(x, y) &= \sup_{\gamma \in \mathbb{R}} \gamma \\
w(A; s; \omega) &= \sup_{\gamma \in \mathbb{R}} \gamma \\
z(\omega) &= \sup_{\gamma \in \mathbb{R}} \gamma \\
\mathcal{L}_f v(t, \omega) &= \sup_{\gamma \in \mathbb{R}} \gamma \\
w \in C(X), \; z \in C(\Omega) \\
v \in C^1 ([0, T] \times \Omega).
\end{align}

Theorem 6.3: Problems (31) and (35) are strongly dual under assumptions A1’–A6’.

Proof: This holds by extending the proof of Theorem 3.3 found in Appendix I and applying Theorem 2.6 of [30].

Remark 11: Program (35) imposes that a chain of lower bounds \( v(t, \omega) \leq z(\omega) \leq w(A; s; \omega) \) holds for all \( s, \omega, t, y \) in \( X \times S \times X_u \) (similar in principle to Remark 4).

Remark 12: We briefly note that the LMI formulation of (31) will converge to \( P^* \) under assumptions A1’–A6’ if all \( [0, T], X, X_u, \Omega_0, \Omega, S \) are Archimedean and if \( f(t, \omega) \) is Archimedean and if \( A(s; \omega) \) is Archimedean (from Theorem 4.1). Constraints (30b) induces a linear expression in moments for \( (\mu_\#, \mu) \) for each \( \alpha \in \mathbb{N}_n \) : \( \langle x^{\alpha}, \eta \rangle = \langle A(s; \omega)^\alpha, \mu_s \rangle \).

Remark 13: If \( A(s; \omega) \) is polynomial with degree \( k \), then the \( d \)-degree relaxation of problem (31) involves moments of \( \mu_s \) up to order \( 2kd \). For a system with \( \mathcal{N}_r \) orientation states and \( n_s \) shape variables, the size of the moment matrix for \( \mu_s \) is then \( (N_s + N_w + n + r \times d) \). LMI constraints associated with \( \mu_s \) can become bottlenecks to computation, surpassing the contributions of \( \mu \) and \( \eta \) as \( k \) increases.

Remark 14: Continuing the discussion Remark 6, the measures \( \mu_s, \eta \) may be combined together into a larger measure \( \eta_s, \eta \) in \( \mathcal{M}_+(S \times X_u \times \Omega) \) with objective \( \inf \langle c(A; s; \omega), \eta_s \rangle \) and constraint \( \pi_\# \pi_\# \eta = \pi_\# \eta \). The moment matrix for \( \eta_s \) would have the generally tractable size \( (N_s + N_w + n + r \times d) \).

VII. NUMERICAL EXAMPLES

All code was written in MATLAB 2021a, and is publicly available at. The SDPs were formulated by Gloptipoly3 [31] through a Yalmip interface [32], and were solved using Mosek [33]. The experimental platform was an Intel i9 CPU with a clock frequency of 2.50 GHz and 64.0 GB of RAM. The squared-\( L_2 \) cost \( c(x, y) = \sum_i (x_i - \bar{y}_i)^2 \) is used in solving Problem (20) unless otherwise specified. The documented bounds are the

\[ \text{[Online]. Available: https://github.com/Jarmill/distance} \]
square roots of the returned quantities, yielding lower bounds to the $L_2$ distance.

### A. Flow System With Moon

The half-circle unsafe set in Fig. 6 is a convex set. The moon-shaped unsafe set $X_u$ in Fig. 9 is the nonconvex region outside the circle with radius 1.16 centered at $(0.6596, 0.3989)$ and inside the circle with radius 0.8 centered at $(0.4, -0.4)$. The dotted red line demonstrates that trajectories of the flow system would be deemed unsafe if $X_u$ was relaxed to its convex hull.

The $L_2$ distance bound of 0.1592 in Fig. 10 was found at the degree-5 relaxation of Problem (20) with $X = [-3, 3]^2$. The moment matrices $M_q(m^p), M_d(m^p), M_d(m^n)$ at $d = 5$ were approximately rank-1, and near-optimal trajectories were successfully extracted. This near-optimal trajectory starts at $x_0^* \approx (1.489, -0.3998)$ and reaches a closest distance between $x_p^* \approx (1.113, -0.4956)$ and $y^* \approx (1.161, -0.6472)$ at time $t_p^* \approx 0.1727$. The distance bounds computed at the first five relaxations are $L_2^{1:5} = [1.487 \times 10^{-4}, 2.433 \times 10^{-4}, 0.1501, 0.1592, 0.1592]$.

### B. Twist System

The Twist system is a three-dimensional dynamical system parameterized by matrices $A$ and $B$

$$\dot{x}_i(t) = \sum_j A_{ij}x_j - B_{ij} \left(4x_j^3 - 3x_j \right)/2$$

(A)

The cyan curves in each panel of Fig. 11 are plots of trajectories of the twist system between times $t \in [0, 5]$. These trajectories start at the $X_0 = \{x | (x_1 + 0.5)^2 + x_2^2 + x_3^2 \leq 0.2^2\}$, which is pictured by the grey spheres. The unsafe set $X_u = \{x | (x_1 - 0.25)^2 + x_2^2 + x_3^2 \leq 0.2^2, x_3 \leq 0\}$ is drawn in the red half-spheres. The underlying space is $X = [-1, 1]^3$.

The red shell in Fig. 11(a) is the cloud of points within an $L_2$ distance of 0.0427 of $X_u$, as found through a degree-5 relaxation of (20). Fig. 11(b) involves an $L_4$ contour of 0.0411, also found at order 5. The first few distance bounds for the $L_2$ distance are $L_2^{1:5} = [0, 0.0.0336, 0.0425, 0.0427]$, and for the $L_4$ distance are $L_4^{2:5} = [0, 0.0298, 0.0408, 0.0413]$. Fourth degree moments are required for the $L_4$ metric, so the $L_4^{2:5}$ sequence starts at order 2.

Tables III and IV lists the $L_2$ bounds and runtimes, respectively, generated by a distance estimation task between trajectories and the half sphere of the above $L_2$ twist system example. The high-degree relaxations (orders 4 and 5) are significantly faster as found by solving the SDP associated with the sparse LMI [dual to the sparse SOS with Putinar expression (26)] as compared to the standard program (20). The certifiable $L_2$ bounds returned are roughly equivalent between relaxations.

### C. Shape Examples

Fig. 12 visualizes a near-optimal trajectory of the shape distance estimation for orientations $\omega \in \mathbb{R}^2$ evolving as the flow system with an initial condition $\Omega_0 = \{\omega : (\omega_1 - 1.5)^2 + \omega_2^2 \leq 0.4^2\}$ in the space $\Omega : (\omega_1, \omega_2) \in [-3, 3]^2$, $\omega_2^2 + \omega_4^2 = 1$ (with

---

**Table III**

| order | 2 | 3 | 4 | 5 | 6 |
|-------|---|---|---|---|---|
| Standard LMI (20) | 0.000 | 0.0313 | 0.0425 | 0.0429 | 0.0429 |
| Sparse LMI with (26) | 0.000 | 0.0311 | 0.0424 | 0.0430 | 0.0429 |

**Table IV**

| order | 2 | 3 | 4 | 5 | 6 |
|-------|---|---|---|---|---|
| Standard LMI (20) | 0.32 | 1.92 | 47.55 | 502.29 | 4028.94 |
| Sparse LMI with (26) | 0.31 | 1.19 | 7.07 | 43.89 | 184.42 |
a state set of \( X = [-3, 3]^2 \). Suboptimal trajectories were suppressed in visualization to highlight the shape structure and attributes of the near-optimal trajectory. The degree-1 coordinate transformation function \( A \) for pure translation with a constant rotation of \( 5\pi/12 \) is

\[
A(s; \omega) = \begin{bmatrix}
\cos(5\pi/12)s_1 - \sin(5\pi/12)s_2 + \omega_1 \\
\cos(5\pi/12)s_1 + \sin(5\pi/12)s_2 + \omega_2
\end{bmatrix}.
\] (38)

This near-optimal trajectory with an \( L_2 \) distance bound of 0.1465 was found at a degree-4 relaxation of Problem (31). The near-optimal trajectory is described by \( \omega_0^* \approx (1.489, -0.3887) \), \( t_p^* \approx 3.090 \), \( \omega_0^* \approx (-0.1225, -0.3704) \), \( s^* \approx (-0.1, 0.1) \), \( x_p^* \approx (0, -0.2997) \), and \( y^* \approx (-0.2261, -0.4739) \). The first five distance bounds are \( L_2^{1:5} = [1.205 \times 10^{-4}, 4.245 \times 10^{-4}, 0.1424, 0.1465, 0.1465] \).

In the following example, the shape \( S \) is now rotating at an angular velocity of \( 1 \text{ rad/s} \), as shown in the right panel of Fig. 8. The orientation \( \omega \in SE(2) \) may be expressed as a semialgebraic lift through \( \omega \in \mathbb{R}^4 \) with trigonometric terms \( \omega_0^2 + \omega_2^2 = 1 \). The dynamics for this system are

\[
\dot{\omega} = \begin{bmatrix}
\omega_2; & -\omega_1 - \omega_2 + \frac{1}{2}\omega_0^2; & -\omega_4; & \omega_3
\end{bmatrix}.
\] (39)

The degree-2 coordinate transformation associated with this orientation is

\[
A(s; \omega) = \begin{bmatrix}
\omega_3s_1 - \omega_4s_2 + \omega_1 \\
\omega_3s_1 + \omega_4s_2 + \omega_2
\end{bmatrix}.
\] (40)

The shape measure \( \mu_s \in M_+(S \times \Omega) \) is distributed over six variables. The size of \( \mu_s \)’s moment matrix with \( h = 2 \) at degrees 1–4 is \([28, 210, 924, 3003]\). The first three distance bounds are \( L_2^{1:3} = [2.9158 \times 10^{-5}, 0.059162, 0.14255] \), and MATLAB runs out of memory on the experimental platform at degree 4. A successful recovery is achieved at the degree-3 relaxation, as pictured in Fig. 13. This rotating-set near-optimal trajectory is encoded by \( \omega_0^* \approx (1.5755, -0.3928, 0.2588, 0.9659) \), \( t_p^* \approx 3.371 \), \( s^* \approx (-0.1, 0.1) \), \( x_p^* \approx (-0.1096, -0.3998) \), \( \omega_0^* \approx (-0.0064, -0.2921, -0.0322, -0.9995) \), and \( y^* \approx (-0.2104, -0.4896) \). Computing this degree-3 relaxation required 75.43 min.

### VIII. Extensions

This section presents modifications to the distance estimation programs in order to handle systems with uncertainties and distance functions \( c \) generated by polyhedral norms.

#### A. Uncertainty

Distance estimation can be extended to systems with uncertainty. For the sake of simplicity, this section is restricted to time-dependent uncertainty. Assume that \( H \subset \mathbb{R}^{n_h} \) is a compact set of plausible values of uncertainty, and that the uncertain process \( h(t), \forall t \in [0, T] \) may change arbitrarily in time within \( H \) [34]. The distance estimation problem with time-dependent uncertain dynamics is

\[
P^* = \inf_{t, x_0, y, h(t)} c(x(t | x_0, h(t)), y)
\] (41)

\[
\dot{x}(t) = f(t, x, h(t)), h(t) \in H \quad \forall t \in [0, T] 
\]

\[
x_0 \in X_0, \ y \in X_y.
\]

The process \( h(t) \) acts as an adversarial optimal control that aims to steer \( x(t) \) as close to \( X_0 \) as possible. The occupation measure \( \mu \) may be extended to a Young measure (relaxed control) \( \mu \in M_+(\{(0, T) \times \mathbb{R} \times H\}) \) [10], [35].

The Liouville (12c) may be replaced by \( \mu_p = \delta_0 \otimes \mu_0 + \pi_{\#}^T \mathcal{L}_f \mu \), which should be understood to read \( \langle v(t, x), \mu_p \rangle = \langle v(0, x), \mu_0 \rangle + \langle h^* v(t, x) \rangle + f(t, x, h) \cdot \nabla v(t, x), \mu \rangle \) for all test functions \( v \in C^1([0, T] \times \mathbb{R}) \). Any trajectory with uncertainty process \( h(t) \) may be represented by a tuple \( (x_0, x_p, t_p, y, h(\cdot)) \). This trajectory admits a measure representation similar to the proof of 3.1, where the measure \( \mu \) is the occupation measure of \( t \to (x(t | x_0, h(t))) \) in times \( [0, t_p] \).

The work in [34] applies a collection of existing uncertainty structures to peak estimation problems (time-independent, time-dependent, switching-type, box-type), and all of these structures may be applied to distance estimation.

To illustrate these ideas, consider the following flow system with time-dependent uncertainty:

\[
\dot{x} = \begin{bmatrix}
x_2 \\
(-1 + h)x_1 - x_2 + \frac{1}{3}x_1^3
\end{bmatrix}, \quad h \in [-0.25, 0.25].
\] (42)

An \( L_2 \) distance bound of 0.1691 is computed at the degree-5 relaxation of the uncertain distance estimation program, as visualized in Fig. 14. The first five distance bounds are \( L_2^{1:5} = [5.125 \times 10^{-5}, 1.487 \times 10^{-4}, 0.1609, 0.1688, 0.1691] \).

#### B. Polyhedral Norm Penalties

The infinite-dimensional LP (12) is valid for all continuous costs \( c(x, y) \in C(\mathbb{R}^2) \), but its LMI relaxation can only handle polynomial costs \( c(x, y) \in \mathbb{R}[x, y] \). The \( L_p \) distance is defined as
The LP objective is equal to the distance of closest approach between points along trajectories and points on the unsafe set under mild compactness and regularity conditions. Finite-dimensional truncations of this LP yield a converging sequence of SDP lower bounds to the minimal distance under further conditions (Archimedean). The distance estimation problem can be modified to accommodate dynamics with uncertainty, piecewise distance functions, and movement of shapes along trajectories. Future work includes formulating and implementing control policies to maximize the distance of closest approach to the unsafe set while still reaching a terminal set within a specified time.

**APPENDIX I**

**PROOF OF STRONG DUALITY IN THEOREM 3.3**

This proof will follow the method used in Theorem 2.6 of [30] to prove duality. The two programs (12) and (18) will be posed as a pair of standard-form infinite-dimensional LPs using notation from [30]. The following spaces may be defined:

\[ \mathcal{X}' = C(X_0) \times C([0, T] \times X)^2 \times C(X \times X_u) \]

\[ \mathcal{X} = \mathcal{M}(X_0) \times \mathcal{M}([0, T] \times X)^2 \times \mathcal{M}(X \times X_u) . \]

(44)

The nonnegative subcones of \( \mathcal{X}' \) and \( \mathcal{X} \), respectively, are:

\[ \mathcal{X}'_+ = C_+(X_0) \times C_+([0, T] \times X)^2 \times C_+(X \times X_u) \]

\[ \mathcal{X}_+ = \mathcal{M}_+(X_0) \times \mathcal{M}_+([0, T] \times X)^2 \times \mathcal{M}_+(X \times X_u) . \]

(45)

The cones \( \mathcal{X}'_+ \) and \( \mathcal{X}_+ \) in (45) are topological duals under assumption A1, and the measures from (12e) and (12f) satisfy \( \mu = (\mu_0, \mu_p, \mu, \eta) \in \mathcal{X}_+ \). The spaces \( \mathcal{Y} \) and \( \mathcal{Y}' \) may be defined as:

\[ \mathcal{Y}' = C(X) \times C^1([0, T] \times X) \times \mathbb{R} \]

\[ \mathcal{Y} = \mathcal{M}(X) \times C^1([0, T] \times X)' \times 0 . \]

(46)

(47)

We express \( \mathcal{Y}_+ = \mathcal{Y} \) and \( \mathcal{Y}'_+ = \mathcal{Y}' \) to maintain a convention with [30] given there are no affine-inequality constraints in (12). We equip \( \mathcal{X} \) with the weak-* topology and \( \mathcal{Y} \) with the (sup-norm bounded) weak topology. The arguments \( \ell = (w, v, \gamma) \) from problem (18) are members of the set \( \mathcal{Y}'_+ \).

The linear operators \( A' : \mathcal{Y}'_+ \to \mathcal{X}'_+ \) and \( A : \mathcal{X}_+ \to \mathcal{Y}_+ \) induced from constraints (12b)-(12d) may be defined as:

\[ A(\mu) = \begin{bmatrix} \pi^\mu \mu_p - \pi^\mu_\eta \delta_0 \otimes \mu_0 + L_f \mu - \mu_p, 1, \mu_0 \end{bmatrix} \]

\[ A'(\ell) = \begin{bmatrix} v(0, x) - \gamma, w(x) - v(t, x), L_f v(t, x) - w(x) \end{bmatrix} . \]

(48)

The last pieces needed to convert (12) into a standard-form LP are the cost vector \( c = [0, 0, 0, c(x, y)] \) and the answer vector \( b = [0, 0, 1] \in \mathcal{Y}' \). Problem (12) is therefore equivalent to (with \( \langle c, \mu \rangle = \langle c, \eta \rangle \))

\[ p^* = \inf_{\mu \in \mathcal{X}_+} \langle c, \mu \rangle b - A(\mu) \in \mathcal{Y}_+. \]

(49)

**IX. CONCLUSION**

This article presented an infinite-dimensional LP in occupation measures to approximate the distance estimation problem.
The dual LP to \( (49) \) in standard form is (with \( \langle \ell, b \rangle = \gamma \))
\[
d^* = \sup_{\ell \in \mathcal{Y}_+} \langle \ell, b \rangle - c \in \mathcal{X}_+.
\] (50)

The operators \( A \) and \( A' \) are adjoints with \( \langle A(\ell), \mu \rangle = \langle \ell, A'(\mu) \rangle \) for all \( \ell \in \mathcal{Y}_+ \) and \( \mu \in \mathcal{X}_+ \).

The sufficient conditions for strong duality and attainment of optimality between \( (49) \) and \( (50) \) as outlined in Theorem 2.6 of [30] are as follows.

**R1:** All support sets are compact (A1).

**R2:** All measure solutions have bounded mass (Lemma 4.1).

**R3:** All functions involved in the definitions of \( c \) and \( A \) are continuous (A2, A3).

**R4:** There exists a \( \mu_{\text{feas}} \in \mathcal{X}_+ \) with \( b - A(\mu_{\text{feas}}) \in \mathcal{Y}_+ \).

The requirements \( R1 \) and \( R2 \) hold by Assumption A1 and Lemma 4.1, respectively. \( R3 \) is valid given that \( c(x, y) \) is \( C^0 \) (A3), the projection map \( \pi^\star \) is continuous, and the mapping \( (t, x) \mapsto L_y v(t, x) \) is \( C^0 \) for \( v \in C^1 \) and \( f \) Lipschitz (continuous) (A2). A feasible measure \( \mu_{\text{feas}} \) may be constructed from the process in Theorem 3.1 from a tuple \( T \), therefore satisfying \( R4 \).

Strong duality between (12) and (18) is therefore proven after satisfaction of all four requirements.

### Appendix II

**MOMENT-SOS HIERARCHY**

The standard form for a measure LP with variable \( \mu \in \mathcal{M}_+(X) \) involving a cost function \( p \in C(X) \) and a (possibly infinite) set of affine constraints \( \{a_j, \mu\} = b_j \) with \( b_j \in \mathbb{R} \) and \( a_j \in C(X) \) for \( j = 1, \ldots, J_{\text{max}} \) is
\[
p^* = \sup_{\mu \in \mathcal{M}_+(X)} \langle p, \mu \rangle \quad (51a)
\]
\[
\langle a_j(x), \mu \rangle = b_j \quad \forall j = 1, \ldots, J_{\text{max}}. \quad (51b)
\]

The dual problem to Program (51) with dual variables \( v_j \in \mathbb{R} : \forall j = 1, \ldots, m \) is
\[
d^* = \inf_{v \in \mathbb{R}^m} \sum_j b_j v_j \quad (52a)
\]
\[
p(x) - \sum_j a_j(x) v_j \geq 0 \quad \forall x \in X. \quad (52b)
\]

The objectives in (51) and (52) will match \( (p^* = d^* \text{ strong duality}) \) if \( p^* \) is finite and if the mapping \( \mu \rightarrow \{\langle a_j(x), \mu \rangle\}_{j=1}^{J_{\text{max}}} \) is closed in the weak-* topology (Theorem 3.10 in [39]).

When \( p(x) \) and all \( a_j(x) \) are polynomial, constraint (52b) is a polynomial nonnegativity constraint. The restriction that a polynomial \( q(x) \in \mathbb{R}[x] \) is nonnegative over \( \mathbb{R}^n \) may be strengthened to finding a set of polynomials \( \{q_i(x)\} \) such that \( \sum q_i(x)^2 \) is a SOS certificate of nonnegativity of \( q(x) \). The set of all polynomials \( q(x) \) with \( q_i(x) \) at each \( i \) and \( x \) is nonnegative. The set of Sum of Squares (SOS) polynomials in indeterminate quantities \( x \) is expressed as \( \Sigma[x] \), with a maximal-degree \( d \)-subset of \( \Sigma[x] \).

The quadratic module \( Q[g] \) formed by the constraints describing the basic semialgebraic set \( \mathbb{K} = \{ x \mid g_i(x) \geq 0, \ i = 1, \ldots, N_c \} \) is the set of polynomials
\[
Q[g] = \left\{ \sigma_0(x) + \sum_{i=1}^{N_c} \sigma_i(x) g_i(x) \right\}. \quad (53)
\]
such that the multipliers \( \sigma \) are SOS
\[
\sigma_i(x) \in \Sigma[x] \quad \forall i = 0, \ldots, N_c. \quad (54)
\]

The basic semialgebraic set \( \mathbb{K} \) is compact if there exists a constant \( 0 < R < \infty \) such that \( \mathbb{K} \) is contained in the ball \( R \leq \|x\|_2 \). \( \mathbb{K} \) satisfies the Archimedean property if the polynomial \( R - \|x\|_2^2 \) is a member of \( Q[g] \). The Archimedean property is stronger than compactness [40], and compact sets may be rendered Archimedean by adding a redundant ball constraint \( R - \|x\|_2 \geq 0 \) to the list of constraints describing in \( \mathbb{K} \) (though finding such an \( R \) may be difficult). When \( \mathbb{K} \) is Archimedean, every polynomial satisfying \( p(x) > 0, \forall x \in \mathbb{K} \) has a representation (Putinar’s Positivestellensatz [41])
\[
p(x) = \sigma_0(x) + \sum_i \sigma_i(x) g_i(x) \quad (55)
\]
\[
\sigma_0(x) \in \Sigma[x] \quad \sigma_i(x) \in \Sigma[x].
\]

The WSOS set \( \Sigma[\mathbb{K}] \) is the set of polynomials that admit a positivity certificate over \( \mathbb{K} \) from (55). Its maximal degree-\( d \)-subset is \( \Sigma[\mathbb{K}]_{\leq d} \). Given a multi-index \( \alpha \in \mathbb{N}^n \), the multi-moment of a measure \( \mu \in \mathcal{M}_+(X) \) is \( m_{\alpha} = \langle x^\alpha, \mu \rangle \) indexed by monomials \( \alpha \in \mathbb{N}^n \) may be constructed from the moment sequence \( m \).

The degree-\( d \)-moment matrix \( M_d[m] \) of size \( (n+d) \times (n+d) \) is the submatrix of \( M[m] \) where the indices \( \{\alpha, \beta\}_d \) have total degree bounded by \( 0 \leq |\alpha|, |\beta| \leq d \). Given a polynomial \( g(x) \in \mathbb{R}[x] \), the localizing matrix associated with \( g \) is a square infinite-dimensional symmetric matrix with entries \( M_{\{g(m)_{\alpha, \beta} = \sum_{j=1}^m g_j m_{\alpha, \beta+j} \}}. \) A moment sequence \( m \) has a representing measure \( \mu \in \mathcal{M}_+(\mathbb{K}) \) if there exists \( \mu \) supported in \( \mathbb{K} \) such that \( m_{\alpha} = \langle x^\alpha, \mu \rangle \forall \alpha \in \mathbb{N}^n \). The LMI conditions that \( M[m] \succeq 0 \) and \( M_{g[m]} \succeq 0 \forall i = 1, \ldots, N_c \) are necessary to guarantee the existence of a representing measure associated with \( m \). The moment matrix \( M[m] \) is a localizing matrix with the function \( g = 1 \). These LMI conditions are sufficient if the set \( \mathbb{K} \) is Archimedean, and all compact sets may be rendered Archimedean through the application of a redundant ball constraint [41].

Assume that each polynomial \( g_i(x) \) in the constraints of \( \mathbb{K} \) has a degree \( d_i \). We define a block-diagonal matrix \( M_d[\mathbb{K} m] \) containing the moment and all localizing matrices as
\[
\text{diag}(M_d[m], \{M_{d-d_i}[g_i m] \forall i = 1, \ldots, N_c \}). \quad (56)
\]

The degree-\( d \)-moment relaxation of Problem (51) with variables \( y \in \mathbb{R}^{n+d} \) is
\[
p^*_d = \max_{m \in \mathbb{K} m} \sum_{\alpha} p_\alpha m_\alpha, \quad M_d[\mathbb{K} m] \succeq 0 \quad (57a)
\]
\[
\sum_{\alpha} a_{j\alpha} m_\alpha = b_j \quad \forall j = 1, \ldots, m. \quad (57b)
\]
The bound $p_d^* \geq p^*$ is an upper bound for the infinite-dimensional measure LP. The decreasing sequence of upper bounds $p_d^* \geq p_{d+1}^* \geq \ldots \geq p^*$ is convergent to $p^*$ as $d \to \infty$ if $\mathcal{K}$ is Archimedean. The dual SDP to (57a) is the degree-$d$ SOS relaxation of (52)

$$d_d^* = \min_{\mathbf{m} \in \mathbb{R}^m} \sum_j b_j v_j$$ (58a)

$$p(x) - \sum_j a_j(x)v_j = \sigma_0(x) + \sum_k \sigma_k(x)g_k(x)$$ (58b)

$$\sigma(x) \in \Sigma[x]_{\leq 2d}$$ (58c)

$$\sigma_k(x) \in \Sigma[x]_{\leq 2d-(\deg g_k/2)} \forall i \in 1, \ldots, N_c.$$ (58d)

We use the convention that the degree-$d$ SOS tightening of (58) involves polynomials of maximal degree $2d$. When the moment sequence $\mathbf{m}_\alpha$ is bounded ($\|\mathbf{m}_\alpha\|_1 < \infty \forall |\alpha| \leq 2d$) and there exists an interior point of the affine measure constraints in (51b), then the finite-dimensional truncations (57a) and (58) will also satisfy strong duality $p_d^* = d_d^*$ at each degree $k$ (by arguments from Appendix D/Theorem 4 of [11] using Theorem 5 of [42], also applied in Corollary 8 of [22]). The sequence of upper bounds (outer approximations) $p_d^* \geq p_{d+1}^* \geq \ldots$ computed from SDP is called the moment-SOS hierarchy.

ACKNOWLEDGMENT

The authors would like to thank Didier Henrion, Victor Magron, Matteo Tacchi, and the POP group at LAAS-CNRS for many technical discussions and suggestions. We are also grateful to the anonymous reviewers for their many suggestions to improve the original manuscript.

REFERENCES

[1] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, Linear Matrix Inequalities in System and Control Theory. Philadelphia, PA, USA: SIAM, 1994, vol. 15.
[2] J. B. Lasserre, Moments, Positive Polynomials And Their Applications, (Series Imperial College Press Optimization Series). Singapore: World Scientific, 2009.
[3] K. Helmes, S. Röhl, and R. H. Stockbridge, “Computing moments of the exit time distribution for Markov processes by linear programming,” Operations Res., vol. 49, no. 4, pp. 516–530, 2001.
[4] M. J. Cho and R. H. Stockbridge, “Linear programming formulation for optimal stopping problems,” SIAM J. Control Optim., vol. 40, no. 6, pp. 1965–1982, 2002.
[5] G. Fantuzzi and D. Goluskin, “Bounding extreme events in nonlinear dynamics using convex optimization,” SIAM J. Appl. Dynamical Syst., vol. 19, no. 3, pp. 1823–1864, 2020.
[6] C. Villani, Optimal Transport: Old and New. Berlin, Germany: Springer, 2008, vol. 338.
[7] F. Santambrogio, “Optimal transport for applied mathematicians,” Calculus Variations, PDEs, Model., vol. 55, no. 58–63, p. 94, 2015.
[8] G. Peyré and M. Cuturi, “Computational optimal transport: With applications to data science,” Foundations Trends Mach. Learn., vol. 11, no. 5/6, pp. 355–607, 2019.
[9] R. Lewis and R. Vinter, “Relaxation of optimal control problems to equivalent convex programs,” J. Math. Anal. Appl., vol. 74, no. 2, pp. 475–493, 1980.
[10] D. Henrion, J. B. Lasserre, and C. Savorgnan, “Nonlinear optimal control synthesis via occupation measures,” in Proc. IEEE 47th Conf. Decis. Control, 2008, pp. 4749–4754.
[11] D. Henrion and M. Korda, “Convex computation of the region of attraction of polynomial control systems,” IEEE Trans. Autom. Control, vol. 59, no. 2, pp. 297–312, Feb. 2014.
[12] M. Korda, D. Henrion, and C. N. Jones, “Inner approximations of the region of attraction for polynomial dynamical systems,” IFAC Proc. Volumes, vol. 46, no. 23, pp. 534–539, 2013.
[13] M. Korda, D. Henrion, and C. Jones, “Convex computation of the maximum controlled invariant set for polynomial control systems,” SIAM J. Control Optim., vol. 52, no. 5, pp. 2944–2960, 2014.
[14] S. Prajna and A. Jadbabaie, “Safety verification of hybrid systems using barrier certificates,” in Proc. Int. Workshop Hybrid Syst., Computation and Control, 2004, pp. 477–492.
[15] S. Prajna, “Barrier certificates for nonlinear model validation,” Automat.ica, vol. 42, no. 1, pp. 117–126, 2006.
[16] A. Rantzer and S. Prajna, “On analysis and synthesis of safe control laws,” in Proc. 42nd Allerton Conf. Communication, Control, Comput., 2004, pp. 1468–1476.
[17] J. Miller, D. Henrion, and M. Sznaier, “Peak estimation recovery and safety analysis,” IEEE Contr. Syst. Lett., vol. 5, no. 6, pp. 1982–1987, Dec. 2021.
[18] J. Miller and M. Sznaier, “Bounding the distance of closest approach to unsafe sets with occupation measures,” in Proc. IEEE 61st Conf. Decis. Control, 2022, pp. 5008–5013.
[19] M. M. Deza and E. Deza, Encyclopedia of Distances. Berlin, Heidelberg: Springer, 2009, pp. 1–583.
[20] H. García, C. Hernández, M. Junca, and M. Velasco, “Approximate super-resolution of positive measures in all dimensions,” Appl. Comput. Harmon. Anal., vol. 52, pp. 251–278, 2021.
[21] J. B. Lasserre, “A joint marginal” approach to parametric polynomial optimization,” SIAM J. Optim., vol. 20, no. 4, pp. 1995–2022, 2010.
[22] M. Tacchi, “Convergence of Lasserre’s hierarchy: The general case,” Optim. Lett., vol. 16, no. 3, pp. 1015–1033, 2022.
[23] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, “Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity,” SIAM J. Optim., vol. 17, no. 1, pp. 218–242, 2006.
[24] J. Wang, V. Magron, and J.-B. Lasserre, “TSSOS: A Moment-SOS hierarchy that exploits term sparsity,” SIAM J. Optim., vol. 31, no. 1, pp. 30–58, 2021.
[25] C. Rienger, T. Theobald, L. J. André, and J. B. Lasserre, “Exploiting symmetries in SDP-Relaxations for polynomial optimization,” Math. Operations Res., vol. 38, no. 1, pp. 122–141, 2013.
[26] C. Schlosser and M. Korda, “Sparse moment-sum-of-squares relaxations for nonlinear dynamical systems with guaranteed convergence,” 2020, arXiv:2012.05572.
[27] J. Wang, V. Magron, and J.-B. Lasserre, “Chordal-TSSOS: A Moment-SOS hierarchy that exploits term sparsity with chordal extension,” SIAM J. Optim., vol. 31, no. 1, pp. 114–141, 2021.
[28] H. Vandenberghe and M. S. Andersen, “Chordal graphs and semidefinite optimization,” Foundations Trends Optim., vol. 1, no. 4, pp. 241–433, 2015.
[29] J. B. Lasserre, “Convergent SDP-relaxations in polynomial optimization with sparsity,” SIAM J. Optim., vol. 17, no. 3, pp. 822–843, 2006.
[30] M. Tacchi, “Moment-SOS hierarchy for large scale set approximation. application to power systems transient stability analysis,” Ph.D. dissertation, INSA, Toulouse, France, 2021.
[31] D. Henrion and J.-B. Lasserre, “GloptiPoly: Global optimization over polynomials with matlab and SeDuMi,” ACM Trans. Math. Softw., vol. 29, no. 2, pp. 165–194, 2003.
[32] J. Löfberg, “YALMIP : A toolbox for modeling and optimization in MATLAB,” in Proc. IEEE Int. Conf. Robot. Automat., 2004, pp. 284–289.
[33] M. ApS, “The MOSEK optimization toolbox for MATLAB manual,” in User’s Guide and Reference Manual, Version 4. 2020.
[34] J. Miller, D. Henrion, M. Sznaier, and M. Korda, “Peak estimation for uncertain and switched systems,” in Proc. IEEE 60th Conf. Decis. Control, 2021, pp. 3222–3228.
[35] L. C. Young, “Generalized surfaces in the calculus of variations,” Ann. Math., vol. 43, pp. 84–103, 1942.
[36] M. Yannakakis, “Expressing combinatorial optimization problems by linear programs,” J. Comput. System Sci., vol. 43, no. 3, pp. 441–466, 1991.
[37] J. Gouveia, P. A. Parrilo, and R. R. Thomas, “Lifts of convex sets and cone factorizations,” Math. Operations Res., vol. 38, no. 2, pp. 248–264, 2013.
[38] D. Andersen and M. Osborne, “Discrete, linear approximation problems in polyhedral norms,” Numerische Mathematik, vol. 26, no. 2, pp. 179–189, 1976.
[39] E. J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces: Theory and Applications*. Hoboken, NJ, USA: Wiley, 1987.

[40] J. Cimprič, M. Marshall, and T. Netzer, “Closures of quadratic modules,” *Automatica*, vol. 183, no. 1, pp. 445–474, 2011.

[41] M. Putinar, “Positive polynomials on compact semi-algebraic sets,” *Indiana Univ. Math. J.*, vol. 42, no. 3, pp. 969–984, 1993.

[42] M. Trnovská, “Strong duality conditions in semidefinite programming,” *J. Elect. Eng.*, vol. 56, no. 12, pp. 1–5, 2005.

Jared Miller (Member, IEEE) received the B.S., M.S., degrees simultaneously, and Ph.D. degree in electrical engineering from Northeastern University, Boston, MA, USA, in 2016, 2018, and 2023, respectively. He is currently a Post-Doctoral Researcher with Robust Systems Laboratory, Northeastern University. His research interests include large-scale convex optimization, semialgebraic geometry, measure theory, and analysis of nonlinear systems.

Dr. Miller was a recipient of the 2020 Chateaubriand Fellowship from the Office for Science Technology of the Embassy of France in the United States, and was a recipient of the Outstanding Student Paper awards at the IEEE Conference on Decision and Control in 2021 and in 2022.

Mario Sznaier (Fellow, IEEE) received the B.S. degree in electrical engineering and the B.S. degree in computer science from the Universidad de la Republica, Montevideo, Uruguay, in 1983 and 1984, respectively, and the M.S. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, WA, USA in 1986 and 1989, respectively. He was a Professor of electrical engineering with Pennsylvania State University, State College, PA, USA. He held visiting positions with the California Institute of Technology, Pasadena, CA, USA. He is currently the Dennis Picard Chaired Professor with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA, USA. His research interests include robust identification and control of hybrid systems, robust optimization, and dynamical vision.

Dr. Sznaier is an Editor-in-Chief of the section on AI and Control of the journal Frontiers in Control Engineering, and Chair of IFAC’s Technical Committee TC 2.5. Additional service includes Associate Editor for the Journal *Automatica* (2005–2021), Chair of the CSS Technical Committee on Computational Aspects of Control System Design (2012–2016), General Co-Chair of the 2016 MSC, and Program Chair of the 2017 CDC. He is a Distinguished Member of the IEEE Control Systems Society.