Uncertainty Quantification of Set-Membership Estimation in Control and Perception: Revisiting the Minimum Enclosing Ellipsoid
Yukai Tang, Jean-Bernard Lasserre, Heng Yang

To cite this version:
Yukai Tang, Jean-Bernard Lasserre, Heng Yang. Uncertainty Quantification of Set-Membership Estimation in Control and Perception: Revisiting the Minimum Enclosing Ellipsoid. 2023. hal-04311534

HAL Id: hal-04311534
https://laas.hal.science/hal-04311534
Preprint submitted on 28 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Uncertainty Quantification of Set-Membership Estimation in Control and Perception: Revisiting the Minimum Enclosing Ellipsoid

Yukai Tang  
Department of Electrical Engineering, Tsinghua University, China  
TANGYK20@MAILS.TSINGHUA.EDU.CN

Jean-Bernard Lasserre  
LAAS-CNRS and Toulouse School of Economics (TSE), Toulouse, France  
LASSERRE@LAAS.FR

Heng Yang  
School of Engineering and Applied Sciences, Harvard University, USA  
HANKYANG@SEAS.HARVARD.EDU

Abstract

Set-membership estimation (SME) outputs a set estimator that guarantees to cover the groundtruth. Such sets are, however, defined by (many) abstract (and potentially nonconvex) constraints and therefore difficult to manipulate. We present tractable algorithms to compute simple and tight overapproximations of SME in the form of minimum enclosing ellipsoids (MEE). We first introduce the hierarchy of enclosing ellipsoids proposed by Nie and Demmel (2005), based on sums-of-squares relaxations, that asymptotically converge to the MEE of a basic semialgebraic set. This framework, however, struggles in modern control and perception problems due to computational challenges. We contribute three computational enhancements to make this framework practical, namely constraints pruning, generalized relaxed Chebyshev center, and handling non-Euclidean geometry. We showcase numerical examples on system identification and object pose estimation.

Keywords: Set-Membership Estimation, Minimum Enclosing Ellipsoid, Semidefinite Relaxations

1. Introduction

Model estimation and learning from measurements is a central task in numerous disciplines (Stengel, 1994; Barfoot, 2017). Let $\theta \in \Theta \subseteq \mathbb{R}^n$ be an unknown model, and $Z = \{z_i\}_{i=1}^N \in \mathbb{Z}^N$ be a set of $N$ measurements such that, if $z_i$ is a perfect (noise-free) measurement, it holds $r(\theta, z_i) = 0, i = 1, \ldots, N$, with $r : \Theta \times Z \rightarrow \mathbb{R}^m$ a given residual function that is typically designed from first principles describing how $z_i$ is generated from $\theta$ (see Examples 1-2 below).

**Maximum Likelihood Estimation.** The most popular approach to estimate $\theta$ from $Z$ is maximum likelihood estimation (MLE). When the measurements are noisy, the residual is adjusted to

$$r(\theta, z_i) = \epsilon_i, \quad i = 1, \ldots, N,$$

with $\epsilon_i \in \mathbb{R}^m$ some small noise. The most common distributional assumption is that $\epsilon_i \sim \mathcal{N}(0, \Sigma)$ follows a Gaussian distribution, leading to the MLE estimator that is the solution of a nonlinear least squares problem (Nocedal and Wright, 1999; Pineda et al., 2022)

$$\theta^*_{\text{MLE}} \in \arg\min_{\theta \in \Theta} \sum_{i=1}^N \|r(\theta, z_i)\|^2_{\Sigma^{-1}}. \quad (2)$$

When the measurement set $Z$ contains outliers, the least squares objective in (2) can be replaced by a robust loss (Huber, 2004; Antonante et al., 2021). Despite significant interests and progress in the

---

1. Or maximum *a posteriori* estimation (MAP) if there is prior knowledge about the distribution of $\theta$. © Y. Tang, J.-B. Lasserre & H. Yang.
literature, two shortcomings of the MLE framework exist. On one hand, the Gaussian assumption is questionable. In fact, in Appendix A we show noises generated by modern neural networks on a computer vision example deviate far from a Gaussian distribution.\(^2\) On the other hand, in safety-critical applications, \textit{provably correct uncertainty quantification} of a given estimator is desired (e.g., how close is \(\theta_{\text{MLE}}^*\) to the groundtruth). The uncertainty of the MLE estimator (2), however, is nontrivial to quantify due to the potential nonlinearity in the residual function \(r\).

\textbf{Set-Membership Estimation.} An alternative framework, known as \textit{set-membership estimation} (SME, or unknown-but-bounded estimation) (Milanese and Vicino, 1991), seeks to resolve the two shortcomings of MLE. Instead of making a distributional assumption on the measurement noise (1), SME only requires the noise to be bounded

\[
\|\epsilon_i\| = \|r(\theta, z_i)\| \leq \beta_i, \quad i = 1, \ldots, N,
\]  

where \(\|\cdot\|\) indicates the \(\ell_2\) vector norm (one can choose \(\ell_1\) or \(\ell_\infty\) norm and our algorithm would still apply). Characterizing the bound of the noise is easier than characterizing the distribution of the noise and can be conveniently done using, e.g., conformal prediction with a calibration dataset (Angelopoulos and Bates, 2021). Given (3), SME returns a \textit{set estimation} of the model

\[
S = \{\theta \in \Theta \mid \|r(\theta, z_i)\| \leq \beta_i, i = 1, \ldots, N\} \tag{SME}
\]

\(i.e., S\) contains all models compatible with the measurements \(Z\) under assumption (3). Clearly, the groundtruth must belong to \(S\), and the “size” of \(S\) informs the uncertainty of the estimated model.

\textbf{Challenges.} It is almost trivial to write down the set \(S\) as in (SME), which, however, turns into a highly nontrivial object to manipulate. The reasons are (a) each of the constraints \(\|r(\theta, z_i)\| \leq \beta_i\) may be a nonconvex constraint, and/or (b) the number of constraints \(N\) may be very large. We use two examples in control and perception, respectively, to illustrate the challenges.

\textbf{Example 1 (System Identification (Kosut et al., 1992; Li et al., 2023))} Consider a discrete-time dynamical system with state \(x \in \mathbb{R}^{n_x}\), control \(u \in \mathbb{R}^{n_u}\), and unknown system parameters \(\theta \in \mathbb{R}^n\)

\[
x_{k+1} = \phi_0(x_k, u_k) + \sum_{i=1}^n \theta_i \phi_i(x_k, u_k) + \epsilon_k = \Phi(x_k, u_k) \hat{\theta} + \epsilon_k, \quad k = 0, \ldots, N - 1, \tag{4}
\]

where \(\Phi(x_k, u_k) = [\phi_0(x_k, u_k), \ldots, \phi_n(x_k, u_k)] \in \mathbb{R}^{n_x \times (n+1)}\) is a (nonlinear) activation function, \(\epsilon_k \in \mathbb{R}^{n_x}, k = 0, \ldots, N - 1\) are unknown noise vectors assumed to satisfy (3),\(^3\) and \(\hat{\theta} = [1, \theta^T]^T\).

Given a system trajectory \(\{x_0, x_0, \ldots, x_{N-1}, x_N\}\), the SME of the parameters \(\theta\) is

\[
S = \left\{ \theta \in \mathbb{R}^n \mid \|x_{k+1} - \Phi(x_k, u_k) \hat{\theta}\| \leq \beta_k, k = 0, \ldots, N - 1 \right\} \tag{5}
\]

\textbf{Example 2 (Object Pose Estimation (Hartley and Zisserman, 2003; Yang and Pavone, 2023))} Consider a 3D point cloud \(\{Y_i\}_{i=1}^N\) and a camera, at an unknown rotation and translation \(\theta = (R, t) \in \text{SE}(3),\(^4\)\) observing the point cloud as a set of 2D image keypoints

\[
y_i = \Pi(RY_i + t) + \epsilon_i, \quad i = 1, \ldots, N, \tag{6}
\]

\(2.\) One can replace the Gaussian assumption with more sophisticated distributional assumptions, but the resulting optimization often becomes more difficult to solve.

\(3.\) Note that our algorithm can allow \(\theta\) to appear nonlinearly in the dynamics, but it is sufficient to restrict to dynamics linear in \(\theta\) as in (4) for system identification tasks in many real applications.

\(4.\) \(\text{SE}(3) := \text{SO}(3) \times \mathbb{R}^3\) with \(\text{SO}(3) := \{R \in \mathbb{R}^{3 \times 3} \mid R^TR = RR^T = I, \det R = +1\}\).
where $\Pi : \mathbb{R}^3 \to \mathbb{R}^2$, $\Pi(v) = [v_1/v_3, v_2/v_3]^T$ denotes the projection of a 3D point onto the 2D image plane and $\epsilon_i \in \mathbb{R}^2$ denotes measurement noise satisfying (3). Given pairs of 3D-2D correspondences $\{y_1, Y_1, \ldots, y_N, Y_N\}$, the SME of the camera pose $\theta$ is

$$
S = \{\theta \in SE(3) \mid \|y_i - \Pi(RY_i + t)\| \leq \beta_i, i = 1, \ldots, N \}.
$$

The SME (5) is convex—defined by quadratic inequality constraints— but $N$ can be large given a long trajectory. The SME (7), shown by (Yang and Pavone, 2023, Proposition 3) to be defined by $N$ quadratic inequalities and $N$ linear inequalities, is unfortunately nonconvex despite that $N < 10$.

**Contributions.** We propose tractable algorithms based on semidefinite programming (SDP) to simplify the set-membership estimator (SME) while maintaining tight uncertainty quantification. Specifically, we seek to find the minimum enclosing ellipsoid (MEE) of (SME), i.e., the ellipsoid that contains $S$ with minimum volume. Such an ellipsoid allows us to (i) use the center of the ellipsoid as a point estimator, (ii) provide a (minimum) worst-case error bound for the point estimator, and (iii) generate samples in $S$ via straightforward rejection sampling (i.e., sample inside the ellipsoid and accept the sample if inside $S$).\(^5\) Our algorithms are based on the sums-of-squares (SOS) relaxation framework proposed in Nie and Demmel (2005), revisited in Kojima and Yamashita (2013), for computing a hierarchy of enclosing ellipsoids that asymptotically converge to the MEE of a basic semialgebraic set, i.e., a set defined by finitely many polynomial (in-)equalities, to which both (5) and (7) belong.\(^6\) When the SME is convex, together with the celebrated Löwner-John’s ellipsoid theorem (Henk, 2012), we give an algorithm that can provide a certificate of convergence when the MEE has been attained by the SOS hierarchy. We show this vanilla algorithm already outperforms the confidence set of least-squares estimation (Abbasi-Yadkori and Szepesvári, 2011) in a simple instance of Example 1. Unfortunately, applying this algorithm to Example 1 with large $N$ and Example 2 encounters three challenges. (C1) A long system trajectory (e.g., $N = 1000$) in Example 1 renders high-order SOS relaxations overly expensive. We therefore introduce a preprocessing algorithm to prune redundant constraints in the set (5) (e.g., over 900 constraints are deemed redundant). (C2) We empirically found for (SME) sets that are nonconvex or defined by many constraints, the SDP solver would encounter serious numerical issues and simply fail. To circumvent this, we propose a two-step approach, where step one draws random samples from (SME) to approximate the shape matrix of the enclosing ellipsoid, and step two minimizes the size of the ellipsoid with a fixed shape. The second step, coincidentally, becomes a strict generalization of the relaxed Chebyshev center method proposed in Eldar et al. (2008). (C3) The last challenge is to handle the non-Euclidean geometry when enclosing an (SME) of SO(3). By leveraging the special geometry of unit quaternions, we show a reduction of computing the MEE on a Riemannian manifold to computing the MEE in Euclidean space. With these computational enhancements, we conduct experiments on system identification (including those that are hard to learn (Tsiamis and Pappas, 2021)) and object pose estimation that were not possible to perform in existing literatures.

**Paper Organization.** We defer a review of related works to Appendix H and first present the SOS-based MEE framework in Section 2. We describe in Section 3 our computational enhancements. We present numerical experiments in Section 4 and conclude in Section 5. Preliminaries on the moment-SOS hierarchy (Lasserre, 2001; Parrilo, 2003) are presented in Appendix B.

---

5. Clearly, this would also allow us to obtain an estimate of the volume of $S$ by counting the acceptance rate.

6. Surprisingly, the SOS-based MEE framework seems unexploited in the context of Examples 1-2.
2. Minimum Enclosing Ellipsoid by SOS Relaxations

Consider a basic semialgebraic set defined by a finite number of polynomial constraints
\[ S = \{ \theta \in \mathbb{R}^n \mid g_i(\theta) \geq 0, i = 1, \ldots, l_g, h_j(\theta) = 0, j = 1, \ldots, l_h \} \]  
(8)
with \( g_i, h_j \in \mathbb{R}[\theta] \) real polynomials in \( \theta \). We use the same notation \( S \) here as in (SME), (5), (7) because the membership sets we consider are all basic semialgebraic sets. Let \( \xi = (\theta_{i_1}, \ldots, \theta_{i_d}) = P\theta \) be a \( d \)-dimensional subvector of \( \theta \) (with \( P \) a selection matrix), \( S_\xi := \{ P\theta \mid \theta \in S \} \) be the restriction of \( S \) on \( \xi \), and consider a \( d \)-dimensional ellipsoid
\[ E = \{ \xi \in \mathbb{R}^d \mid 1 - (\xi - \mu)^T E (\xi - \mu) \geq 0 \}, \]  
(9)
with \( \mu \in \mathbb{R}^d \) the center and \( E \in \mathbb{S}^d_{++} \) the shape matrix \( (E \) is positive definite). We want to find the ellipsoid \( E \) with minimum volume that encloses \( S_\xi \)
\[ \max_{\mu \in \mathbb{R}^d, E \in \mathbb{S}^d_{++}} \{ \log \det E \mid \xi \in E, \forall \xi \in S_\xi \}, \]  
(MEE)
where \( \det E \) is inversely proportional to the volume of \( E \) (i.e., (MEE) minimizes the volume of \( E \)). We remark that considering an ellipsoid in the subvector \( \xi \) is general and convenient as (i) \( P = I \) recovers the \( n \)-dimensional ellipsoid, (ii) in Example 2 it is desired to enclose the rotation \( R \) and translation \( t \) separately, and (iii) having \( \xi = \theta_i \in \mathbb{R} \) for some dimension \( i \) allows the ellipsoid \( E_i \) to be a line segment that encloses \( \theta_i \) and hence \( E_1 \times \ldots \times E_n \) forms an enclosing box of \( S \).

Problem \( \text{(MEE)} \) is generally intractable even when \( S \) is convex.\(^7\) However, denoting \( e(\xi) := 1 - (\xi - \mu)^T E (\xi - \mu) \) in (9) as the polynomial in \( \xi \), we observe the constraint in (MEE) simply asks \( e(\xi) \) to be nonnegative on the set \( S \). This observation allows a hierarchy of convex relaxations of \( \text{(MEE)} \) based on sums-of-squares (SOS) programming (Lasserre, 2001; Blekherman et al., 2012).

Theorem 1 (MEE Approximation by SOS Programming) Assume \( S \) is Archimedean,\(^8\) consider the SOS program with an integer \( \kappa \) such that \( 2\kappa \geq \max \{ \log \det E, \{ \deg(g_i) \}_{i=1}^{l_g}, \{ \deg(h_j) \}_{j=1}^{l_h} \} \)
\[ \max_{E, b, c} \log \det E \]  
subject to
\[ 1 - (\xi^T E \xi + 2b^T \xi + c) = \sum_{i=0}^{l_g} \sigma_i(\theta) g_i(\theta) + \sum_{j=1}^{l_h} \lambda_j(\theta) h_j(\theta) \]  
(10b)
\[ \sigma_i(\theta) \in \text{SOS}[\theta], \quad \deg(\sigma_i g_i) \leq 2\kappa, \quad i = 0, \ldots, l_g \]  
(10c)
\[ \lambda_j(\theta) \in \mathbb{R}[\theta], \quad \deg(\lambda_j h_j) \leq 2\kappa, \quad j = 1, \ldots, l_h \]  
(10d)
\[ \begin{bmatrix} E & b \\ b^T & c \end{bmatrix} \succeq 0 \]  
(10e)
where \( g_0(\theta) := 1, \text{SOS}[\theta] \) is the set of SOS polynomials in \( \theta \), and \( \deg(\cdot) \) denotes the degree of a polynomial. Let \( (E_*, b_*, c_*) \) be an optimal solution of (10), then,

(i) for any \( \kappa \), we have \( \xi^T E_* \xi + 2b_*^T \xi + c_* = (\xi - \mu_*)^T E_*(\xi - \mu_*) \) with \( \mu_* = -E_*^{-1}b_* \), and the ellipsoid \( E_\kappa = \{ \xi \in \mathbb{R}^d \mid (\xi - \mu_*)^T E_*(\xi - \mu_*) \leq 1 \} \) encloses \( S_\xi \);

---

7. When \( S \) is a set of points, then \( \text{(MEE)} \) is easy to solve (Gärtner, 1999; Magnani et al., 2005; Moshtagh et al., 2005).
8. The definition of Archimedeaness is given in (Blekherman et al., 2012, Def. 3.137). One can make the Archimedean condition trivially hold by adding a constraint \( M_\theta - \theta^T \theta \geq 0 \) to (8), which is easy for Examples 1-2.
Vol(\(\mathcal{E}_\kappa\)) decreases as \(\kappa\) increases, and \(\mathcal{E}_\kappa\) tends to the solution of (MEE) as \(\kappa \to \infty\).

The proof of Theorem 1 is given in Appendix C and is inspired by Nie and Demmel (2005); Magnani et al. (2005). Problem (10) is convex and can be readily implemented by YALMIP (Lofberg, 2004) and solved by MOSEK (ApS, 2019). Its intuition is simple: (10b)-(10d) ensures 1 - (\(\xi^T E \xi + 2b^T \xi + c\)) is nonnegative on \(S\) and hence every feasible solution is an enclosing ellipsoid. (10e) uses a lifting technique to convexify the bilinearity of \(\mu\) and \(E\) in the original ellipsoid parametrization (9), and the objective (10a) seeks to minimize the volume. The convergence of the hierarchy follows from Putinar’s Positivstellensatz (Putinar, 1993).

Certifying Convergence. Unlike applying SOS relaxations to polynomial optimization, where a certificate of convergence is known (Henrion and Lasserre, 2005), detecting the convergence of (10) is in general difficult (Lasserre, 2015). However, when the set \(S\) is convex, it is possible to leverage Löwner-John’s ellipsoid theorem (Xie, Miaolan, 2016) to derive a simple convergence certificate. We present such a result in Appendix D and a numerical example in Appendix G.

3. Computational Enhancement

The SOS-based algorithm (10) works very well on simple examples. In this section, we describe three challenges of applying (10) and present three techniques to enhance its performance.

3.1. Pruning Redundant Constraints

The first challenge arises when the number of inequalities \(l_g\) is very large, in which case the convex optimization (10) has \(l_g + 1\) positive semidefinite (PSD) variables (whose sizes grow quickly with the relaxation order \(\kappa\)). In system identification (Example 1), a large \(l_g = N\) is common as practitioners often collect a long system trajectory to accurately identify the system.

Nevertheless, since the dimension \(n\) of the parameter \(\theta\) is usually much smaller than the trajectory length \(N\), one would expect many constraints in (5) to be redundant. The difficultly lies in how to identify and prune the redundant constraints. When all the constraints are linear, algorithms from linear programming can identify redundancy (Caron et al., 1989; Telgen, 1983; Paulraj et al., 2010; Cotorruelo et al., 2020). To handle the quadratic constraints in (5), we propose Algorithm 1.

**Algorithm 1: Prune Redundant Constraints in (5)**

1. **Input:** Constraint set \(\{g_k(\theta)\}_{k \in [N-1]}\) with \(g_k(\theta) := \beta_k^2 - \|x_{k+1} - \Phi(x_k, u_k)\bar{\theta}\|^2\) (cf. (5))
2. **Output:** Pruned set of constraints \(\{g_k(\theta)\}_{k \in \mathcal{I}}\)
3. Initialize \(\mathcal{I} = [N - 1]\)
4. for \(k \leftarrow 0\) to \(N - 1\) do
5. \hspace{1em} Solve \(g_k \leq g_k^* = \min_{\theta} \{g_k(\theta) \mid g_i(\theta) \geq 0, i \in \mathcal{I}\setminus\{k\}\}\) \hspace{1em} First-order moment-SOS hierarchy
6. \hspace{1em} if \(g_k^* \geq 0\) then \(\mathcal{I} \leftarrow \mathcal{I}\setminus\{k\}\); 
7. end
8. Return: \(\mathcal{I}\)

Intuitively, line 5 of Algorithm 1 seeks to minimize \(g_k(\theta)\) when the rest of the constraints hold. If \(g_k^* \geq 0\), then \(g_k(\theta)\) is redundant as it is implied by the rest of the constraints. The optimization in line 5, however, is nonconvex because \(g_k(\theta)\) is a concave polynomial. Therefore, we use the first-order moment-SOS hierarchy to obtain a lower bound \(g_k \leq g_k^*\): if \(g_k \geq 0\), then \(g_k^* \geq 0\) must
hold and \(g_\kappa(\theta)\) is deemed redundant. In practice, the first-order moment-SOS hierarchy is very efficient and easily scales to \(N = 1000\). In a pendulum system identification experiment presented in Section 4.1.4, we show Algorithm 1 effectly prunes over 90% constraints.

### 3.2. Generalized Relaxed Chebyshev Center

The second challenge comes from the “\(\log \det E\)” objective in (10). Maximizing the “\(\log \det\)” of \(E\) is convex (Vandenberghe et al., 1998), yet it cannot be written as a standard linear SDP. Consequently, it is often replaced by the geometric mean (Lofberg, 2004) and modelled with an exponential cone constraint (ApS, 2019), causing serious numerical issues in our experiments. Even after pruning redundant constraints using Algorithm 1, we found problem (10) still difficult to solve. Similar observations have been reported in (Lasserre, 2023, page 950).

We can use three different algorithms to uniformly sample \(S\). (i) When \(S\) is convex, we can use the hit-and-run algorithm (Bélisle et al., 1993), where each iteration involves solving a convex optimization. (ii) For Example 2, Yang and Pavone (2023) proposed a RANSAG algorithm, where each iteration involves solving a geometry problem. (iii) We can first solve (GCC) with rejection sampling  

This motivates solving the following generalized Chebyshev center (GCC) problem

\[
\eta^* = \min_{\mu \in \mathbb{R}^d} \max_{\xi \in S_\xi} \|\mu - \xi\|_Q^2 \tag{GCC}
\]

with a given \(Q > 0\) (recall \(\xi = P\theta\) is a subvector of \(\theta\)). When \(Q = I\), (GCC) reduces to the usual Chebyshev center problem (Milanese and Vicino, 1991; Eldar et al., 2008). In (GCC), given any \(\mu\), the inner “\(\max_{\xi}\)” computes the maximum \(Q\)-weighted distance from \(\mu\) to the set \(S_\xi\), denoted as \(\eta_\mu^*\). Therefore, the ellipsoid \(E_{Q,\mu} = \{\xi \mid \|\xi - \mu\|^2_Q \leq \eta_\mu^*\}\) must enclose the set \(S_\xi\). Via the outer “\(\min_{\mu}\)”\), (GCC) seeks the \(\mu\) such that \(E_{Q,\mu}\) is the smallest enclosing ellipsoid with a given shape matrix \(Q\). We found (GCC) to be much easier to solve than (10) due to removing the “\(\log \det E\)” objective.

Estimating \(Q\) from Samples. Intuitively, (GCC) may be more conservative than (MEE) because (GCC) assumes \(Q\) is given while (MEE) optimizes the shape matrix. However, if we can uniformly draw samples \((\xi^1, \ldots, \xi^{N_s})\) from \(S_\xi\), then a good estimate of \(Q\) can be obtained as

\[
Q = \frac{1}{N_s} \sum_{i=1}^{N_s} (\xi^i - \bar{\xi})(\xi^i - \bar{\xi})^T, \quad \bar{\xi} = \frac{1}{N_s} \sum_{i=1}^{N_s} \xi^i. \tag{11}
\]

We can use three different algorithms to uniformly sample \(S_\xi\). (i) When \(S\) is convex, we can use the hit-and-run algorithm (Bélisle et al., 1993), where each iteration involves solving a convex optimization. (ii) For Example 2, Yang and Pavone (2023) proposed a RANSAG algorithm, where each iteration involves solving a geometry problem. (iii) We can first solve (GCC) with \(Q = I\) to obtain an enclosing ball, then perform rejection sampling inside the enclosing ball. We use (iii) in our experiments as it is general and does not require convexity.

Convex Relaxations for (GCC). Problem (GCC) is still nonconvex, but we can design a hierarchy of SDP relaxations that asymptotically converges to \(\eta^*\).

**Theorem 2 (Generalized Relaxed Chebyshev Center)** Let \(S\) in (8) be Archimedean, \(\kappa\) be an integer such that \(2\kappa \geq \max\{2, \{\deg(g_i)\}_{i=1}^{l_g}, \{\deg(h_j)\}_{j=1}^{l_h}\}\), and \(s_n(d) := \binom{n+d}{d}\), consider the following convex quadratic SDP

\[
-\eta_\kappa^* = \min_{z \in \mathbb{R}^{s_n(2\kappa)}} \quad L_z(\theta)^T P Q P L_z(\theta) - \langle C, M_\kappa(z) \rangle \tag{12a}
\]

subject to

\[
M_\kappa(z) \succeq 0, \quad M_0(z) = 1 \tag{12b}
\]

\[
M_{\kappa - \deg(g_i)/2}(g_iz) \succeq 0, \quad i = 1, \ldots, l_g \tag{12c}
\]

\[
L_z(h_j[\theta]_{2\kappa - \deg(h_j)}) = 0, \quad j = 1, \ldots, l_h \tag{12d}
\]

9. SDPT3 (Toh et al., 1999) natively supports \(\log \det\) but empirically we found it performs worse than MOSEK.

10. Particularly, MOSEK returns a solution with large duality gap.
where \( z \in \mathbb{R}^{s_n(2\kappa)} \) is the pseudomoment vector in \( \theta \) of degree up to \( 2\kappa \), \( L_z(u) \) and \( M_{\kappa}(uz) \) are both linear functions of \( z \) given coefficients of a certain polynomial \( u(\theta) \) (cf. Appendix B), and \( C \) is a constant matrix whose expression is given in Appendix E. Let \( \bar{z}_{\kappa}^* \) be an optimal solution to (12) and \( \mu_{\kappa} = \bar{L}_{z_{\kappa}}(\theta) \), then,

(i) for any \( \kappa \), the ellipsoid \( \mathcal{E}_{Q,\kappa} := \{ \| \xi - \mu_{\kappa}^* \|_Q^2 \leq \eta_{\kappa}^* \} \) encloses \( \mathcal{S}_\xi \);

(ii) \( \eta_{\kappa}^* \geq \eta^* \) for any \( \kappa \) and \( \eta_{\kappa}^* \) converges to \( \eta^* \) as \( \kappa \to \infty \).

The proof of Theorem (2), together with a lifting technique to write (12) as a standard linear SDP, is given in Appendix E. The basic strategy is to first apply the moment-SOS hierarchy, with order \( \kappa \), to relax the inner \( \max_{\xi \in \mathcal{S}_\xi} \) (a polynomial optimization) as a convex SDP whose decision variable is \( z \in \mathbb{R}^{s_n(2\kappa)} \), the pseudomoment vector constrained by (12b)-(12d). Then, by invoking Sion’s minimax theorem \( Sion, 1958 \), one can switch \( \min_{L_z(u)} \) and \( \max_{\mu_{\kappa}^*} \) admits a closed-form solution \( \mu = L_z(\theta) \) and leads to a single-level optimization with cost (12a).

It is worth noting that letting \( \kappa = 1 \) (and \( Q = I \)) in (12) recovers the relaxed Chebyshev center in Eldar et al. (2008) for bounded error estimation. For this reason, we call \( \mu_{\kappa} \), the generalized relaxed Chebyshev center of order \( \kappa \) (GRCC-\( \kappa \)). As we will see in Section 4.1.2, the enclosing balls at GRCC-\( \kappa \) with \( \kappa \geq 2 \) is significantly smaller than those of the RCC.

3.3. Handling the non-Euclidean Geometry of \( SO(3) \)

With the (GCC) formulation, we can approximate the minimum ellipsoid that encloses \( \mathcal{S}_t \), the translation part of the SME in Example 2. The last challenge lies in enclosing \( \mathcal{S}_R \), the rotation part of (7). Since the concept of an ellipsoid is not well defined on a Riemannian manifold such as \( SO(3) \), it is natural to seek an enclosing geodesic ball. This is closely related to the problem of finding the Riemannian minimax center of \( \mathcal{S}_R \), defined as (Arnaudon and Nielsen, 2013)

\[
\mu_{R, \kappa} = \arg \min_{\mu \in SO(3)} \max_{R \in \mathcal{S}_R} d_{SO(3)}(\mu, R),
\]

where, given two rotations \( \mu, R \in SO(3) \), \( d_{SO(3)}(\mu, R) \) denotes the geodesic distance on \( SO(3) \) defined by \( d_{SO(3)}(\mu, R) = \arccos((\mu^T R - 1)/2) \). Clearly, if we can find \( \mu_{R, \kappa} \), then we have found the best point estimate from which the worst-case error bound is the smallest.

Unfortunately, problem (13) cannot be directly solved using Theorem 2 because (i) \( \mu \in SO(3) \) is constrained (instead of in (GCC) \( \mu \in \mathbb{R}^d \) is unconstrained), and (ii) the objective \( d_{SO(3)}(\mu, R) \) is not a polynomial. The next result resolves these issues by leveraging unit quaternions.

Theorem 3 (Minimum Enclosing Ball for Rotations) Suppose there exists a 3D rotation \( \tilde{R} \) such that \( d_{SO(3)}(\tilde{R}, R) \leq \frac{\pi}{2} \), \( \forall R \in \mathcal{S}_R \), and let \( \tilde{q} \in S^3 := \{ q \in \mathbb{R}^4 \mid \| q \| = 1 \} \) (or \( -\tilde{q} \)) be any one of the two unit quaternions associated with \( \tilde{R} \). Consider the following Chebyshev center problem in \( \mathbb{R}^4 \)

\[
\mu_* = \arg \min_{\mu \in \mathbb{R}^4} \max_{q \in \mathbb{R}^4} \left\{ \| \mu - q \| \mid q \in S^3, \mathcal{R}(q) \in \mathcal{S}_R, \tilde{q}^T q \geq 0 \right\},
\]

where \( \mathcal{R} : S^3 \to SO(3) \) maps a unit quaternion to a 3D rotation matrix (whose expression is given in Appendix F). Then the projection of \( \mu_* \) onto \( S^3 \) is the solution to (13), i.e., \( \mu_{R, \kappa} = \mu_*/\| \mu_* \| \).
The proof of Theorem 3 is given in Appendix F. Theorem 3 states that we can solve the original Riemannian minimax problem (13) by solving (14), which is easily verified to be an instance of (GCC) and hence can be approximated by the SDP relaxations in Theorem 2. Theorem 3 requires the existence of \( \bar{R} \) such that all rotations in \( S_R \) are reasonably close to \( \bar{R} \) and similar assumptions are needed in Arnaudon and Nielsen (2013). In other words, \( S_R \) needs to have relatively small uncertainty. This assumption can be numerically verified by first using the RANSAG algorithm in Yang and Pavone (2023) to compute \( \bar{R} \) and then compute the maximum distance between \( \bar{R} \) and \( S_R \), which we found empirically to be below \( \frac{\pi}{2} \) for a majority of the test problems.

4. Experiments

4.1. System Identification (Example 1)

4.1.1. Comparison with Least Squares Estimation

Consider a one-dimensional linear system

\[
x_{k+1} = A_* x_k + B_* u_k + w_k
\]

with \( A_* = 0.8, B_* = 1.0 \), and \( w_k \sim \mathcal{N}(0, \sigma_w^2) \). Suppose we collected a system trajectory \( x_0, u_0, \ldots, x_N \), we wish to find a set \( S \) such that we can guarantee \( \mathbb{P}[(A_*, B_*) \in S] \geq 80\% \). (Abbasi-Yadkori and Szepesvári, 2011, Theorem 1) provides an ellipse centered at the least-squares estimator (LSE) that satisfies the probabilistic coverage guarantee, which requires knowing a bound such that

\[
A_*^2 + B_*^2 \leq K^2
\]

Let \( K_* = \sqrt{A_*^2 + B_*^2} \) be the best bound. We can get much smaller coverage sets via the SME (5) and its enclosing ellipsoids without knowledge on \( K_* \). Let \( \beta > 0 \) be the smallest number such that \( \mathbb{P}[w_k^2 \leq \beta] \geq (80\%)^{1/N} \), implying \( \mathbb{P}[w_k^2 \leq \beta, \ldots, w_{N-1}^2 \leq \beta] \geq 80\% \) (assume \( w_k \)'s are independent). Consequently, the SME (5) covers \( (A_*, B_*) \) with probability at least 80%. Let \( E_4 \) be the enclosing ellipse of (5) computed via Theorem 1 using \( \kappa = 4 \) and \( E_{4, \text{LSE}} \) be the confidence ellipse from Abbasi-Yadkori and Szepesvári.

---

11. This is the same setup as in Li et al. (2023).
and \( \theta \) is the length of the pole, and \( (2011) \). We compute \( \text{Vol}(\mathcal{E})/\text{Vol}(\mathcal{E}_{\text{LSE}}) \) as the volume ratio, which indicates SME finds a smaller enclosing ellipse when it is below 1.

Fig. 1(a) boxplots the volume ratio (in log scale) under different choices of \( \sigma_w \) and \( K \) with fixed \( N = 50 \) (we perform 20 random experiments in each setup). We observe that LSE outperforms SME only when (i) LSE has access to a precise upper bound \( K \) (which we think is often unrealistic) and (ii) the noise \( w_k \) is large. In all other cases, the enclosing ellipse of SME is orders of magnitude smaller than that of LSE. A sample visualization is provided in Fig. 1(a).

4.1.2. Identification of Random Linear Systems

Consider a linear system \( x_{k+1} = A_x x_k + B u_k + w_k, k = \ldots, N - 1 \). We generate random \( A_x \in \mathbb{R}^{n_x \times n_x}, B \in \mathbb{R}^{n_x \times n_u} \) with each entry following a standard Gaussian distribution \( \mathcal{N}(0, 1) \), then truncate the singular values of \( A_x \) that are larger than 1 to 1. \( u_k \) follows \( \mathcal{N}(0, 1) \), and \( w_k \) is random inside a ball with radius \( \beta \). We treat \( B \) as known. We compute enclosing balls of the SME (5) using two algorithms: the RCC algorithm in Eldar et al. (2008), and our GRCC algorithm presented in Theorem 2. Fig. 1(b) plots the radii of the enclosing balls with two algorithms: the RCC algorithm in Eldar et al. (2008), and our GRCC algorithm presented in Theorem 2. Fig. 1(b) boxplots the radii of the enclosing balls with \( N \in \{20, 40, 60, 80, 100\} \), \( n_x = 2 \), and \( \beta \in \{0.01, 0.1, 1.0\} \) (each boxplot summarizes 20 random experiments). We observe that (i) GRCC (\( \kappa = 1 \)) leads to exactly the same result as RCC, verifying that our GRCC algorithm recovers the RCC algorithm with \( \kappa = 1 \). (ii) The enclosing balls get much smaller with high-order relaxations (i.e., larger \( \kappa \)). More results with \( n_x = 3 \) and \( n_x = 4 \) are presented in Appendix G.

4.1.3. Linear Systems that Are Hard to Learn

Tsiamis and Pappas (2021) presented examples of linear systems that are hard to learn. We show that SME and its enclosing balls are adaptive to the hardness of identification, i.e., one gets large enclosing balls when the system is hard to learn. Consider the system \( x_{k+1} = A x_k + H w_k \) with

\[
A = \begin{bmatrix}
0 & \theta_1 & 0 \\
0 & 0 & \theta_2 \\
0 & 0 & 0
\end{bmatrix}, \quad H = \begin{bmatrix}
1 & 0 \\
0 & 0 \\
0 & 1
\end{bmatrix}
\]

where \( w_k \) is a random disturbance with noise bound 0.1. Consider (i) an easy-to-learn system with \( \theta_2 = 1 \), and (ii) a hard-to-learn system with \( \theta_2 = 10^{-5} \). Let \( \theta_1 = 1 \). Fig. 1(c) boxplots the radii of the enclosing balls for the SME of \((\theta_1, \theta_2)\) with \( \kappa = 2 \) and increasing \( N \). Clearly, we observe that the radii for hard-to-learn systems are much larger than that for easy systems.

4.1.4. Constraints Pruning in a Long Trajectory

Consider the continuous-time dynamics of a simple pendulum

\[
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} = \begin{bmatrix}
-\frac{b}{m l^2} x_2 + \frac{1}{l^2} u - \frac{g}{l} \sin x_1 \\
\frac{1}{l^2} x_1 - \frac{b}{m l^2} \sin x_1
\end{bmatrix} = \begin{bmatrix}
0 & \frac{1}{l^2} & 0 & 0 \\
\frac{1}{l^2} & 0 & 0 & \frac{g}{l}
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2 \\
u \\
\sin x_1
\end{bmatrix}
\]

where \( x = (x_1, x_2) \) is the state (angle and angular velocity), \( b \) is the damping ratio, \( m \) is the mass, \( l \) is the length of the pole, and \( g \) is the gravity constant. We wish to identify \( \theta_1 = \frac{b}{m l^2}, \theta_2 = \frac{1}{l^2} \) and \( \theta_3 = \frac{g}{l} \). To do so, we discretize the dynamics using Euler method with \( dt = 0.01 \), add random
disturbance $w_k$ that has bounded norm 0.1, and collect a single trajectory of length $N = 1000$. Without constraint pruning, we can only run the GRCC algorithm with $\kappa = 2$ and the resulting enclosing ellipsoid and ball are visualized in Fig. 1(d). We then prune the constraint set using Algorithm 1, which leads to a much smaller set of 19 constraints (only 1.9% of the original number of constraints). We can then increase the relaxation order of GRCC and the resulting enclosing ellipsoid and ball become much smaller as visualized in Fig. 1(d). The volume of the enclosing ball with $\kappa = 7$ is only $1.08 \times 10^{-6}$, indicating the SME has almost converged to a single point.

4.2. Object Pose Estimation (Example 2)

We follow the same procedure as Yang and Pavone (2023), i.e., we use conformal prediction to calibrate the norm bounds $\beta_i$ for the noise vectors $\epsilon_i$ (cf. (6)) generated by the pretrained neural network in Pavlakos et al. (2017) and form the SME (7). We then use the GRCC algorithm in Theorem 2 with $\kappa = 2$ to compute enclosing balls (for the rotation we apply GRCC in Theorem 3 to compute enclosing geodesic balls). We compare the radii of the enclosing balls obtained by GRCC with the radii of the enclosing balls obtained by RANSAG of Yang and Pavone (2023). Fig. 2(a) shows the empirical cumulative distribution function (CDF) of the translation bounds and rotation bounds, respectively (there are 7035 translation problems and 6667 rotation problems). We observe that the error bounds obtained by GRCC are smaller than those obtained by RANSAG. Examples of enclosing balls and ellipsoids for the translation SME are shown in Fig. 2(b), where we observe that the enclosing ellipsoid precisely captures the shape of the SME. Fig. 2(b) also plots the enclosing balls of the rotation SME using stereographic projection.

![CDF plots. Left: translation, right: rotation](image1)

![Enclosing balls/ellipsoids. Left: translation, right: rotation.](image2)

Figure 2: Experimental results on object pose estimation (Example 2).

5. Conclusions

We introduced a suite of computational algorithms based on semidefinite programming relaxations to compute minimum enclosing ellipsoids of set-membership estimation in system identification and object pose estimation. Three computational enhancements are highlighted, namely constraints pruning, generalized relaxed Chebyshev center, and handling non-Euclidean geometry. These algorithms are still limited to small- and medium-sized problems (though these problems are already interesting) due to computational challenges in semidefinite programming. Multiple future research directions are possible, e.g., applying SME to system identification with partial observations, extending SME on object pose estimation to more perception problems, and integrating SME with adaptive control and reinforcement learning.

---

12. RANSAG first estimates an average rotation and translation, then uses SDP relaxations to compute the inner “max” problem in (GCC). In other words, RANSAG does not seek to find a better estimate with smaller error bounds.
Appendix A. Non-Gaussian Measurement Noise from Neural Networks

In this section, we provide, to the best of our knowledge, the first numerical evidence that the noise in measurements generated by neural networks in object pose estimation (Example 2) does not follow a Gaussian distribution.

Our strategy is to compute the noise vector $\epsilon_i$ from (6) as

$$\epsilon_i = y_i - \Pi(R_{\text{gt}}Y_i + t_{\text{gt}}), \quad i = 1, \ldots, N,$$

where $(R_{\text{gt}}, t_{\text{gt}})$ is the groundtruth camera pose, and $y_i$’s are neural network detections of the object keypoints. We use the LineMOD Occlusion (LM-O) dataset (Brachmann et al., 2014) that includes 2D images picturing a set of household objects on a table from different camera perspectives. The groundtruth camera poses are annotated and readily available in the dataset. We use the 2D keypoint detector in Pavlakos et al. (2017) that was trained on the LM-O dataset. At test time, the trained model outputs a heatmap of the pixel location of each semantic keypoint, and following Pavlakos et al. (2017) we set $y_i$ as the peak (maximum likelihood) location in the heatmap for each keypoint. This procedure produces a set of $N = 83,347^{13}$ 2D noise vectors $\{\epsilon_i\}_{i=1}^N$ and we want to test if they are drawn from a multivariate Gaussian (normal) distribution. To do so, we use the R package MVN (Korkmaz et al., 2014) that provides a suite of popular multivariate normality tests well established in statistics. To consider potential outliers in the noise vectors $\{\epsilon_i\}_{i=1}^N$ (i.e., maybe only the smallest noise vectors satisfy a Gaussian distribution (Antonante et al., 2021)), we run MVN on $\alpha\%$ of the noise vectors with smallest magnitudes, and we sweep $\alpha$ from 1 (i.e., keep only the 1% smallest noise vectors) up to 100 (i.e., keep all noise vectors).

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Mardia Skewness</th>
<th>Mardia Kurtosis</th>
<th>Henze-Zirkler</th>
<th>Royston</th>
<th>Doornik-Hansen</th>
<th>Energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>5%</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>10%</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>20%</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>40%</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>100%</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>

Table A1: Results of MVN tests on the noise vectors generated by neural network measurements in the LM-O dataset. We use the pretrained network from (Pavlakos et al., 2017).

Table A1 shows the test results. We can see that among the 36 tests performed, the data passed the test only once. This gives strong evidence that the noise vectors do not follow a Gaussian distribution, even after filtering potential outliers. Fig. A1 shows the perspective plots (top) and the Chi-square quantile-quantile (Q-Q) plots of the empirical density functions under different inlier ratios, in comparison to that of a Gaussian distribution. We can see that the empirical density functions deviate far away from a Gaussian distribution, and are difficult to characterize. This motivates the set membership estimation framework in Section 1.

13. There are 8 different objects in the dataset, each with about 8 semantic keypoints. The dataset has 1214 images.
Appendix B. SOS Relaxations and the Moment-SOS Hierarchy

B.1. Monomials, Polynomials, and Sum-of-Squares (SOS)

Monomials and Polynomials. Given \( x = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^n \), denote the monomial \( \Pi_i x_i^{\alpha_i} \) as \( x^\alpha \) where \( \alpha = [\alpha_1, \ldots, \alpha_n] \in \mathbb{N}^n \). The degree of a monomial is defined as \( \sum_{i=1}^{n} \alpha_i \), and the degree of a polynomial, \( \sum_{\alpha} c_\alpha x^\alpha \), is the maximum degree of all monomials. We use the notation \( \mathbb{R}[x]_d \) to denote all the polynomials with degree less than or equal to \( d \).

Monomial Basis and Sum-of-Squares Polynomials. Denote the monomial basis \( [x]_\kappa \) as the column vector of all monomials of degree up to \( \kappa \). For example, if \( n = 2 \) and \( \kappa = 2 \), the corresponding monomila basis is \( [x]_2 = [1; x_1; x_2; x_1^2; x_1 x_2; x_2^2] \). If we define \( s_n(d) := \binom{n+d}{d} \), the length of the monomial basis \( [x]_\kappa \) is \( s_n(\kappa) \).

A polynomial \( g(x) \) is called a sum-of-squares (SOS) polynomial if it can be written as a sum of squares of polynomials, i.e., \( g(x) = \sum_{i=1}^{m} q_i(x)^2 \) for some \( q_i(x) \in \mathbb{R}[x] \). Next proposition gives a necessary and sufficient condition for a polynomial to be SOS, which is more convenient for us to check whether a polynomial is a SOS polynomial.

Proposition A1 (Condition for SOS Polynomials Lasserre (2009)) A polynomial \( g \in \mathbb{R}_{2d}[x] \) is a SOS polynomial if and only if there exists a symmetric and positive semidefinite matrix \( Q \in \mathbb{R}^{s_n(d) \times s_n(d)} \) such that \( g(x) = [x]_d^T Q [x]_d \) where \( [x]_d \) is the monomial basis.

B.2. Nonnegativity Certificate

We consider a basic semialgebraic set \( \mathcal{S} := \{ x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \ldots, l_g, h_j(x) = 0, j = 1, \ldots, l_h \} \), where \( g_i(x), h_j(x) \) are polynomials. How to describe the nonnegativity of a polynomial on \( \mathcal{S} \) is a crucial technique in the following proofs. We first introduce a property called Archimedean.

Assumption A1 (Archimedean Property) For the basic semialgebraic set \( \mathcal{S} \), there exists \( u \in \mathbb{R}[x] \) of the form:

\[
  u = \sigma_0 + \sum_{j=1}^{l_g} \sigma_j g_j + \sum_{k=1}^{l_h} \lambda_k h_k
\]
where \(\sigma_j \in \mathbb{R}[x]\) are sum of squares and \(\lambda_k \in \mathbb{R}[x]\) are arbitrary polynomials, such that the level set \(\{x \in \mathbb{R}^n \mid u(x) \geq 0\}\) is compact.

It’s notable that, if the set is compact, then by adding an enclosing ball constraint, the Archimedean property can be satisfied trivially. If \(S\) is Archimedean, then for strictly nonnegative polynomials on \(S\), we have the following theorem.

**Theorem A2 (Putinar’s Positivstellensatz ([Putinar, 1993])** Let \(S\) be a basic semialgebraic set defined by \(\{x \mid g_i(x) \geq 0, i = 1, \ldots, l_g, h_j(x) = 0, j = 1, \ldots, l_h\}\), and \(f\) be a polynomial in \(\mathbb{R}[x]\). If assumption A1 holds, then \(f\) is strictly nonnegative on \(S\) if and only if there exists a sequence of sum-of-squares polynomials \(\sigma_i\) and polynomials \(\lambda_j\), such that \(f = \sigma_0 + \sum_{i=1}^{l_g} \sigma_i g_i + \sum_{j=1}^{l_h} \lambda_j h_j\).

There exist other types of nonnegativity certificates, for which we refer the interested reader to Blekherman et al. (2012).

**B.3. Polynomial Optimization and the Moment-SOS Hierarchy**

We introduce basic concepts of the moment-SOS hierarchy which are used in the derivation in the main text, specifically Theorem 2. For further information, we refer to Lasserre (2009); Blekherman et al. (2012).

**Polynomial optimization.** Consider the following polynomial optimization problem (POP):

\[
\begin{align*}
\min_{x} & \quad f(x) \\ \text{subject to} & \quad g_i(x) \geq 0 & i = 1 \cdots l_g \\
& \quad h_j(x) = 0 & j = 1 \cdots l_h
\end{align*}
\]

where \(g_i(x), h_j(x)\) are polynomials. Denote the basic semialgebraic set (feasible set of (A1)) \(S := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \ldots, l_g, h_j(x) = 0, j = 1, \ldots, l_h\}\). Generally speaking, (A1) is a nonconvex problem, and it is NP-hard to solve. However, we can relax problem (A1) into a sequence of convex semidefinite programming (SDP) problems, which guarantees to monotonically converge to the optimal value of (A1). The relaxation is called Lasserre’s moment-SOS hierarchy (Lasserre, 2001). Next we will introduce this technique briefly.

First, we can recast the polynomial optimization problem as an infinite dimensional linear programming problem.

\[
\begin{align*}
\min_{\mu \in \mathcal{M}(S)_+} & \quad \int f(x) d\mu \\
\text{subject to} & \quad \int_S d\mu = 1
\end{align*}
\]

where \(\mathcal{M}(S)_+\) is the space of all finite Borel measures supported on \(S\). Then we introduce a linear functional \(L_z(f)\)

\[
f(x) = \sum_{\alpha \in \mathbb{N}^n} f_\alpha x^\alpha \rightarrow L_z(f) = \sum_{\alpha \in \mathbb{N}^n} f_\alpha z_\alpha,
\]

where \(\{z_\alpha, \alpha \in \mathbb{N}^n\}\) is sequence of real numbers (often known as the pseudomoment vector). For a sequence \(\{z_\alpha, \alpha \in \mathbb{N}^n\}\), we say it has a representing measure on \(S\) if there exists a measure \(\mu\).
supported on $S$ satisfying: $z_\alpha = \int_S x^\alpha d\mu$. Then the optimization problem (A2) is equivalent to:

$$\min_z \quad L_z(f)$$

subject to

$$z_0 = 1$$

$$z \text{ has a representing measure on } S$$

If we introduce a new notation of moment matrix and localizing matrix, we can cast a necessary and sufficient condition for the existence of representing measure as infinite semidefinite constraints.

**Moment matrix.** Given a $s_n(2r)$-sequence $z = (z_\alpha)$, the moment matrix $M_r(z)$ is defined as follows:

$$M_r(z)(\alpha, \beta) = L_z(x^{\alpha + \beta}) = z_{\alpha + \beta} \quad \forall \alpha, \beta \in \mathbb{N}^n_r$$

where $\mathbb{N}^n_r$ is the set of all multi-indices $\alpha \in \mathbb{N}^n$ with $|\alpha| = \sum_{i=1}^{n} \alpha_i \leq r$.

**Localizing matrix.** Given a polynomial $g(x) \in \mathbb{R}[x]$, the localizing matrix with respect to $z, g$ is defined as:

$$M_r(g, z)(\alpha, \beta) = L_z(x^{\alpha + \beta}g(x)) = \sum_{\gamma \in \mathbb{N}^n} z_{\alpha + \beta + \gamma} g_\gamma.$$

If the set satisfies the Archimedean property, we have the following theorem.

**Theorem A3 (Representing measure (Lasserre, 2009, Theorem 3.8))** Let $z = (z_\alpha)_{\alpha \in \mathbb{N}^n}$ be a given infinite sequence in $\mathbb{R}$ and assume the basic semialgebraic set $S$ is compact. Then $z$ has a representing measure on $S$ if and only if for all $r \in \mathbb{N}$

$$M_r(z) \succeq 0$$

$$M_r(g_i z) \succeq 0, \quad i = 1 \ldots l_g$$

$$M_r(h_j z) = 0, \quad j = 1 \ldots l_h$$

This theorem, however, is not practical since it requires infinite number of constraints. In practice we solve the relaxation of the above theorem by only using constraints with a finite $r$.

Thus the final problem for Lasserre’s hierarchy, with a relaxation order $\kappa$ is:

$$\min_z \quad L_z(f)$$

subject to

$$z_0 = 1$$

$$M_\kappa(z) \succeq 0$$

$$M_\kappa(g_i z) \succeq 0, \quad i = 1 \ldots l_g$$

$$M_\kappa(h_j z) = 0, \quad j = 1 \ldots l_h$$

Using monomial basis $[x]_\kappa$, which is the column vector of all monomials of degree up to $\kappa$, the last part constraints can also be interpreted as:

$$L_z(h_j [\theta]_{2\kappa-d_{eg}(h_j)}) = 0, \quad j = 1, \ldots, l_h.$$
Appendix C. Proof of Theorem 1

Proof  First let’s prove (i). We show that the following problem is equivalent to (MEE)

\[
\begin{align*}
\max_{E,b,c} & \quad \log \det E \\
\text{subject to} & \quad 1 - (\xi^T E \xi + 2 b^T \xi + c) \geq 0, \quad \forall \theta \in S \quad (A5a) \\
& \quad c - b^T E^{-1} b = 0 \quad (A5c)
\end{align*}
\]

because with the last equality, we can see that the first inequality constraint is equivalent to:

\[ (\xi + E^{-1} b)^T E (\xi + E^{-1} b) \leq 1. \]

With \( \mu := -E^{-1}b \), problem (A5) recovers (MEE).

Next, we only need to show the following problem is equivalent to problem (A5).

\[
\begin{align*}
\max_{E,b,c} & \quad \log \det E \\
\text{subject to} & \quad 1 - (\xi^T E \xi + 2 b^T \xi + c) \geq 0, \quad \forall \theta \in S \quad (A6b) \\
& \quad \begin{bmatrix} E & b \\ b^T & c \end{bmatrix} \succeq 0 \quad (A6c)
\end{align*}
\]

The first constraint (A6b) indicates enclosure. We can equivalently write it as:

\[ (\xi + E^{-1} b)^T E (\xi + E^{-1} b) \leq 1 - (c - b^T E^{-1} b), \quad \forall \theta \in S \]

By \( E \) is positive definite, we can conclude that \( c - b^T E^{-1} b < 1 \). The second constraint (A6c) implies \( c - b^T E^{-1} b \geq 0 \) by the Schur complement theorem. We will show that in fact we always have \( c - b^T E^{-1} b = 0 \) at the optimal solution of (A6).

Consider the ellipsoid representation \( \xi^T E \xi + 2 b^T \xi + c \leq 1 \). Note that one ellipsoid can correspond to different pairs of \( (E, b, c) \) due to

\[ \xi^T E \xi + 2 b^T \xi + c \leq 1 \iff m(\xi^T E \xi + 2 b^T \xi + c) \leq m, \forall m > 0. \]

All the equivalent class can be written as: \( (mE, mb, mc - m + 1) \) for any \( m > 0 \).

Suppose now we have a feasible solution \( (E_0, b_0, c_0) \) which satisfies \( c_0 - b_0^T E_0^{-1} b_0 > 0 \). We claim that this solution must not be optimal, because we can construct a better solution by maintaining the same ellipsoid, as follows.

For \( m > 1 \), consider \( (E_1, b_1, c_1) = (mE_0, mb_0, mc_0 - (m - 1)) \). Clearly, if \( (mE_0, mb_0, mc_0 - (m - 1)) \) is a feasible solution, then it will always be a better solution than \( (E_0, b_0, c_0) \) due to \( \log \det(mE_0) > \log \det(E_0) \) when \( m > 1 \).

Now we show that indeed we can make \( (E_1, b_1, c_1) \) feasible for (A6). It suffices to construct \( (E_1, b_1, c_1) \) feasible for (A6c). To do so, we write

\[
\begin{align*}
c_1 - b_1^T E_1^{-1} b_1 &= mc_0 - (m - 1) - (mb_0)^T (mE_0)^{-1} (mb_0) \\
&= m(c_0 - b_0^T E_0^{-1} b_0) - (m - 1) \\
&= (m - 1)(c_0 - b_0^T E_0^{-1} b_0 - 1) + c_0 - b_0^T E_0^{-1} b_0.
\end{align*}
\]
Note that $c_0 - b_0^T E_0^{-1} b_0 > 0$ by our hypothesis, and $c_0 - b_0^T E_0^{-1} b_0 - 1 < 0$ by enclosure. Therefore, we can always choose some $m - 1 > 0$ that is sufficiently small such that $c_1 - b_1^T E_1^{-1} b_1 \geq 0$ and hence $(E_1, b_1, c_1)$ can be made feasible.

Therefore, any optimal solution of problem (A6) must satisfy (A5).

Then we leverage the sum-of-squares technique to relax the problem. By substituting constraints (A6b) with (10b),(10c) and (10d). Because of the certificate of nonnegativity (cf. Section B.2), we can always get an enclosing ellipsoid. Also from the derivation above, we shall always have $c - b^T E^{-1} b = 0$. Therefore, the ellipsoid we get can be written as: $(\xi + E^{-1} b)^T E(\xi + E^{-1} b) \leq 1$.

Then we prove (ii). First, it’s easy to see that when $\kappa$ increases, the feasible region of the problem (10) gets larger, so the optimal value of the optimization problem will get larger, which means that Vol($E$) decreases.

Suppose the optimal solution of the minimum enclosing ellipsoid problem (MEE) is: $E^* := (\xi - \mu^*)^T E^*(\xi - \mu^*) \leq 1$. Then for any sufficiently small $\varepsilon > 0$ making sure that $E^* - \varepsilon I_n \succeq 0$, we can have $1 - (\xi - \mu^*)^T (E^* - \varepsilon I_n)(\xi - \mu^*)$ strictly positive on $S$. Then leveraging the Putinar’s Positivstellensatz (cf. Section B.2), we know that there exists sum of squares polynomials $\sigma_i(\theta)$ satisfies:

$$1 - ((\xi - \mu^*)^T (E^* - \varepsilon I_n)(\xi - \mu^*)) = \sum_{i=0}^{l_n} \sigma_i(\theta) g_i(\theta) + \sum_{j=1}^{l_h} \lambda_j(\theta) h_j(\theta),$$

which shows that the point $(E^* - \varepsilon I_n, \mu^*)$ is a feasible point of the problem (10) for some $\kappa$.

From the continuity property of the determinant function, we can know that $\det(A - \varepsilon I_n) = \det(A) + O(\varepsilon)$. By letting $\varepsilon \to 0$, we can show that the optimal value of the problem (10) converges to the optimal value of the problem (MEE) when $\kappa \to \infty$.

Appendix D. Certificate of MEE

We present a result to check the convergence of the SOS-based MEE algorithm in Theorem 1 when the set $S$ is convex.

**Proposition A4 (Certificate of MEE)** Let $S_\xi$ be convex, and $(E_*, b_*, c_*)$ be an optimizer of (10) with certain $\kappa$. Consider the following (nonconvex) polynomial optimization problem

$$\min_{\eta \in S_\xi} 1 - (\xi^T E_* \xi + 2b_*^T \xi + c_*).$$

Denote the optimal value of the $\nu$-th order moment-SOS relaxation of (A7) as $p_\nu^*$ (cf. Section B.3). Suppose at a certain relaxation order $\nu$, we have (i) $p_\nu^* = 0$, (ii) a finite number of global minimizers of (A7) can be extracted as $\{\xi_1^*, \ldots, \xi_K^*\}$, and (iii) there exist positive $\{\alpha_1, \ldots, \alpha_K\}$ such that

$$\sum_{k=1}^K \alpha_k (\xi_k^* + E_*^{-1} b_*) = 0, \quad \sum_{k=1}^K \alpha_k (\xi_k^* + E_*^{-1} b_*) (\xi_k^* + E_*^{-1} b_*)^T = E_*^{-1},$$

then $E = \{\xi \in \mathbb{R}^d \mid \xi^T E_* \xi + 2b_*^T \xi + c_* \leq 1\}$ is the unique solution of (MEE).
Before presenting the proof of Proposition A4, we briefly describe the intuition. By (Xie, Miaolan, 2016, Theorem 3.2.5), $E$ is the minimum enclosing ellipsoid of $S_\xi$ if and only if there exists a finite number of contact points at the intersection of $E$ and $S_\xi$ that satisfy (A8). Since the contact points lie on the boundary of $E$, they must attain zero value in (A7), and hence must be its global minimizers. Therefore, we can use the moment-SOS hierarchy to find the contact points. Section G.1 gives a concrete numerical example of computing the certificate of convergence.

**Proof** We first present John’s theorem of minimum enclosing ellipsoid.

**Lemma A5 (John’s Theorem)** Let $C \subseteq \mathbb{R}^n$ be a convex body and $C \subseteq B_H(\bar{x}, 1)$, where $B_H(\bar{x}, 1) := \{x \in \mathbb{R}^n | (x - \bar{x})^T H (x - \bar{x}) \leq 1\}$, $H > 0$. Then the following statements are equivalent:

- $B_H(\bar{x}, 1)$ is the minimum volume ellipsoid enclosing $C$.
- There exists some $r \in \{n, n + 1, \ldots, n(n + 3)/2\}$ contact points $p_1, \ldots, p_r \in \partial(C) \cap \partial(B_H(\bar{x}, 1))$ and $\alpha \in \mathbb{R}_{++}^r$ such that

\[
H^{-1} = \sum_{i=1}^r \alpha_i (p_i - \bar{x})(p_i - \bar{x})^T \quad \text{and} \quad \sum_{i=1}^r \alpha_i (p_i - \bar{x}) = 0.
\] (A9)

Then we leverage this Lemma to prove Proposition A4.

From Theorem 1, the bounding ellipsoid of $S_\xi$ is $E = (\xi + E_*^{-1}b_*)^T E_* (\xi + E_*^{-1}b_*)^T \leq 1$. Based on Lemma A5, if the ellipsoid is the minimum volume enclosing ellipsoid, there must exists contact points satisfying (A9). Then the optimal value of the optimization problem (A7) is 0 and the solutions are the contact points.

Suppose by solving the optimization problem (A7) at certain relaxation order $\nu$, we get the optimal value $\rho_\nu^* = 0$, then the solutions are the contact points. Suppose we can extract the optimal solutions $\xi_1^*, \ldots, \xi_K^* \in \partial(S_\xi) \cap \partial(E)$ using the result in Henrion and Lasserre (2005). Then from John’s theorem, if there exist $\alpha \in \mathbb{R}_{++}^K$ such that,

\[
\sum_{k=1}^K \alpha_k (\xi_k^* + E_*^{-1}b_*) = 0, \quad \sum_{k=1}^K \alpha_k (\xi_k^* + E_*^{-1}b_*)(\xi_k^* + E_*^{-1}b_*^T) = E_*^{-1}
\] (A10)

the ellipsoid $E$ is the minimum volume enclosing ellipsoid. Note that (A10) is a system of linear equations in $\alpha$ and can be easily solved.

**Appendix E. Proof of Theorem 2**

We first show that there exists $R > 0$ such that the optimization problem (GCC) is equivalent to the following problem

\[
\eta^* = \min_{\mu \in B_R(O)} \max_{\xi \in S_\xi} \|\mu - \xi\|^2_Q,
\] (A11)

where $B_R(O)$ denotes a ball of radius $R$ centered at $O$. 

17
Theorem A6 (Equivalence of the problem) There exists $R > 0$ such that the original problem

$$
\eta^* = \min_{\mu \in \mathbb{R}^d} \max_{\xi \in S_\xi} \|\mu - \xi\|^2_Q
$$

is equivalent to the problem:

$$
\eta^* = \min_{\mu \in B_R(O)} \max_{\xi \in S_\xi} \|\mu - \xi\|^2_Q
$$

Proof By Sain et al. (2016), the Chebyshev center of the bounded set in $\mathbb{R}^n$ exists and is unique. We only need to show that the solution lies inside some ball $B_R(O)$.

Denote the Chebyshev center as $\mu^*$ and the corresponding radius as $\eta^*$. Due to the assumption of boundedness, there exists $R'$ such that for all $\xi \in S_\xi$,

$$
\|\mu^* - O\| \leq \|\mu^* - s\| + \|s - O\| \leq \eta^* + R',
$$

where $s$ is the point in $S_\xi$ that attains the maximum distance to $\mu^*$, and $R'$ is the distance from $s$ to $O$. Thus we can choose $R = R' + \eta^*$ to complete the proof.

Then using Lasserre’s hierarchy, for relaxation order $\kappa$ satisfying $2\kappa \geq \max\{\deg(g_1), \ldots, \deg(g_l), 2\}$, we can relax the inner maximization problem

$$
\eta(\mu) = \max_{\theta \in S} \mu^T Q \mu - 2 \mu^T Q P \theta + \theta^T P^T Q P \theta \tag{A12}
$$

Denote the pseudomoment vector of degree up to $2\kappa$ as $[z]_{2\kappa}$ and shorthand it as $z$, and we denote the moment matrix with respect to $z$ of order $\kappa$ as $M_\kappa(z)$, then the moment relaxation of (A12) at order $\kappa$ reads:

$$
\eta_\kappa(\mu) = \max_z \mu^T Q \mu - 2 \mu^T Q P L_z(\theta) + \langle C, M_\kappa(z) \rangle \tag{A13}
$$

subject to

$$
M_\kappa(z) \succeq 0, \quad M_0(z) = 1 \tag{A14}
$$

$$
M_{\kappa - \lceil \deg(g_i)/2 \rceil}(g_i z) \succeq 0, \quad i = 1, \ldots, l_g \tag{A15}
$$

$$
L_z(h_j[z]_{2\kappa - \deg(h_j)}) = 0, \quad j = 1, \ldots, l_h \tag{A16}
$$

where $L_z(\theta)$ denotes the order-one moment vector in $z$ (i.e., the sub-vector of $z$ corresponding to the order-one monomials $\theta$), $M_{\kappa - \lceil \deg(g_i)/2 \rceil}(g_i z)$ are the localizing matrices corresponding to inequalities $\{g_i\}_{i=1}^l$, $M_0(z) = 1$ enforces the zero-order pseudomoment to be 1, and the matrix $C$ is

$$
C = \begin{bmatrix}
0 & 0 & 0 \\
0 & A & 0 \\
0 & 0 & 0
\end{bmatrix}
$$

where $A$ satisfies

$$
A_{ij} = \begin{cases} 
(P^T Q P)_{ij} & \text{if } i = j \\
(P^T Q P)_{ij}/2 & \text{else}
\end{cases}
$$

Thus, $\langle C, M_\kappa(z) \rangle$ corresponds to the moment relaxation of $\theta^T P^T Q P \theta$.

Denote $\Omega_\kappa$ as the feasible set of the relaxation (A13). Note that $\Omega_\kappa$ does not depend on $\mu$. 

18
Now that we have relaxed the inner “max” as a convex problem, we bring back the outer “min”

\[
\eta^*_\kappa = \min_{\mu \in B_R(O)} \left\{ \eta_\kappa(\mu) = \max_{z \in \Omega_\kappa} \left\{ \mu^T Q \mu - 2 \mu^T Q P L_z(\theta) + \langle C, M_\kappa(z) \rangle \right\} \right\}
\] (A17)

from Lasserre’s hierarchy, we have \( \eta_\kappa(\mu) \geq \eta(\mu) \) for any \( \mu \) and any relaxation order \( \kappa \), thus we have

\[
\eta^*_\kappa = \min_{\mu \in B_R(O)} \eta_\kappa(\mu) \geq \min_{\mu \in B_R(O)} \eta(\mu) = \eta^*
\]

i.e., the optimum of (A17) always provides an upper bound for \( \eta^* \).

Notice that now the objective function of the relaxed problem (A17) is affine in \( \theta \) and convex in \( \mu \). And note the \( B_R(O) \) is a compact set. So that we can apply Sion’s minimax theorem to switch the order of “min” and “max”.

**Theorem A7 (Sion’s minimax theorem (Sion, 1958, Corollary 3.3))** Let \( M, N \) be convex spaces one of which is compact, and \( f \) a function on \( M \times N \), quasi-concave-convex and upper semicontinuous - lower semicontinuous. Then \( \inf_{x \in N} \sup_{y \in M} f(x, y) = \sup_{y \in M} \inf_{x \in N} f(x, y) \).

Then we get the equivalent problem:

\[
\eta^*_\kappa = \max_{z \in \Omega_\kappa} \min_{\mu \in B_R(O)} \mu^T Q \mu - 2 \mu^T Q P L_z(\theta) + \langle C, M_\kappa(z) \rangle
\] (A18)

Note that by taking the derivative of the objective function with respect to \( \mu \), now the inner minimization problem has an explicit solution:

\[
\mu^*_\kappa = PL_z(\theta)
\]

Next we will show that: we can choose the \( R_1 \) big enough such that for all \( \kappa, \mu^*_\kappa \in B_{R_1}(O) \).

**Theorem A8 (Attainability of the minimum)** There exists \( R_1 > 0 \) which doesn’t rely on the relaxation order \( \kappa \), such that \( \mu^*_\kappa \in B_{R_1}(O) \).

**Proof** \( P \) does not rely on \( \kappa \), thus it suffices to show that \( L_z(\theta) \) is bounded independent of \( \kappa \).

Due to the assumption of compactness of \( S \), there exists \( R_1 \) such that \( S \subset B_{R_1}(O) \), which is independent from the optimization procedure. Thus, we can add an extra inequality constraint \( g_{l+1}(\theta) = -\theta^T \theta + R_1^2 \geq 0 \) without affecting the original problem. So without loss of generality, we can assume that we include the constraint \( g_{l+1}(\theta) = -\theta^T \theta + R_1^2 \geq 0 \) in the original problem.

Then we examine the constraint of the localizing matrix. The corresponding constraint for the inequality \( g_{l+1}(\theta) = -\theta^T \theta + R_1^2 \geq 0 \) is \( M_{\kappa-1}(g_{l+1}y) \geq 0 \). The principal minors of a positive semidefinite matrix is nonnegative, so the first principal minor of \( M_{\kappa-1}(g_{l+1}y) \) is nonnegative, which implies \( L_z(g_{l+1}) \geq 0 \). And this is equivalent to the constraint that \( \langle T, M_\kappa(z) \rangle \leq R_1^2 \) where

\[
T = \begin{bmatrix}
0 & 0 & 0 \\
0 & I_n & 0 \\
0 & 0 & 0
\end{bmatrix}
\]

Then let’s pass this constraint to \( L_z(\theta) \) via the moment matrix. Denote the \((n + 1) \times (n + 1)\) principal submatrix as:

\[
\begin{bmatrix}
1 & L_z(\theta)^T \\
L_z(\theta) & A
\end{bmatrix}
\]
where \( \text{tr} (A) = \langle T, M_\kappa(z) \rangle \leq R_1^2 \).

The moment matrix \( M_\kappa(z) \succeq 0 \), so that the \( (n+1) \times (n+1) \) principal submatrix is also positive semidefinite. By Schur complement, we know that \( A - L_z(\theta)L_z(\theta)^T \succeq 0 \). Thus \( R_1^2 \geq \text{tr} (A) \geq \text{tr} (L_z(\theta)L_z(\theta)^T) = L_z(\theta)^T L_z(\theta) \). This implies \( L_z(\theta) \) lies inside \( B_{R_1}(O) \), which is independent of the choice of \( \kappa \).

Thus we can know that we can solve the inner minimization exactly. Plugging the \( \mu^*_\kappa \) back into the objective function, we can get another equivalent problem:

\[
\eta^*_\kappa = \max_{z \in \Omega_\kappa} -L_z(\theta)^T P^T Q P L_z(\theta) + \langle C, M_\kappa(z) \rangle
\]

or equivalently in minimization form

\[
-\eta^*_\kappa = \min_{z \in \Omega_\kappa} L_z(\theta)^T P^T Q P L_z(\theta) - \langle C, M_\kappa(z) \rangle
\]

This is already a convex problem. But it is a convex quadratic SDP problem. By leveraging lifting technique, we can also convert it into a standard form SDP as follows:

\[
-\eta^*_\kappa = \min_{z \in \Omega_\kappa} t - \langle C, M_\kappa(z) \rangle
\]

subject to \[
\begin{bmatrix}
Q^{-1}
& P L_z(\theta)

L_z(\theta)^T P^T
& t
\end{bmatrix} \succeq 0
\] (A22)

Proof of convergence. Because the basic semialgebraic set \( S \) is compact, we can always assume that the Assumption A1 holds. Then we will introduce two theorems which are the main ingredients in our proof of the convergence result.

Theorem A9 (Degree bound of Putinar’s Positivstellensatz Nie and Schweighofer (2007)) Let \( K \) be a compact basic semialgebraic set defined by:

\[
\{ x \in \mathbb{R}^n | g_i(x) \geq 0 \quad i = 1 \ldots m \}
\]

satisfying the Archimedean property. Assume \( \emptyset \neq K \subset (-1,1)^n \). Then there is some \( c > 0 \) such that for all \( f \in \mathbb{R}[x] \) of degree \( d \), and positive on \( K \) (i.e. such that \( f^* := \min f(x) : x \in K \)), the representation in Putinar’s Positivstellensatz A2 holds with

\[
\deg(f_j g_j) \leq c \exp((d^2 n^d \| \|f\|_0^* \|^c)
\]

where \( \| f \|_0 := \max_{\alpha} \frac{f_{\alpha}}{\alpha!} \) and \( C_{|\alpha|} = \frac{|\alpha|!}{\alpha_1! \ldots \alpha_n!} \).

We have the following theorem to guarantee the convergence of the Lasserre’s hierarchy in this problem, which is the extension of the proof in Lasserre (2009). Before the proof of our main theorem, we shall first show the following lemma which summarizes the relationship between optimization problems which we will leverage in the proof of the main theorem.

Lemma A10 (Relationship between optimization problems Lasserre (2009)) Given a basic semialgebraic set \( S := \{ x \in \mathbb{R}^n | g_i(x) \geq 0 \quad i = 1 \ldots l_g \} \), and a polynomial \( f(x) \), there are four related optimization problems:
1. Moment problem:

\[ \eta_{\text{mom}} = \max_x f(x) \quad \text{(A23)} \]

subject to \(x \in S\) \quad \text{(A24)}

2. Dual problem to the moment problem:

\[ \eta_{\text{pop}} = \min_\lambda \lambda \quad \text{(A25)} \]

subject to \(\lambda - f(x) \geq 0 \quad \forall x \in S\) \quad \text{(A26)}

3. Primal relaxation problem

(with relaxation order \(\kappa\) satisfying \(2\kappa \geq \max\{\deg(g_i), \deg(f)\}\)):

\[ \eta_\kappa = \max_z L_z(f) \quad \text{(A27)} \]

subject to \(z_0 = 1\) \quad \text{(A28)}

\[ M_\kappa(z) \succeq 0 \quad \text{(A29)} \]

\[ M_\kappa - \lceil \frac{\deg(g_i)}{2} \rceil (g_i z) \succeq 0 \quad \text{(A30)} \]

4. Dual relaxation problem:

(with relaxation order \(\kappa\) satisfying \(2\kappa \geq \max\{\deg(g_i), \deg(f)\}\)):

\[ \eta^*_\kappa = \min_\lambda \lambda \quad \text{(A31)} \]

subject to \(\lambda - f = \sigma_0 + \sum_{j=1}^{l_g} \sigma_j g_j\) \quad \text{(A32)}

\(\sigma_j\) is sum of squares polynomial \quad \text{(A33)}

\[ \deg(\sigma_j) \leq 2\kappa - \deg(g_j), j = 1 \cdots l_g \quad \text{(A34)} \]

We can conclude that :

\[ \eta_{\text{mom}} = \eta_{\text{pop}} \leq \eta_\kappa \leq \eta^*_\kappa \quad \forall \kappa \text{ satisfying } 2\kappa \geq \max\{\deg(g_i), \deg(f)\} \quad \text{(A35)} \]

We prove the Theorem 2 by proving the following equivalent theorem:

**Theorem A11** Let \(S\) be a compact basic semialgebraic set \(S := \{g_i(x) \geq 0, i = 1 \cdots l_g, h_j(x) = 0, j = 1 \cdots l_h\}\) which has non-empty interior, such that the Assumption A1 holds. Then the sequence of optimal values of the Lasserre’s hierarchy (A21) converges to the optimal value of the original problem when the relaxation order \(\kappa \to \infty\). Plus, for any relaxation order \(\kappa\) satisfying \(2\kappa \geq \max\{\deg(g_i), \deg(h_i), 2\}\), we have \(\eta^*_\kappa \geq \eta^*_{\kappa+1} \geq \eta^*\)

**Proof** First we fixed \(\mu\) in the outer minimization and only focus on the inner maximization problem. Here we denote \(f(\theta) = \|\mu - P\theta\|_2^2\) as the objective function by holding \(\mu\) as a constant. From lemma A10, we can know that when the relaxation order \(\kappa\) satisfies \(2\kappa \geq \max\{\deg(g_i), \deg(h_i), 2\}\), we have:

\[ \eta_{\text{mom}} = \eta_{\text{pop}} \leq \eta_i \leq \eta^*_i \]
Using the definition of the dual problem (A25), if we fix an arbitrary \( \epsilon > 0 \), we can know that there exists a feasible point for problem (A25) namely \( \lambda \) which satisfies:

\[ \eta_{\text{mom}} = \eta_{\text{pop}} \leq \lambda \leq \eta_{\text{mom}} + \epsilon = \eta_{\text{pop}} + \epsilon \]

Next, consider \( \bar{\lambda} = \lambda + \epsilon \). Thus we can know that \( \bar{\lambda} - f \geq \epsilon > 0 \ \forall x \in S \), then we can use Putinar’s Positivstellensatz A2:

\[ \bar{\lambda} - f = \sigma_0 + \sum_{i=1}^{l_\theta} \sigma_i g_i + \sum_{j=1}^{l_h} \lambda_j h_j \quad \sigma_j \text{ is sum of squares polynomial} \]

which shows that once \( 2\kappa \geq \max\{\deg(\sigma_j g_j), \deg(\lambda_j h_j)\} \), we can have

\[ \eta_{\text{pop}} = \eta_{\text{mom}} \leq \eta_{\kappa} \leq \eta_{\kappa}^* \leq \bar{\lambda} = \lambda + \epsilon \leq \eta_{\text{mom}} + 2\epsilon \]

and as \( \epsilon \) is arbitrary, we can let \( \epsilon \to 0 \).

In the following discussion we will only focus on the inequality constraints, for all the equality constraints can be interpreted as two inequality constraints.

The only barrier we will encounter when we want to prove the minimax version is that, the bound of degree of \( \sigma_1 g_i \) may not be the same for different \( \mu \). So we need a uniform convergence result with respect to \( \mu \) (the outer min). Thus we want to apply the degree bound of Putinar’s Positivstellensatz. Remind that we have shown the equivalence of problem by restricting the feasible set of \( \mu \) to a ball \( B_{R_1}(O) \) for all \( \kappa \). Also consider the ball of radius \( R \) contains \( S \). Then \( S \) is also contained in a box \([-R, R]^n\).

Consider the polynomial \( \tilde{f}(\tilde{\theta}) = \bar{\lambda} - f(\tilde{\theta} \cdot R), \) where \( \tilde{\theta} \in S/R \subseteq [-1, 1]^n \). Suppose \( \lambda_{\text{max}} \) is the maximum eigenvalue of \( Q \), then we apply Theorem A9. The parameters are as follows: \( d = 2, n \) is the dimension of vector \( \theta \). Note that: \( \|f\|_0 = \max\{2\lambda_{\text{max}} R_1 \tilde{\theta}(i), \lambda_{\text{max}} R^2\} \), which can be bounded by a constant independent of \( \mu \). Also note that: \( f^* = \min\{f(\theta) : \theta \in S\} > \epsilon \), thus we can choose \( \epsilon \) to bound \( f^* \). Now we as long as \( 2\kappa \geq c \exp((d^2 n^d \|f\|_0 c^d) \epsilon) \), we can guarantee all the \( \sigma_i g_i \) will have degree less than \( 2\kappa \). Thus we can guarantee the inner maximization is exact, and the convergence of the minimax problem follows.

\[ \text{Appendix F. Proof of Theorem 3} \]

\[ \text{Proof} \] We first introduce some basics about unit quaternions.

\textbf{Unit quaternion.} There are multiple equivalent representations of 3D rotations (Barfoot, 2017). The most popular choice is a 3D rotation matrix \( R \in SO(3) \). Another popular choice is a unit quaternion (Yang and Carlone, 2019). A unit quaternion is a 4-D vector with unit norm, i.e., \( q \in S^3 := \{q \in \mathbb{R}^4 \mid q^T q = 1\} \). Given a unit quaternion \( q = (q_0, q_1, q_2, q_3) \), its corresponding rotation matrix reads

\[ R(q) = \begin{bmatrix} 2(q_0^2 + q_1^2) - 1 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & 2(q_0^2 + q_2^2) - 1 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & 2(q_0^2 + q_3^2) - 1 \end{bmatrix} . \tag{A36} \]
Clearly, \( q \) and \( -q \) corresponds to the same rotation matrix. The inverse of a rotation matrix \( R \) is \( R^{-1} = R^\top \), and the inverse of a unit quaternion \( q \) is \( q^{-1} \), which simply flips the sign of \( q_1, q_2, q_3 \)
\[
q^{-1} = [q_0, -q_1, -q_2, -q_3]^\top.
\] (A37)

The product of two quaternions \( q_c = q_a q_b \) is defined as
\[
q_c = q_a q_b = \begin{bmatrix}
q_{a,0} & -q_{a,1} & -q_{a,2} & -q_{a,3} \\
q_{a,1} & q_{a,0} & -q_{a,3} & q_{a,2} \\
q_{a,2} & q_{a,3} & q_{a,0} & -q_{a,1} \\
q_{a,3} & -q_{a,2} & q_{a,1} & q_{a,0}
\end{bmatrix} q_b.
\] (A38)

The next proposition characterizes the equivalence of distance metrics on \( SO(3) \) and \( S^3 \).

**Proposition A12 (Equivalent expression of distance on \( SO(3) \) and \( S^3 \))** Given two rotations \( R, Q \), the geodesic distance is defined as:
\[
\theta_1 = \arccos \left( \frac{\text{tr} \left( R^\top Q \right) - 1}{2} \right).
\]

Let \( q_R \) and \( q_Q \) be the unit quaternions corresponding to \( R \) and \( Q \), the geodesic distance on \( S^3 \) is defined as
\[
\theta_2 = 2 \arccos(\langle q_R, q_Q \rangle),
\]
where \( \langle \cdot, \cdot \rangle \) denotes the inner product viewing \( q_R, q_Q \) as vectors in \( \mathbb{R}^4 \). It holds that \( \theta_1 = \theta_2 \).

**Proof** Consider the rotation \( A = R^\top Q = R^{-1} Q \). The rotation angle of \( A \) is
\[
\theta_1 = \arccos \left( \frac{\text{tr}(A) - 1}{2} \right).
\]

Let \( q_A = q_{R}^{-1} q_{Q} \) \( (q_{R}^{-1} \) is the inverse quaternion), by direct computation using (A37) and (A38), we know
\[
q_{A,0} = \langle q_R, q_Q \rangle.
\]
where \( \langle \cdot, \cdot \rangle \) denotes the inner product viewing \( q_R, q_Q \) as vectors in \( \mathbb{R}^4 \). For \( A \), \( \cos \frac{\theta_2}{2} = q_{A,0} \). By convention, we want \( \theta_2 \in [0, \pi) \), so we need to carefully select the quaternions to guarantee the inner product is positive. For simplicity, we can just take the absolute value for \( q_{A,0} \). Thus, the rotation angle of \( q_A \) is:
\[
\theta_2 = 2 \arccos(\langle q_R, q_Q \rangle).
\]
These two angle refers to the same rotation, thus \( \theta_1 = \theta_2 \), which completes the proof. \( \blacksquare \)

The fact that for quaternions, \( q \) and \( -q \) represent the same rotation is the reason why we need to take the absolute value of the inner product. Geometrically, this is saying the same rotation \( R \) corresponds to two quaternions on \( S^3 \) and these two quaternions are actually antipodal. When considering the distance of two rotations on \( S^3 \), we shall always pick the nearest pair of points.

Now think of the set \( S_R \), it will get mapped to two antipodal sets on \( S^3 \). When the set \( S_R \) is reasonably small, we can get rid of the absolute value when calculating the minimum enclosing ball.
Figure A2: Analogy of Proposition A13 in $\mathbb{R}^2$. The black line segment is a subset of $S^1$ (the dotted circle). The solid blue circle is the minimum enclosing ball of the black line segment in $\mathbb{R}^2$. The center of the solid blue circle, projected onto $S^1$, is the center of the minimum enclosing geodesic ball on $S^1$.

Precisely speaking, if there exists a rotation $\bar{R}$ such that $d_{SO(3)}(\bar{R}, R) \leq \frac{\pi}{2}, \forall R \in S_R$, we can know that on $S^3$, there exists a quaternion $\bar{q}$ such that the maximum angle difference from $\bar{q}$ to the set is smaller than $\frac{\pi}{4}$. Thus by triangle inequality, in each branch on the $S^3$, the angle between any two points in the set will be less than $\frac{\pi}{2}$. Under this assumption, we can completely get rid of the absolute value when we are calculating the distance of any two quaternions.

Thus the problem of finding the minimum enclosing ball of $S_R$ on $SO(3)$ can be conveniently translated into finding a minimum enclosing ball for one branch of the counterpart of $S_R$ on $S^3$. We claim finding the minimum enclosing ball on $S^3$ is equivalent to finding the minimum enclosing ball in $\mathbb{R}^4$.

**Proposition A13 (Finding miniball on $S^3$ is equivalent to finding miniball in $\mathbb{R}^4$)** Given a set of rotations $S_R$ with angle span less than $\frac{\pi}{2}$ defined on $S^3$, the miniball of the set on $S^3$ is just the restriction of the miniball of the set in $\mathbb{R}^4$ onto $S^3$.

An analogy of Proposition A13 is shown in Fig. A2. The underlying geometric intuition is that: the center of the minimum enclosing ball in $\mathbb{R}^4$ must line in the affine space defined by the contact points. And the minimum enclosing ball on $S^3$ is just the equator of the minimum enclosing ball in $\mathbb{R}^4$. Thus by projecting the corresponding center of the minimum enclosing ball in $\mathbb{R}^4$ onto $S^3$, we can get the center of the minimum enclosing ball on $S^3$.

**Proof** We’ll show that the restriction of the miniball of the set in $\mathbb{R}^4$ is the miniball on $S^3$. Suppose $B$ is the miniball of the set in $\mathbb{R}^4$, then denote $C = B \cap S^3$. Clearly, $C$ is a geodesic ball that encloses the set $S_R$. We claim that $C$ is also the smallest enclosing geodesic ball.

A key observation is that, the center of $B$ must lie in the affine space, else the radius wouldn’t be the smallest. The miniball must have contact points, which must lie on the intersection of $S^3$ and $B$, i.e., $C$. 

24
Now Suppose $C$ is not the miniball in $S^3$ and $C_1$ is. Then the radius of $C_1$ on its affine space is smaller than the radius of $C$ on $T$. Then consider the ball $B_1$ in $\mathbb{R}^4$ whose equator is $C_1$.

We only need to show that $B_1$ contains the set. If we prove this, then $B_1$ a smaller miniball in $\mathbb{R}^4$ than $B$, which causes a contradiction.

We only need to show that $B_1$ contains the ball crown enclosed by $C_1$ (in $\mathbb{R}^4$). Suppose $a_1 \in \mathbb{R}^4$ is the center of $C_1$, then the radius of $C_1$ is $r_1 = \sqrt{1 - \|a_1\|^2}$. The farthest distance from the ball crown enclosed by $C_1$ to the $a_1$ is $1 - \|a_1\|$. By simple mathematics, we know that $1 - \|a_1\| < \sqrt{1 - \|a_1\|^2}$ for $0 < \|a_1\| < 1$, so $B_1$ contains the ball crown, which completes the proof.

Since the set $S_R$ has two branches on $S^3$ that are antipodal to each other, we need to get only one branch and compute its minium enclosing ball. This is easily done by adding $\bar{q}^T q \geq 0$ to restrict the quaternions to one branch. This leads to the problem in (14) and concludes the proof.

Appendix G. Extra Experiments

We provide extra experiments here to supplement the main experiments in Section 4.

G.1. Simple Geometric Sets

In this section, we apply the GRCC algorithm (12) to simple geometric shapes. We choose $Q = I$ so that we are only interested in computing the Chebyshev center.

- Sphere: $\{x \in \mathbb{R}^3 | \|x\| = r\}$. The origin is the Chebyshev center and $r$ is the radius.
- Ball: $\{x \in \mathbb{R}^3 | \|x\|_2 \leq r\}$. The origin is the Chebyshev center and $r$ is the radius.
- Cube: $\{x \in \mathbb{R}^3 | \|x\|_\infty \leq r\}$. The origin is the Chebyshev center and $\sqrt{3}r$ is the radius.
- Boundary of ellipsoid: $\{x \in \mathbb{R}^3 | x^T P^{-1} x = 1, P \in S_{++}\}$. The origin is the Chebyshev center and the root square of the largest eigenvalue of $P$ is the radius.
- TV screen: $\{x \in \mathbb{R}^3 | x_1^4 + x_2^4 + x_3^4 \leq 1\}$. The origin is the Chebyshev center and $3^{1/4}$ is the radius.

It is worth mentioning that, even if the shapes above are simple, the RCC algorithm (Eldar et al., 2008) actually can’t handle them. For example, some of the sets are not convex (e.g., sphere and boundary of ellipsoid) or not defined by quadratic constraints (e.g., cube, TV screen).

We choose $r = 3$ for the experiment, and the relaxation order $\kappa = 1$ ($\kappa = 2$ for the TV screen problem). And we choose $P$ to be a random positive definite matrix. The results are shown in Table A2. Center difference is the Frobenius norm of the difference between the real Chebyshev center and the result of the SDP relaxation. Radius difference is the difference between the solution of our algorithms and the ground truth radius.

We can see that for simple geometrical shapes, the GRCC relaxation is tight even when $\kappa = 1$.

Certificate for the finite convergence of the MEE. Here we present the example of TV Screen and leverage the certificate of MEE (A4) proposed in Appendix D to show the finite convergence. For this symmetric shape, the results of GRCC and MEE coincide.
Table A2: Computing the Chebyshev center of simple geometric shapes.

<table>
<thead>
<tr>
<th>Example</th>
<th>Can RCC handle?</th>
<th>Center difference</th>
<th>Radius difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>NO</td>
<td>4.7924e-17</td>
<td>6.2866e-10</td>
</tr>
<tr>
<td>Ball</td>
<td>YES</td>
<td>9.9467e-18</td>
<td>6.4283e-09</td>
</tr>
<tr>
<td>Cube</td>
<td>YES</td>
<td>1.0080e-17</td>
<td>2.6002e-09</td>
</tr>
<tr>
<td>Ellipsoid</td>
<td>NO</td>
<td>1.1486e-16</td>
<td>9.6512e-10</td>
</tr>
<tr>
<td>TV screen</td>
<td>NO</td>
<td>1.8918e-17</td>
<td>4.1497e-09</td>
</tr>
</tbody>
</table>

Figure A3: The certificate of finite convergence for TV Screen Example.
We first get the radius of the ball by letting $\kappa = 2$. Then by solving the nonconvex optimization problem (A7), we can extract eight contact points:

\[
\begin{pmatrix}
0.7598 \\
0.7598 \\
0.7598
\end{pmatrix},
\begin{pmatrix}
-0.7598 \\
0.7598 \\
-0.7598
\end{pmatrix},
\begin{pmatrix}
-0.7598 \\
0.7598 \\
0.7598
\end{pmatrix},
\begin{pmatrix}
0.7598 \\
-0.7598 \\
0.7598
\end{pmatrix},
\begin{pmatrix}
0.7598 \\
0.7598 \\
-0.7598
\end{pmatrix},
\begin{pmatrix}
-0.7598 \\
0.7598 \\
0.7598
\end{pmatrix},
\begin{pmatrix}
0.7598 \\
-0.7598 \\
0.7598
\end{pmatrix},
\begin{pmatrix}
-0.7598 \\
0.7598 \\
-0.7598
\end{pmatrix}
\]

By solving the feasibility problem (A8) with a slackness of $10^{-15}$, we can get the following $\alpha$:

\[
0, 0.75, 0.75, 0, 0.75, 0, 0.75, 0
\]

which certifies the finite convergence of the MEE algorithm. The figure is shown in Figure A3.

G.2. Random Linear system

To supplement Section 4.1.2, here we provide more results on random linear system identification where the matrix $A_\star$ has size $n_x = 3$ and $n_x = 4$.

Appendix H. Related Work

We provide a brief review of set-membership estimation (SME) in control and perception (mainly related to system identification and object pose estimation), as well as existing algorithms that seek simpler enclosing sets of SME. For related work about system identification, we refer to Oymak and Ozay (2019); Bakshi et al. (2023), for related work about object pose estimation, we refer to Yang and Pavone (2023) and references therein.

As mentioned in Section 1, set-membership estimation (SME) is a framework for model estimation which doesn’t place any assumptions on the statistical distribution of the noise. It only assumes the noises are unknown but bounded. SME dates back to at least Milanese and Vicino (1991) and Kosut et al. (1992) in control theory where the main interest was to estimate unknown parameters of dynamical systems. Li et al. (2023) proves that the SME converges to a single point estimate (i.e., the groundtruth) as the system trajectory becomes long enough when the noise satisfies bounded Gaussian distributions. Recently, Yang and Pavone (2023) applied SME to a computer vision example known as object pose estimation.

Because an SME is defined by complex constraints and the number of constraints increases as the horizon of the system gets longer, an important task in set-membership estimation is to find simpler and more intuitive sets that enclose the original sets (Walter and Pronzato, 1994). The main idea of applying SME framework to system identification is to maintain an easily-computed outer approximation of the uncertainty set. There are two main ways to get the outer approximation, namely batch computation or recursive computation.

For batch calculations, we observe a sequence of inputs and outputs of the system, and the uncertainty set is all the possible parameters that can generate the sequence under the assumption of the bounded noise. Cerone et al. (2011) propose an algorithm leveraging Lasserre’s hierarchy to calculate the bounding interval on every dimension of the parameters. In robotics, similar bounding
Figure A4: Identification of random linear systems with $n_x = 3$ and $n_x = 4$. From left to right: increasing noise bounds $\beta = 0.01, \beta = 0.1, \beta = 1.0$. See Fig. 1(b) in the main text for results with $n_x = 2$.

intervals have been proposed in the context of simultaneous localization and mapping (Ehambram et al., 2022; Mustafa et al., 2018). Casini et al. (2014) propose relaxation algorithms to approximate minimum volume enclosing orthotope. An extension to bounding polytope is also shown by calculating the interval bounds on different directions. Eldar et al. (2008) propose a relaxation algorithm based on Shor’s semidefinite relaxation to calculate the enclosing ball of the sets defined by quadratic constraints. Durieu et al. (2001) consider the ellipsoidal approximation of the intersection of $K$ ellipsoids, and solved the minimum enclosing ellipsoid in a certain parametric class. Dabbene and Henrion (2013) proposed an asymptotically convergent algorithm using sublevel sets of polynomials to enclose the uncertainty set.

For recursive calculations, we sequentially observe new observations and update the approximation according to new observations. Fogel and Huang (1982) propose a recursive formula on how to update the bounding ellipsoid minimizing the volume and trace after each observation. Walter and Piet-Lahanier (1989) propose the algorithm to update the polyhedral description of the feasible parameter set. Valero and Paulen (2019) propose how to update the parallelootope by intersecting the current one with the observation strips, and then find the minimum one. In favor of computational efficiency, recursive methods find an outer approximation of the feasible set at every time step, thus are always not as exact as the batch methods.
**Position of our algorithms.** Our SOS-MEE algorithm (Theorem 1) and GRCC algorithm (Theorem 2) both belong to batch computation algorithms (although it is easy to see that our algorithms can also be used in a recursive fashion). Our GRCC formulation is more general and contains the minimum enclosing interval studied in Cerone et al. (2011) by choosing $\xi$ as just one dimension of $\theta$. Durieu et al. (2001) calculates the minimum enclosing ellipsoids only in a certain parametric set, so it could be sub-optimal, while our SOS-MEE algorithm has asymptotic convergence to the true MEE and the convergence can be certified. Moreover it can only handle the case of intersections of ellipsoids while we can handle arbitrary basic semialgebraic sets. Our GRCC algorithm strictly generalizes (Eldar et al., 2008), as it’s just the simplest case ($\kappa = 1$) of our algorithm. Lastly, although computing the enclosing ellipsoids is not as as general as computing polynomial level sets as in Dabbene and Henrion (2013), our algorithms are more computational efficient and implementable, as demonstrated by numerical experiments in Section 4. Our contributions lie in the three computational enhancements (Section 3) that make the SOS-based framework tractable and applicable to modern system identification and object pose estimation problems.

**References**


Debmalya Sain, Vladimir Kadets, Kallol Paul, and Anubhab Ray. Chebyshev centers that are not farthest points, 2016.


