Towards rigorous robust optimal control via generalized high-order moment expansion

Boris Houska | Jiaqi C. Li | Benoît Chachuat

Summary

This study is concerned with the rigorous solution of worst-case robust optimal control problems having bounded time-varying uncertainty and nonlinear dynamics with affine uncertainty dependence. We propose an algorithm that combines existing uncertainty set-propagation and moment-expansion approaches. Specifically, we consider a high-order moment expansion of the time-varying uncertainty, and we bound the effect of the infinite-dimensional remainder term on the system state, in a rigorous manner, using ellipsoidal calculus. We prove that the error introduced by the expansion converges to zero as more moments are added. Moreover, we describe a methodology to construct a conservative, yet more computationally tractable, robust optimization problem, whose solution values are also shown to converge to those of the original robust optimal control problem. We illustrate the applicability and accuracy of this approach with the robust time-optimal control of a motorized robot arm.

KEYWORDS

moment expansion, robust optimal control, set propagation, time-varying uncertainty

1 | INTRODUCTION

Classical robust control theory has a long history, with enormous practical relevance that has led to countless contributions over the last few decades. In essence, the need for robust control arises whenever an uncertain process or system is subject to critical safety or operational constraints. Worst-case robust optimal control aims to minimize or maximize a function of the states and controls measuring the performance of a dynamic system, while satisfying the terminal and path constraints for all possible realizations of all uncertain quantities. For dynamic processes described by differential equations, one may distinguish two types of uncertainty: (1) parametric uncertainty, entering either the initial condition or the right-hand side function in the form of time-invariant parameters; and (2) time-varying uncertainty, describing either exogenous disturbances or endogenous disturbances such as structural plant-model mismatch. The main focus of this paper is on worst-case robust optimal control methods that are applicable to bounded time-varying uncertainty, although these methods can be readily extended to encompass time-invariant uncertainty as a special case.

Existing approaches to worst-case robust optimal control with bounded time-invariant or time-varying uncertainty may be divided into two broad classes, namely, discretization and set-propagation methods. The former includes direct simultaneous optimal control methods, which proceed by discretizing all the control and state trajectories along with the time-varying uncertainties, and often neglects the discretization errors. This way, the original worst-case robust optimal control problem is approximated by a finite-dimensional min-max optimization problem. The robust counterpart methodology by Ben-Tal et al.
addresses min-max problems under certain convexity assumptions, and in particular, it encompasses optimal control problems with linear dynamics and convex objective and constraint functions. Other approaches to solving min-max problems arising from linear control systems are based on convex analysis.\textsuperscript{13,14} For more general, nonconvex robust optimization problems, several tractable approaches involve approximating the nonlinear uncertainty dependencies and then bound the approximation error either rigorously or heuristically. One such heuristic approach based on a first-order Taylor expansion entails a linear approximation of the uncertainty dependency and uses convex analysis techniques to compute a near-robust solution.\textsuperscript{15-17} Because linearization may only be accurate enough for small uncertainty sets, where higher-order terms happen to be negligible, Nagy and Braatz\textsuperscript{16,18} introduced higher-order polynomial chaos expansions, yet without bounding the truncation error rigorously. In the special case that all the functions in the robust counterpart problem are polynomial, fully rigorous techniques based on hierarchies of linear matrix inequality (LMI) relaxations\textsuperscript{19,20} can be applied too. Other robust optimization approaches can be found in the semi-infinite programming (SIP) literature.\textsuperscript{21-23} Computationally tractable, yet approximate, approaches to SIP either discretize the uncertainty set to arrive at a finite-dimensional optimization problem,\textsuperscript{21,24,25} construct convex relaxations to determine a feasible SIP point,\textsuperscript{26,27} or use a tailored min-max Lagrangian relaxation.\textsuperscript{28} Finally, complete-search techniques have been developed based on the decomposition algorithm by Blankenship and Falk,\textsuperscript{29} which offer global optimality guarantees,\textsuperscript{30} but often come at the price of considerable computational effort.

Unlike the previous discretization methods, set-propagation methods compute the direct effect of the uncertainty on the continuous-time state trajectories of the dynamic system, thereby exploiting the particular structure of an optimal control problem. They find their origins in Aubin’s viability theory,\textsuperscript{31} which provides a sound and mature mathematical foundation for set-theoretic methods in control theory. Existing numerical approaches to characterizing the exact reachable set include level-set techniques,\textsuperscript{32} which compute the viscosity solution of the Hamilton-Jacobi-Isaacs equations.\textsuperscript{33} Other more tractable, yet inherently conservative, numerical approaches compute an outer approximation of the reachable set in the form of polytopes\textsuperscript{34} or ellipsoids.\textsuperscript{1,2,35} Similar set-propagation techniques can be found in the robust model predictive control literature.\textsuperscript{36} Rigorous methods for bounding the reachable set of ordinary differential equations can also be found in the validated integration literature.\textsuperscript{37-39} A current limitation of many such set-propagation techniques, however, is that the computed enclosures typically fail to be differentiable with respect to other participating variables, thus precluding the application of efficient, gradient-based optimization algorithms for robust optimization. Initial attempts to develop differentiable bounding techniques and apply them for the solution of robust optimal control problems with nonlinear dynamics can be found for instance in Houska et al.\textsuperscript{5,40}

This study presents an algorithm that combines, for the first time, high-order moment expansion and set-propagation techniques. Specifically, a high-order expansion with respect to the first few moments of the time-varying uncertainty is constructed, and the infinite-dimensional remainder term is bounded in a rigorous manner using ellipsoidal calculus. The idea is to retain the robustness certificates offered by set-propagation methods, while reducing the overall conservatism by applying a high-order moment expansion. The study starts in Section 2 with a discussion on how to formulate worst-case robust optimal control problems, along with a review of existing solution techniques based on moment expansion and set propagation in Sections 2.1 and 2.2, respectively. The main contribution of the study is presented in Sections 3 and 4, which introduce the new method for combining a high-order moment expansion of the time-varying uncertainty with set propagation to account for the moment-expansion truncation error. These theoretical results are supported by Appendix A summarizing a numerical solution approach to computing a rigorous and convergent enclosure of the truncation error. Finally, Section 5 illustrates the benefits of this new approach by solving a robust time-optimal control problem for a constrained pick-and-place maneuver of a robot arm with two motors.

## 2 Problem Formulation and Preliminaries

We consider problems in optimal control under uncertainty in the form

$$\inf_{(x,u) \in \mathcal{F}} m(x(T))$$

subject to

$$\forall t \in [0, T],$$

$$\dot{x}(t) = f(x(t), u(t), w(t)) = a(x(t), u(t)) + B(t)w(t)$$

$$x(0) = x_0$$

$$0 \succeq g(x(t), u(t)).$$

The optimization variables \((x, u) \in \mathcal{F} := \mathbb{W}_2^{\nu_x} \times \mathcal{U},\) with \(\mathcal{U} := \{u \in \mathbb{L}_2^{\nu_u} \mid \forall t \in [0, T]. u(t) \in U(t) \subseteq \mathbb{R}^{\nu_u}\},\) denote the differential state and control input of dynamic system, respectively. Here, \(\mathbb{L}_2^{\nu_u}\) denotes the set of \(L_2\)-integrable vector-valued functions of dimension \(\nu_u\) and \(\mathbb{W}_2^{\nu_x}\) the set of \(n_x\)-dimensional weakly differentiable functions whose weak derivative is
\(L_2\)-integrable. In a closed-loop control setup, \(u(t)\) could represent a parameterization of the control law at time \(t\). The problem (1) is uncertain in the sense that its solution depends on a time-varying \(w \in W := \{w \in L_2^n \mid \forall t \in [0, T], w(t) \in W(t) \subseteq R^{n_w}\}\). The right-hand side function \(f : R^n \times R^{n_x} \times R^{n_u} \rightarrow R^n\) presents a linear dependence in this time-varying uncertainty \(w\). It should be noted that the ensuing developments and results readily extend to the case of a state- and control-dependent matrix-valued function \(B(x(t), u(t))\), but we omit this dependence here to keep the notation as simple as possible. A generalization of the results to a nonlinear dependence of \(f\) in \(w\) is not possible nonetheless, which constitutes a limitation of the proposed approach; see, eg, Houska and Chachuat\(^{41}\) for a related discussion. Another straightforward extension concerns the presence of a time-invariant uncertain parameter, either in the initial condition \(x_0\) or in the right-hand side function \(f\) of the dynamic system. The cost function \(m : R^n \rightarrow R\) is in the Mayer form, and \(g : R^n \times R^{n_x} \rightarrow R^n\) denotes general path constraints. If necessary, integral terms in the cost function can be reformulated in the Mayer form by appending extra quadrature equations to the dynamic system. Explicit dependence in time of the cost, constraint, and right-hand side functions can also be handled by appending an auxiliary state describing time to the dynamic system. Moreover, the following developments trivially encompass problems with terminal or interior-point constraints, which we omit in (1) for the sake of brevity.

We make the following blanket assumptions throughout the paper:

**Assumption 1.** The sets \(W(t)\) are compact for each \(t \in [0, T]\).

**Assumption 2.** The functions \(f, g,\) and \(m\) are smooth with respect to all their arguments.

**Assumption 3.** The function \(f(\cdot, u, w)\) is globally Lipschitz-continuous, for all \((u, w) \in R^{n_u+n_w}\).

We introduce Assumption 2 mainly for ease of exposition, and it could be replaced by the weaker requirement that the functions \(f, g,\) and \(m\) are “sufficiently often” differentiable without loss of generality. Assumption 3 could also be replaced by a weaker local Lipschitz-continuity condition on a “sufficiently large” domain. Taken together, these two assumptions ensure the existence of a unique solution to the differential equations over \([0, T]\), for any given control \(u \in U\) and any given uncertainty \(w \in W\), which we denote by \(\xi[\cdot, u, w]\) throughout.

Next, the robust counterpart to problem (1) can be stated as

\[
\inf_{t \in U} \sup_{w \in W} m(\xi[T, u, w])
\]

s.t. \(\sup_{t \in [0, T]} g_i(\xi[t, u, w], u(t)) \leq 0, i = 1, \ldots, n_g\).

(2)

Such robust optimal control problems are hard to solve in general, because of their min-max structure and infinite-dimensional nature. The following subsections provide a high-level formalism for both discretization and set-propagation approaches to robust optimal control, based on which the proposed high-order moment expansion methodology will be developed in the rest of the paper.

### 2.1 Approximate robust optimal control via polynomial expansion

Polynomial expansion methods were first developed for finite-dimensional min-max problems of the form

\[
\min_u \max_{\sigma \in \Omega} F_0(u, \sigma) \quad \text{s.t.} \quad \max_{\sigma \in \Omega} F_i(u, \sigma) \leq 0, \quad i = 1, \ldots, n_F,
\]

(3)

where the \(F_i\)'s are sufficiently smooth functions, and \(\sigma\) is a finite-dimensional vector in a given compact set \(\Omega \subseteq R^{n_\sigma}\). The robust optimal control counterpart problem (2) can be approximated by such a min-max problem upon (1) discretizing the time-varying disturbance as \(w(t) := \sum_{j=0}^{N} \sigma_j \Phi_j(t)\) for some \(N \geq 0\) and a given set of (orthogonal) basis functions \(\Phi_0, \ldots, \Phi_N \in \mathbb{L}_2;\) and (2) enforcing the path constraints on a finite grid \(0 = t_0 < t_1 < \ldots < t_K = T\) for some \(K > 0\). That is, the uncertain parameters are given by

\[
\sigma := (\sigma_0, \ldots, \sigma_N)
\]
the uncertainty set by
\[
\Omega := \left\{ (\sigma_0, \ldots, \sigma_N) \in \mathbb{R}^{(N+1)n_u} \left| \sum_{i=0}^{N} \sigma_i \Phi_i \in \mathcal{W} \right. \right\},
\]
and the cost and constraint functions by
\[
F_0(u, \sigma) := m \left( \xi \left[ T, u, \sum_{i=0}^{N} \sigma_i \Phi_i \right] \right),
\]
\[
F_i(u, \sigma) := g_i \left( \xi \left[ t_k, u, \sum_{i=0}^{N} \sigma_i \Phi_i \right], u(t_k) \right),
\]
where the index \( i := kn_g + j \) runs over both the constraints \( j \in \{1, \ldots, n_g\} \) and the time-grid \( k \in \{0, \ldots, K\} \).

The application of polynomial expansion to compute a suboptimal, yet feasible, solution to the min-max problem (3) can be summarized with the following four steps:

1. Consider a polynomial expansion of given order \( q \geq 0 \) and the corresponding remainder function for each objective and constraint function \( F_i \) with respect to the finite-dimensional uncertainty \( \sigma \),
\[
F_i(u, \sigma) = \sum_{|k| \leq q} a_{i,k}(u) \sigma^k + r_i(u, \sigma).
\]
2. Compute suitable range bounders \( B_i(u) \) overestimating the polynomial part,
\[
B_i(u) \geq \max_{\sigma \in \Omega} \sum_{|k| \leq q} a_{i,k}(u) \sigma^k.
\]
3. Determine upper bounds \( R_i(u) \) on the remainder part,
\[
R_i(u) \geq \max_{\sigma \in \Omega} r_i(u, \sigma).
\]
4. Compute a solution to the conservative, single-level optimization problem
\[
\min_u T_0(u) + R_0(u) \quad \text{s.t.} \quad T_i(u) + R_i(u) \leq 0, \quad i = 1 \ldots n_F. \tag{4}
\]

In the field of robust optimal control, the emphasis so far has been on finite-dimensional uncertainty, and existing approaches are based on Taylor expansion, yet without computing rigorous bounds on the remainder term.\textsuperscript{15,18} Obtaining the Taylor coefficients \( a_{i,k}(u) \) in a multivariate Taylor expansion calls for a differentiation of the state trajectory \( \xi \) with respect to the uncertain parameter \( \sigma \). Adjoint (backward) differentiation is advantageous over forward differentiation whenever the uncertain discretization is larger than the number of discretized constraints, \( n_\sigma \gg K n_g \).\textsuperscript{42} An important limitation, however, is that both the computation time and the storage requirement may become very large when a fine discretization is used in combination with a high-order expansion order \( q \). Another key limitation is that neither the accuracy nor the robustness of the discretized solution can be guaranteed for a time-varying uncertainty.

Alternative approaches to Taylor expansion include the use of interpolating polynomials such as Legendre or Chebyshev interpolation. Well-developed arithmetics exist for Taylor and Chebyshev expansion of multivariate functions,\textsuperscript{43-45} which allow a rigorous enclosure of the truncation terms. Moreover, conservative bounds \( T_i(u) \) on the resulting multivariate polynomials can be computed by using LMI techniques\textsuperscript{19} or more conservative approaches based on Bernstein polynomials.\textsuperscript{46} Finally, if the resulting single-level optimization problem (4) is solved using a gradient-based algorithm, special care must be taken to ensure that the bounds \( B_i \) and \( R_i \) are differentiable with respect to the decision variable \( u \), as many existing bounding techniques based on interval analysis provide Lipschitz-continuous bounds only.

*We use multi-index notation for conciseness.
2.2 Rigorous robust optimal control using set propagation

These methods account for the direct effect of the uncertainty on the state trajectories, by considering the set of reachable states $X(t, u) \subseteq \mathbb{R}^{n_x}$ at any time $t \in [0, T]$ and for any given control $u \in \mathbb{U}$, given by

$$X(t, u) := \{ \xi[t, u, w] \in \mathbb{R}^{n_x} | w \in \mathcal{W} \}.$$ 

Next, we define the robust counterpart functional $M: \Pi(\mathbb{R}^{n_x}) \rightarrow \mathbb{R}$ to the Mayer cost function $m$,

$$M(X) := \max_{\xi \in X} m(\xi),$$

and likewise, the robust counterpart functionals $G_i: \Pi(\mathbb{R}^{n_x}) \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ to each constraint function $g_i$, $i = 1 \ldots n_g$,

$$G_i(X, u) := \max_{\xi \in X} g_i(\xi, u).$$

**Proposition 1.** Let Assumptions 1–3 be satisfied. Then, problem (1) is equivalent to the following set-based optimization problem,

$$\min_{u \in \mathcal{U}} M(X(T, u))$$

s.t. $G_i(X(t, u), u(t)) \leq 0, \ \forall t \in [0, T], \ \ i = 1 \ldots n_g,$

(5)

in the sense that any optimal solution function $u^*$ of problem (5) corresponds to a robustly optimal control input solving problem (2).

**Proof.** See Houska et al.\textsuperscript{40} \hfill \square

A practical implementation for nonlinear dynamic systems would typically rely on the computation of parameterized enclosures $\overline{X}(t, u) \supseteq X(t, u)$, for all $t \in [0, T]$ and all admissible controls $u \in \mathcal{U}$. For instance, such enclosures can be propagated using the theory of differential inequalities or ellipsoidal calculus.\textsuperscript{39,40,47,48} It follows that the following optimal control problem,

$$\min_{u} M(\overline{X}(T, u))$$

s.t. $G_i(\overline{X}(t, u), u(t)) \leq 0, \ \forall t \in [0, T], \ \ i = 1 \ldots n_g,$

(6)

is conservative in the sense that all of its minimizers are also feasible solutions of problem (5).

The main computational cost for solving problem (6) is often dominated by the cost of propagating the reachability tubes $X(\cdot, u)$, especially for a high state-space dimension $n_x$. One would nonetheless expect this approach to be competitive with the solution of the discretized min-max problem (4) as the uncertainty discretization and/or the polynomial expansion are refined, driven by the need to compute $O(\min \{ n_x, n_w \})$ directional derivatives thereof. Computational burden aside, set-propagation methods come with a certificate of robust feasibility, even in the presence of time-varying uncertainty, and are therefore more rigorous than discretization methods for such problems. An important limitation though is that the reachable-set enclosures may be subject to large overestimation because of the so-called wrapping effect in set-valued integrators. Despite recent progress on bound stabilization techniques,\textsuperscript{37} the computed set enclosures are for many practical examples limited by finite escape-time phenomena, where the computed bounds $\overline{X}(\cdot, u)$ are divergent on finite horizons, although the actual reachable tube $X(\cdot, u)$ remains bounded on the whole horizon. Another difficulty is that the bounds computed with state-of-the-art set-valued integrators often fail to be differentiable with respect to other participating variables, which can impair the convergence of gradient-based algorithms as applied to problem (6).

The conclusion from this cursory analysis is that integrating high-order moment expansion within set-propagation methods could provide an improved approach to robust optimal control by reducing conservatism, yet without sacrificing robustness.
3 | GENERALIZED MOMENT EXPANSION WITH RIGOROUS ERROR ESTIMATE

Similar to the discretization method in Section 2.1, we consider a finite parameterization of the time-varying uncertainty in the form $w(t) := \sum_{i=0}^{N} \sigma_{i} \Phi_{i}(t)$, for a given expansion order $N \geq 0$. Our focus in the rest of the paper is on orthogonal basis functions $\{\Phi_{i}\}_{i\in\mathbb{N}}$, such as the Legendre or Chebyshev polynomials, which satisfy

$$\frac{1}{\sigma_{i}} \int_{0}^{T} \Phi_{i}(\tau)\Phi_{j}(\tau)\mu(\tau) d\tau = \delta_{ij} := \begin{cases} 0 \text{ if } i \neq j, \\ 1 \text{ otherwise,} \end{cases}$$

for all $i, j \in \{0, \ldots, N\}$, with $\mu : [0, T] \to \mathbb{R}_{++}$ a bounded weighting function and $\{\sigma_{i}\}_{i\in\mathbb{N}}$ a sequence of strictly positive scalars. To retain the infinite-dimensional nature of the uncertainty $w$, we also introduce the set-valued function

$$\mathcal{W}_{N}(\sigma) := \left\{ w \in \mathcal{W} \left| \int_{0}^{T} w(t)\Phi_{i}(t)\mu(t) dt = \sigma_{i}, i = 0 \ldots N \right\} \right.,$$

mapping all the admissible time-varying uncertainty trajectories whose projections onto the basis subset $\{\Phi_{i}\}_{i\in\mathbb{N}}$ match the parameterization $\sigma$, and we denote its domain by $\Omega_{N} := \{ \sigma \mid \mathcal{W}_{N}(\sigma) \neq \emptyset \}$. The corresponding set of all state-response defects for a given control $u \in \mathcal{U}$ along the time horizon $[0, T]$, called discretization error set hereafter, is given by

$$E_{N}(t, u, \sigma) := \{ \zeta_{N}[t, u, \sigma] - \xi[t, u, w] \mid w \in \mathcal{W}_{N}(\sigma) \}, \quad (7)$$

where $\zeta_{N}[t, u, \sigma] := \xi[t, u, \sum_{i=0}^{N} \sigma_{i} \Phi_{i}(\cdot)]$ denotes the response of the dynamic system for a time-varying disturbance parameterized finitely by $\sigma$. In particular, $E_{N}(t, u, \sigma)$ is nonempty for each $\sigma \in \Omega_{N}$.

The expectation behind introducing such discretization error sets is that the diameter of $E_{N}(t, u, \sigma)$ will converge to zero upon increasing the moment expansion order $N$. Although this conjecture turns out to be wrong for nonlinear dynamic systems in general, the following theorem establishes that convergence is indeed possible for uncertainty-affine dynamic systems under mild assumptions. This is the consequence of an important result from functional analysis in infinite-dimensional Hilbert spaces that has been derived within a global optimization context by the authors.\cite{Houska2014} Here, we present this result in a such way that it becomes accessible for use in a robust optimal control context.

**Theorem 1.** Let Assumptions 1–3 be satisfied. Then, for every given admissible control $u \in \mathcal{U}$, there exists a constant $C < \infty$ such that

$$\max_{t \in [0, T]} \text{diam} \left( E_{N}(t, u, \sigma) \right) \leq \frac{C}{\sqrt{N}} \quad (8)$$

for all $\sigma \in \Omega_{N}$ and all $N \in \mathbb{N}$. In particular, the diameter of the discretization error set $E_{N}(t, u, \sigma)$ converges to 0 for $N \to \infty$ since $C$ is independent of $N$.

**Proof.** The main idea of the proof is to introduce the auxiliary functional $\mathcal{F}(c, w) : \mathbb{R}^{n} \times \mathcal{W} \to \mathbb{R}$ defined by $\mathcal{F}(c, w) := c^{T}\xi[t, u, w]$, for a given admissible control $u \in \mathcal{U}$. An immediate consequence of Theorem 1 and Example 5 in Houska and Chachuat\cite{Houska2014} is that, under Assumptions 1–3, $\mathcal{F}(c, \cdot)$ satisfies an inequality of the form

$$\forall (c, w) \in \mathbb{R}^{n} \times \mathcal{W}_{N}(\sigma), \quad \| \mathcal{F}(c, w) - \mathcal{F}(c, \sum_{i=0}^{N} \sigma_{i} \Phi_{i}) \| \leq \frac{L}{\sqrt{N}} c^{T} c,$$

for a sufficiently large (strong Lipschitz) constant $L < \infty$. It follows that the support function $h_{N} : \mathbb{R}^{n} \to \mathbb{R}$ of the discretization error set $E_{N}(t, u, \sigma)$ satisfies

$$h_{N}(c) := \max_{w \in E_{N}(t, u, \sigma)} \| c^{T} w \| = \max_{w \in \mathcal{W}_{N}(\sigma)} \mathcal{F}(c, w) - \mathcal{F}(c, \sum_{i=0}^{N} \sigma_{i} \Phi_{i}) \leq \frac{L\|c\|^{2}}{\sqrt{N}}.$$
Moreover, we have
\[
\text{diam} \left( E_N(t, u, \sigma) \right) \leq 2 \max_{c, e^T c \leq 1} h_N(c) \leq \frac{2L}{\sqrt{N}},
\]
and therefore, the result in (8) holds with \( C = 2L \).

It is important to keep in mind that Theorem 1 yields a conservative bound on the convergence rate of the discretization error sets \( E_N(\cdot, u, \sigma) \), and one may observe faster convergence on particular instances. Under slightly stronger regularity assumptions on the uncertainty function \( \sigma \), one can even establish exponential convergence rates for \( E_N(\cdot, u, \sigma) \), as discussed in Houska and Chachuat.\(^{49}\) In practice, one may also wish to determine tighter bounds on the set \( E_N(\cdot, u, \sigma) \) than the one given in Theorem 1. This involves computing a tube enclosure \( \overline{E}_N(\cdot, u, \sigma) \supseteq E_N(\cdot, u, \sigma) \), for instance with pointwise-in-time interval or ellipsoidal cross-sections, by exploiting state-of-the-art set-valued integration capabilities. Of course, the minimum requirement on the accuracy of the tubes \( \overline{E}_N(\cdot, u, \sigma) \) is that they should inherit the convergence properties of \( \overline{E}_N(\cdot, u, \sigma) \) for every \((u, \sigma) \in U \times W_N \) as the expansion order \( N \) is increased:

**Assumption 4.** A computationally tractable outer-approximation procedure is available that determines enclosures \( \overline{E}_N(t, u, \sigma) \supseteq E_N(t, u, \sigma) \) such that, for every admissible control \( u \in U \) and every admissible uncertainty parameterization \( \sigma \in \Omega_N \), we have

\[
\exists \mathcal{C} < \infty : \forall N \in \mathbb{N}, \max_{t \in [0, T]} \text{diam} \left( \overline{E}_N(t, u, \sigma) \right) \leq \frac{\mathcal{C}}{\sqrt{N}}. \tag{9}
\]

A tailored bounding approach satisfying to the previous assumption is outlined in Appendix A. This approach is inspired from recent developments in the area of branch-and-lift for global optimal control.\(^{41}\) Notice that the discretization error sets \( E_N(\cdot, u, \sigma) \), and the tube enclosures \( \overline{E}_N(\cdot, u, \sigma) \) likewise, are now dependent on the time-invariant uncertain parameter \( \sigma \), and its effect must therefore be bounded to ensure robust feasibility. A novel methodology for rigorous robust optimal control based on these ideas is detailed in the following section.

4 | **ROBUST OPTIMAL CONTROL VIA GENERALIZED MOMENT EXPANSION**

We describe a solution methodology for the robust optimal control problem (2) to local optimality, which builds upon the generalized moment expansion approach in Section 3 and provides both feasibility and convergence guarantees.

Consider the auxiliary robust optimization problem

\[
\inf_{u} \sup_{\sigma \in \Omega_u} F^N_0(u, \sigma) \quad \text{s.t.} \quad \sup_{t \in [0, T]} \max_{\sigma \in \Omega} F^N_i(t, u, \sigma) \leq 0, \quad i = 1 \ldots n_y, \tag{10}
\]

with the cost and constraint functions \( F_i \) given by

\[
F^N_0(u, \sigma) := \max_{e \in E_N(T, u, \sigma)} m(\zeta_N[T, u, \sigma] + e)
\]
\[
F^N_i(t, u, \sigma) := \max_{e \in E_N(t, u, \sigma)} g_i(\zeta_N[t, u, \sigma] + e, u(t)).
\]

By definition of the error discretization set \( E_N(T, u, \sigma) \), we have

\[
\max_{e \in E_N(T, u, \sigma)} m(\zeta_N[T, u, \sigma] + e) = \max_{w \in \mathcal{V}_N} m(\zeta[T, u, w]) \quad \text{and}
\]
\[
\max_{e \in E_N(t, u, \sigma)} g_i(\zeta_N[t, u, \sigma] + e, u(t)) = \max_{w \in \mathcal{V}_N} g_i(\zeta[t, u, w], u(t)),
\]

for all \( t \in [0, T] \) and each \( i = 1 \ldots n_y \). Then, since the \( F_i \)'s account for the worst-case discretization errors, thus providing rigorous upper-bounds on the original objective and constraint functions, it follows that the robust optimization problems (2) and (10) are equivalent.
Now, assuming that Assumption 4 holds, we may use the enclosures $\overline{E}_N(t, u, \varpi)$ to construct conservative, yet tractable, bounds $\overline{F}_i^N$ such that
\[
\begin{align*}
F_0^N(u, \varpi) &\leq \max_{e \in \overline{E}_N(t, u, \varpi)} m(\zeta_N[T, u, \varpi] + e) \\
&\leq \overline{F}_0^N(u, \varpi), \\
\forall t \in [0, T], \quad F_i^N(t, u, \varpi) &\leq \max_{e \in \overline{E}_N(t, u, \varpi)} g_i(t, \zeta_N[t, u, \varpi] + e) \\
&\leq \overline{F}_i^N(t, u, \varpi).
\end{align*}
\] (11)

In turn, a conservative approximation of problem (2) or (10) is obtained as
\[
\inf_{u} \sup_{\varpi \in \Omega_N} \overline{F}_0^N(u, \varpi) \quad \text{s.t.} \quad \sup_{t \in [0, T] \atop m \in \Omega_N} \overline{F}_i^N(t, u, \varpi) \leq 0,
\] (11)

whereby the lower-level optimization problem is now finite-dimensional in the uncertainty. Practically, provided that the functions $f$, $g$, and $m$ are all factorable, one can apply existing set-propagation techniques based on a variety of computer algebra systems, including interval, ellipsoidal, Taylor model and Chebyshev model arithmetics, to construct the desired upper bounding functions $\overline{F}_i^N$. Apart from being computationally tractable, bounding functions computed with the said techniques lead to enclosures that satisfy the following linear Hausdorff convergence conditions under Assumption 2,
\[
\begin{align*}
\overline{F}_0^N(u, \varpi) - \max_{e \in \overline{E}_N(t, u, \varpi)} m(\zeta_N[T, u, \varpi] + e) \\
&= \mathcal{O} \left( \text{diam} \left( \overline{E}_N(T, u, \varpi) \right) \right), \\
\forall t \in [0, T], \quad \overline{F}_i^N(t, u, \varpi) - \max_{e \in \overline{E}_N(t, u, \varpi)} g_i(\zeta_N[t, u, \varpi] + e) \\
&= \mathcal{O} \left( \text{diam} \left( \overline{E}_N(t, u, \varpi) \right) \right).
\end{align*}
\] (12) (13)

We introduce the following regularity condition to analyze the gap between the original and relaxed robust optimization problems (2) and (11), respectively.

**Definition 1.** The robust optimal control problem (2) is said to admit a regular solution if (1) it is feasible and (2) there exists a constant $\Lambda < \infty$ such that the optimal value function
\[
\mathcal{V}(\epsilon) \eqdef \inf_{u \in \mathcal{U}} \sup_{w \in \mathcal{W}} m(\zeta[T, u, w]) \\
\text{s.t.} \sup_{w \in \mathcal{W}} g(t, \zeta[t, u, w], u(t)) \leq -\epsilon
\] (14)

satisfies the condition $\mathcal{V}(\epsilon) \leq \mathcal{V}(0) + \Lambda \epsilon$ for all sufficiently small $\epsilon > 0$.

**Remark 1.** We note that if an optimal point with a corresponding set of finite constraint multipliers exists for the upper-level optimization problem in Equation (14), then this point is regular in the sense of Definition 1 since $\mathcal{V}(\epsilon) \leq \mathcal{V}(0) + \Lambda \epsilon$ is satisfied by setting $\Lambda$ to be any upper bound on the largest multiplier. For instance, if the linear independence constraint qualification holds at an optimum of the upper-level minimization problem, then there exist a unique set of multipliers, and the above regularity condition is satisfied.

**Theorem 2.** Let Assumptions 1–4 be satisfied, and let the upper-bounding functions $\overline{F}_i^N$, $i = 0 \ldots N$, satisfy the conditions (12)–(13) for all sufficiently large $N \geq 0$. If problem (2) admits a regular solution in the sense of Definition 1, then every $\epsilon$-suboptimal solution of problem (11) is also an $\epsilon + \mathcal{O}(1/\sqrt{N})$-suboptimal solution of problem (2). Moreover, such solutions exist for sufficiently small $\epsilon > 0$ and sufficiently large $N$.

**Proof.** Substituting the convergence rate bounds (9) into (12)–(13) gives the error bound estimates
\[
\begin{align*}
\left| F_0^N(u, \varpi) - \overline{F}_0^N(u, \varpi) \right| &= \mathcal{O}(1/\sqrt{N}) \quad \text{and} \\
\left| F_i^N(t, u, \varpi) - \overline{F}_i^N(t, u, \varpi) \right| &= \mathcal{O}(1/\sqrt{N}).
\end{align*}
\]
Then, it follows from the regularity condition in Definition 1 and Remark 1 that the overestimation error in any constraint leads to an increase of the objective value that is no larger than $O(1/\sqrt{N})$, which corresponds to the statement of the theorem.

Notice that problem (11) can be solved accurately by using the polynomial-expansion approach introduced earlier in Section 2.1, with the only difference that the functions $F_i$ thereof are replaced by the upper-bounding functions $F^N_i$ for some $N \geq 0$. In this context, it is important to keep in mind that, to retain differentiability with respect to $\sigma$ and (any parameterization of) $u$, one must use a set-bounding method that guarantees a differentiable representation of the enclosure functions $E^N_{\sigma}(\cdot, \sigma, u)$. Fortunately, it is neither a theoretical nor a practical problem to guarantee differentiability of these enclosures. From a theoretical viewpoint, Theorem 1 asserts the existence of a bound on $E_N$ that is independent of both $\sigma$ and $u$. Consequently, there also exist bounding sets $\overline{E}_N$ that are independent of, and thus are differentiable in, $\sigma$ and $u$. From a more practical viewpoint, one may compute $\overline{E}_N$ using Taylor or Chebyshev model arithmetic, by making sure that any nondifferentiability is independent of $\sigma$ and $u$. For instance, in a first-order Taylor or Chebyshev model with interval remainder $^4,w^5$ both the reference point and the first-order terms of any smooth functions could be dependent on $\sigma$ or $u$. But the remainder term would need bounding with respect to $\sigma$ and $u$ to prevent possible nondifferentiability, as interval bounds are typically Lipschitz-continuous, but may be nondifferentiable with respect to their parametric dependencies. $^5,w^6$ Another possible implementation entails the development of an ellipsoidal set-propagation strategy in combination with differentiable remainder estimates, as pursued in Houska et al. $^4,w^5$

A clear advantage of the approximate robust optimization problem (11) is that it provides a means of balancing the computational complexity associated to the high-order expansion with the conservatism introduced by the reformulation. For $N = 0$ in particular, we have $E_0(t, u, \sigma) = X(t, u) \cup \{ \xi_0[t, u, \sigma] \}$, and the proposed approach is thus identical to the original set-propagation approach described in Section 2.2. A serious limitation previously mentioned in Section 2.2 is the wrapping effect, which can lead to large overestimation or even cause a finite escape-time for the reachability tube enclosures. In contrast, Theorem 1 asserts that the cross-section of the discretization error tube $E_N(\cdot, u, \sigma)$ vanishes for a sufficiently large moment expansion order $N$, at a (worst case) rate of order $O(1/\sqrt{N})$. Naturally, choosing a higher moment-expansion order $N$ may significantly increase the computational burden. More precisely, applying the ellipsoidal-based approach in Appendix A alongside set-propagation techniques based on $q$-th order polynomial models for computing upper bounds on the suprema in problem (11) has an overall run-time complexity of $O((N + 1)^q + 1n^2_D)$. Nonetheless, it is acceptable to select a lower expansion in the proposed approach compared with the truncated expansion method in Section 2.1, since robust feasibility is indeed guaranteed by accounting for the effect of the parameterization error. The practical implications of this trade-off are further analyzed in the following numerical case study.

## 5 NUMERICAL CASE STUDY

We consider a motorized robot arm, whose equations of motion come in the form $\dot{x}(t) = a(x(t), u(t)) + B(t)w(t)$, with

$$a(x, u) := \begin{bmatrix} \omega \\ \psi \\ \omega \psi \\ \frac{\omega}{m_{eff}} - 2 \cot(\theta) \gamma y \psi \\ \frac{\omega}{m_{eff}} + \sin(\theta) g + \sin(\theta) \cos(\theta) \omega^2 \\ \frac{\omega}{m_{eff}} + \sin(\theta) g + \sin(\theta) \cos(\theta) \omega^2 \\ B(t) := \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. $$

This model is written in spherical coordinates, with the state vector $x$ given by

$$x(t) = (\phi(t), \theta(t), \omega(t), \psi(t))^\top \in \mathbb{R}^4,$$

where $\phi(t)$ and $\theta(t)$ denote the arm angles, and $\omega(t)$ and $\psi(t)$ are the corresponding angular velocities, at a given time $t$. The arm is equipped with $t$ controllable motors, which act on both angles $\phi$ and $\theta$ and are denoted accordingly by

$$u(t) = (u_\phi(t), u_\theta(t))^\top \in \mathbb{R}^2.$$

Also, acting on the angle $\theta$ is a torque disturbance, in the form of a bounded time-varying uncertainty, denoted by $w(t) \in [w_l, w_u] \subseteq \mathbb{R}$. The dynamic model is highly nonlinear in the state $x$ because of the action of centrifugal, gravitational, and Coriolis forces, but it is nonetheless affine in the uncertainty $w$ and in both controls $u_\phi$ and $u_\theta$. Both the robot arm length $l$ and effective mass $m_{eff}$ are given constants.
TABLE 1 Physical parameters, initial values, path and terminal constraint bounds, and uncertainty bounds

| Description                                      | Symbol | Value          |
|--------------------------------------------------|--------|----------------|
| gravitational constant                           | \( g \) | \( 9.81 \, \text{ms}^{-2} \) |
| effective mass                                    | \( m_{\text{eff}} \) | \( 1 \, \text{kg} \) |
| length of the robot arm                           | \( l \)  | \( 1 \, \text{m} \) |
| initial angle \( \theta \)                       | \( \theta_0 \) | \( 0 \, \text{rad} \) |
| initial angle \( \phi \)                        | \( \phi_0 \) | \( \pi/2 \, \text{rad} \) |
| initial angular velocity in \( \phi \)-direction | \( \omega_0 \) | \( 0 \, \text{rad s}^{-1} \) |
| initial angular velocity in \( \theta \)-direction | \( \psi_0 \) | \( 0 \, \text{rad s}^{-1} \) |
| bounds on torque in \( \phi \)-direction         | \([u_{\phi_l}, u_{\phi_u}]\) | \([-10, 10]\, \text{Nm} \) |
| bounds on torque in \( \theta \)-direction       | \([u_{\theta_l}, u_{\theta_u}]\) | \([-20, 20]\, \text{Nm} \) |
| bounds on angle \( \theta \)                     | \([\theta_l, \theta_u]\) | \([1.4, 3.0]\, \text{rad} \) |
| bounds on angular velocity in \( \phi \)-direction| \([\omega_l, \omega_u]\) | \([0, 6]\, \text{rad s}^{-1} \) |
| bounds on angular velocity in \( \theta \)-direction| \([\psi_l, \psi_u]\) | \([{-5.5}, {5.5}]\, \text{rad s}^{-1} \) |
| terminal range of angle \( \phi \)               | \([\phi_{T_l}, \phi_{T_u}]\) | \([\pi/2, \pi]\, \text{rad} \) |
| terminal range of angle \( \theta \)             | \([\theta_{T_l}, \theta_{T_u}]\) | \([1.4, \pi/2]\, \text{rad} \) |
| bounds on uncertain torque \( w \)               | \([w_l, w_u]\) | \([-0.5, 0.5]\, \text{Nm} \) |

The goal of the optimization is to bring the arm position \((\phi(t), \theta(t))\) from an initial resting position at \((\phi_0, \theta_0)\) to the final range \([\phi_{T_l}, \phi_{T_u}] \times [\theta_{T_l}, \theta_{T_u}]\), in a minimum time \( T \). Moreover, position, velocity, and torque constraints given by \( \phi(t) \in [\phi_l, \phi_u], \theta(t) \in [\theta_l, \theta_u], \omega(t) \in [\omega_l, \omega_u], \psi(t) \in [\psi_l, \psi_u], u_{\phi}(t) \in [u_{\phi_l}, u_{\phi_u}], \) and \( u_{\theta}(t) \in [u_{\theta_l}, u_{\theta_u}] \) are imposed at each time along the path. As already mentioned in Section 2, the terminal constraints can be treated in the exact same way as the path constraints in the proposed algorithm, and standard manipulations can be applied to reformulate a minimum-time problem in the form of problem (1). The numerical values for all the physical parameters, initial conditions, and constraints limits are all collected in Table 1.

For numerical optimization, we follow the procedure outlined in Section 2.1 to outer-approximate the conservative inf-sup problem (11) by a more tractable single-level optimization problem. The latter is solved using a multiple-shooting approach based on exact Hessian sequential quadratic programming (SQP), where a piecewise constant discretization of the controls \( u_{\phi} \)
and $u_0$ over 20 identical time subintervals is considered. Moreover, we use Taylor models to construct the polynomial expansions and keep track of the truncation errors for the cost and constraint functions thereof. Since the dynamic system is affine in the time-varying uncertainty $w$, the generalized high-order moment expansion approach introduced in Section 3 provides a means of computing convergent enclosures of the discretization error tube $\bar{E}_N$ as the number of moments is increased. In particular, we implement the procedure outlined in Appendix A, where all the remainder bounds can be precomputed since the domain of the optimal control problem is compact. This ensures that the parameterization of the set-valued function $\bar{E}_N$ is differentiable, and therefore, the standard algorithmic differentiation tools from ACADO toolkit can be used to compute the gradients that are needed for solving the single-level robust optimization problem.

Numerical results for the conservative reformulation of the robust time-optimal maneuver control are shown in Figure 1. The gray-shaded regions in the upper-left and upper-center plots correspond to an outer-approximation of the reachable set, in projection onto the $\phi$ and $\theta$ directions, for a moment expansion order of $N = 15$. Note that each SQP iteration in ACADO takes about 2 hours with $N = 15$, compared to less than a second for a more conservative approximation with $N = 2$ or $N = 3$, as shown in the upper-right plot. Moreover, depending on the initial guess, the total number of SQP iterations varies between 10 and 200 to satisfy a KKT-tolerance of less than $10^{-10}$. For $N = 1$, however, the reformulation is too conservative and turns out to be infeasible. This shows that it is not possible to solve this robust maneuver control problem rigorously with the current set-propagation techniques as described in Section 2.2, hence justifying the need for more accurate, higher-order expansion. The lower-right plot, reporting the minimum maneuver time corresponding to moment expansion orders $N = 2$ ... 15, quantifies the extra conservatism introduced by a low expansion order and illustrates the convergence of the approach as the expansion order is increased. For this case study, one could argue that an expansion order of $N = 5$ provides a good compromises between the accuracy and the computational burden. Upper and lower position bounds for $N = 5$ and $N = 2$ are depicted with red dashed and blue dotted lines, respectively, in the lower-left and lower-center plots, where they can also be compared with those for $N = 15$ depicted with solid lines (same bounds as in the upper-left and upper-center plots). Notice that neither of the solution tubes for $N = 5$ and $N = 2$ contain the solution tube for $N = 15$ as a subset, since more conservative reformulations lead to significantly different control maneuvers in this problem. However, the duration of the corresponding maneuvers are monotonically decreasing with $N$, as seen on the lower-right plot.

6 | CONCLUSIONS

After reviewing existing, both approximate and rigorous, approaches to robust optimal control with time-varying uncertainty, this paper has introduced a new method for worst-case robust optimal control of nonlinear dynamic systems with affine uncertainty dependence. A high-order expansion of the states with respect to the moments of the time-varying uncertainty is constructed, and the infinite-dimensional discretization-error term is bounded in a rigorous manner using ellipsoidal calculus. In particular, it is shown that the discretization-error term convergences to zero as the moment expansion order is increased for uncertainty-affine systems (Theorem 1), and a practical approach to constructing these convergent bounds is outlined in Appendix A. It is also shown how to use these discretization-error bounds for constructing conservative, yet computationally tractable and convergent, approximations of the original robust optimal control problem (Theorem 2). This new robust optimization approach has been illustrated with the case study of a constrained pick-and-place maneuver of a robot arm with 4 states, 2 control inputs, and 1 time-varying uncertainty. An important aspect here is the selection of the moment expansion order to compromise between the accuracy and computational burden of the algorithm. As part of future work, special emphasis will be devoted to developing a rigorous and fully automated implementation of the proposed algorithm, that can exploit sparsity and structure of the problem to address larger-scale problems.

ACKNOWLEDGMENTS

The work of Boris Houska and Jiaqi C. Li was supported by the National Science Foundation China (NSFC) under Grant 61473185 and ShanghaiTech University under Grant F-0203-14-012. Benoit Chachuat gratefully acknowledges financial support by the Engineering and Physical Research Council (EPSRC) under Grant EP/J006572/1.

Data Statement: No new data was collected in the course of this research.

REFERENCES

1. Glover JD, Schewepp FC. Control of linear dynamic systems with set constrained disturbances. IEEE Trans Autom Control. 1971;16:411–423.
2. Schewepp FC. Uncertain Dynamic Systems. Englewood Cliffs, NJ: Prentice Hall; 1973.
3. Zames G. Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses. IEEE Trans Autom Control. 1981;26:301–320.
Pivotal to the practical implementation of generalized high-order moment expansion technique in Section 3 is the ability to compute the set-valued enclosures

\[
\overline{E}_N(\cdot, u, \sigma) \supseteq E_N(\cdot, u, \sigma)
\]

such that Assumption 4 is satisfied. The approach outlined in this appendix is inspired from the branch-and-lift developments by Houska and Chachuat\textsuperscript{41} for computing globally optimal controls in input-affine systems. Starting from the definition of the discretization error set \(E_N\) in (7), we consider the response defect \(\hat{\delta}(t) \equiv \zeta_N(t, u, \sigma) - \zeta(t, u, w)\), for a given control \(u\), uncertainty parameterization \(\sigma\), and a compatible time-varying uncertainty function \(w \in \mathcal{W}_N(\sigma)\). This defect is described by the differential equation

\[
\dot{\delta}(t) = f \left( \zeta_N[t, u, \sigma], u(t), \sum_{i=0}^{N} \sigma \Phi_i(t) \right) - f(\zeta[t, u, w], u(t), w(t))
\]

\[
= [a(\zeta_N[t, u, \sigma], u(t)) - a(\zeta[t, u, w], u(t))] - B(t) \left[ w(t) - \sum_{i=0}^{N} \sigma \Phi_i(t) \right]
\]

\[
= [A(t, u, \sigma) + \Delta(t, u, \sigma)] \delta(t) - B(t) \left[ w(t) - \sum_{i=0}^{N} \sigma \Phi_i(t) \right],
\]

(\text{A1})

with

\[
A[t, u, \sigma] := \frac{\partial a}{\partial x} (\zeta_N[t, u, \sigma], u(t))
\]

and where \(\Delta(t, u, \sigma) \in \mathbb{R}^{n, m_t}\) is a suitable matrix-valued function collecting the higher-order terms. In particular, a bound \(\overline{\Delta}(t)\) on the matrix \(\Delta(t, u, \sigma)\) can be obtained by invoking Taylor’s theorem,

\[
\|\Delta(t, u, \sigma)\| \leq \max_{d', d'' \in \mathbb{R}^{m_t}} \left\| \frac{\partial^2 a}{\partial x^2}(\zeta_N[t, u, \sigma] + d') \cdot d'' \right\| \leq \overline{\Delta}(t),
\]

where \(\overline{\Delta}(t)\) can be any a priori enclosure of the set \(E_N[t, u, \sigma]\), as computed for instance with standard set-valued integration tools.\textsuperscript{39,50} In the special case of a dynamic systems where the function \(a\) affine in \(u\)—as for instance in the numerical example in Section 5—the second derivative \(\frac{\partial^2 a}{\partial x^2} \) is independent of \(u\), which avoids introducing unnecessary conservatism in bounding the effect of all admissible controls in \(U\).

It should be clear at this point that the response defect can be split as \(\delta(t) = \delta_1(t) + \delta_2(t)\), where the linear component \(\delta_1\) and the nonlinear component \(\delta_2\) are given by the following cascaded system of differential equations,

\[
\dot{\delta}_1(t) = A(t, u, \sigma) \delta_1(t) + B(t) \left[ w(t) - \sum_{i=0}^{N} \sigma \Phi_i(t) \right],
\]

(\text{A2})
\[
\dot{\delta}_2(t) = (A[t, u, \sigma] + \Delta[t, u, \sigma]) \delta_2(t) + \Delta[t, u, \sigma] \dot{\delta}_1(t),
\]
with initial conditions \(\delta_1(0) = \delta_2(0) = 0\). The solution sets to these differential equations can then be bounded sequentially:

1. The solutions to the differential equation (A2) can be expressed in the form

\[
\delta_1(t) = \int_0^T \mathcal{H}[t, u, \sigma](\tau) \left( w(\tau) - \sum_{i=0}^N \sigma_i \Phi_i(\tau) \right) d\tau,
\]

with \(\mathcal{H}[t, u, \sigma]\) denoting the linear Hankel operator associated with the pair \((A[\cdot, u, \sigma], B(\cdot))\). It has been established in Houska and Chachuat\(^4\) that an ellipsoidal enclosure of \(\delta_1(t)\) for all \(w \in \mathcal{W}\) can be obtained in the form

\[
\mathcal{E}(Q_1[t, u, \sigma]) := \left\{ Q_1[t, u, \sigma]^{1/2} v \in \mathbb{R}^n, v^T v \leq 1 \right\},
\]

with the shape matrix

\[
Q_1[t, u, \sigma] = \gamma^2 \int_0^T \left( \mathcal{H}[t, u, \sigma](\tau) - \sum_{i=0}^N \int_0^T \frac{\mathcal{H}[t, u, \sigma](\tau) \Phi_i(\tau) \mu(\tau) d\tau}{\sigma_i} \right) \Phi_i(\tau) d\tau,
\]

where \(\gamma > 0\) is any upper bound on the \(\mathbb{L}_2\)-norm of \(w(\tau) - \sum_{i=0}^N \sigma_i \Phi_i(\tau)\), which could be eliminated by a rescaling of the dynamic system. The advantage of having a closed-form expression for \(Q_1\) is that it is more readily amenable to numerical computation. In particular, an upper bound on the term \(\mathcal{H}[t, u, \sigma](\tau) - \sum_{i=0}^N \int_0^T \frac{\mathcal{H}[t, u, \sigma](\tau) \Phi_i(\tau) \mu(\tau) d\tau}{\sigma_i} \Phi_i(\tau)\) can be directly obtained from the remainder term of a polynomial model of order \(N\) of \(\mathcal{H}[t, u, \sigma](\tau)\), for instance a Chebyshev model\(^5\) if the basis functions \(\Phi_i\) are chosen as the Chebyshev polynomials of the first kind. Denoting these bounds by \(r_H[\cdot, u, \sigma](\tau) \in \mathbb{R}^n\), the (upper- or lower-triangular) entries of the shape matrix \(Q_1[\cdot, u, \sigma]\) can then be propagated as quadrature variables along \([0, T]\) using a standard numerical integrator,

\[
\dot{Q}_1[t, u, \sigma] = \gamma^2 r_H[t, u, \sigma] \cdot r_H[t, u, \sigma]^T,
\]

with \(Q_1[0, u, \sigma] = 0\).

2. Ellipsoidal bounds \(\mathcal{E}(Q_2[\cdot, u, \sigma])\) on the nonlinear defect \(\delta_2\) described by the differential equation (A3) can be propagated by regarding both \(\delta_1\) and \(\Delta\) as time-varying uncertainties and using the ellipsoidal bounds \(\mathcal{E}(Q_1[\cdot, u, \sigma])\) computed through step (1) together with the bounds \(\overline{\Delta}()\). This may be done by applying well-established techniques based on linearization and additive ellipsoidal uncertainties,\(^4\) or recent advances in the field of robust control of linear systems with multiplicative ellipsoidal uncertainties,\(^4\) which provide tighter enclosures based on a novel tight version of the S-procedure for structured quadratically constrained quadratic programming problems with two convex constraints.

Overall, a convex enclosure of the discretization error set is obtained as the Minkowski sum of the \(t\) bounding ellipsoids for the linear and nonlinear parts of the differential equation (A1),

\[
\overline{E}_N(t, u, \sigma) = \mathcal{E}(Q_1[t, u, \sigma]) \oplus \mathcal{E}(Q_2[t, u, \sigma]).
\]

In the implementation of the case study in Section 5, both \(Q_1\) and \(Q_2\) are differentiable in \(u\), since the Hankel operator \(\mathcal{H}\) itself is differentiable in \(u\) and all the nonlinearity bounds are independent of \(u\).