J Glob Optim (2009) 44:395–407
DOI 10.1007/s10898-008-9334-6

A numerical approach to design control invariant sets for constrained nonlinear discrete-time systems with guaranteed optimality

Jian Wan · Josep Vehi · Ningsu Luo

Received: 21 July 2007 / Accepted: 6 August 2008 / Published online: 28 August 2008

Abstract A numerical approach to design control invariant sets for constrained nonlinear discrete-time systems with guaranteed optimality is proposed in this paper. The addressed approach is based on the fact that zonotopes are more flexible for representing sets than boxes in interval analysis. Then the solver of set inversion via interval analysis is extended to set inversion via zonotope geometry by introducing the novel idea of bisecting zonotopes. The main feature of the extended solver of set inversion is the bisection and the evolution of a zonotope rather than a box. Thus the shape of admissible domains for set inversion can be broadened from boxes to zonotopes and the wrapping effect can be reduced as well by using the zonotope evolution instead of the interval evolution. Combined with global optimization via interval analysis, the extended solver of set inversion via zonotope geometry is further applied to design control invariant sets for constrained nonlinear discrete-time systems in a numerical way. Finally, the numerical design of a control invariant set and its application to the terminal control of the dual-mode model predictive control are fulfilled on a benchmark Continuous-Stirred Tank Reactor example.

Keywords Set inversion · Global optimization · Interval analysis · Zonotope geometry · Control invariant sets

1 Introduction

Set invariance plays a significant role in many practical problems concerning constrained control, robustness synthesis and constrained optimization [1,2]. For example, in the dual-mode model predictive control (MPC), a terminal control invariant set along with a local stabilizing linear feedback control law is usually needed to be designed in advance for the terminal control mode once the system state has been driven into the terminal set through the receding horizon control mode [3,4]. The optimality of the designed control invariant

J. Wan (✉) · J. Vehi · N. Luo
Institut d'Informàtica i Aplicacions, Universitat de Girona, Campus Montilivi, Girona 17071, Spain
e-mail: jwan@eia.udg.es

set can be judged according to its volume, where a larger control invariant set is preferred for a shorter control horizon and an easier stabilization [5]. The analytical design of the maximal control invariant set along with a local stabilizing linear feedback control law for a constrained nonlinear discrete-time system is quite challenging and usually the obtained control invariant set is restricted to be an ellipsoid or a low-complexity polytope [5–10]. Approaches based on linear dynamic approximation together with Lipschitz bounds on the errors of approximation were discussed in [7,8] to obtain terminal control invariant ellipsoids for constrained nonlinear systems. Another approach to obtain a terminal control invariant ellipsoid was based on a linear difference inclusion of the original nonlinear system [9]. In [10], the related local stabilizing linear feedback control law was designed using the LQ method for the linearized system of the original nonlinear system and the associated terminal control invariant ellipsoid was obtained through an optimization. In [5], a low-complexity control invariant polytope together with a local stabilizing linear feedback control law was also designed in an optimal way for an input-affine nonlinear system, where the advantage of a control invariant polytope over a control invariant ellipsoid in terms of volume was demonstrated. The design of a control invariant low-complexity polytope with respect to a feedback linearizing control law for input-affine nonlinear systems was further proposed in [6], where the designed invariant set along with the feedback linearizing control law was used in the terminal control of the dual-mode MPC. In fact, polytopes are usually the natural expression of physical constraints on state and control variables leading to a more flexible and pertinent description of corresponding control invariant sets [1]. However, the design of a relatively complex control invariant polytope, which is more likely to have a bigger volume, for a general constrained nonlinear discrete-time system is still an open problem.

The design of a control invariant set for a general constrained nonlinear discrete-time system is closely related to the test of control invariance for a candidate set in the viewpoint of set computation. A set is control invariant for a discrete-time control system means that the dynamic evolution of the system under the related control law is always within the given set. Thus the test of control invariance for a given set is a specific set inversion problem [11], where the admissible domain and the range for the dynamic function are the same. The solver of set inversion via interval analysis was proposed in [12] as a method of set computation and it has been widely applied to find feasible solutions concerning nonlinear inequalities [13,14]. The solver needs an initial admissible interval vector or a box from which to search all feasible solutions through bisections and selections. However, the admissible domains of practical problems such as the test of control invariance are not always boxes and additional linear constraints are usually imposed on function variables. In such a scenario, the resulting admissible domains are polytopes or zonotopes rather than boxes and the existing solver of set inversion via interval analysis are not suitable to deal with such kind of set inversion problems in a direct way.

This paper aims to provide a numerical method to design control invariant sets for general constrained nonlinear discrete-time systems with guaranteed optimality. It is realized by extending the solver of set inversion via interval analysis to set inversion via zonotope geometry, where both the bisection and the evolution are based on zonotopes instead of interval vectors or boxes. Thus the admissible domains for set inversion can be broadened from boxes to zonotopes, which are more flexible than boxes for representing sets and fulfilling set computation [15]. Global optimization via interval analysis is also applied to find the control invariant set with the maximal volume through searching among all the feasible control invariant sets. The paper is organized as follows: the problem considered is stated in Sect. 2; interval analysis is introduced briefly in Sect. 3; basic concepts and operations of interval analysis are extended to zonotope geometry in Sect. 4; the extended solver of set inversion via zonotope geometry is presented in Sect. 5; using set inversion via zonotope geometry as well as global optimization via interval analysis for the design of control invariant sets is addressed in Sect. 6; an illustrative example for the proposed numerical approach to design a control invariant set for a constrained nonlinear discrete-time system as well as the application of the designed control invariant set to the terminal control of the dual-mode MPC is given in Sect. 7; and finally, some conclusions are drawn in Sect. 8.

2 Problem statement

Consider the general constrained nonlinear discrete-time system:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k), \quad k = 0, 1, \ldots, \tag{1}$$

where $\mathbf{x}_k \in \mathbf{X} \subset \mathbb{R}^n$ is the system state and $\mathbf{u}_k \in \mathbf{U} \subset \mathbb{R}^m$ is the control input. The set $\mathbf{X}$ and the set $\mathbf{U}$ are compact. The target is to design an optimal control invariant zonotope $\mathcal{Z} \subseteq \mathbf{X}$ along with a local stabilizing linear feedback control law $\mathbf{u}_k = \mathbf{k}\mathbf{x}_k$, which satisfies the following condition:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{k}\mathbf{x}_k) \in \mathcal{Z}, \quad \mathbf{k}\mathbf{x}_k \in \mathbf{U}, \quad \forall \mathbf{x}_k \in \mathcal{Z}. \tag{2}$$

The related local stabilizing linear feedback control law $\mathbf{u}_k = \mathbf{k}\mathbf{x}_k$ is to be designed in advance using the LQ method for the linearized system of the original nonlinear system [10].

The test of control invariance for a given set and a related control law can be transformed to be a set inversion problem concerning set computation. For a possibly nonlinear function $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ with the admissible domain $\mathbf{x} \in \mathbf{X}$ and the known range $\mathbb{T} \subset \mathbb{R}^m$, set inversion is the characterization [11]:

$$\mathbb{X} = \{\mathbf{x} \in \mathbf{X} \subset \mathbb{R}^n | \mathbf{f}(\mathbf{x}) \in \mathbb{T}\} = \mathbf{f}^{-1}(\mathbb{T}). \tag{3}$$

According to the condition of control invariance in (2), the control invariance of a given zonotope $\mathcal{Z}$ along with a related local stabilizing linear feedback control law $\mathbf{u}_k = \mathbf{k}\mathbf{x}_k$ for the system (1) is implied by the following result of the set inversion problem (3):

$$\mathbf{X} = \mathcal{Z} = \mathbb{T} = \mathbb{X}, \tag{4}$$

where $\mathbf{u}_k = \mathbf{k}\mathbf{x}_k \in \mathbf{U}, \forall \mathbf{x}_k \in \mathcal{Z}$. This means that the given zonotope is treated as the admissible domain and the solution set for (3) is the same to the admissible domain as well as the known range. Thus the design of an optimal control invariant zonotope $\mathcal{Z}$ can be transformed to be the test of control invariance and the selection of an optimal control invariant zonotope among a number of feasible control invariant zonotopes, where the bisection and the selection of zonotopes are concerned instead of boxes during the test of control invariance of a given zonotope using set computation.

3 Interval analysis

The initial idea of interval analysis is to enclose real numbers in intervals and real vectors in boxes as a method of considering the imprecision of representing real numbers by finite digits in numerical computers. Interval analysis has become a fundamental numerical tool for representing uncertainties or errors, proving properties of sets, solving sets of nonlinear equations or inequalities and designing some global optimization algorithms. The key concepts of interval analysis are interval arithmetic and inclusion function, whose definitions are as follows [11]:

Interval arithmetic: Interval arithmetic is a special case of computation on sets, which includes real compact intervals $[a, b] = \{x \in \mathbb{R} | a \leq x \leq b\}$, where $a \in \mathbb{R}, b \in \mathbb{R}$ and $a \leq b$, real compact interval vectors $\mathbf{X}_{n\times 1}$ and real compact interval matrices $\mathbf{X}_{n\times m}$. The four elementary arithmetic operations $(+, -, \times, \div)$ are extended to intervals. Concretely, for any such binary operator, denoted by $\circ$, performing the operation associated with $\circ$ on the intervals $[a, b]$ and $[c, d]$ means computing $[a, b] \circ [c, d] = [\{x \circ y \in \mathbb{R} | x \in [a, b], y \in [c, d]\}]$, where $[\cdot]$ denotes the convex hull of $\{x \circ y \in \mathbb{R} | x \in [a, b], y \in [c, d]\}$. Correspondingly, the set of all interval vectors in the domain of $\mathbb{R}^n$ is denoted to be $\mathbb{I}(\mathbb{R}^n)$.

Inclusion function: Consider a function $\mathbf{f}$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, the interval function $\mathbf{F}$ from $\mathbb{I}(\mathbb{R}^n)$ to $\mathbb{I}(\mathbb{R}^m)$ is an inclusion function for $\mathbf{f}$ if $\forall \mathbf{X} \in \mathbb{I}(\mathbb{R}^n), \mathbf{f}(\mathbf{X}) \subseteq \mathbf{F}(\mathbf{X})$. The natural inclusion function of $\mathbf{f}(\mathbf{X})$ can be obtained by replacing each occurrence of every variable with the corresponding interval variable, by executing all operations according to interval arithmetic and by computing intervals enclosing the range of standard functions.

The fundamental concepts of interval analysis can be integrated to set up various algorithms for solving set inversion, global optimization and minimax optimization problems in a guaranteed numerical way [11,16]. A basic operation within these solvers is to bisect an interval vector into two sub-interval vectors. Taking the interval vector $\mathbf{X} = [a_1, b_1] \times \cdots \times [a_n, b_n]$ as an example, its width is denoted to be:

$$\texttt{Width}(\mathbf{X}) = \max_{i=1,\ldots,n} |a_i - b_i|, \tag{5}$$

and the index j is denoted to be:

$$j = \min_{i=1,\ldots,n} \{i | (|a_i - b_i|) = \texttt{Width}(\mathbf{X})\}, \tag{6}$$

then the bisection $\texttt{Bisect}(\mathbf{X})$ returns two sub-interval vectors $\texttt{L}\mathbf{X}$ and $\texttt{R}\mathbf{X}$ ($\texttt{L}$ for Left and $\texttt{R}$ for Right):

$$\begin{cases} \texttt{L}\mathbf{X} = [a_1, b_1] \times \cdots \left[a_j, \frac{(a_j+b_j)}{2}\right] \times \cdots [a_n, b_n] \\ \texttt{R}\mathbf{X} = [a_1, b_1] \times \cdots \left[\frac{(a_j+b_j)}{2}, b_j\right] \times \cdots [a_n, b_n]. \end{cases} \tag{7}$$

The solver of set inversion via interval analysis is an algorithm of set computation aiming to find all feasible solutions $\mathbb{X}$ satisfying $\mathbf{f}(\mathbb{X}) \subseteq \mathbb{T}$ in (3), where the admissible domain $\mathbf{X}$ of $\mathbf{f}(\mathbf{x})$ is usually assumed to be an interval vector or a box. It is fulfilled in a numerical way by bisecting the admissible domain $\mathbf{X}$ into subboxes, computing the inclusion function $\mathbf{F}$ of $\mathbf{f}(\mathbf{x})$ for each subbox via interval arithmetic, and finally selecting all feasible subboxes through comparing the resulting image of the inclusion function $\mathbf{F}$ for each subbox with the known range $\mathbb{T}$ [11].

4 Zonotope geometry

This section introduces zonotope geometry with comparison to interval analysis, i.e., the basic concepts of zonotope geometry are derived by extending their counterpart concepts of interval analysis and zonotope geometry is treated as an extension of interval analysis in geometry.

4.1 Zonotope definition

A zonotope is a centrally symmetric convex polytope and it is closely related to interval analysis. Given a vector $\mathbf{p} \in \mathbb{R}^n$ and a matrix $H \in \mathbb{R}^{n\times m}$, the zonotope $\mathcal{Z}$ of order $n \times m$ is the set:

$$\mathbf{p} \oplus H\mathbf{B}^m = \{\mathbf{p} + H\mathbf{z} | \mathbf{z} \in \mathbf{B}^m\}, \tag{8}$$

where $\mathbf{B}^m$ is an interval vector (i.e., aligned box) composed of m unitary intervals $\mathbf{B} = [-1, 1]$ and $\oplus$ is the Minkowski sum of sets. Assume that $H = [\mathbf{h}_1 \cdots \mathbf{h}_m]$, then the zonotope can also be regarded as a set spanned by the column vectors of H, which are also called line segment generators:

$$\mathcal{Z} = \left\{\mathbf{p} + \sum_{i=1}^{m} \alpha_i \mathbf{h}_i | -1 \le \alpha_i \le 1\right\}. \tag{9}$$

Geometrically, the zonotope $\mathcal{Z}$ is the transferred Minkowski sum of the line segments defined by the columns of the matrix H to the central point $\mathbf{p}$. The zonotope $\mathcal{Z}$ becomes a simple parallelotope when the matrix $H \in \mathbb{R}^{n\times m}$ is invertible. More specifically, the zonotope $\mathcal{Z}$ is simplified to be an interval vector as well as a box when H is a diagonal matrix or when $m = 1$. The mathematical concept of zonotopes has not yet been widely used explicitly in the control literature although zonotopes have previously been applied to bound system states in [17–19]. However, its implicit form, i.e., the Minkowski sum of sets, has already been widely applied to approximate various kinds of invariant sets such as the minimal robust positively invariant set and the minimal disturbance invariant set in a recursive way for linear discrete-time systems [2,20]. Here zonotopes are to be explored further to obtain the maximal control invariant zonotope in a numerical way for general constrained nonlinear discrete-time systems.

4.2 Zonotope construction

The list of line segment generators is an efficient implicit representation of a zonotope. Using the implicit representation, set operations such as the Minkowski sum and difference are then trivial. However, the explicit representation of a zonotope is needed for some operations such as the judgement of inclusion and exclusion of a polytope to a zonotope. The explicit representation of a zonotope is the zonotope construction problem aiming to list all extreme points of a zonotope defined by its line segment generators, which could be computationally intensive for high-dimensional cases. A relatively efficient algorithm was proposed in [21] to address the zonotope construction problem, where the addition of line segments was replaced by the addition of convex polytopes. For example, the construction of the zonotope $\mathcal{Z} = \mathbf{p} \oplus H\mathbf{B}^6$, where $\mathbf{p} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $H = \begin{bmatrix} 0.4414 & -0.5855 & -0.0484 & 0.2570 & 0.2293 & 0.1498 \\ -0.0016 & -0.3930 & 0.3526 & -0.2396 & 0.4257 & -0.3117 \end{bmatrix}$, can be transformed to be the Minkowski sum of three simpler zonotopes, i.e., $\mathcal{Z} = \mathcal{Z}_1 \oplus \mathcal{Z}_2 \oplus \mathcal{Z}_3$, where $\mathcal{Z}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 0.4414 & -0.5855 \\ -0.0016 & -0.3930 \end{bmatrix} \mathbf{B}^2$, $\mathcal{Z}_2 = \begin{bmatrix} -0.0484 & 0.2570 \\ 0.3526 & -0.2396 \end{bmatrix} \mathbf{B}^2$ and $\mathcal{Z}_3 = \begin{bmatrix} 0.2293 & 0.1498 \\ 0.4257 & -0.3117 \end{bmatrix} \mathbf{B}^2$. Thus the zonotope $\mathcal{Z}$ can be constructed explicitly with reduced dimensionality and it can be plotted as well using polytope geometry softwares such as Multi-Parametric Toolbox [22], which is shown in Fig. 1.

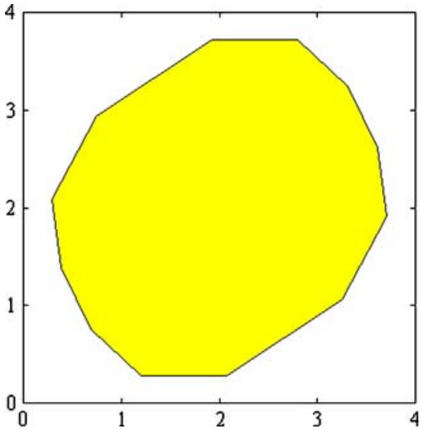


Fig. 1 An example of zonotope construction

4.3 Zonotope bisection

Similar to an interval vector or a box, a method is also proposed to bisect a zonotope in this subsection. Taking the zonotope $\mathcal{Z} = \mathbf{p} \oplus H\mathbf{B}^m$ as an example, where $\mathbf{p} \in \mathbb{R}^n$, $H \in \mathbb{R}^{n\times m}$ and $m \geq n$, the maximum absolute value among all elements h_{ij} in H is denoted to be:

$$\mathtt{Max}(H) = \max_{i=1,\ldots,n,\, j=1,\ldots,m} |h_{ij}|, \tag{10}$$

and the index k is denoted to be:

$$k = \min_{j=1,\ldots,m} \{j | (|h_{ij}|) = \mathtt{Max}(H), i = 1, \ldots, n\}. \tag{11}$$

Then the zonotope $\mathcal{Z}$ can be bisected along the line segment generator $\mathbf{h}_k$, which is addressed in Theorem 1.

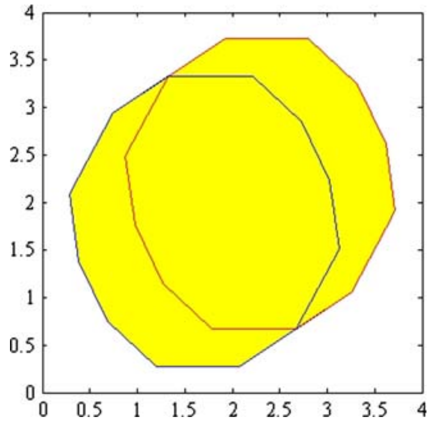
Theorem 1 (Zonotope bisection) *The bisection* $\mathtt{Bisect}(\mathcal{Z})$ *along the line segment generator* $\mathbf{h}_k$ *returns two sub-zonotopes* $\mathtt{L}\mathcal{Z} = (\mathbf{p} - \frac{\mathbf{h}_k}{2}) \oplus [\mathbf{h}_1 \cdots \frac{\mathbf{h}_k}{2} \cdots \mathbf{h}_m]\mathbf{B}^m$ *and* $\mathtt{R}\mathcal{Z} = (\mathbf{p} + \frac{\mathbf{h}_k}{2}) \oplus [\mathbf{h}_1 \cdots \frac{\mathbf{h}_k}{2} \cdots \mathbf{h}_m]\mathbf{B}^m$.

Proof Since $\mathcal{Z} = \mathbf{p} \oplus [\mathbf{h}_1 \cdots \mathbf{h}_k \cdots \mathbf{h}_m]\mathbf{B}^m$, then $\mathcal{Z} = \mathbf{p} \oplus [\mathbf{h}_1 \cdots \mathbf{h}_k \cdots \mathbf{h}_m]\ [[-1, 1]_1 \cdots [-1, 0]_k \cdots [-1, 1]_m]^T \cup \mathbf{p} \oplus [\mathbf{h}_1 \cdots \mathbf{h}_k \cdots \mathbf{h}_m][[-1, 1]_1 \cdots [0, 1]_k \cdots\ [-1, 1]_m]^T = \mathtt{L}\mathcal{Z} \cup \mathtt{R}\mathcal{Z}$, where $\mathtt{L}\mathcal{Z} = (\mathbf{p} - \frac{\mathbf{h}_k}{2}) \oplus [\mathbf{h}_1 \cdots \frac{\mathbf{h}_k}{2} \cdots \mathbf{h}_m]\mathbf{B}^m$ and $\mathtt{R}\mathcal{Z} = (\mathbf{p} + \frac{\mathbf{h}_k}{2}) \oplus [\mathbf{h}_1 \cdots \frac{\mathbf{h}_k}{2} \cdots \mathbf{h}_m]\mathbf{B}^m$. □

Taking the zonotope shown in Fig. 1 as an example, the proposed bisection $\mathtt{Bisect}(\mathcal{Z})$ returns two sub-zonotopes, which are shown in Fig. 2. It can be seen that the bisection is not complete for the zonotope of order 2×6. The reason for the overlapping of $\mathtt{L}\mathcal{Z}$ and $\mathtt{R}\mathcal{Z}$ is that the line segment generators $\mathbf{h}_1, \ldots, \mathbf{h}_6$ are redundant or not linearly independent and then the rank of the intersection $\mathtt{L}\mathcal{Z} \cap \mathtt{R}\mathcal{Z}$ can be 2. However, for a zonotope $\mathcal{Z}$ with linearly independent line segment generators, the bisection is complete, which is addressed in Theorem 2.

Theorem 2 (Complete bisection) *For a zonotope* $\mathcal{Z} = \mathbf{p} \oplus H\mathbf{B}^n$, *where* $\mathbf{p} \in \mathbb{R}^n$, $H \in \mathbb{R}^{n\times n}$ *and* $\mathtt{Rank}(H) = n$, *the defined bisection is complete, i.e.,* $\mathtt{L}\mathcal{Z}$ *and* $\mathtt{R}\mathcal{Z}$ *only share a face of dimension* $n - 1$.

Fig. 2 Bisection of a zonotope with redundant line segment generators



Proof For $\mathcal{Z} = \{\mathbf{p} + \sum_{i=1}^{n} \alpha_i \mathbf{h}_i \,|\, -1 \le \alpha_i \le 1\}$, assume that there exist $-1 \le \alpha_1^L \le 1, \ldots, -1 \le \alpha_k^L \le 0, \ldots, -1 \le \alpha_n^L \le 1$ and $-1 \le \alpha_1^R \le 1, \ldots, 0 \le \alpha_k^R \le 1, \ldots, -1 \le \alpha_n^R \le 1$, s.t. $\mathbf{p} - \frac{\mathbf{h}_k}{2} + \sum_{i=1}^{k-1} \alpha_i^L \mathbf{h}_i + \alpha_k^L \frac{\mathbf{h}_k}{2} + \sum_{i=k+1}^{n} \alpha_i^L \mathbf{h}_i = \mathbf{p} + \frac{\mathbf{h}_k}{2} + \sum_{i=1}^{k-1} \alpha_i^R \mathbf{h}_i + \alpha_k^R \frac{\mathbf{h}_k}{2} + \sum_{i=k+1}^{n} \alpha_i^R \mathbf{h}_i$, then $\sum_{i=1}^{k-1} (\alpha_i^L - \alpha_i^R)\mathbf{h}_i + (\alpha_k^L - \alpha_k^R - 2)\frac{\mathbf{h}_k}{2} + \sum_{i=k+1}^{n} (\alpha_i^L - \alpha_i^R)\mathbf{h}_i = 0$ while $\texttt{Rank}(H) = n$, so $\alpha_k^L = 1, \alpha_k^R = -1$, i.e., $\texttt{L}\mathcal{Z}$ and $\texttt{R}\mathcal{Z}$ only share a face of dimension $n - 1$. □

4.4 Zonotope inclusion

Using zonotopes, Kühn developed a procedure to bound the orbits of discrete-time dynamic systems with a guaranteed sub-exponential overestimation. The following definition introduces the zonotope inclusion operator of Kühn's method [15].

Definition 1 (*Zonotope inclusion*) Consider a family of zonotopes represented by $\mathbb{Z} = \mathbf{p} \oplus \mathbf{M}\mathbf{B}^m$, where $\mathbf{p} \in \mathbb{R}^n$ is a real vector and $\mathbf{M} \in \mathbb{I}(\mathbb{R}^{n\times m})$ is an interval matrix. A zonotope inclusion, denoted by $\diamond(\mathbb{Z})$, is defined by:

$$\diamond(\mathbb{Z}) = \mathbf{p} \oplus [\texttt{Mid}(\mathbf{M}) \;\; G] \begin{bmatrix} \mathbf{B}^m \\ \mathbf{B}^n \end{bmatrix}, \tag{12}$$

where $\texttt{Mid}(\mathbf{M})$ is the centered-point matrix of $\mathbf{M}$ and $G \in \mathbb{R}^{n\times n}$ is a diagonal matrix that satisfies:

$$G_{ii} = \sum_{j=1}^{m} \frac{\texttt{Diam}(\mathbf{M}_{ij})}{2}, \quad i = 1, \ldots, n, \tag{13}$$

where $\texttt{Diam}(\mathbf{M}_{ij})$ is the length of the interval $\mathbf{M}_{ij}$. Under these definitions, it results that: $\mathbb{Z} \subseteq \diamond(\mathbb{Z})$.

Given a possibly nonlinear function $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{x} \in \mathcal{X} = \mathbf{p} \oplus M\mathbf{B}^m$, its centered inclusion function $\mathbf{F}_c(\mathcal{X}) : \mathbf{f}(\mathcal{X}) \subseteq \mathbf{F}_c(\mathcal{X})$ can be deduced by the mean-value theorem [11], i.e.,

$$\mathbf{F}_c(\mathcal{X}) \triangleq \mathbf{f}(\mathbf{p}) + \triangledown_{\mathbf{x}} \mathbf{f}(\mathcal{X})(\mathcal{X} - \mathbf{p}), \tag{14}$$

where $\triangledown_{\mathbf{x}} \mathbf{f}(\mathcal{X})$ is the Jacobian of $\mathbf{f}(\mathcal{X})$ and $\mathcal{X} - \mathbf{p} = M\mathbf{B}^m$. Thus the centered inclusion function $\mathbf{F}_c(\mathcal{X})$ of $\mathbf{f}(\mathbf{x})$ turns out to be a family of zonotopes represented by $\mathbb{Z} = \mathbf{p}_c \oplus \mathbf{M}_c\mathbf{B}^m$,

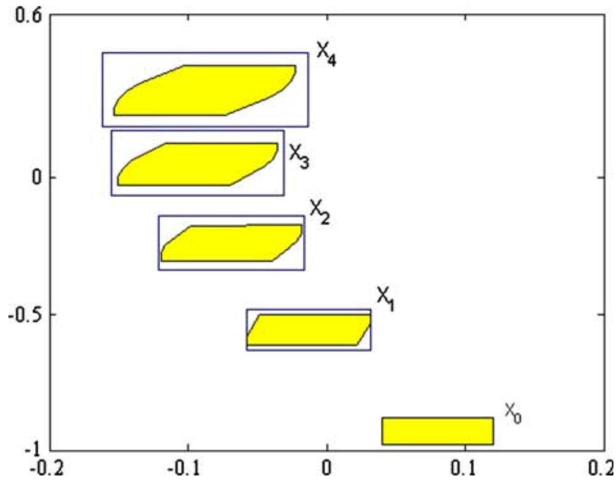


Fig. 3 The interval evolution vs the zonotope evolution

where $\mathbf{p}_c = \mathbf{f}(\mathbf{p})$ and $\mathbf{M}_c = \nabla_{\mathbf{x}}\mathbf{f}(\mathcal{X})M$, which can be further bounded by its corresponding zonotope inclusion $\diamond(\mathbb{Z})$. This is the primary principle of Kühn's method to bound the evolution of dynamic systems using zonotopes, where centered inclusion functions are applied instead of natural inclusion functions. The reduced wrapping effect can be seen in an illustrative example shown in Fig. 3, where the interval evolution and the zonotope evolution of four steps for the nonlinear discrete-time system discussed in [5] with the same control sequence and the same initial state domain are compared.

5 Set inversion via zonotope geometry

Given a possibly nonlinear dynamic system $\mathbf{f}(\mathbf{x})$, a zonotope $\mathcal{X}$ as the admissible domain and the range $\mathbb{T}$, the solver of set inversion via zonotope geometry is listed in **Algorithm I**, where ε is the bound of error tolerance and $\Sigma_{\mathcal{X}}$, $\Sigma^b_{\mathcal{X}}$ are to store the feasible sub-zonotopes and the neighboring sub-zonotopes with $\mathtt{Max}(H_i) < \varepsilon$ to all the feasible sub-zonotopes, respectively.

Algorithm I: Set Inversion Via Zonotope Geometry (SIVZG)
1. Initialize **Stack** $= \mathcal{X}$ and $\Sigma_{\mathcal{X}} = \Sigma^b_{\mathcal{X}} = \emptyset$;
2. while **Stack** $\neq \emptyset$
3. Pop out a zonotope $\mathcal{X}_i = \mathbf{p}_i \oplus H_i\mathbf{B}^m$ from **Stack**;
4. *if* $\diamond(\mathbf{F}_c(\mathcal{X}_i)) \subseteq \mathbb{T}$, $\Sigma_{\mathcal{X}} = \Sigma_{\mathcal{X}} \cup \mathcal{X}_i$ and return to 2;
5. *elseif* $\mathtt{Max}(H_i) < \varepsilon$, $\Sigma^b_{\mathcal{X}} = \Sigma^b_{\mathcal{X}} \cup \mathcal{X}_i$ and return to 2;
6. *elseif* $\diamond(\mathbf{F}_c(\mathcal{X}_i)) \cap \mathbb{T} = \emptyset$, discard $\mathcal{X}_i$ and return to 2;
7. *else*
8. Bisect $\mathcal{X}_i$ to L$\mathcal{X}_i$ and R$\mathcal{X}_i$, push them on **Stack**;
9. *endif*
10. endwhile

The solver of set inversion via zonotope geometry is similar to the solver of set inversion via interval analysis and only the bisection and the evolution of interval vectors are

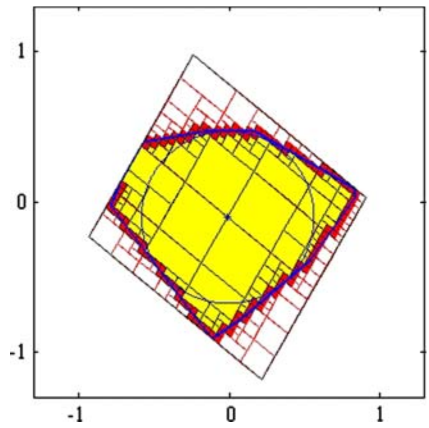


Fig. 4 Zonotope approximation of a polytope

replaced by the bisection and the evolution of zonotopes [11], where the initial admissible domain is broadened from boxes to zonotopes and the wrapping effect is reduced by using the zonotope evolution instead of the interval evolution. It is worthy to note that the zonotopes used are transformed to be the format of polytopes using zonotope construction and thus the test of inclusion in `Step` 4 and the test of intersection in `Step` 5 can be fulfilled by polytope geometry softwares such as Multi-Parametric Toolbox [22]. Since the bisection of a zonotope with redundant line segment generators is not complete, an alternative approach is to bound the zonotope with redundant line segment generators by a zonotope with linearly independent line segment generators at first and thus the bisection of the bounding zonotope is complete. The bounding of a zonotope with redundant line segment generators can be realized by using singular value decomposition of the matrix H [23]. Furthermore, an initial admissible polytope can also be bounded by a zonotope with linearly independent line segment generators using the center of the largest ball inscribed in the polytope as the center of the bounding zonotope [22]. A direct application of the solver of set inversion via zonotope geometry in Table I is shown in Fig. 4, where a polytope is approximated innerly by a union of zonotopes using a bounding zonotope of the polytope as the initial admissible domain.

6 Global optimization for set inversion via zonotope geometry

Given an initial interval matrix $\mathbf{M} \in \mathbb{I}(\mathbb{R}^{n\times m})$, its width is denoted to be:

$$\mathtt{Width}(\mathbf{M}) = \max_{i,j} \mathtt{Width}(\mathbf{M}_{ij}), \quad i = 1, \ldots, n,\, j = 1, \ldots, m, \tag{15}$$

where $\mathtt{Width}(\mathbf{M}_{ij})$ is denoted to be the width of the interval $\mathbf{M}_{ij}$. The interval matrix $\mathbf{M}$ can also be bisected into two sub-interval matrix $\mathtt{L}\mathbf{M}$ and $\mathtt{R}\mathbf{M}$ by bisecting the widest member of its components. Using the bisection and the selection of $\mathbf{M}$, the solver of global optimization via interval analysis for set inversion via zonotope geometry can be built to search the maximal control invariant zonotope for a constrained nonlinear discrete-time system with a related local stabilizing linear feedback control law $\mathbf{u} = \mathbf{kx}$, just as shown in **Algorithm II**, where $\mathtt{Vol}(\cdot)$ stands for the volume of a zonotope and ε is the bound of error tolerance.

```
Algorithm II: Global Optimization For Set Inversion (GOFSI)
1. Initialize Stack = M  and  𝒵 = ∅;
2. while Stack ≠ ∅
3.     Pop out an interval matrix M_k from Stack;
4.     Set ℤ_k = M_k B^m and test if ◇(ℤ_k) is control invariant or no by
SIVZG;
5.     if ◇(ℤ_k) is control invariant and Vol(◇(ℤ_k)) > Vol(𝒵), 𝒵 = ◇(ℤ_k)
and return to 2;
6.     elseif Width(M) ≤ ε, return to 2;
7.     else
8.           Bisect M_k to LM_k and RM_k, push them on Stack;
9.     endif
10.endwhile
```

The solver of global optimization via interval analysis for set inversion via zonotope geometry in **Algorithm II** searches the maximal control invariant zonotope starting from an initial zonotope derived from an initial interval matrix $\mathbf{M}$, where the input of the algorithm is $\mathbf{M}$ and the output of the algorithm is $\mathcal{Z}$. The selection of such an initial interval matrix $\mathbf{M}$ for the search should include the stability domain of the dynamic system under the related local stabilizing control law. The zonotope inclusion of each family of zonotopes represented by $\mathbb{Z}_k = \mathbf{M}_k\mathbf{B}^m$ is passed to the solver of set inversion via zonotope geometry for the test of control invariance with the related local stabilizing linear feedback control law, where the input of **Algorithm I** is $\mathcal{X} = \mathbb{T} = \diamond(\mathbb{Z}_k)$. Through the bisection and the selection of the initial interval matrix $\mathbf{M}$ globally, the algorithm can return the control invariant zonotope with the maximal volume, where the volumes of zonotopes can be computed analytically once zonotopes are transformed to be the format of polytopes [24]. It is worthy to note that the complexity and the volume of the obtained maximal control invariant zonotope $\mathcal{Z}$ is closely related to the dimension and the value of the selected initial interval matrix $\mathbf{M}$. The complexity of the algorithm increases with the dimension and the range of the selected initial interval matrix $\mathbf{M}$. Furthermore, other kinds of local stabilizing feedback control laws can be applied as well because the solver of set inversion via zonotope geometry is applicable to a whatever nonlinear autonomous system with first-order differentiability for deriving its centered inclusion function.

7 An illustrative example

This section gives an illustrative example for the application of the solver of set inversion via zonotope geometry and the solver of global optimization for set inversion to design the maximal control invariant zonotope for a constrained nonlinear discrete-time system, i.e., the solver of global optimization for set inversion via zonotope geometry searches from an initial zonotope derived from an interval matrix $\mathbf{M}$ to find the control invariant zonotope $\mathcal{Z}$ with the maximal volume.

Taking the highly nonlinear model of a Continuous Stirred-Tank Reactor (CSTR) as the example [10], assuming constant liquid volume, the CSTR for an exothermic, irreversible reaction, $\mathbf{A} \rightarrow \mathbf{B}$, is described by the following dynamic model based on a component balance for the reactant $\mathbf{A}$ and an energy balance:

$$\begin{cases} \dot{C}_A = \frac{q}{V}\left(C_{Af} - C_A\right) - k_0 \exp\left(-\frac{E}{RT}\right) C_A, \\ \dot{T} = \frac{q}{V}(T_f - T) + \frac{(-\triangle H)}{\rho C_p} k_0 \exp\left(-\frac{E}{RT}\right) C_A + \frac{UA}{V\rho C_p}\left(T_c - T\right), \end{cases} \tag{16}$$

where C_A is the concentration of **A** in the reactor, T is the reactor temperature and T_c is the temperature of the coolant stream. The constraints are $280\,\mathrm{K} \leq T_c \leq 370\,\mathrm{K}$, $280\,\mathrm{K} \leq T \leq 370\,\mathrm{K}$ and $0 \leq C_A \leq 1$ mol/l. The nominal operating conditions, which correspond to an unstable equilibrium $C_A^{eq} = 0.5$ mol/l, $T^{eq} = 350$ K, $T_c^{eq} = 300$ K are: $q = 100$ l/ min, $C_{Af} = 1$ mol/l, $T_f = 350$ K, $V = 100$ l, $\rho = 1000$ g/l, $C_p = 0.239$ J/g K, $\Delta H = -5 \times 10^4$ J/mol, $E/R = 8750$ K, $k_0 = 7.2 \times 10^{10}\,\mathrm{min}^{-1}$, $UA = 5 \times 10^4$J/ min K. The corresponding nonlinear discrete-time state-space model is obtained by defining the state vector $\mathbf{x} = [C_A - C_A^{eq} \quad (T - T^{eq})/100]^T$, the manipulated input $u = (T_c - T_c^{eq})/100$ and by discretizing the ODE with a sampling time $\triangle t = 0.03$ min. The purpose of adding the scaling factor for the temperature is to make the value of $C_A - C_A^{eq}$ and $(T - T^{eq})/100$ comparable and thus the error tolerance for the bisections of them could be the same. A local stabilizing linear feedback control law $u = [-0.0690\, -4.3387]\mathbf{x}$ is designed in advance according to the linearized model and the LQ method discussed in [10]. With the designed local stabilizing linear feedback control law, an optimal control invariant zonotope $\mathcal{Z} = \begin{bmatrix} 0.03 & -0.01 & 0.02 & 0 \\ 0.01 & 0.01 & 0 & 0.02 \end{bmatrix} \mathbf{B}^4$ is obtained through the solver of global optimization for set inversion via zonotope geometry, where the initial interval matrix for the search is selected to be $\mathbf{M} = \begin{bmatrix} [0, 0.04] & [-0.04, 0] \\ [0, 0.04] & [0, 0.04] \end{bmatrix}$; the bound of error tolerance for global optimization is selected to be $\varepsilon = 0.005$ while the bound of error tolerance for set inversion is selected to be $\varepsilon = 0.05$. The overall computation time for the simulation using MATLAB is 2244.7 seconds on a Pentium Centrino 1.4 GHz Notebook. The optimal control invariant zonotope along with its bisection is shown in Fig. 5, where the coordinate system is transformed to be original so as to compare with the control invariant ellipse designed in [10]. It is worthy to note that the volume of the designed control invariant zonotope is 0.8, which is much bigger than the control invariant ellipse (Its volume is 0.1606) designed in [10]. The obtained zonotope can be demonstrated geometrically to be control invariant using the solver of set inversion via zonotope geometry, where the evolutions of the sub-zonotopes under the related local stabilizing feedback control law are within the original zonotope, just as shown in Fig. 6. The designed control invariant zonotope and the related local stabilizing linear feedback control law can be utilized as the terminal set and the terminal control law in the terminal control of the dual-mode MPC, where the control target is to drive the system state into the terminal set using receding horizon control

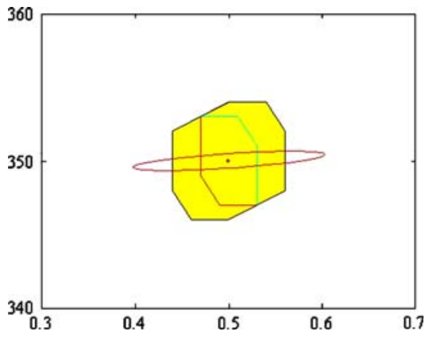


Fig. 5 The bisection of the optimal zonotope

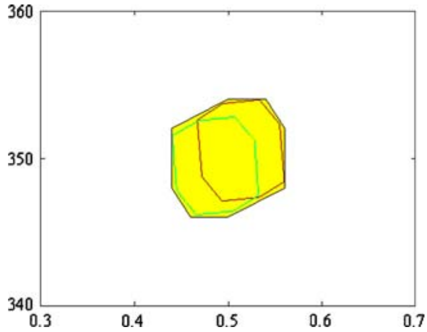


Fig. 6 The evolution of each sub-zonotope of the optimal zonotope

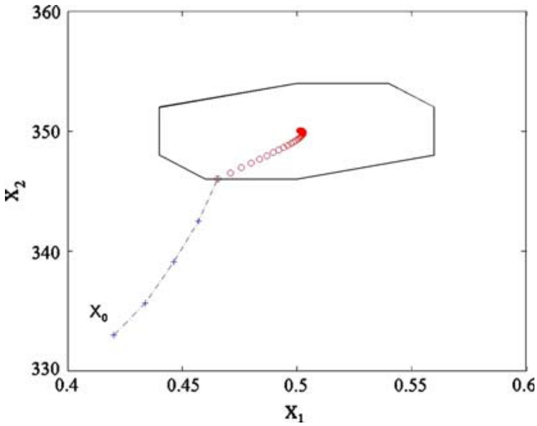


Fig. 7 The control process of the dual-mode MPC

strategy when the system state is outside the terminal set and the related local stabilizing linear feedback control law is applied instead once the system state enters into the terminal set, just as demonstrated in Fig. 7 for the initial state $\mathbf{x}_0 = [0.42 \;\; 333]^T$.

8 Conclusions

Zonotopes are centrally symmetric convex polytopes with implicit representations of line segment generators. The solver of set inversion via interval analysis has been extended to set inversion via zonotope geometry using the proposed method of bisecting zonotopes. The extended solver is further combined with the solver of global optimization via interval analysis to design control invariant sets for constrained nonlinear discrete-time systems in a numerical way. Using zonotopes, numerical tools for nonlinear systems such as interval analysis and numerical tools for linear systems such as polytope geometry are unified in the same framework of convex sets and it is anticipated that the proposed generalization from interval analysis to zonotope geometry could be helpful for more practical problems than the design of control invariant sets for constrained nonlinear discrete-time systems.

Acknowledgements This work was partially funded by the European Union and the Spanish government through the coordinated research projects DPI2004-07167-C02-02 and DPI2005-08668-C03-02 and by the government of Catalonia through SGR00296. The authors are grateful for the uncommercial Matlab Toolboxs of INTerval LABoratory (INTLAB) by Prof. Dr. Siegfried M. Rump and Multi-Parametric Toolbox (MPT) by

M. Kvasnica, P. Grieder and M. Baotic used for the simulations of the paper. The authors are also very thankful to the anonymous reviewers for their pertinent comments and valuable suggestions, which are significant for correcting and improving the paper further.

References

1. Blanchini, F.: Set invariance in control. Automatica J. IFAC **35**(11), 1747–1767 (1999)
2. Rakovic, S.V., Kerrigan, E.C., Kouramas, K.I., Mayne, D.Q.: Invariant approximations of the minimal robust positively invariant set. IEEE Trans. Automat. Control **50**(3), 406–410 (2005)
3. Mayne, D.Q., Rawlings, J.B., Rao, C.V., Scokaert, P.O.M.: Constrained model predictive control: stability and optimality. Automatica J. IFAC **26**(6), 789–814 (2000)
4. Kerrigan, E.C.: Robust constraint satisfaction: invariant sets and predictive control. Ph.D. thesis, University of Cambridge (2000)
5. Cannon, M., Deshmukh, V., Kouvaritakis, B.: Nonlinear model predictive control with polytopic invariant sets. Automatica J. IFAC **39**(8), 1487–1494 (2003)
6. Bacic, M., Cannon, M., Kouvaritakis, B.: Invariant sets for feedback linearisation based nonlinear predictive control. IEE Proc. Control Theory Appl. **152**(3), 259–265 (2005)
7. Michalska, H., Mayne, D.Q.: Robust receding horizon control of constrained nonlinear systems. IEEE Trans. Automat. Control **38**(11), 1623–1633 (1993)
8. Chen, H., Allgower, F.: A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability. Automatica J. IFAC **34**(10), 1205–1218 (1998)
9. Kothare, M.V., Balakrishnan, V., Morari, M.: Robust constrained model predictive control using Linear Matrix Inequalities. Automatica J. IFAC **32**(10), 1361–1379 (1996)
10. Magni, L., De Nicolao, G., Magnani, L., Scattolini, R.: A stabilizing model-based predictive control algorithm for nonlinear systems. Automatica **37**(9), 1351–1362 (2001)
11. Jaulin, L., Kieffer, M., Didrit, O., Walter, E.: Applied Interval Analysis: with Examples in Parameter and State Estimation, Robust Control and Robotics. Springer, London (2001)
12. Jaulin, L., Walter, E.: Set inversion via interval analysis for nonlinear bounded-error estimation. Automatica J. IFAC **29**(4), 1053–1064 (1993)
13. Walter, E., Jaulin, L.: Guaranteed characterization of stability domains via set inversion. IEEE Trans. Automat. Control **39**(4), 886–889 (1994)
14. Meizel, D., Leveque, O., Jaulin, L., Walter, E.: Initial localization by set inversion. IEEE Trans. Robotics Automat. **18**(6), 966–971 (2002)
15. Kühn, W.: Rigorously computed orbits of dynamical systems without the wrapping effect. Computing **61**(1), 47–67 (1998)
16. Hansen, E.: Global Optimization Using Interval Analysis. Marcel Dekker, New York (1992)
17. Alamo, T., Bravo, J.M., Camacho, E.F.: Guaranteed state estimation by zonotopes. Automatica J. IFAC **41**(6), 1035–1043 (2005)
18. Combastel, C.: A state bounding observer for uncertain non-linear continuous-time systems based on zonotopes. In: IEEE ECC-CDC, Seville, Spain (2005)
19. Stancu, A., Puig, V., Cuguero, P., Quevedo, J.: Benchmarking on approaches to interval observation applied to robust fault detection. In: European Control Conference, Cambridge, UK (2003)
20. Ong, C.J., Gilbert, E.G.: The minimal disturbance invariant set: Outer approximations via its partial sums. Automatica J. IFAC **42**(9), 1563–1568 (2006)
21. Fukuda, K.: From the zonotope construction to the Minkowski addition of convex polytopes. J. Symb. Comput. **38**(4), 1261–1272 (2004)
22. Kvasnica, M., Grieder, P., Baotic, M.: Multi-Parametric Toolbox (MPT), http://control.ee.ethz.ch/~mpt/. Cited July 2007
23. Ingimundarson, A., Bravo, J. M., Puig, V., Alamo, T.: Robust fault diagnosis using parallelotope-based set-membership consistency tests. In: IEEE ECC-CDC, Seville, Spain (2005)
24. Lasserre, J.B.: An analytical expression and an algorithm for the volume of a convex polyhedron in R^n. J. Optim. Theory Appl. **39**(3), 363–377 (1983)