Automatica 69 (2016) 126–136

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Constrained zonotopes: A new tool for set-based estimation and fault detection☆

Joseph K. Scott [a], Davide M. Raimondo [b], Giuseppe Roberto Marseglia [b], Richard D. Braatz [c]

[a] Department of Chemical and Biomolecular Engineering, Clemson University, Clemson, SC, USA
[b] Identification and Control of Dynamic Systems Laboratory, University of Pavia, Italy
[c] Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

ARTICLE INFO

Article history:
Received 25 March 2015
Received in revised form
21 December 2015
Accepted 15 February 2016

Keywords:
State estimation
Fault detection
Set-based computing
Zonotopes
Reachability analysis

ABSTRACT

This article introduces a new class of sets, called *constrained zonotopes*, that can be used to enclose sets of interest for estimation and control. The numerical representation of these sets is sufficient to describe arbitrary convex polytopes when the complexity of the representation is not limited. At the same time, this representation permits the computation of exact projections, intersections, and Minkowski sums using very simple identities. Efficient and accurate methods for computing an enclosure of one constrained zonotope by another of lower complexity are provided. The advantages and disadvantages of these sets are discussed in comparison to ellipsoids, parallelotopes, zonotopes, and convex polytopes in halfspace and vertex representations. Moreover, extensive numerical comparisons demonstrate significant advantages over other classes of sets in the context of set-based state estimation and fault detection.

1. Introduction

Many modern control algorithms make use of sets (e.g., intervals, ellipsoids, zonotopes, polytopes) as basic computational objects, with the aim of characterizing some sets of interest, such as reachable or invariant sets of dynamical systems, or sets of states or parameters consistent with a bounded-error model (Althoff, Stursberg, & Buss, 2010; Ingimundarson, Bravo, Puig, Alamo, & Guerra, 2009; Le, Stoica, Alamo, Camacho, & Dumur, 2013; Mayne, Rakovic, Findeisen, & Allgower, 2006; Scott & Barton, 2013). The true set of interest is often difficult or impossible to represent exactly with finite data, so its enclosure by an element of a class of simpler sets is sought instead. The choice of class for a given application is based on a tradeoff between (i) the accuracy with which a member of the class can represent the set of interest, and (ii) the complexity of the required computations. For linear estimation and control problems, the required computations typically involve standard set operations such as Minkowski sums, linear mappings, intersections, and Pontryagin differences. Another important consideration in (ii) is the convenience of the approximating set for its end use, which may involve checking for the inclusion of given points, as in model invalidation and fault diagnosis (Rosa, Silvestre, Shamma, & Athans, 2010), checking for intersection with another set, as in system verification and safety analysis (Althoff et al., 2010), or using the set as a constraint in an optimization problem, as in robust optimal control and active fault diagnosis (Mayne et al., 2006; Raimondo, Marseglia, Braatz, & Scott, 2013).

This article introduces a new class of sets, *constrained zonotopes*, and demonstrates that this class provides a better tradeoff between accuracy and efficiency than existing classes for some representative problems of interest. Although these new sets potentially have broad applicability, their performance is demonstrated here by considering the classical set-based state estimation problem for discrete-time linear systems with bounded noise (Schweppe, 1968), and its application to set-based fault diagnosis (Scott, Findeisen, Braatz, & Raimondo, 2014). Some notable advantages of the constrained zonotope representation are:

- (*Accuracy*) When the complexity of the representation is not limited, it can describe arbitrary convex polytopes;
- (*Efficiency*) Standard set operations, including intersections, can be computed exactly through simple identities;
- (*Tunability*) Effective techniques are provided to conservatively reduce the complexity of a given set, enabling a highly tunable tradeoff between efficiency and accuracy.

☆ BP is acknowledged for financial support of this project. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Mario Sznaier under the direction of Editor Richard Middleton.

E-mail addresses: jks9@clemson.edu (J.K. Scott), davide.raimondo@unipv.it (D.M. Raimondo), groberto.marseglia@gmail.com (G.R. Marseglia), braatz@mit.edu (R.D. Braatz).

http://dx.doi.org/10.1016/j.automatica.2016.02.036
0005-1098/

To motivate this new class of sets, common set representations are reviewed in Section 2 and their advantages and disadvantages are discussed with respect to common set operations. Constrained zonotopes are introduced in Section 3, and associated computations are described in Sections 3.1–4. Numerical results are presented in Sections 5–6, and Section 7 concludes the paper.

2. Set representations and operations

Definition 1. Let $P, Z, E \subset \mathbb{R}^n$. P is a *convex polytope* if it is bounded and (1) holds; Z is a *zonotope* if (2) holds, and E is an *ellipsoid* if (3) holds:

$$\exists(\mathbf{H}, \mathbf{k}) \in \mathbb{R}^{n_h \times n} \times \mathbb{R}^{n_h} : \quad P = \{\mathbf{z} \in \mathbb{R}^n : \mathbf{H}\mathbf{z} \leq \mathbf{k}\}, \tag{1}$$

$$\exists(\mathbf{G}, \mathbf{c}) \in \mathbb{R}^{n \times n_g} \times \mathbb{R}^n \quad : \quad Z = \{\mathbf{G}\boldsymbol{\xi} + \mathbf{c} : \|\boldsymbol{\xi}\|_\infty \leq 1\}, \tag{2}$$

$$\exists(\mathbf{Q}, \mathbf{c}) \in \mathbb{R}^{n \times n} \times \mathbb{R}^n \quad : \quad E = \{\mathbf{Q}\boldsymbol{\xi} + \mathbf{c} : \|\boldsymbol{\xi}\|_2 \leq 1\}. \tag{3}$$

Z is a *parallelotope* if (2) holds with $n_g = n$ and an *interval* if (2) holds with $\mathbf{G} = \mathbf{I}_{n \times n}$.

Eq. (1) is called the *halfspace-representation* (H-rep) of P. P can also be represented as the convex hull of its vertices (V-rep). Zonotopes are convex polytopes that are *centrally symmetric*; every chord through $\mathbf{c}$ is bisected by $\mathbf{c}$. Moreover, a convex polytope is a zonotope if and only if every 2-face is centrally symmetric (McMullen, 1971). This symmetry makes the representation (2) possible. The vector $\mathbf{c}$ is called the *center*, the columns of $\mathbf{G}$ are called the *generators*, and (2) is called the *generator-representation* (G-rep). The G-rep of a zonotope is often much more compact than the equivalent H- or V-rep. Both zonotopes and ellipsoids are affine images of a unit ball. However, zonotopes use the ∞-norm and n_g need not equal n. The representation (3) captures degenerate ellipsoids when $\mathbf{Q}$ is singular and is equivalent to the familiar form $E = \{\mathbf{z} : (\mathbf{z} - \mathbf{c})^{\mathrm{T}}(\mathbf{Q}\mathbf{Q}^{\mathrm{T}})^{-1}(\mathbf{z} - \mathbf{c}) \leq 1\}$ whenever $\mathbf{Q}$ is invertible.

Note that intervals, parallelotopes, and ellipsoids all have fixed complexity for fixed n. In contrast, convex polytopes and zonotopes can be made arbitrarily complex by increasing the number of halfspaces and generators, respectively, which makes these sets more flexible, but also more cumbersome. The complexity of a zonotope is described by its *order*, n_g/n.

For the estimation and fault diagnosis problems considered in Sections 5–6, as well as many other problems in linear control theory, the accuracy and efficiency of the below set operations are of primary concern:

Definition 2. Let $Z, W \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^k$, $\mathbf{R} \in \mathbb{R}^{k \times n}$, and define

$$\mathbf{R}Z \equiv \{\mathbf{R}\mathbf{z} : \mathbf{z} \in Z\}, \tag{4}$$

$$Z + W \equiv \{\mathbf{z} + \mathbf{w} : \mathbf{z} \in Z,\ \mathbf{w} \in W\}, \tag{5}$$

$$Z \cap_{\mathbf{R}} Y \equiv \{\mathbf{z} \in Z : \mathbf{R}\mathbf{z} \in Y\}. \tag{6}$$

Eq. (4) is a linear mapping of Z, (5) is the *Minkowski sum*, and (6) is a generalized intersection that arises in state estimation (e.g., with Z containing the current state and Y a bounded-error measurement; see Section 5). Note that $\cap_{\mathbf{R}}$ is the standard intersection when $k = n$ and $\mathbf{R} = \mathbf{I}$.

A class of sets is *closed* under a set operation if performing the operation on members of the class results in another member of the class. The convex polytopes are closed under (4)–(6) and, using H-rep, both (4) and (6) can be computed efficiently if $\mathbf{R}$ is invertible. However, the complexity of (5) is exponential in n, as is the worst-case number of halfspaces describing $Z + W$ (Hagemann, 2015; Tiwary, 2008). The same is true of (4) and (6) when $\mathbf{R}$ is not invertible (e.g., polytope projection) (Jones, Kerrigan, & Maciejowski, 2008). In V-rep, (4)–(5) are much simpler, but (6) is NP-hard (Tiwary, 2008), and existing algorithms for interconversion between H- and V-rep have worst-case exponential run-time. Consequently, working with convex polytopes is very costly and numerically unstable when n exceeds about 5 or the number of halfspaces or vertices is large.

In contrast, intervals, parallelotopes, and ellipsoids all provide low-complexity set representations and relatively low-cost set operations. However, the intervals are not closed under (4) unless $\mathbf{R}$ is diagonal, the parallelotopes and ellipsoids are not closed under (5), and none of these classes are closed under (6) except intervals when $\mathbf{R}$ is diagonal. Thus, the results of these operations must be conservatively enclosed, which can ultimately lead to very inaccurate enclosures of the set of interest. The optimal interval enclosures of these operations are easily computed (Neumaier, 1990), but are often very weak enclosures of the true sets. For ellipsoids, cheap heuristic enclosure methods are given in Schweppe (1968). Optimal enclosures are given in Chernousko (1980), Durieu, Walter, and Polyak (2001) and Fogel and Huang (1982), but (6) requires the solution of a convex optimization when $k > 1$. Cheap heuristic enclosures for parallelotopes are given in Chisci, Garulli, and Zappa (1996), and numerical results there show that these are tighter than even the optimal ellipsoidal enclosures in the context of state estimation.

Over the past decade, zonotopes have gained popularity within the control community, particularly because (4)–(5) can be computed exactly and efficiently in G-rep (Kuhn, 1998). Define the shorthand $Z = \{\mathbf{G}, \mathbf{c}\} \subset \mathbb{R}^n$ for Z defined by (2). Then, with $Z = \{\mathbf{G}_z, \mathbf{c}_z\}$ and $W = \{\mathbf{G}_w, \mathbf{c}_w\}$,

$$\mathbf{R}Z = \{\mathbf{R}\mathbf{G}_z, \mathbf{R}\mathbf{c}_z\}, \tag{7}$$

$$Z + W = \{[\mathbf{G}_z\ \mathbf{G}_w], \mathbf{c}_z + \mathbf{c}_w\}. \tag{8}$$

Clearly, these computations can be done efficiently and reliably, even in high dimensions. Like general convex polytopes, these operations are nonconservative, but lead to an increase in the complexity of the set representation. However, in contrast to the worst-case exponential increase in the size of the H-rep under (4)–(5), the increase in the complexity of the G-rep is modest; $\mathbf{R}Z$ has the same n_g as Z, while the n_g of $Z + W$ is simply the sum of the n_g's of Z and W. Moreover, conservative *order reduction* techniques are available that enclose a given zonotope within a zonotope of lower order (Althoff et al., 2010; Combastel, 2003). Similar techniques have also been proposed for convex polytopes, but the required computations are much more complex (Hagemann, 2015). For zonotopes, these techniques provide a tunable mechanism for balancing accuracy and complexity that has proven to be effective in reachability analysis (Althoff et al., 2010; Kuhn, 1998), identification (Bravo, Alamo, & Camacho, 2006), state estimation (Alamo, Bravo, & Camacho, 2005), and fault detection (Ingimundarson et al., 2009; Scott et al., 2014).

However, zonotopes are not closed under intersection, and tight enclosures are difficult to compute, which leads to serious complications in many applications, such as state estimation and hybrid systems verification (Althoff & Krogh, 2011; Bravo et al., 2006). Indeed, the symmetry of zonotopes, as well as intervals, parallelotopes, and ellipsoids, implies that they cannot accurately represent sets that are strongly centrally asymmetric, which are readily generated by (6). This has led some researchers to use a combination of G- and H-rep, although the conversion from G- to H-rep can be costly; it scales as $n_g\binom{n_g}{n-1}$ (Althoff & Krogh, 2011).

Set representations based on collections of sets have also been proposed, such as unions of intervals (Neumaier, 1990) and intersections of ellipsoids (Kurzhanski, 2011) or zonotopes (Althoff & Krogh, 2011). These can be very accurate, but the associated cost increases with the number of sets required, which can be large. We restrict the scope of the comparisons herein to 'single-set' representations.

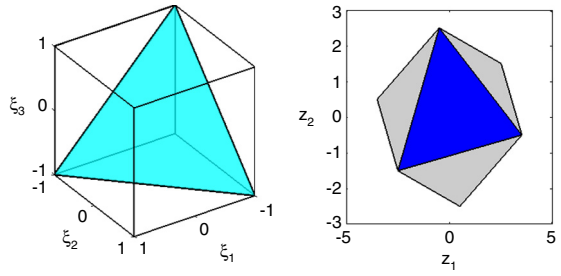


Fig. 1. Left: The unit hypercube B_∞ restricted by $\mathbf{1}^T\boldsymbol{\xi} = -1$. Right: The constrained zonotope $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$ defined in (10) (blue) overlaid on the unconstrained zonotope $Z' = \{\mathbf{G}, \mathbf{c}\}$ (gray). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Constrained zonotopes

Motivated by the above discussion, we define *constrained zonotopes*.

Definition 3. A set $Z \subset \mathbb{R}^n$ is a *constrained zonotope* if there exists $(\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}) \in \mathbb{R}^{n\times n_g} \times \mathbb{R}^n \times \mathbb{R}^{n_c\times n_g} \times \mathbb{R}^{n_c}$ such that

$$Z = \{\mathbf{G}\boldsymbol{\xi} + \mathbf{c} : \|\boldsymbol{\xi}\|_\infty \le 1,\ \mathbf{A}\boldsymbol{\xi} = \mathbf{b}\}. \tag{9}$$

In contrast to standard zonotopes, Definition 3 permits linear equality constraints on $\boldsymbol{\xi}$. Denote the unit hypercube in $\mathbb{R}^{n_g}$ by B_∞ and define $B_\infty(\mathbf{A}, \mathbf{b}) \equiv \{\boldsymbol{\xi} \in B_\infty : \mathbf{A}\boldsymbol{\xi} = \mathbf{b}\}$. Thus, a set is a constrained zonotope iff it is the image of some linearly constrained unit hypercube $B_\infty(\mathbf{A}, \mathbf{b})$ under an affine mapping. We call (9) the *constrained generator representation* (CG-rep) and introduce the shorthand $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\} \subset \mathbb{R}^n$.

Clearly, every zonotope is a constrained zonotope. A constrained zonotope that is not a zonotope is

$$Z = \left\{ \begin{bmatrix} 1.5 & -1.5 & 0.5 \\ 1 & 0.5 & -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, -1 \right\}. \tag{10}$$

Fig. 1 shows that, due to the constraint $\mathbf{1}^T\boldsymbol{\xi} = -1$, Z is not centrally symmetric and hence not a zonotope. Thus, constrained zonotopes are more flexible than zonotopes. Indeed, we will shortly prove that $Z \subset \mathbb{R}^n$ is a constrained zonotope iff it is a convex polytope (Theorem 1). Thus, the novelty of the constrained zonotopes can be stated in two ways:

- When n_g and n_c are not limited, the CG-rep provides a new representation of convex polytopes that will be shown to confer many of the computational advantages of zonotopes to this larger class of sets;
- With n_g and n_c limited, the CG-rep describes a new class of sets that significantly extends the zonotopes of order n_g/n while maintaining computational efficiency, which is possible due to the reduction methods developed in Section 4.

3.1. Basic set operations with constrained zonotopes

This subsection shows that constrained zonotopes are closed under (4)–(6), and that these operations can be accomplished through simple identities that follow almost immediately from Definition 3. Thus, the main contribution of Section 3 is not the simple matter of formulating these identities, but rather the definition of the CG-rep itself, and the observation that it is simultaneously flexible (Theorem 1) and easily propagated through basic set operations.

Proposition 1. *For every* $Z = \{\mathbf{G}_z, \mathbf{c}_z, \mathbf{A}_z, \mathbf{b}_z\} \subset \mathbb{R}^n$, $W = \{\mathbf{G}_w, \mathbf{c}_w, \mathbf{A}_w, \mathbf{b}_w\} \subset \mathbb{R}^n$, $Y = \{\mathbf{G}_y, \mathbf{c}_y, \mathbf{A}_y, \mathbf{b}_y\} \subset \mathbb{R}^k$, *and* $\mathbf{R} \in \mathbb{R}^{k\times n}$, *the three identities hold:*

$$\mathbf{R}Z = \{\mathbf{R}\mathbf{G}_z, \mathbf{R}\mathbf{c}_z, \mathbf{A}_z, \mathbf{b}_z\}, \tag{11}$$

$$Z + W = \left\{ [\mathbf{G}_z\ \mathbf{G}_w], \mathbf{c}_z + \mathbf{c}_w, \begin{bmatrix} \mathbf{A}_z & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_w \end{bmatrix}, \begin{bmatrix} \mathbf{b}_z \\ \mathbf{b}_w \end{bmatrix} \right\}, \tag{12}$$

$$Z \cap_{\mathbf{R}} Y = \left\{ [\mathbf{G}_z\ \mathbf{0}], \mathbf{c}_z, \begin{bmatrix} \mathbf{A}_z & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_y \\ \mathbf{R}\mathbf{G}_z & -\mathbf{G}_y \end{bmatrix}, \begin{bmatrix} \mathbf{b}_z \\ \mathbf{b}_y \\ \mathbf{c}_y - \mathbf{R}\mathbf{c}_z \end{bmatrix} \right\}. \tag{13}$$

Proof. Let Z_R be the right-hand side of (11). For any $\mathbf{z} \in Z$, $\exists\boldsymbol{\xi} \in B_\infty(\mathbf{A}_z, \mathbf{b}_z)$ such that $\mathbf{z} = \mathbf{G}_z\boldsymbol{\xi} + \mathbf{c}_z$, and hence $\mathbf{R}\mathbf{z} = \mathbf{R}\mathbf{G}_z\boldsymbol{\xi} + \mathbf{R}\mathbf{c}_z$. By the definition of Z_R, this implies that $\mathbf{R}\mathbf{z} \in Z_R$, and since $\mathbf{z}$ is arbitrary, $\mathbf{R}Z \subset Z_R$. Conversely, for any $\mathbf{r} \in Z_R$, $\exists\boldsymbol{\xi} \in B_\infty(\mathbf{A}_z, \mathbf{b}_z)$ such that $\mathbf{r} = \mathbf{R}(\mathbf{G}_z\boldsymbol{\xi} + \mathbf{c}_z)$. It follows that $\exists\mathbf{z} \in Z$ with $\mathbf{r} = \mathbf{R}\mathbf{z}$. Thus, $\mathbf{r} \in \mathbf{R}Z$, and since $\mathbf{r}$ is arbitrary, $Z_R \subset \mathbf{R}Z$. We conclude that $Z_R = \mathbf{R}Z$.

Let X denote the right-hand side of (12) and choose any $\mathbf{z} \in Z$ and $\mathbf{w} \in W$. Then

$$\exists\boldsymbol{\xi} \in B_\infty(\mathbf{A}_z, \mathbf{b}_z)\ :\ \mathbf{z} = \mathbf{G}_z\boldsymbol{\xi} + \mathbf{c}_z, \tag{14}$$

$$\exists\boldsymbol{\delta} \in B_\infty(\mathbf{A}_w, \mathbf{b}_w)\ :\ \mathbf{w} = \mathbf{G}_w\boldsymbol{\delta} + \mathbf{c}_w. \tag{15}$$

Letting $\boldsymbol{\gamma} = (\boldsymbol{\xi}, \boldsymbol{\delta})$, this implies that $\|\boldsymbol{\gamma}\|_\infty \le 1$ and

$$\begin{bmatrix} \mathbf{A}_z & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_w \end{bmatrix} \boldsymbol{\gamma} = \begin{bmatrix} \mathbf{b}_z \\ \mathbf{b}_w \end{bmatrix}, \tag{16}$$

$$\mathbf{z} + \mathbf{w} = [\mathbf{G}_z\ \mathbf{G}_w]\boldsymbol{\gamma} + (\mathbf{c}_z + \mathbf{c}_w). \tag{17}$$

Thus, $\mathbf{z} + \mathbf{w} \in X$ and $Z + W \subset X$. Conversely, choose any $\mathbf{x} \in X$. Then $\exists\boldsymbol{\gamma}$ such that $\|\boldsymbol{\gamma}\|_\infty \le 1$, (16) holds, and $\mathbf{x} = [\mathbf{G}_z\ \mathbf{G}_w]\boldsymbol{\gamma} + (\mathbf{c}_z + \mathbf{c}_w)$. Again, letting $\boldsymbol{\gamma} = (\boldsymbol{\xi}, \boldsymbol{\delta})$ shows that there exist $\mathbf{z} \in Z$ and $\mathbf{w} \in W$ such that $\mathbf{x} = \mathbf{z} + \mathbf{w}$. Thus, $\mathbf{x} \in Z + W$ and $X \subset Z + W$.

Let $\tilde{Z}$ denote the right-hand side of (13) and choose any $\mathbf{z} \in Z \cap_{\mathbf{R}} Y$. Then $\exists\boldsymbol{\xi} \in B_\infty(\mathbf{A}_z, \mathbf{b}_z)$ such that $\mathbf{z} = \mathbf{G}_z\boldsymbol{\xi} + \mathbf{c}_z$, and $\mathbf{R}\mathbf{z} = \mathbf{R}\mathbf{G}_z\boldsymbol{\xi} + \mathbf{R}\mathbf{c}_z \in Y$. By this last condition, $\exists\boldsymbol{\gamma} \in B_\infty(\mathbf{A}_y, \mathbf{b}_y)$ such that $\mathbf{R}\mathbf{G}_z\boldsymbol{\xi} + \mathbf{R}\mathbf{c}_z = \mathbf{G}_y\boldsymbol{\gamma} + \mathbf{c}_y$. Let $\boldsymbol{\delta} = (\boldsymbol{\xi}, \boldsymbol{\gamma})$. Then $\boldsymbol{\delta} \in B_\infty$, $\mathbf{z} = [\mathbf{G}_z\ \mathbf{0}]\boldsymbol{\delta} + \mathbf{c}_z$, and

$$\begin{bmatrix} \mathbf{A}_z & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_y \\ \mathbf{R}\mathbf{G}_z & -\mathbf{G}_y \end{bmatrix} \boldsymbol{\delta} = \begin{bmatrix} \mathbf{b}_z \\ \mathbf{b}_y \\ \mathbf{c}_y - \mathbf{R}\mathbf{c}_z \end{bmatrix}. \tag{18}$$

Thus, $\mathbf{z} \in \tilde{Z}$ and $Z \cap_{\mathbf{R}} Y \subset \tilde{Z}$. Conversely, let $\mathbf{z} \in \tilde{Z}$. Then $\exists\boldsymbol{\delta} \in B_\infty$ such that $\mathbf{z} = [\mathbf{G}_z\ \mathbf{0}]\boldsymbol{\delta} + \mathbf{c}_z$ and (18) holds. Partitioning $\boldsymbol{\delta}$ as $\boldsymbol{\delta} = (\boldsymbol{\xi}, \boldsymbol{\gamma})$, it follows that $\boldsymbol{\xi} \in B_\infty(\mathbf{A}_z, \mathbf{b}_z)$, $\boldsymbol{\gamma} \in B_\infty(\mathbf{A}_y, \mathbf{b}_y)$, $\mathbf{z} = \mathbf{G}_z\boldsymbol{\xi} + \mathbf{c}_z$, and $\mathbf{R}\mathbf{G}_z\boldsymbol{\xi} - \mathbf{G}_y\boldsymbol{\gamma} = \mathbf{c}_y - \mathbf{R}\mathbf{c}_z$. These conditions imply that $\mathbf{z} \in Z$ and $\mathbf{R}\mathbf{z} = \mathbf{R}\mathbf{G}_z\boldsymbol{\xi} + \mathbf{R}\mathbf{c}_z \in Y$. Thus, $\mathbf{z} \in Z \cap_{\mathbf{R}} Y$ and $\tilde{Z} \subset Z \cap_{\mathbf{R}} Y$. □

Despite their simplicity, (11)–(13) increase the complexity of the CG-rep. This problem is more serious than with zonotopes because both n_g and n_c are increased, but considerably less serious than the growth of the H-rep discussed in Section 2. Nonetheless, accurate and efficient methods for conservatively reducing the complexity of the CG-rep are essential to computing with constrained zonotopes. These methods are developed in Section 4, and form the second major contribution of this article. We close this section by noting that constrained zonotopes can be empty, and that checking this requires the solution of a linear program (LP). Like zonotopes, checking the inclusion $\mathbf{z} \in Z$ also requires an LP.

Proposition 2. *For every* $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\} \subset \mathbb{R}^n$,

$$Z \ne \emptyset \iff \min\{\|\boldsymbol{\xi}\|_\infty : \mathbf{A}\boldsymbol{\xi} = \mathbf{b}\} \le 1, \tag{19}$$

$$\mathbf{z} \in Z \iff \min \left\{ \|\boldsymbol{\xi}\|_\infty : \begin{bmatrix} \mathbf{G} \\ \mathbf{A} \end{bmatrix} \boldsymbol{\xi} = \begin{bmatrix} \mathbf{z} - \mathbf{c} \\ \mathbf{b} \end{bmatrix} \right\} \le 1. \tag{20}$$

Proof. By Definition 3, $\mathbf{z} \in Z$ iff $\exists \boldsymbol{\xi}$ such that $\|\boldsymbol{\xi}\|_\infty \leq 1$, $\mathbf{A}\boldsymbol{\xi} = \mathbf{b}$, and $\mathbf{z} = \mathbf{G}\boldsymbol{\xi} + \mathbf{c}$. Equivalently, $\mathbf{z} \in Z$ iff $\exists \boldsymbol{\xi}$ that is feasible in the right-hand side of (20) and yields an objective value less than 1. Thus, (20) holds. Similarly, $\exists \mathbf{z} \in Z$ iff $\exists \boldsymbol{\xi}$ such that $\|\boldsymbol{\xi}\|_\infty \leq 1$ and $\mathbf{A}\boldsymbol{\xi} = \mathbf{b}$, which implies (19). □

3.2. Relation to polytopes in halfspace representation

This section proves the equivalence of constrained zonotopes and polytopes, elaborates on the advantages of CG-rep, and discusses inter-conversion between CG- and H-rep.

Theorem 1. *$Z \subset \mathbb{R}^n$ is a constrained zonotope iff it is a convex polytope.*

Proof. Clearly, every constrained zonotope is a convex polytope. To prove the converse, let $P = \{\mathbf{z} : \mathbf{H}\mathbf{z} \leq \mathbf{k}\}$ be a convex polytope. By compactness, we may choose $Z_0 = \{\mathbf{G}, \mathbf{c}\} \subset \mathbb{R}^n$ and $\boldsymbol{\sigma} \in \mathbb{R}^n$ such that $P \subset Z_0$ and $\mathbf{H}\mathbf{z} \in [\boldsymbol{\sigma}, \mathbf{k}]$, $\forall \mathbf{z} \in P$ (our use of $\subset$ includes the possibility of equality throughout). Then $P = \{\mathbf{z} \in Z_0 : \mathbf{H}\mathbf{z} \in [\boldsymbol{\sigma}, \mathbf{k}]\}$. But $[\boldsymbol{\sigma}, \mathbf{k}]$ can be written in G-rep. as $\left\{\operatorname{diag}\left(\frac{\mathbf{k}-\boldsymbol{\sigma}}{2}\right), \frac{\mathbf{k}+\boldsymbol{\sigma}}{2}\right\}$, so (13) gives

$$P = \left\{[\mathbf{G} \quad \mathbf{0}], \mathbf{c}, \left[\mathbf{H}\mathbf{G} \quad \operatorname{diag}\left(\frac{\boldsymbol{\sigma} - \mathbf{k}}{2}\right)\right], \frac{\mathbf{k} + \boldsymbol{\sigma}}{2} - \mathbf{H}\mathbf{c}\right\}. \tag{21}$$

Thus, P satisfies Definition 3. □

With the aid of slack variables, the construction (21) essentially reproduces the halfspaces defining P in the constraints of the CG-rep. Thus, there is cause for some healthy skepticism about the advantages of the CG-rep. The key difference, however, is that the constraints in (21) act on the underlying variables $\boldsymbol{\xi}$ rather than on $\mathbf{z}$, and are therefore unaffected by Minkowski sums and linear mappings. Consider, e.g., a singular linear mapping of P, $\mathbf{R}P$, which is a worst-case exponential computation in H-rep, but is easily computed in CG-rep via (11). The enabling feature of the CG-rep in this context is that the mapping $\mathbf{R}$ in (11) does not affect the underlying set $B_\infty(\mathbf{A}, \mathbf{b})$, but only changes the mapping from this set into $\mathbb{R}^n$, $\boldsymbol{\xi} \mapsto \mathbf{G}\boldsymbol{\xi} + \mathbf{c}$.

The proof of Theorem 1 shows that the conversion $H \to CG$ requires only the computation of a bounding box for P, from which Z_0 and $\boldsymbol{\sigma}$ are easily computed and (21) can be applied. The conversion $CG \to H$ rests on the below result.

Proposition 3. *Let $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$ and consider any partition $[\mathbf{A} \quad \mathbf{b}] = \begin{bmatrix}\mathbf{A}_1 & \mathbf{b}_1\\ \mathbf{A}_2 & \mathbf{b}_2\end{bmatrix}$. For every $\mathbf{z} \in \mathbb{R}^n$,*

$$\mathbf{z} \in Z \iff \begin{bmatrix}\mathbf{z}\\ \mathbf{0}\end{bmatrix} \in Z^+ \equiv \left\{\begin{bmatrix}\mathbf{G}\\ \mathbf{A}_1\end{bmatrix}, \begin{bmatrix}\mathbf{c}\\ -\mathbf{b}_1\end{bmatrix}, \mathbf{A}_2, \mathbf{b}_2\right\}. \tag{22}$$

Proof. $\mathbf{z} \in Z$ iff $\exists \boldsymbol{\xi} \in B_\infty(\mathbf{A}, \mathbf{b})$ such that $\mathbf{z} = \mathbf{G}\boldsymbol{\xi} + \mathbf{c}$, which is in turn true iff $\exists \boldsymbol{\xi} \in B_\infty(\mathbf{A}_2, \mathbf{b}_2)$ such that $\begin{bmatrix}\mathbf{z}\\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}\mathbf{G}\\ \mathbf{A}_1\end{bmatrix}\boldsymbol{\xi} + \begin{bmatrix}\mathbf{c}\\ -\mathbf{b}_1\end{bmatrix}$. □

With the trivial partition $[\mathbf{A} \quad \mathbf{b}] = [\mathbf{A}_1 \quad \mathbf{b}_1]$, Z^+ is a zonotope, which we refer to as the *lifted zonotope* for Z. This result permits some algorithms developed for standard zonotopes to be applied to constrained zonotopes. In particular, the H-rep of Z can be computed by first computing the H-rep of the lifted zonotope using, e.g., the method in Althoff et al. (2010), and observing that $[\mathbf{H}_1 \quad \mathbf{H}_2]\begin{bmatrix}\mathbf{z}\\ \mathbf{0}\end{bmatrix} \leq \mathbf{k} \iff \mathbf{H}_1\mathbf{z} \leq \mathbf{k}$. However, due to the complexity of the conversion $G \to H$ (see Section 2), $CG \to H$ scales as $n_g\binom{n_g}{n+n_c-1}$, which can become prohibitive. Thus, computing in CG-rep is most advantageous when a result in CG-rep is acceptable. This is often not problematic and can be beneficial. Consider, for example, enforcing $\mathbf{z} \in P$ in an optimization problem. In H-rep, this requires n_h linear constraints, while in CG-rep it requires n_g dummy variables $\boldsymbol{\xi}$ and $n + n_c$ linear constraints. Although the addition of variables is undesirable, the worst-case exponential increase of n_h under projections and Minkowski sums implies that the CG-rep may often require far fewer constraints.

Remark 1. Polytopes can also be represented as symbolic orthogonal projections (SOPs) (Hagemann, 2015), which are similar to constrained zonotopes with $\mathbf{G}$ always equal to the orthogonal projection onto the first n elements of $\boldsymbol{\xi}$. The operations (4)–(6) can be done efficiently on SOPs, but cause a linear increase in the set complexity. At present, the lack of efficient and accurate reduction techniques limits the utility of SOPs (Hagemann, 2015). Constrained zonotopes differ in the use of a general $\mathbf{G}$ matrix and in the canonical form of the constraints, $B_\infty(\mathbf{A}, \mathbf{b})$. The former yields a more compact representation of, e.g., affine mappings, while both play a role in enabling highly effective reduction techniques in Section 4.

Remark 2. Constrained zonotopes are also closely related to the *zonotope bundles* proposed in Althoff and Krogh (2011) and defined as $\mathcal{Z} = \cap_{i=1}^{n_b} Z_i$ with each Z_i in G-rep. Both classes of sets are closed under intersection. However, constrained zonotopes are advantageous because (4)–(5) are done conservatively with zonotope bundles, rather than exactly as in (11)–(12), and complexity reduction has not yet been extensively developed for zonotope bundles.

4. Complexity reduction for constrained zonotopes

For many operations on constrained zonotopes, the result involves more generators (n_g) and/or constraints (n_c) than the arguments. Thus, the complexity of the CG-rep increases as operations are applied. For standard zonotopes, this problem is addressed by applying reduction techniques that overapproximate a given zonotope by another with fewer generators. Here, reduction techniques are developed for constrained zonotopes. Two new complications arise: (i) reducing n_g is complicated by the presence of constraints, so zonotopic methods cannot be applied directly; (ii) a method is required for reducing n_c.

For a zonotope, the *order* is defined as $o = n_g/n$, and reduction is typically done whenever the result of some operation has o greater than a prescribed value $\hat{o}$. The complexity of a constrained zonotope is conveniently described by n_c and the *degrees-of-freedom order*, $o_d = (n_g - n_c)/n$. Note that $o_d = o$ whenever $n_c = 0$. Given a constrained zonotope Z and target values $\hat{n}_c$ and $\hat{o}_d$, reduction is performed in three steps. First, the CG-rep of Z is *rescaled* as described in Section 4.1, which does not change the set Z, but reduces the conservatism of the subsequent reductions. Next, constraints are eliminated until $n_c = \hat{n}_c$ as described in Section 4.2. The proposed method for this step eliminates one generator for each eliminated constraint, so that o_d is unchanged. Finally, o_d is reduced to $\hat{o}_d$ by eliminating generators as described in Section 4.3. The final result is a reduced constrained zonotope $\tilde{Z}$ satisfying $Z \subset \tilde{Z}$. Our aim is to make this inclusion as tight as possible.

4.1. Rescaling

The conservatism of constraint reduction (Section 4.2) can be significantly reduced by first transferring some information from the constraint data $(\mathbf{A}, \mathbf{b})$ to the generator data $(\mathbf{G}, \mathbf{c})$. To do this, the constraints are first used to tighten the bounds $\boldsymbol{\xi} \in [-\mathbf{1}, \mathbf{1}]$. For

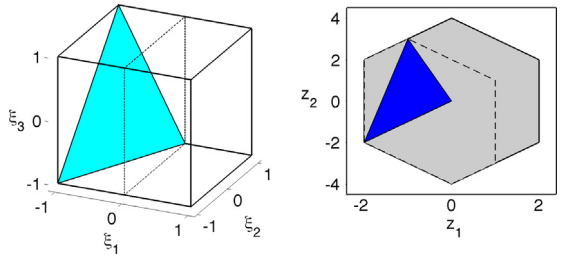


Fig. 2. Left: B_∞ restricted by the constraints in (23) (cyan) and the tightened bound $\xi_1 \leq 0$ (dashed). Right: $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$ defined in (23) (blue) overlaid on the images under $\boldsymbol{\xi} \mapsto \mathbf{G}\boldsymbol{\xi} + \mathbf{c}$ of B_∞ (gray) and $B_\infty \cap \{\boldsymbol{\xi} : \xi_1 \leq 0\}$ (dashed). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

example, consider the set

$$Z = \left\{ \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 & 1 & -1 \end{bmatrix}, 2 \right\}. \tag{23}$$

Fig. 2 shows that the standard inequality $\xi_1 \leq 1$ can be tightened to $\xi_1 \leq 0$ without modifying Z, and that this reduces the error induced by dropping the constraints $\mathbf{A}\boldsymbol{\xi} = \mathbf{b}$. Next, the CG-rep is rescaled to recover the standard form $\boldsymbol{\xi} \in [-\mathbf{1}, \mathbf{1}]$.

Proposition 4. *Let* $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$. *If* $\boldsymbol{\xi}^L, \boldsymbol{\xi}^U \in \mathbb{R}^n$ *satisfy* $B_\infty(\mathbf{A}, \mathbf{b}) \subset [\boldsymbol{\xi}^L, \boldsymbol{\xi}^U] \subset [-\mathbf{1}, \mathbf{1}]$, *then an equivalent CG-rep is*

$$Z = \left\{ \mathbf{G}\mathrm{diag}(\boldsymbol{\xi}_r), \mathbf{c} + \mathbf{G}\boldsymbol{\xi}_m, \mathbf{A}\mathrm{diag}(\boldsymbol{\xi}_r), \mathbf{b} - \mathbf{A}\boldsymbol{\xi}_m \right\}, \tag{24}$$

where $\boldsymbol{\xi}_m = \frac{1}{2}(\boldsymbol{\xi}^U + \boldsymbol{\xi}^L)$ *and* $\boldsymbol{\xi}_r = \frac{1}{2}(\boldsymbol{\xi}^U - \boldsymbol{\xi}^L)$.

Proof. Note that $\boldsymbol{\xi} \in [\boldsymbol{\xi}^L, \boldsymbol{\xi}^U]$ iff $\exists \boldsymbol{\delta} \in B_\infty$ such that $\boldsymbol{\xi} = \boldsymbol{\xi}_m + \mathrm{diag}(\boldsymbol{\xi}_r)\boldsymbol{\delta}$. Thus, the result follows from

$$\begin{aligned} \mathbf{z} \in Z &\iff \exists \boldsymbol{\xi} \in B_\infty(\mathbf{A}, \mathbf{b}) : \mathbf{z} = \mathbf{G}\boldsymbol{\xi} + \mathbf{c}, \\ &\iff \exists \boldsymbol{\xi} \in [\boldsymbol{\xi}^L, \boldsymbol{\xi}^U] : \mathbf{z} = \mathbf{G}\boldsymbol{\xi} + \mathbf{c},\ \mathbf{0} = \mathbf{A}\boldsymbol{\xi} - \mathbf{b}, \\ &\iff \exists \boldsymbol{\delta} \in B_\infty : \mathbf{z} = \mathbf{G}(\boldsymbol{\xi}_m + \mathrm{diag}(\boldsymbol{\xi}_r)\boldsymbol{\delta}) + \mathbf{c}, \\ &\qquad\qquad \mathbf{0} = \mathbf{A}(\boldsymbol{\xi}_m + \mathrm{diag}(\boldsymbol{\xi}_r)\boldsymbol{\delta}) - \mathbf{b}. \quad \square \end{aligned}$$

The process of computing the interval $[\boldsymbol{\xi}^L, \boldsymbol{\xi}^U]$ and replacing $\{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$ by (24) is termed *rescaling*. The best possible interval is given by solving the $2n_g$ LPs

$$\xi_j^L \equiv \min \left\{ \xi_j : \mathbf{A}\boldsymbol{\xi} = \mathbf{b},\ \|\boldsymbol{\xi}\|_\infty \leq 1 \right\}, \tag{25}$$

$$\xi_j^U \equiv \max \left\{ \xi_j : \mathbf{A}\boldsymbol{\xi} = \mathbf{b},\ \|\boldsymbol{\xi}\|_\infty \leq 1 \right\}. \tag{26}$$

However, this may be too expensive when rescaling is done very often. Instead, we refine the initial bounds $[-\mathbf{1}, \mathbf{1}]$ using an iterative method based on interval arithmetic with a complexity of $O(n_c n_g^2)$. See Appendix for details.

4.2. Constraint reduction

The proposed constraint reduction method requires the below proposition, which is inspired by similar results for approximating zonotope intersections in Alamo et al. (2005), Bravo et al. (2006), Combastel (2003), and Ocampo-Martinez, Guerra, Puig, and Quevedo (2007).

Proposition 5. *Let* $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$. *The set*

$$\tilde{Z} \equiv \{\mathbf{G} - \Lambda_G\mathbf{A}, \mathbf{c} + \Lambda_G\mathbf{b}, \mathbf{A} - \Lambda_A\mathbf{A}, \mathbf{b} - \Lambda_A\mathbf{b}\} \tag{27}$$

satisfies $Z \subset \tilde{Z}$ *for every* $\Lambda_G \in \mathbb{R}^{n \times n_c}$ *and* $\Lambda_A \in \mathbb{R}^{n_c \times n_c}$.

Proof. $\mathbf{z} \in Z$ iff $\exists \boldsymbol{\xi} \in B_\infty$ such that $\begin{bmatrix} \mathbf{z} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \\ \mathbf{A} \end{bmatrix} \boldsymbol{\xi} + \begin{bmatrix} \mathbf{c} \\ -\mathbf{b} \end{bmatrix}$. For any such $\boldsymbol{\xi}$, $\begin{bmatrix} \mathbf{z} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{G} \\ \mathbf{A} \end{bmatrix} \boldsymbol{\xi} + \begin{bmatrix} \mathbf{c} \\ -\mathbf{b} \end{bmatrix} + \begin{bmatrix} \Lambda_G(\mathbf{b} - \mathbf{A}\boldsymbol{\xi}) \\ \Lambda_A(\mathbf{b} - \mathbf{A}\boldsymbol{\xi}) \end{bmatrix}$, so $\mathbf{z} \in \tilde{Z}$. $\square$

If Λ_A has the ith unit vector as its ith row, then $\tilde{Z}$ has the trivial ith constraint $\mathbf{0}^T\boldsymbol{\xi} = 0$, which can be removed. We say that this constraint has been *dualized*. With $\Lambda_A = \mathbf{I}$, all constraints are dualized and $\tilde{Z}$ is a zonotope. In any case, Λ_G and any unspecified rows of Λ_A can be used to modify $\tilde{Z}$ in order to compensate for the eliminated constraints.

Below is a strategy for eliminating a single constraint at a time. This does not imply that each eliminated constraint is treated as if the others did not exist, because the remaining constraints are modified by Λ_A. After experimenting with many strategies, the most effective was found to be a heuristic we call *partial solve dualization*. The key idea is to solve one of the constraint equations, say the first, for a single ξ_j:

$$\xi_j = a_{1j}^{-1} b_j - a_{1j}^{-1} \sum_{k \neq j} a_{1k}\xi_k. \tag{28}$$

Next, (28) is used to eliminate ξ_j from the remaining constraints and the equations $\mathbf{z} = \mathbf{G}\boldsymbol{\xi} + \mathbf{c}$. Finally, the first constraint is removed. Straightforward algebra shows that this is accomplished by choosing

$$\Lambda_G \equiv \mathbf{G}\mathbf{E}_{j1}a_{1j}^{-1}, \qquad \Lambda_A \equiv \mathbf{A}\mathbf{E}_{j1}a_{1j}^{-1}, \tag{29}$$

where $\mathbf{E}_{j1} \in \mathbb{R}^{n_g \times n_c}$ is zero except for a one in the $(j, 1)$ position, and evaluating (27). This yields $\tilde{Z} = \{\tilde{\mathbf{G}}, \tilde{\mathbf{c}}, \tilde{\mathbf{A}}, \tilde{\mathbf{b}}\}$ such that $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{A}}$ have identically zero jth columns due to the elimination of $\boldsymbol{\xi}_j$, and $[\tilde{\mathbf{A}}|\tilde{\mathbf{b}}]$ has an identically zero first row, which arises from the substitution of (28) into the constraint $\mathbf{a}_1^T\boldsymbol{\xi} = b_1$ to obtain the trivial constraint $\mathbf{0}^T\boldsymbol{\xi} = 0$. Removing these columns and rows, the dualization has eliminated one constraint and one generator.

Choosing which ξ_j to eliminate is important, but our decision to use the first constraint in (28) is not. In fact, using any constraint with $a_{ij} \neq 0$ gives the same result (see (A.4)) To select j, we consider the Hausdorff error introduced by dualization, $H_j \equiv \max_{\tilde{\mathbf{z}} \in \tilde{Z}} \min_{\mathbf{z} \in Z} \|\tilde{\mathbf{z}} - \mathbf{z}\|_2$. It would be prohibitive to solve this bilevel program for each j. Thus, a major contribution of the proposed strategy is an effective method for approximating $H_1, \ldots, H_{n_g}$ with a total complexity of only $O((n_g + n_c)^3)$. See Appendix for details.

4.3. Generator reduction

For a standard zonotope, generator reduction can be done using the simple and inexpensive method in Combastel (2003), or more accurately using the method proposed in Althoff et al. (2010). In brief, each method selects a subset of $k > n$ generators that are then replaced with n generators whose Minkowski sum overestimates that of the original k. In the first method, the new n generators form an interval. In the second, they describe a parallelotope.

For a constrained zonotope $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$, the presence of the constraints prevents either of these zonotopic methods from being applied directly. Instead, we propose a *lift-then-reduce* strategy, shown schematically as:

$$Z \to Z^+ \equiv \left\{ \begin{bmatrix} \mathbf{G} \\ \mathbf{A} \end{bmatrix}, \begin{bmatrix} \mathbf{c} \\ -\mathbf{b} \end{bmatrix} \right\} \to \tilde{Z}^+ \equiv \left\{ \begin{bmatrix} \tilde{\mathbf{G}} \\ \tilde{\mathbf{A}} \end{bmatrix}, \begin{bmatrix} \mathbf{c} \\ -\mathbf{b} \end{bmatrix} \right\} \to \tilde{Z}. \tag{30}$$

First, we form the lifted zonotope Z^+ corresponding to Z (see Proposition 3). Next, Z^+ is reduced using a zonotopic method to yield $\tilde{Z}^+$, where $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{A}}$ have fewer columns than $\mathbf{G}$ and $\mathbf{A}$. Finally,

we set $\tilde{Z} = \left\{\tilde{\mathbf{G}}, \mathbf{c}, \tilde{\mathbf{A}}, \mathbf{b}\right\}$. Since $\tilde{Z}^+ \supset Z^+$, two applications of Proposition 3 show that $\tilde{Z} \supset Z$ as desired.

Our numerical experiments showed that the overestimation of Z^+ by the zonotopic method in Combastel (2003) frequently led to severe overestimation of Z. The method in Althoff et al. (2010) performed much better, but scales as $\binom{n+\kappa}{n}$, where $\kappa \leq n_g - n$ is a heuristic integer. With $\kappa = 8$, this proved to be too costly for higher dimensional experiments (e.g., $n = n_c = 10$). Thus, a new method was developed based on the results in Chisci et al. (1996) for reducing a zonotope with $n_g = n + 1$ to a parallelotope, which we have observed to have comparable performance to the method in Althoff et al. (2010) in numerical experiments. However, the complexity of reducing n_g to $n_g - k$ in our new method is only $O(n^2 n_g + k n n_g)$. See Appendix for details.

The proposed lift-then-reduce strategy has one major limitation. Because $Z^+ \subset \mathbb{R}^{n+n_c}$, it is generally not possible to reduce $\tilde{Z}^+$ to fewer than $n + n_c$ generators (i.e., a parallelotope in $\mathbb{R}^{n+n_c}$). It follows that $\tilde{Z}$ has at least $n + n_c$ generators, even though $Z \subset \mathbb{R}^n$, which reflects the additional complexity imparted by the constraints and shows that further reductions can only be made by eliminating constraints.

5. Application to set-based state estimation

This section considers the classical set-based state estimation problem for discrete-time linear systems and demonstrates the advantages of constrained zonotope computations as compared to existing methods. Consider the system

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}_w\mathbf{w}_{k-1}, \qquad \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}_v\mathbf{v}_k, \tag{31}$$

with state $\mathbf{x}_k \in \mathbb{R}^{n_x}$, output $\mathbf{y}_k \in \mathbb{R}^{n_y}$, disturbance $\mathbf{w}_{k-1} \in \mathbb{R}^{n_w}$, and measurement error $\mathbf{v}_k \in \mathbb{R}^{n_v}$. The input $\mathbf{u}_k$ is not required and is omitted for brevity. Assuming that $\mathbf{x}_0 \in X_0$ and $(\mathbf{w}_k, \mathbf{v}_k) \in W \times V$, $\forall k \in \mathbb{N}$, with X_0, W, and V compact, the objective is to compute an enclosure of the set $\hat{X}_k$ of states at each k that is consistent with (31) and a measured output sequence $\{\mathbf{y}_k\}$. Using the definitions (4)–(6), $\hat{X}_k$ is given exactly by the recursive formula

$$\hat{X}_k = (\mathbf{A}\hat{X}_{k-1} + \mathbf{B}_w W) \cap_{\mathbf{C}} (\mathbf{y}_k - \mathbf{D}_v V), \tag{32}$$

with $\hat{X}_0 = X_0 \cap_{\mathbf{C}} (\mathbf{y}_0 - \mathbf{D}_v V)$. In general, the set operations in (32) cannot be computed exactly. However, enclosures $\mathcal{O}_k \supset \hat{X}_k$ can be recursively computed using simple set representations by outer-approximating each operation in (32):

$$\mathcal{O}_k \supset (\mathbf{A}\mathcal{O}_{k-1} + \mathbf{B}_w W) \cap_{\mathbf{C}} (\mathbf{y}_k - \mathbf{D}_v V), \tag{33}$$

with $\mathcal{O}_0 \supset X_0 \cap_{\mathbf{C}} (\mathbf{y}_0 - \mathbf{D}_v V)$.

This problem was first solved using ellipsoidal computations in Bertsekas and Rhodes (1971) and Schweppe (1968). Subsequently, more accurate but less efficient algorithms were developed using minimum-volume enclosures of the basic operations on ellipsoids (Chernousko, 1980; Durieu et al., 2001; Fogel & Huang, 1982). Parallelotope computations were first applied to state estimation in Chisci et al. (1996). The resulting estimator, called the *recursive optimal bounding parallelotope* (ROBP) estimator, was shown to be superior to even the optimal ellipsoidal estimator in terms of volume, while being significantly more efficient. Set-based estimators based on convex polytopes have also been extensively developed (Blanchini & Miani, 2008; Shamma & Tu, 1999; Walter & Piet-Lahanier, 1989). However, the complexity of polytope computations severely limits these methods as shown below. Following the work of Kuhn (1998), zonotopic estimators have received significant attention as an alternative means to improve the accuracy of ROBP while maintaining efficient and scalable computations (Alamo et al., 2005; Bravo et al., 2006; Combastel, 2003; Le et al., 2013).

Here, we investigate the advantages of using constrained zonotope computations in (33). For comparison, X_0, W, and V are assumed to be parallelotopes. Given $\mathcal{O}_{k-1} \supset \hat{X}_{k-1}$ in CG-rep at time k, the proposed estimator computes $\mathcal{O}_k$ by evaluating the right-hand side of (33) exactly using Proposition 1, and subsequently reducing $\mathcal{O}_k$ to a fixed number of constraints n_c and degrees-of-freedom order o_d as described in Section 4. This estimator is denoted by $CZ(n_c, o_d)$ and is compared to the *ROBP* estimator and two zonotopic estimators that differ in their over-approximations of the intersection in (33). The first is described in Section 3.3 of Combastel (2003), and the second in Section III.B of Bravo et al. (2006), which improves the method in Alamo et al. (2005). These methods are designated by $ZCo(o)$ and $ZBr(o)$ respectively, where $\mathcal{O}_k$ is reduced to order o after each evaluation of (33). For both methods, order reduction is performed by the method described in Section 4.3 for consistency with $CZ(n_c, o_d)$. We also compare against the H-rep polytopic estimator in Shamma and Tu (1999) (*Poly*), which computes each $\hat{X}_k$ exactly but requires a polytope projection from $2n_x + n_w + n_y + n_v$ dimensions to n_x in each step. We do not compare with classical ellipsoidal methods because comparisons in Chisci et al. (1996) demonstrate the superior performance of ROBP. Recent ellipsoidal approaches that use multiple ellipsoids at each time point are also omitted, in contrast to the 'single-set' estimators considered here (Durieu et al., 2001; Kurzhanski, 2011). Finally, we do not compare against the recent extension of $ZBr(o)$ in Le et al. (2013), which boasts reduced complexity compared to Alamo et al. (2005) but provides enclosures of similar or slightly larger volume. In contrast, our concern here is the ability to compute significantly tighter enclosures at modest additional cost.

The methods described above were implemented in `MATLAB` and compared on 500 random LTI systems (31) and parallelotopes (X_0, W, V) with dimensions $d \equiv n_x = n_y = n_w = n_v = 2$ and $d = 10$. All computations with polytopes in H-rep were done using MPT (Kvasnica, Grieder, Baotić, & Morari, 2004). The matrices $\mathbf{A}$, $\mathbf{B}_w$, and $\mathbf{C}$ were generated using the `MATLAB` routine `drss` and $\mathbf{D}_v = \mathbf{I}$. To generate X_0, W, and V in G-rep, a matrix $[\mathbf{G}|\mathbf{c}]$ with normally distributed elements was generated, and each column was normalized and scaled by a random number selected uniformly from $[0, 10]$. For all systems generated, V had $\mathbf{G} = \mathbf{I}$. Accuracy was evaluated in terms of the estimator volumes and radii, normalized to those of the exact estimator computed using constrained zonotopes with no reduction. Due to the difficulty of computing minimal enclosing balls for general polytopes, the *radius* is defined here as half the length of the longest edge of the interval hull, which is easily computed by solving $2n$ LPs. Volume computations were done by converting to H-rep and using the volume routine in MPT (Kvasnica et al., 2004). Volumes are not provided for $d = 10$ because this computation became intractable.

Fig. 3 compares the accuracy of ROBP, $ZCo(5)$, $ZBr(5)$, and $CZ(0, 5)$. Note that the latter three methods use fifth-order zonotopes for all computations. ROBP is the weakest by a large margin, followed by $ZCo(5)$. Volumes and radii for $ZBr(5)$ and $CZ(0, 5)$ are comparable with a slight but consistent advantage to $CZ(0, 5)$. Since $CZ(0, 5)$ is a zonotopic method, this demonstrates that the new reduction methods of Section 4 provide a means to compute zonotopic enclosure of intersections that can reduce conservatism compared to the best available methods. Moreover, Table 1 shows that $CZ(0, 5)$ is slightly more efficient than $ZBr(5)$ with $d = 2$, and scales more favorably to higher dimensions. Although $ZCo(5)$ and ROBP are not competitive in terms of accuracy, they are cheaper than $CZ(0, 5)$ by factors of about $1/3$ for $d = 2$ and $1/10$ for $d = 10$.

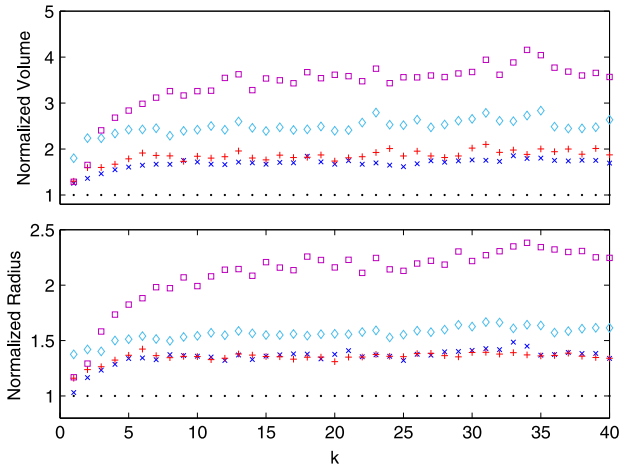


Fig. 3. Average estimator volumes and radii for *ROBP* (□), *ZCo*(5) (◇), *ZBr*(5) (+), *CZ*(0, 5) (×), and *Poly* (·) in two dimensions.

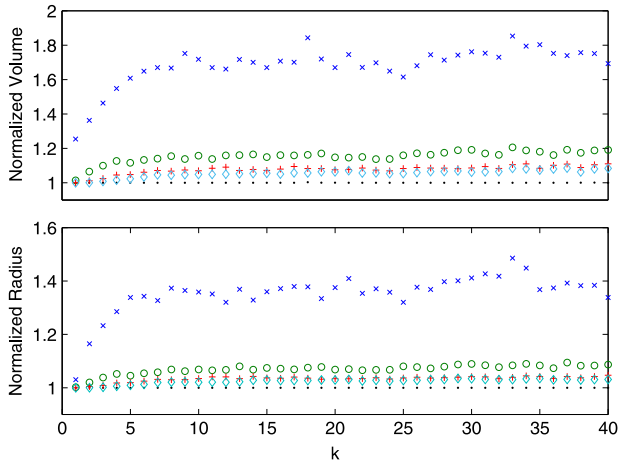


Fig. 4. Average estimator volumes and radii for *CZ*(0, 5) (×), *CZ*(1, 5) (○), *CZ*(2, 5) (+), *CZ*(3, 5) (◇), and *Poly* (·) in two dimensions.

Table 1
Average time per step in ms for state estimators applied to 2000 random systems with $n_x = n_w = n_y = n_v = d$.

d	*ROBP*	*ZCo*(5)	*ZBr*(5)	*CZ*(0, 5)
2	0.55	0.60	1.60	1.50
10	2.80	3.40	51.8	34.0
d	*CZ*(1, 5)	*CZ*(2, 5)	*CZ*(3, 5)	*Poly*
2	1.80	2.10	2.30	148.1
10	37.0	44.7	47.1	–

Fig. 4 investigates the advantage of using constraints in the CG-rep by comparing *CZ*(0, 5), *CZ*(1, 5), *CZ*(2, 5), *CZ*(3, 5), and *Poly*. A trend of increasing accuracy with increasing number of constraints is clear. Moreover, the volumes and radii for *CZ*(3, 5) are within 5% of the figures for the exact observer *Poly*. Table 1 shows that the computational cost of *CZ*(3, 5) is over 60× less than *Poly*, and only 1.5× more than *ZBr*(5) and 4× more than *ROBP* ($d = 2$).

Fig. 5 shows very similar accuracy trends for $d = 10$. *CZ*(0, 5) compares less favorably to *ZBr*(5) than for $d = 2$. But again, *CZ*(3, 5) produces radii within 5% of the exact figures at a cost of <50 ms per step. This is faster than *ZBr*(5) and about 17× slower than ROBP. At the same time, our experiments showed that *Poly* could not proceed beyond $k = 4$ within 20 min for most random systems with $d \geq 4$ due to the exponential scaling of the required projection.

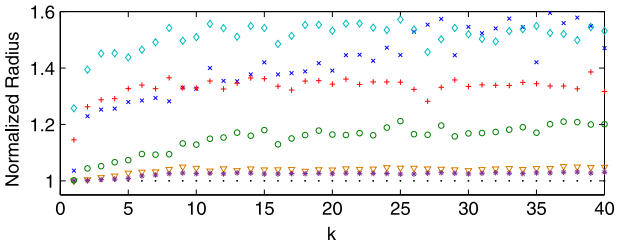


Fig. 5. Average estimator radii for *ZCo*(5) (◇), *ZBr*(5) (+), *CZ*(0, 5) (×), *CZ*(1, 5) (○), *CZ*(2, 5) (▽), *CZ*(3, 5) (∗), and *Poly* (·) in ten dimensions. *ROBP* averages ∼2.5 (not shown).

To better understand the performance of the *CZ* observers, Fig. 6 shows a single step of the computation (32). The upper left panel shows a randomly generated zonotope representing $(\mathbf{A}\hat{X}_{k-1} + \mathbf{B}_w W)$. The solid lines are bounded error measurements and the red set is the intersection in (32). The top-right and bottom-left panels show enclosures of this intersection by zonotopes, the first computed as in Bravo et al. (2006), and the second by applying Proposition 1 and subsequently eliminating all constraints. The latter enclosure is sharper, suggesting that the reduction methods in Section 4.2 can provide improved zonotopic enclosures of intersections. However, the intersection is strongly centrally asymmetric and cannot be accurately enclosed by any zonotope. In contrast, the bottom-right panel shows a sharper enclosure computed using a constrained zonotope of intermediate complexity, computed by applying Proposition 1 and subsequently eliminating one of the two constraints from the result. This demonstrates the second major advantage of *CZ* observers; because basic set operations can be done efficiently on the CG-rep., it is possible to propagate superior enclosures of intermediate complexity through further computations.

6. Application to set-based fault diagnosis

The results of Section 5 demonstrate that constrained zonotope computations can provide set-based state estimates with a significantly better compromise between accuracy and efficiency than existing methods. However, the widespread use of highly efficient ellipsoidal and parallelotopic estimators suggest that their accuracy is often sufficient, and additional cost and complexity is unwarranted. The purpose of this section is to highlight an important application, namely fault detection, in which estimation accuracy is critical, and hence constrained zonotopes offer significant new capabilities.

The use of set-based state estimators for fault detection has received much attention recently (Ingimundarson et al., 2009; Rosa et al., 2010; Scott et al., 2014; Tabatabaeipour, Odgaard, Bak, & Stoustrup, 2012; Tornil-Sin, Ocampo-Martinez, Puig, & Escobet, 2012). Consider an LTI model of the form (31) representing the nominal plant dynamics under linear feedback. The basic fault detection approach is to use a set-based estimator of the form (33), and to check the inclusion

$$\mathbf{y}_k \in \mathbf{C}(\mathbf{A}\mathcal{O}_{k-1} + \mathbf{B}_w W) + \mathbf{D}_v V \tag{34}$$

online for each k. Failure indicates that the system is no longer described by the nominal model; i.e., a fault has occurred. In this context, one set-based state estimator is superior to another if it results in a failure of (34) in fewer time-steps after the onset of a fault.

Below is a comparison of the performance of several observers for this task for a model for a low-frequency permanent-magnet

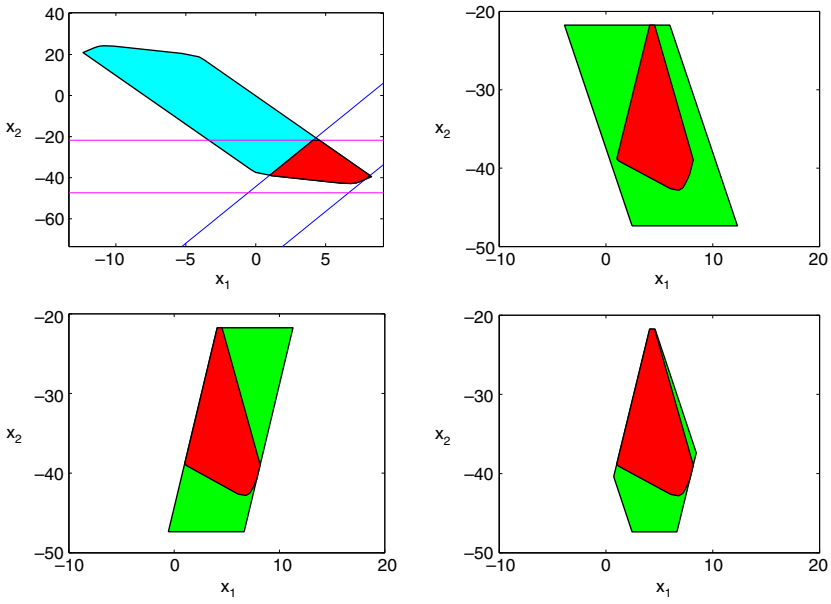


Fig. 6. Random zonotope (cyan) intersected (red) with bounded-error measurements (solid lines) (top-left), and zonotopic enclosures (green) of the intersection (red) computed as in Bravo et al. (2006) (top-right) and by using Proposition 1 and subsequently eliminating all constraints (bottom-left) or only one constraint (bottom-right) as in Section 4.2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Nominal (1) and Faulty (2) model parameters.

Model	R_a (Ω)	$L \times 10^{-3}$ (H)	$K_e \times 10^{-2}$ (V rad/s)	$J_1 \times 10^{-4}$ (N m s^2/rad)	$f_r \times 10^{-4}$ (N m s/rad)
1	1.2030	5.5840	8.5740	1.4166	2.4500
2	1.5030	5.5840	8.5740	1.4166	2.4500

DC motor (Liu, Zhang, Liu, & Yang, 2000):

$$\begin{bmatrix} \frac{di(t)}{dt} \\ \frac{dn(t)}{dt} \end{bmatrix} = \begin{bmatrix} -R_a/L & -K_e/L \\ K_t/J_1 & -f_r/J_1 \end{bmatrix} \begin{bmatrix} i(t) \\ n(t) \end{bmatrix} + \begin{bmatrix} 1/L \\ 0 \end{bmatrix} u(t)$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

where the input u is the armature voltage, the states are the current i and motor speed n, and the parameters R_a, L, K_e, K_t, J_1, and f_r are, respectively, the resistance, inductance, torque constant, back EMF constant, motor inertia, and friction coefficient. The torque constant K_t (Nm/amp) is related to K_e by $K_t = 1.0005K_e$. Table 2 gives values for one nominal and one faulty model, which differ by a 0.3 Ω increase of the armature resistance (see Liu et al., 2000).

Both models were discretized by forward Euler with a sampling interval of 1 ms to obtain models of the form:

$$\mathbf{x}_k = \mathbf{A}(i)\mathbf{x}_{k-1} + \mathbf{B}(i)\mathbf{u}_{k-1} + \mathbf{B}_w(i)\mathbf{w}_{k-1}, \tag{35}$$

$$\mathbf{y}_k = \mathbf{C}(i)\mathbf{x}_k + \mathbf{D}_v(i)\mathbf{v}_k, \tag{36}$$

where $i = 1, 2$ distinguishes the nominal and faulty models. Specifically, $\mathbf{A}(i)$, $\mathbf{B}(i)$, and $\mathbf{C}(i)$ were obtained from discretization, and the measurement and process noise terms were added with $\mathbf{D}_v(i) = \mathbf{I}$, $i = 1, 2$. The matrices $\mathbf{B}_w(i)$ were obtained assuming 5% uncertainty in R_a, K_e, J_1, and f_r, and computing the worst-case additive error when the current and the motor speed are bounded in, respectively, $[-2, 2]$ V and $[-150, 150]$ rad/s:

$$\mathbf{B}_w(1) = \begin{bmatrix} -0.0085 & -0.0006 \\ -0.0603 & 0.0002 \end{bmatrix},$$

$$\mathbf{B}_w(2) = \begin{bmatrix} -0.0101 & -0.0006 \\ -0.0595 & 0.0002 \end{bmatrix}.$$

We assume $\mathbf{x}_0 \in X_0 \equiv \left\{ \begin{bmatrix} 0.06 & 0 \\ 0 & 0.6 \end{bmatrix}, \begin{bmatrix} 0.6 \\ 70 \end{bmatrix} \right\}$, $\mathbf{w}_k \in W \equiv \{\mathbf{I}, \mathbf{0}\}$, and $\mathbf{v}_k \in V \equiv \left\{ \begin{bmatrix} 0.06 & 0 \\ 0 & 0.6 \end{bmatrix}, \mathbf{0} \right\}$, $\forall k \in \mathbb{N}$. Finally, we apply $u_k = u_N - \mathbf{K}(\mathbf{y}_k - \mathbf{x}_N)$ with saturation limits $u_k \in [0, 12]$ V, where $u_N = 6$ V maintains the nominal steady-state $\mathbf{x}_N = (0.2$ A, 70.3 rad/s) and $\mathbf{K}$ is the LQ gain with $\mathbf{Q} = \mathbf{I}$ and $\mathbf{R} = 0.1$.

To compare the fault detection capabilities of the set-based state estimators discussed in Section 5, each estimator was applied to the nominal model ($i = 1$), using online measurements $\mathbf{y}_k$ generated by simulating the faulty model ($i = 2$) with random $\mathbf{x}_0$, $\mathbf{w}_k$, and $\mathbf{v}_k$ uniformly distributed in X_0, W, and V, respectively. We did not compare the exact observer *Poly* because it was found in Section 5 to have much higher complexity than the competing methods, although it has been applied for fault detection in low dimensional systems previously (Rosa et al., 2010; Tabatabaeipour, 2015; Tabatabaeipour et al., 2012). Fig. 7 shows the number of time steps required for each estimator to detect the fault through a failure of (34), averaged over 500 simulations. Among past methods, $ZBr(5)$ is the most effective by far. However, $ZBr(5)$ requires 2.5 more steps than $ZC(0, 5)$ on average, and the constrained zonotope observers show a clear trend of decreasing detection time with increasing n_c. Notably, $CZ(3, 5)$ detects the fault nearly 8 steps before $ZBr(5)$ and 18 steps before ROBP on average. These results demonstrate that constrained zonotopes can provide significantly faster fault detection, thereby reducing the potential for serious harm following a fault.

7. Conclusions

Constrained zonotopes provide a new set representation that combines the flexibility of convex polytopes with the efficiency and scalability of zonotopes, with regard to several key set operations. In addition, the complexity of these sets can be accurately

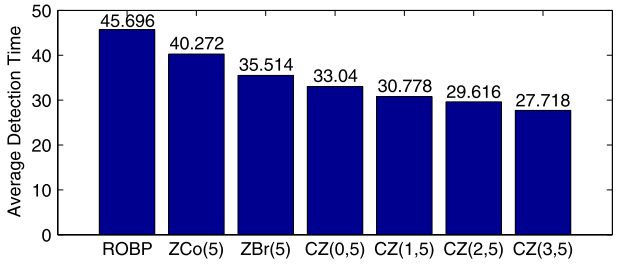


Fig. 7. Average time steps needed to detect the fault.

and efficiently reduced, making them practical for long sequences of set-based computations. Numerical studies demonstrate that constrained zonotopes provide significant improvements over the existing state-of-the-art in the context of set-based estimation and fault detection.

Appendix. Reduction implementation details

Here are the implementation details for rescaling $Z = \{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\}$ as described in Section 4.1. Algorithm 1 below is used to compute an interval $E \equiv [\boldsymbol{\xi}^L, \boldsymbol{\xi}^U]$ such that $B_\infty(\mathbf{A}, \mathbf{b}) \subset E \subset [-\mathbf{1}, \mathbf{1}]$ for use in Proposition 4. Algorithm 1 also returns $R \equiv [\rho^L, \rho^U] \subset \mathbb{R}^{n_g}$ satisfying

$$R_j \supset \left\{\xi_j : \mathbf{A}\boldsymbol{\xi} = \mathbf{b},\ |\boldsymbol{\xi}_i| \leq 1,\ \forall i \neq j\right\}, \quad \forall j. \tag{A.1}$$

R is used for constraint reduction as described below.

Consider the rearrangements of $\mathbf{A}\boldsymbol{\xi} = \mathbf{b}$:

$$\xi_j = a_{ij}^{-1} b_i - \sum_{k \neq j} a_{ij}^{-1} a_{ik} \xi_k, \quad \forall i, j : a_{ij} \neq 0. \tag{A.2}$$

Beginning with $E = [-\mathbf{1}, \mathbf{1}]$, each iteration of Algorithm 1 attempts to refine each E_j by bounding the right-hand side of (A.2) with $\xi_k \in E_k, \forall k \neq j$, using interval arithmetic.

Algorithm 1.

(1) Assign $E := [-\mathbf{1}, \mathbf{1}], R := [-\infty, +\infty], i = j = 1$.

(2) If $a_{ij} \neq 0$, assign

$$R_j := R_j \cap \left(a_{ij}^{-1} b_i - \sum_{k \neq j} a_{ij}^{-1} a_{ik} E_k\right), \qquad E_j := E_j \cap R_j.$$

(3) If $j < n_g$, assign $j := j + 1$ and go to Step 2. Otherwise, if $j = n_g$ and $i < n_c$, assign $(i, j) := (i + 1, 1)$ and go to Step 2. If $(i, j) = (n_c, n_g)$, terminate.

Algorithm 1 can be applied iteratively to further refine E. If $E \cap R = \emptyset$ in some iteration, then $Z = \emptyset$. Thus, Algorithm 1 can potentially detect an empty constrained zonotope without solving an LP as in Section 3.1.

The bounds obtained from Algorithm 1 can be greatly improved by preconditioning the constraints (Neumaier, 1990), which does not affect Z, since $\{\mathbf{G}, \mathbf{c}, \mathbf{A}, \mathbf{b}\} = \{\mathbf{G}, \mathbf{c}, \mathbf{PA}, \mathbf{Pb}\}$ for any invertible $\mathbf{P}$. In our implementation, $[\mathbf{A}|\mathbf{b}]$ is taken to reduced row echelon form by Gauss–Jordan elimination with full pivoting prior to applying Algorithm 1. In each step of elimination, the pivot element is chosen as the element in the unreduced submatrix (omitting the final column $\mathbf{b}$) that is largest relative to the infinity norm of the row it occupies (see Neumaier, 1990 for a discussion of preconditioning strategies). Since column pivoting changes the ordering of $\boldsymbol{\xi}$, column pivots must be carried out on $\mathbf{G}$ as well. The total complexity of preconditioning and Algorithm 1 is $O(n_c^2 n_g + n_c n_g^2)$.

Upon termination of Algorithm 1, the CG-rep of Z is rescaled as per Proposition 4. For (A.1) to remain valid, R must also be rescaled as

$$R_j := [(\rho_j^L - \xi_{m,j})/\xi_{r,j}, (\rho_j^U - \xi_{m,j})/\xi_{r,j}], \quad \forall j. \tag{A.3}$$

Next are the details of forming $\tilde{Z}$ by eliminating one constraint from the CG-rep of Z by the method in Section 4.2. In particular, it remains to estimate the Hausdorff error $H_j \equiv \max_{\tilde{\mathbf{z}} \in \tilde{Z}} \min_{\mathbf{z} \in Z} \|\tilde{\mathbf{z}} - \mathbf{z}\|_2$ used for selecting j. First, note that eliminating ξ_j using (28) actually preserves the constraint $\mathbf{a}_1^{\mathrm{T}} \boldsymbol{\xi} = b_1$ implicitly, although it is not present in the CG-rep of $\tilde{Z}$. However, the ability to enforce the bound $|\xi_j| < 1$ has been lost because this variable no longer appears in the CG-rep of $\tilde{Z}$. From this observation, it can be shown that

$$\tilde{Z} = \{\mathbf{G}\boldsymbol{\xi} + \mathbf{c} : \mathbf{A}\boldsymbol{\xi} = \mathbf{b},\ |\xi_i| \leq 1,\ \forall i \neq j\}. \tag{A.4}$$

It follows that

$$H_j = \max_{\boldsymbol{\xi}} \min_{\delta \in B_\infty(\mathbf{A}, \mathbf{b})} \|\mathbf{G}(\boldsymbol{\xi} - \boldsymbol{\delta})\|_2, \tag{A.5}$$
$$\text{s.t. } \mathbf{A}\boldsymbol{\xi} = \mathbf{b},\ |\xi_i| \leq 1,\ i \neq j.$$

To estimate H_j, let $\mathbf{d}^* \equiv (\boldsymbol{\xi}^* - \delta^*)$, where $(\boldsymbol{\xi}^*, \delta^*)$ is an optimal solution of (A.5). Note that $\mathbf{A}\mathbf{d}^* = \mathbf{0}$ and $|d_i^*| \leq 2$ for all $i \neq j$. Moreover, recall that we have available $R = [\rho^L, \rho^U]$ satisfying (A.1). If $|\rho_j^U|, |\rho_j^U| < 1$, then (A.1) shows that the bound $|\xi_j| \leq 1$ is redundant in the definition of $B_\infty(\mathbf{A}, \mathbf{b})$. Otherwise, $\max(|\rho_j^L|, |\rho_j^U|) - 1$ provides an indication of how far $B_\infty(\mathbf{A}, \mathbf{b})$ would extend outside of $[-\mathbf{1}, \mathbf{1}]$ in the δ_j-direction if the constraint $|\delta_j| \leq 1$ were omitted. Thus, the action of the outer program in (A.5) is approximated by the requirement that

$$d_j^* = r_j \equiv \max\left\{0, \max(|\rho_j^L|, |\rho_j^U|) - 1\right\}. \tag{A.6}$$

To simplify further, the inequality constraints $|d_i^*| \leq 2$ are relaxed by penalizing the norm of $\mathbf{d}$ in the objective function, which gives the estimate $\hat{H}_j \approx H_j^* \equiv \min\{\|\mathbf{G}\mathbf{d}\|_2^2 + \|\mathbf{d}\|_2^2 : \mathbf{A}\mathbf{d} = \mathbf{0},\ d_j = r_j\}$. Since this is an equality constrained quadratic program, its solution $\hat{\mathbf{d}}$ and duality multipliers $\hat{\boldsymbol{\lambda}}$ are known explicitly as the solutions of the linear system

$$\begin{bmatrix} \mathbf{G}^{\mathrm{T}}\mathbf{G} + \mathbf{I}_{n_g \times n_g} & \mathbf{A}^{\mathrm{T}} & \mathbf{e}_j \\ \mathbf{A} & \mathbf{0}_{n_c \times n_c} & \mathbf{0}_{n_c \times 1} \\ \mathbf{e}_j^{\mathrm{T}} & \mathbf{0}_{1 \times n_c} & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{d}} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{n_g \times 1} \\ \mathbf{0}_{n_c \times 1} \\ r_j \end{bmatrix}, \tag{A.7}$$

where $\mathbf{e}_j \in \mathbb{R}^{n_g}$ is the jth standard unit vector. Direct solution of (A.7) for each j has a complexity of $O(n_g(n_g + n_c + 1)^3)$. However, in our implementation, it is done with $O((n_g + n_c)^3)$ complexity by factoring the upper-left $(n_g + n_c) \times (n_g + n_c)$ submatrix in (A.7) only once. Letting $\mathbf{Q}$ denote this submatrix, (A.7) is equivalent to

$$\begin{bmatrix} \mathbf{I}_{(n_g+n_c)\times(n_g+n_c)} & \mathbf{Q}^{-1}\mathbf{e}_j \\ \mathbf{e}_j^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{d}} \\ \hat{\boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{(n_g+n_c)\times 1} \\ r_j \end{bmatrix} \tag{A.8}$$

where now $\mathbf{e}_j \in \mathbb{R}^{n_g+n_c}$. Once $\mathbf{Q}^{-1}$ is computed, solving (A.8) for each j requires computing $\mathbf{Q}^{-1}\mathbf{e}_j$, executing a single elementary row operation to eliminate the 1 in the bottom row, and solving the resulting upper triangular system. We implement this using the LU-factors of $\mathbf{Q}$ rather than $\mathbf{Q}^{-1}$.

To select j, we first check if $r_j = 0$ for any j. If this holds, then (A.1) and (A.4) imply that $Z = \tilde{Z}$ when j is selected. Failing this, we choose the j that minimizes $\hat{H}_j$.

We now give the details of the generator reduction method outlined in Section 4.3. In particular, it remains to describe the new zonotopic order reduction method alluded to there. Let $Z = \{\mathbf{G}, \mathbf{c}\}$. The first step is to reorder the columns of $\mathbf{G}$ as $\mathbf{G} \to \begin{bmatrix} \mathbf{T} & \mathbf{V} \end{bmatrix}$ such

that $\mathbf{T} \in \mathbb{R}^{n\times n}$ is invertible, which is done by taking $\mathbf{G}$ to the reduced row echelon form $\begin{bmatrix}\mathbf{I}_{n\times n} & \mathbf{R}\end{bmatrix}$ using Gauss–Jordan elimination with the full pivoting. If $\mathbf{G}$ is found to be rank deficient, then Z is simply reduced using the method in Combastel (2003). Otherwise, the sequence of column pivots performed during elimination gives the desired reordering $\begin{bmatrix}\mathbf{T} & \mathbf{V}\end{bmatrix}$, and $\mathbf{R} = \mathbf{T}^{-1}\mathbf{V}$. In each iteration of elimination, the pivot element is chosen as the element in the unreduced submatrix that is largest relative to the infinity norm of the row it occupies. As a consequence, each element of $\mathbf{R}$ is less than or equal to 1.

In the second step, a column $\mathbf{v}$ is chosen from $\mathbf{V}$ as described below, and Z is written as follows, where $\mathbf{V}_-$ denotes the matrix formed by removing $\mathbf{v}$ from $\mathbf{V}$:

$$Z = X + Y \equiv \{[\mathbf{T}\,\mathbf{v}], \mathbf{c}\} + \{\mathbf{V}_-, \mathbf{0}\}. \tag{A.9}$$

Next, the $(n+1)/n$ order zonotope X is conservatively reduced to a parallelotope $\tilde{X}$ using the method in Chisci et al. (1996). Finally, $\tilde{Z} = \tilde{X} + Y$ is defined, which has one fewer generator than Z.

Let $\mathbf{r}$ be the column of $\mathbf{R}$ corresponding to $\mathbf{v}$, so that $\mathbf{r} = \mathbf{T}^{-1}\mathbf{v}$. Using the fact that $\|\mathbf{r}\|_\infty \leq 1$, Theorem 3 in Chisci et al. (1996) states that the minimum volume parallelotope containing X is $\tilde{X} = \{[\mathbf{T}(\mathbf{I} + \operatorname{diag}|\mathbf{r}|)], \mathbf{c}\}$. Thus, $\tilde{Z} = \{[\mathbf{T}(\mathbf{I} + \operatorname{diag}|\mathbf{r}|)\,\mathbf{V}_-], \mathbf{c}\}$.

It remains to find an effective heuristic for choosing the column $\mathbf{v}$. To do so, we investigate the volumes of X and $\tilde{X}$. Using the standard formula for the volume of a zonotope (Bravo et al., 2006) and standard properties of the determinant,

$$\begin{aligned} v(\tilde{X}) &= 2^n |\det \mathbf{T} \det(\mathbf{I} + \operatorname{diag}|\mathbf{r}|)|, \\ &= 2^n |\det \mathbf{T}| \Pi_{i=1}^n (1 + |r_i|), \end{aligned}$$

and, letting $\mathbf{t}_i$ denote the ith column of $\mathbf{T}$,

$$\begin{aligned} v(X) &= 2^n \left(|\det \mathbf{T}| + \sum_{i=1}^n |\det[\mathbf{t}_1 \cdots \mathbf{t}_{i-1} \quad \mathbf{v} \quad \mathbf{t}_{i+1} \cdots \mathbf{t}_n]| \right), \\ &= 2^n \left(|\det \mathbf{T}| + \sum_{i=1}^n |\det(\mathbf{T}[\mathbf{e}_1 \cdots \mathbf{e}_{i-1} \quad \mathbf{r} \quad \mathbf{e}_{i+1} \cdots \mathbf{e}_n])| \right), \\ &= 2^n |\det \mathbf{T}| \left(1 + \sum_{i=1}^n |r_i| \right). \end{aligned}$$

We choose the column $\mathbf{v}$ such that $\mathbf{r} = \mathbf{T}^{-1}\mathbf{v}$ minimizes the volume error $v(\tilde{X}) - v(X)$. Note that determining this does not require computing $\det \mathbf{T}$.

Direct application of this method to eliminate k generators has complexity $O(kn^2 n_g)$ stemming from repeated factorization of $\mathbf{G}$. In our implementation, this factorization is done only once, and both $\mathbf{T}$ and $\mathbf{R}$ are updated directly via

$$\mathbf{T} := \mathbf{T}(\mathbf{I} + \operatorname{diag}|\mathbf{r}|), \qquad \mathbf{R} := (\mathbf{I} + \operatorname{diag}|\mathbf{r}|)^{-1}\mathbf{R}_-, \tag{A.10}$$

where $\mathbf{R}_-$ is formed by removing $\mathbf{r}$ from $\mathbf{R}$. $\tilde{Z}$ can be recovered after the update as $\tilde{Z} = \{\begin{bmatrix}\mathbf{T} & \mathbf{TR}\end{bmatrix}, \mathbf{c}\}$. At the same time, the updated $\mathbf{R}$ matrix retains the property that every element is less than one, since the elements of $(\mathbf{I} + \operatorname{diag}|\mathbf{r}|)$ are greater than one. Thus, all of the information necessary for further reduction is available. The complexity for reducing k generators is $O(n^2 n_g + kn_g n)$, where $n^2 n_g$ results from the one time factorization of $\mathbf{G}$, and knn_g results from kn_g computations of $v(\tilde{X}) - v(X)$ and k updates as per (A.10).

References

Alamo, T., Bravo, J. M., & Camacho, E. F. (2005). Guaranteed state estimation by zonotopes. *Automatica*, *41*(6), 1035–1043.

Althoff, M., & Krogh, B.H. (2011). Zonotope bundles for the efficient computation of reachable sets. In *Proc. 50th IEEE conference on decision and control* (pp. 6814–6821).

Althoff, M., Stursberg, O., & Buss, M. (2010). Computing reachable sets of hybrid systems using a combination of zonotopes and polytopes. *Nonlinear Analysis-Hybrid Systems*, *4*(2), 233–249.

Bertsekas, D. P., & Rhodes, I. B. (1971). Recursive state estimation for a set-membership description of uncertainty. *IEEE Transactions on Automatic Control*, *16*(2), 117–128.

Blanchini, F., & Miani, S. (2008). *Set-theoretic methods in control* (2nd ed.). Springer International Publishing.

Bravo, J. M., Alamo, T., & Camacho, E. F. (2006). Bounded error identification of systems with time-varying parameters. *IEEE Transactions on Automatic Control*, *51*, 1144–1150.

Chernousko, F. L. (1980). Optimal guaranteed estimates of indeterminacies with the aid of ellipsoids. Parts I–II. *Engineering Cybernetics*, *18*, 1–9.

Chisci, L., Garulli, A., & Zappa, G. (1996). Recursive state bounding by parallelotopes. *Automatica*, *32*(7), 1049–1055.

Combastel, C. (2003). A state bounding observer based on zonotopes. In *Proc. European control conference. Cambridge, UK.*

Durieu, C., Walter, E., & Polyak, B. (2001). Multi-input multi-output ellipsoidal state bounding. *Journal of Optimization Theory and Applications*, *111*(2), 273–303.

Fogel, E., & Huang, F. (1982). On the value of information in system identification—Bounded noise case. *Automatica*, *18*, 229–238.

Hagemann, W. (2015). Efficient geometric operations on convex polyhedra, with an application to reachability analysis of hybrid systems. *Mathematics in Computer Science*, *9*(3), 283–325.

Ingimundarson, A., Bravo, J. M., Puig, V., Alamo, T., & Guerra, P. (2009). Robust fault detection using zonotope-based set-membership consistency test. *International Journal of Adaptive Control and Signal Processing*, *23*(4), 311–330.

Jones, C. N., Kerrigan, E. C., & Maciejowski, J. M. (2008). On polyhedral projection and parametric programming. *Journal of Optimization Theory and Applications*, *138*(2), 207–220.

Kuhn, W. (1998). Rigorously computed orbits of dynamical systems without the wrapping effect. *Computing*, *61*(1), 47–67.

Kurzhanski, A. B. (2011). Hamiltonian techniques for the problem of set-membership state estimation. *International Journal of Adaptive Control and Signal Processing*, *25*(3), 249–263.

Kvasnica, M., Grieder, P., Baotić, M., & Morari, M. (2004). Multi-parametric toolbox (MPT). In *Hybrid systems: computation and control* (pp. 448–462). Springer.

Le, V. T. H., Stoica, C., Alamo, T., Camacho, E. F., & Dumur, D. (2013). Zonotopic guaranteed state estimation for uncertain systems. *Automatica*, *49*(11), 3418–3424.

Liu, X. Q., Zhang, H. Y., Liu, J., & Yang, J. (2000). Fault detection and diagnosis of permanent-magnet DC motor based on parameter estimation and neural network. *IEEE Transactions on Industrial Electronics*, *47*(5), 1021–1030.

Mayne, D. Q., Rakovic, S. V., Findeisen, R., & Allgower, F. (2006). Robust output feedback model predictive control of constrained linear systems. *Automatica*, *42*(7), 1217–1222.

McMullen, P. (1971). On zonotopes. *Transactions of the American Mathematical Society*, *159*, 91–109.

Neumaier, A. (1990). *Interval methods for systems of equations*. Cambridge: Cambridge University Press.

Ocampo-Martinez, C., Guerra, P., Puig, V., & Quevedo, J. (2007). Actuator fault-tolerance evaluation of linear constrained model-predictive control using zonotope-based set computations. *Journal of Systems and Control Engineering*, *221*(6), 915–926.

Raimondo, D.M., Marseglia, G.R., Braatz, R.D., & Scott, J.K. (2013). Fault-tolerant model predictive control with active fault isolation. In *Proc. 2nd international conference on control and fault tolerant systems*. Paper ThB1.4.

Rosa, P., Silvestre, C., Shamma, J.S., & Athans, M. (2010). Fault detection and isolation of LTV systems using set-valued observers. In *Proc. 49th IEEE conference on decision and control* (pp. 768–773).

Schweppe, F. (1968). Recursive state estimation: Unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, *13*(1), 22–28.

Scott, J. K., & Barton, P. I. (2013). Bounds on the reachable sets of nonlinear control systems. *Automatica*, *49*, 93–100.

Scott, J. K., Findeisen, R., Braatz, R. D., & Raimondo, D. M. (2014). Input design for guaranteed fault diagnosis using zonotopes. *Automatica*, *50*(6), 1580–1589.

Shamma, J. S., & Tu, K.-Y. (1999). Set-valued observers and optimal disturbance rejection. *IEEE Transactions on Automatic Control*, *44*, 253–264.

Tabatabaeipour, S. M. (2015). Active fault detection and isolation of discrete-time linear time-varying systems: a set-membership approach. *International Journal of Systems Science*, *46*(11), 1917–1933.

Tabatabaeipour, S. M., Odgaard, P. F., Bak, T., & Stoustrup, J. (2012). Fault detection of wind turbines with uncertain parameters: A set-membership approach. *Energies*, *5*, 2424–2448.

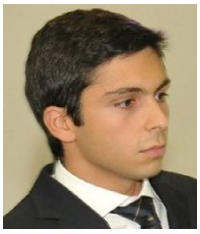
Tiwary, H. R. (2008). On the hardness of computing intersection, union and Minkowski sum of polytopes. *Discrete & Computational Geometry*, *40*(3), 469–479.

Tornil-Sin, S., Ocampo-Martinez, C., Puig, V., & Escobet, T. (2012). Robust fault detection of non-linear systems using set-membership state estimation based on constraint satisfaction. *Engineering Applications of Artificial Intelligence*, *25*(1), 1–10.

Walter, E., & Piet-Lahanier, H. (1989). Exact recursive polyhedral description of the feasible parameter set for bounded-error models. *IEEE Transactions on Automatic Control*, *34*(8), 911–915.

Joseph K. Scott is an Assistant Professor in the Department of Chemical and Biomolecular Engineering at Clemson University. He received his B.S. (2006) in Chemical Engineering from Wayne State University, and his M.S. (2008) and Ph.D. (2012) in Chemical Engineering from MIT. His honors include the 2012 Best Paper Award for the Journal of Global Optimization and the 2016 Air Force Young Investigator Research Program Award. His research interests include dynamical systems, optimization theory, simulation and optimization of chemical processes, advanced process control, and fault diagnosis. Current applications focus on optimization and control of renewable energy systems.

Davide M. Raimondo is an associate professor in the Department of Electrical, Computer and Biomedical Engineering at University of Pavia, Italy. He received the Ph.D. in Electronics, Computer Science and Electrical Engineering from the University of Pavia, Italy, in 2009. As a Ph.D. student he held a visiting position at the Department of Automation and Systems Engineering, University of Seville, Spain. From January 2009 to December 2010 he was a postdoctoral fellow in the Automatic Control Laboratory, ETH Zürich, Switzerland. In 2012 (March to June), 2013 (August to September) and 2014 (September to November) he was visiting scholar in Prof. Braatz Group, Department of Chemical Engineering, MIT, USA. From December 2010 to May 2015 he was assistant professor at University of Pavia. He is the author or co-author of more than 60 papers published in refereed journals, edited books, and refereed conference proceedings. He is subject editor of the Journal of Optimal Control Applications and Methods. His current research interests include model predictive control, active fault diagnosis, fault-tolerant control, distributed control, renewable energy, autonomous surveillance, and control of glycemia in diabetics.



Giuseppe Roberto Marseglia received the Ph.D. in Electronics, Computer Science and Electrical Engineering from the University of Pavia, Italy, in 2016. As a Ph.D. student he held a visiting position at the Braatz Group, Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA. He is the author or co-author of diverse papers published in refereed journals, edited books, and refereed conference proceedings. His current research interests include active fault diagnosis, fault-tolerant control and renewable energy.

Richard D. Braatz is the Edwin R. Gilliland Professor at the Massachusetts Institute of Technology (MIT) where he does research in applied mathematics and control theory and its application to manufacturing processes, biomedical systems, and nanotechnology. He received M.S. and Ph.D. degrees from the California Institute of Technology and was on the faculty at the University of Illinois at Urbana–Champaign and was a Visiting Scholar at Harvard University before moving to MIT. He has consulted or collaborated with more than 20 companies including United Technologies Corporation, IBM, BP, and Novartis. Honors include the AACC Donald P. Eckman Award, the Antonio Ruberti Young Researcher Prize, the IEEE Control Systems Society Transition to Practice Award, and best paper awards from IEEE- and IFAC-sponsored control journals. He is a Fellow of IEEE and IFAC.