

A set-membership state estimation algorithm based on DC programming^{*}

T. Alamo^a, J.M. Bravo^b, M.J. Redondo^b and E.F. Camacho^a

^a*Departamento de Ingeniería de Sistemas y Automática. Universidad de Sevilla.
Camino de los Descubrimientos s/n. 41092 Sevilla. Spain.*

^b*Departamento de Ingeniería Electrónica, Sistemas Informáticos y Automática. Universidad de Huelva.
Carretera Huelva - La Rábida. Palos de la Frontera, 21071 Huelva. Spain.*

Abstract

This paper presents a new approach to guaranteed state estimation for non-linear discrete-time systems with a bounded description of noise and parameters. The sets of states that are consistent with the evolution of the system, the measured outputs and bounded noise and parameters are represented by zonotopes. DC programming and intersection operations are used to obtain a tight bound. An example is given to illustrate the proposed algorithm.

Key words: Nonlinear observers, Set-membership state estimation, DC programming.

1 Introduction

The purpose of this paper is to present a new state estimator for uncertain discrete time nonlinear dynamic systems. The Kalman filter theory provides an estimation of the state of a given process based on output measurements. This estimation is optimal with respect to the error variance. A different alternative is to consider a norm-bounded uncertainty. This hypothesis is used by the set-membership approach (Milanese, Norton, Piet-Lahanier & Walter 1996, Garulli, Tesi & Vicino 1999, Calafiore 2005) and it is adopted in this paper. This strategy builds a compact set that bounds the states of the system that are consistent with the measured output and the norm-bounded uncertainty.

In the set-membership approach, several geometric figures have been used to bound the consistent state set. The application of ellipsoidal sets to the state estimation problem has been introduced in pioneering works (Schweppe 1968) and by different authors. See, for example, (Kurzanski & Valyi 1996, Savkin & Petersen 1998, El Ghaoui & Calafiore 2001, Durieu, Walter &

Polyak 2001).

The use of polyhedrons was proposed by (Kuntsevich & Lychak 1985) to obtain an increased estimation accuracy. In (Kieffer, Jaulin & Walter 2002), a guaranteed recursive nonlinear estimator based on an interval branch and bound algorithm is given. To improve the exponential complexity, consistency techniques are considered in (Jaulin 2002). The complexity of these representations grows considerably with the number of observations and the order of the system. An alternative approach based on parallelotopes was presented in (Chisci, Garulli & Zappa 1996, Chisci, Garulli, Vicino & Zappa 1998), where minimum-volume bounding parallelotopes are used to estimate the state of a discrete-linear dynamic system. Also, the state estimation problem for piecewise affine systems is addressed in (Raković & Mayne 2004) using polyhedrons.

A zonotope is a linear transformation of a unitary box (Montgomery 1989, Shephard 1974). They have been used in (Puig, Cugueró & Quevedo 2001, Combastel 2003) to build a worst-case state estimator. In (Puig et al. 2001) the measured output is used to estimate the state by means of a gain K . In (Combastel 2003), a singular value decomposition is used to obtain the consistent region of the state space. Interval arithmetic and zonotopes are combined in (Alamo, Bravo & Camacho 2005) to obtain a guaranteed nonlinear state estimator.

^{*} This paper was not presented at any IFAC meeting. Corresponding author T. Alamo. Tel. +034-605631404.

Email addresses: alamo@cartuja.us.es (T. Alamo), caro@uhu.es (J.M. Bravo), caro@uhu.es (M.J. Redondo), eduardo@cartuja.us.es (E.F. Camacho).

In this paper, a new method for guaranteed state estimation in the case of nonlinear discrete processes with bounded uncertain parameters and noise is presented. The goal is to apply a DC programming approach to the state estimation problem. Zonotopes and DC programming are used by the proposed method to obtain a guaranteed bound of the uncertain trajectory of the nonlinear system each sample time. An example illustrates that the proposed method improves the results obtained in (Alamo et al. 2005).

A DC function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a function that can be expressed as the difference of two convex functions, that is, $f(x) = g(x) - h(x)$ where $g(x)$ and $h(x)$ are convex functions. The class of DC functions is close under a good number of basic operations. For example, if $f_1(x)$ and $f_2(x)$ are DC functions then $f_1(x) + f_2(x)$, $f_1(x) - f_2(x)$, $f_1(x) \cdot f_2(x)$, $\max\{f_1(x), f_2(x)\}$ and $\min\{f_1(x), f_2(x)\}$ are DC functions (Tuy 1995, Horst & Thoai 1999). It is also worth remarking that any continuous piecewise affine function is a DC function. DC programming problems are mathematical programming problems dealing with functions that can be represented as a difference of two convex functions. Several techniques have been developed using DC programming to solve non convex global optimization problems.

The paper is organized as follows: The problem formulation and the general lines of the algorithm are presented in section 2. In section 3, a brief introduction to DC programming is given. New proposed methods to bound the evolution of the uncertain system and the set of states consistent with the measurements are presented in section 4 and 5. The full version of the set-membership state estimation algorithm appears in section 6. Finally, an example is used to illustrate the new algorithm.

2 Problem formulation

In what follows, some preliminary notations are introduced. An interval $[a, b]$ is the set $\{x : a \leq x \leq b\}$. The unitary interval is $\mathbf{B} = [-1, 1]$. A box is an interval vector. A unitary box in \mathbb{R}^m , denoted as \mathbf{B}^m , is a box composed by m unitary intervals. The Minkowski sum of two sets X and Y is defined by $X \oplus Y = \{x + y : x \in X, y \in Y\}$. Given a vector $p \in \mathbb{R}^n$ and a matrix $H \in \mathbb{R}^{n \times m}$, the set:

$$p \oplus HB^m = \{p + Hz : z \in \mathbf{B}^m\}$$

is called a zonotope of order m . Note that this is the Minkowski sum of the segments defined by the columns of matrix H . A parallelotope is a zonotope with $n = m$. Given the parallelotope $P = p \oplus HB^n$, where $H \in \mathbb{R}^{n \times n}$ is invertible, P can be rewritten as $P = \{x : \|H^{-1}x - H^{-1}p\|_\infty \leq 1\}$.

Consider an uncertain nonlinear discrete-time system of the form:

$$\begin{cases} x_{k+1} = \mathbf{f}(x_k, w_k) \\ y_k = \mathbf{d}(x_k, v_k) \end{cases} \quad (1)$$

where $x_k \in X \subseteq \mathbb{R}^n$ with $k \geq 0$ is the state of the system and $y_k \in \mathbb{R}^p$ is the measured output vector at sample time k . The vector $w_k \in W \subseteq \mathbb{R}^{n_w}$ with $k \geq 0$ represents the time varying process parameters and process perturbation vector and $v_k \in V \subseteq \mathbb{R}^{p_v}$ with $k \geq 0$ is the measurement noise vector. It is assumed that the uncertainties and the initial state are bounded by zonotopes: $w_k \in W = c_w \oplus M_w \mathbf{B}^{n_w}$, $v_k \in V = c_v \oplus M_v \mathbf{B}^{p_v}$ and $x_0 \in \mathcal{X}_0 = p_0 \oplus H_0 \mathbf{B}^n$ where $c_w \in \mathbb{R}^{n_w}$, $c_v \in \mathbb{R}^{p_v}$ and $p_0 \in \mathbb{R}^n$.

It will be assumed that $\mathbf{f}(\cdot)$ and $\mathbf{d}(\cdot)$ are continuous functions, and that each component of $\mathbf{f}(\cdot)$ and $\mathbf{d}(\cdot)$ have DC representations, that is,

$$\begin{aligned} f_i(x, w) &= g_i(x, w) - h_i(x, w), \quad i = 1, \dots, n \\ d_i(x, w) &= a_i(x, w) - b_i(x, w), \quad i = 1, \dots, p \end{aligned}$$

where $f_i(\cdot, \cdot)$, $d_i(\cdot, \cdot)$ represent the i -th component of functions $\mathbf{f}(x, w)$ and $\mathbf{d}(x, w)$ respectively and where the functions $h_i(x, w)$, $g_i(x, w)$, $i = 1, \dots, n$ and $a_i(x, w)$, $b_i(x, w)$, $i = 1, \dots, p$ are convex in (X, W) and (X, V) respectively. This is not a very restrictive assumption because every continuous function can be approximated by a difference of two convex functions (DC function) (Horst & Thoai 1999) and every \mathcal{C}^2 -function is a DC function (Tuy 1995). In section 3 an example is given.

Given a continuous function $\phi(\cdot)$ and a set $X \subset \mathbb{R}^n$, $\phi(X)$ denotes the set $\{\phi(x) : x \in X\}$. With this notation, the consistent state set and the exact uncertain set are defined as follows:

Definition 1 (Consistent state set) *Given system (1) and a measured output y_k , the consistent state set at time k is defined as $\mathcal{X}_{y_k} = \{x \in \mathbb{R}^n : y_k \in \mathbf{d}(x, V)\}$.*

Definition 2 (Exact uncertain state set) *Consider a system given by equation (1). The exact uncertain state set \mathcal{X}_k is equal to the set of states that are consistent with the measured outputs y_1, y_2, \dots, y_k and the initial state set \mathcal{X}_0 :*

$$\mathcal{X}_k = \mathbf{f}(\mathcal{X}_{k-1}, W) \bigcap \mathcal{X}_{y_k}, \quad k \geq 1$$

The exact computation of these sets is a difficult task. In order to reduce the complexity of the computations, these sets are bounded by means of conservative outer bounds. Then, at sample time k , the objective is to find an outer approximation of the corresponding exact uncertain set \mathcal{X}_k .

This paper presents a new set-membership state estimation algorithm for nonlinear systems. Suppose that an outer bound of the exact uncertain state set is available at time $k - 1$ (this bound will be denoted as $\hat{\mathcal{X}}_{k-1}$). Suppose also that a measured output y_k is obtained at sample time k . Under these assumptions, this is the general outline of the algorithm:

Algorithm 1

Step 1: Use DC programming to bound the uncertain trajectory of the non-linear system: $\bar{\mathcal{X}}_k \supseteq \mathbf{f}(\hat{\mathcal{X}}_{k-1}, W)$.

Step 2: Compute an outer bound of the consistent state set \mathcal{X}_{y_k} . Denote it $\bar{\mathcal{X}}_{y_k}$.

Step 3: Compute an outer bound of $\bar{\mathcal{X}}_k \cap \bar{\mathcal{X}}_{y_k}$. Denote it $\hat{\mathcal{X}}_k$.

End of algorithm 1

The proposed algorithm is similar to the Kalman filter: the first step can be considered as a prediction step while the second and third steps constitute a correction step. In the first step, zonotopes (Montgomery 1989, Shephard 1974) and DC programming are used to obtain an outer bound of the evolution of the system. This outer bound is improved using the information provided by the new measurement and DC programming (second and third steps). The full version of the algorithm is detailed in section 6.

3 DC programming

This section presents essential results about DC programming. These concepts are required to introduce the proposed state estimation algorithm. References (Horst & Thoai 1999, Tuy 1998) and (Tuy 1995) are excellent surveys about DC programming.

Definition 3 Let S be a convex polytope (bounded polyhedral set) of \mathbb{R}^n . A real-valued function $f : S \rightarrow \mathbb{R}$ is called DC on S , if there exists two convex functions $g, h : S \rightarrow \mathbb{R}$ such that f can be expressed in the form: $f(x) = g(x) - h(x)$.

It is known that the set of DC functions defined on a compact convex set of \mathbb{R}^n is dense in the set of continuous functions of this set (Tuy 1995, Horst & Thoai 1999). Therefore, every continuous function on a compact convex set can be approximated by a DC function with any desired precision. Moreover, given a \mathcal{C}^2 -function, it is always possible to obtain a DC-representation. In effect, suppose that $f : S \rightarrow \mathbb{R}$ satisfies $\frac{\partial^2}{\partial x^2} f(x) > -2\alpha I$, $\forall x \in S$ with $\alpha \geq 0$. Recall now that a \mathcal{C}^2 -function is convex in S if and only if $\frac{\partial^2}{\partial x^2} f(x) \geq 0$, $\forall x \in S$. Bearing this in mind, it is easy to see that $f(x) = g(x) - h(x)$, with $g(x) = f(x) + \alpha x^\top x$ and $h(x) = \alpha x^\top x$ constitutes a DC representation of $f(x)$. A systematic method to obtain

(by means of interval arithmetic) an appropriate value of α for a given \mathcal{C}^2 -function can be found in (Adjiman & Floudas 1996). The following example illustrates this idea. Consider the function $f(x) = x^3 + x^2 + 1$ in the domain $x \in [-1, 1]$. Since $\frac{\partial^2}{\partial x^2} f(x) = 6x + 2$, it results that $\frac{\partial^2}{\partial x^2} f(x) \geq -4$, $\forall x \in [-1, 1]$. Thus, $f(x) + 2x^2$ satisfies $\frac{\partial^2}{\partial x^2} (f(x) + 2x^2) \geq 0$ for all $x \in [-1, 1]$. Defining $g(x) = f(x) + 2x^2$ and $h(x) = 2x^2$, the equivalent function $f(x) = g(x) - h(x)$ is a DC function in $x \in [-1, 1]$.

Definition 4 Programming problems dealing with DC functions are called DC programming problems. A general form of DC programming problem is given by:

$$\min_{x \in S} f(x)$$

where $f(x) = g(x) - h(x)$ and $g(x)$ and $h(x)$ are convex in S .

Note that it is not necessary to restrict S to the class of polytopes. For a more general definition of DC Programming see (Pinter 1996). The following definitions are standard in the convex optimization literature. See for example, (Rockafellar 1970, Boyd & Vandenberghe 2004).

Definition 5 The subdifferential of a convex function $g : S \rightarrow \mathbb{R}$ at point x_0 (also denominated the set of subgradients of g at point x_0) denoted $\partial g(x_0)$ is defined by:

$$\partial g(x_0) = \{ u_0 \in \mathbb{R}^n : g(x) \geq g(x_0) + u_0^\top (x - x_0), \forall x \in S \}$$

If the function g is differentiable in S , the vector u_0 can be computed by the gradient of the function: $u_0 = \frac{\partial}{\partial x} g(x_0)$. This stems directly from the convexity of g .

Definition 6 Given a convex function $g : S \rightarrow \mathbb{R}$ and a subgradient u_0 of g at point $x_0 \in S$, a linear minorant of g is the linear function:

$$\bar{g}(x) = g(x_0) + u_0^\top (x - x_0)$$

By definition, it is clear that $g(x) \geq \bar{g}(x)$, $\forall x \in S$. In the same way, given the convex function $h : S \rightarrow \mathbb{R}$, $\bar{h}(x)$ denotes a linear minorant of h (obtained by means of the concept of subgradient).

Denoting as $vert(S)$ the set of vertices of S , and bearing in mind that $\bar{g}(x) - h(x)$ is a concave function and $g(x) - \bar{h}(x)$ is a convex function, it is possible to obtain an approximated solution of the DC programming problem by:

$$\min_{x \in S} f(x) \geq \min_{x \in vert(S)} \bar{g}(x) - h(x)$$

$$\max_{x \in S} f(x) \leq \max_{x \in \text{vert}(S)} g(x) - \bar{h}(x)$$

Therefore, in order to obtain lower and upper bounds for a global solution, all the vertices of set S must be visited.

Using these ideas, DC programming can be used to bound the range of a function. Next, a simple example is provided. Consider the function $f(x) = x^2 - \exp(x)$ in the domain $S = [0, 2]$. Clearly, $f(x)$ is a DC function ($g(x) = x^2$ and $h(x) = \exp(x)$). The exact range of the function is $f(S) = [4 - \exp(2), -\exp(0)] = [-3.3891, -1]$. The range obtained by interval arithmetic (Moore 1966) is $f([0, 2]) = [0, 2]^2 - \exp([0, 2]) = [-7.3891, 3.0000]$. Using $x_0 = 1$ to obtain the linear minorants of x^2 and $\exp(x)$, the approximated range obtained by DC programming is $[-4.3891, 0]$. The overestimation is considerably reduced. Therefore, the use of DC programming potentially improves previous results based in interval arithmetics (Alamo et al. 2005). The bounds obtained by DC functions are based on a linear approximation of a convex function providing a second order approximation (in a Taylor sense). That is, the error diminishes quadratically with the distance to the linearization point. We think that this property assures a good trade off between overestimation and computational cost.

4 Bounding the evolution of the system

This section presents a new method to bound the evolution of the nonlinear system (1). First, a linear approximation of the functional form of the system is used to obtain an approximation of the evolution of the system. Next, the proposed method takes advantage of the DC structure of system (1) to bound the error produced by the linear approximation in a guaranteed way. Combining the linear approximation and the bounded error, an outer bound of the evolution of the nonlinear system is obtained.

Consider the function $\mathbf{f}(x, w) : \mathbb{R}^n \times \mathbb{R}^{n_w} \rightarrow \mathbb{R}^n$, where $x \in X = p \oplus H\mathbf{B}^m$ and $w \in W = c_w \oplus M_w\mathbf{B}^{r_w}$. As commented before, it is assumed that each component of $\mathbf{f}(x, w)$ is a DC function, that is $f_i(x, w) = g_i(x, w) - h_i(x, w)$ with $i = 1, \dots, n$. The functions $g_i(x, w)$ and $h_i(x, w)$ are convex functions in (X, W) .

The objective of the method is to obtain an outer bound of set $\mathbf{f}(X, W)$. A linear function

$$\mathbf{f}^L(x, w) = \mathbf{f}(p, c_w) + G_x(x - p) + G_w(w - c_w)$$

is used to approximate the original function $\mathbf{f}(x, w)$. This function can be obtained by different ways, for example, when $\mathbf{f}(x, w)$ is a differentiable function, matrices G_x and G_w can be set equal to $\frac{\partial}{\partial x}\mathbf{f}(p, c_w)$ and

$G_w = \frac{\partial}{\partial w}\mathbf{f}(p, c_w)$ respectively. In the following subsection, function $\mathbf{f}^L(x, w)$ and the error produced by the linear approximation are stated in a precise way.

4.1 Bounding the error term

In this subsection, a guaranteed bound of the error incurred when approximating the nonlinear system by the linearization $\mathbf{f}^L(x, w)$ is provided. For this purpose, the following definition is introduced.

Definition 7 *The error set \mathcal{E} is defined by:*

$$\mathcal{E} = \{ e \in \mathbb{R}^n : e = \mathbf{f}(x, w) - \mathbf{f}^L(x, w), x \in X, w \in W \}$$

where

$$\mathbf{f}^L(x, w) = \mathbf{f}(p, c_w) + G_x(x - p) + G_w(w - c_w).$$

In what follows, a way to compute an outer bound $\bar{\mathcal{E}}$ of the error set \mathcal{E} is presented. This bound is obtained using the DC programming concepts presented in section 3. Firstly, it will be assumed that a parallelotope $P = t \oplus Q\mathbf{B}^n \subset \mathbb{R}^n$ that bounds set X ($X \subseteq P$) is available (this parallelotope can be obtained by means of the result presented in appendix A). Under this assumption, consider now the following affine functions in x and w :

$$\begin{aligned} \bar{g}_i(x, w) &= g_i(p, c_w) + u_{g_i}^\top \begin{bmatrix} x - p \\ w - c_w \end{bmatrix}, \quad i = 1, \dots, n \\ \bar{h}_i(x, w) &= h_i(p, c_w) + u_{h_i}^\top \begin{bmatrix} x - p \\ w - c_w \end{bmatrix}, \quad i = 1, \dots, n \end{aligned}$$

where u_{g_i} , u_{h_i} are subgradients at $(x, w) = (p, c_w)$ of $g_i(x, w)$ and $h_i(x, w)$ respectively. Due to the convexity of $g_i(\cdot, \cdot)$ and $h_i(\cdot, \cdot)$ it results that $\bar{g}_i(x, w) \leq g_i(x, w)$ and $\bar{h}_i(x, w) \leq h_i(x, w)$, $\forall (x, w)$, $i = 1, \dots, n$. That is, they are linear minorants.

Denote now with $f_i^L(x, w)$ the i -th component of $\mathbf{f}^L(x, w)$. With this notation:

$$\begin{aligned} f_i(x, w) - f_i^L(x, w) &= g_i(x, w) - h_i(x, w) - f_i^L(x, w) \\ &\leq g_i(x, w) - \bar{h}_i(x, w) - f_i^L(x, w) \end{aligned}$$

That is, $g_i(x, w) - \bar{h}_i(x, w) - f_i^L(x, w)$ is a convex majorant of $f_i(x, w) - f_i^L(x, w)$. Denoting now as $\text{vert}(P, W)$ the set of vertices of (P, W) it is concluded that:

$$\begin{aligned}
& \max_{(x,w) \in (X,W)} f_i(x,w) - f_i^L(x,w) \leq \\
& \max_{(x,w) \in (X,W)} g_i(x,w) - \bar{h}_i(x,w) - f_i^L(x,w) \leq \\
& \max_{(x,w) \in (P,W)} g_i(x,w) - \bar{h}_i(x,w) - f_i^L(x,w) = \\
& \max_{(x,w) \in \text{vert}(P,W)} g_i(x,w) - \bar{h}_i(x,w) - f_i^L(x,w)
\end{aligned}$$

Reasoning along the same lines, it can be affirmed that:

$$\begin{aligned}
& \min_{(x,w) \in (X,W)} f_i(x,w) - f_i^L(x,w) \geq \\
& \min_{(x,w) \in \text{vert}(P,W)} \bar{g}_i(x,w) - h_i(x,w) - f_i^L(x,w)
\end{aligned}$$

What has preceded proves the following result:

Lemma 1 *Suppose that the parallelotope P contains X and define the parallelotope $\bar{\mathcal{E}}$ as:*

$$\bar{\mathcal{E}} = \{ x \in \mathbb{R}^n : \gamma_i^- \leq x_i \leq \gamma_i^+, i = 1, \dots, n \}$$

where:

$$\begin{aligned}
\gamma_i^+ &= \max_{(x,w) \in \text{vert}(P,W)} (g_i(x,w) - \bar{h}_i(x,w) - f_i^L(x,w)) \\
\gamma_i^- &= \min_{(x,w) \in \text{vert}(P,W)} (\bar{g}_i(x,w) - h_i(x,w) - f_i^L(x,w))
\end{aligned}$$

then, the parallelotope $\bar{\mathcal{E}}$ is an outer bound of the set \mathcal{E} , this is: $\mathcal{E} \subseteq \bar{\mathcal{E}}$.

Remark 1 *Note that, in order to compute the parallelotope $\bar{\mathcal{E}}$, it is necessary, in principle, to visit the 2^{n+r_w} vertices of (P,W) . Suppose that w enters in an additive way into the model of the system, that is, $\mathbf{f}(x,w) = \tilde{\mathbf{f}}(x) + Ew$. In this case, making G_w equal to E it results that $f_i(x,w) - f_i^L(x,w)$ does not depend on w and only 2^n vertices have to be considered. This is the complexity order associated to bound the evolution of the uncertain system. Note that this complexity is affordable for low order systems and provides a good trade off between computational complexity and accuracy.*

4.2 Initial guaranteed bound of the evolution of the system

Now, a theorem that provides a first operator to bound the evolution of the system is given. This operator supposes a known outer bound of the error set \mathcal{E} (obtained by means of lemma 1).

Theorem 1 *Consider the zonotopes $X = p \oplus H\mathbf{B}^m$ and $W = c_w \oplus M_w\mathbf{B}^{r_w}$. Suppose that the parallelotope $\bar{\mathcal{E}} = \bar{t} \oplus \bar{Q}\mathbf{B}^n$ satisfies $\mathcal{E} \subseteq \bar{\mathcal{E}}$. Obtain now the zonotope $Z = p_z \oplus H_z\mathbf{B}^{m+r_w+n}$ where:*

- $p_z = \mathbf{f}(p, c_w) + \bar{t}$
- $H_z = [G_x H \quad G_w M_w \quad \bar{Q}]$

then, under these definitions:

$$\mathbf{f}(X, W) \subseteq Z$$

Proof. By definition 7:

$$\begin{aligned}
\mathbf{f}(X, W) &\subseteq \mathbf{f}^L(X, W) \oplus \mathcal{E} \subseteq \mathbf{f}^L(X, W) \oplus \bar{\mathcal{E}} = \\
&\mathbf{f}(p, c_w) \oplus G_x H \mathbf{B}^m \oplus G_w M_w \mathbf{B}^{r_w} \oplus \bar{\mathcal{E}} = \\
&(\mathbf{f}(p, c_w) + \bar{t}) \oplus G_x H \mathbf{B}^m \oplus G_w M_w \mathbf{B}^{r_w} \oplus \bar{Q} \mathbf{B}^n = \\
&p_z \oplus H_z \mathbf{B}^{m+r_w+n} = Z
\end{aligned}$$

■

4.3 Improving the obtained bound

Before introducing the main result of this subsection, the following definition is enunciated:

Definition 8 *Given matrix $E \in \mathbb{R}^{n \times n}$ and the DC functions: $f_i(x,w) = g_i(x,w) - h_i(x,w)$, $i = 1, \dots, n$, functions $g_i^E(x,w)$, $h_i^E(x,w)$, $i = 1, \dots, n$ are defined as follows:*

$$g_i^E(x,w) = \sum_{j=1}^n g_i^j(x,w), \quad h_i^E(x,w) = \sum_{j=1}^n h_i^j(x,w)$$

where

$$g_i^j(x,w) = \begin{cases} E_{i,j} g_j(x,w) & \text{if } E_{i,j} \geq 0 \\ -E_{i,j} h_j(x,w) & \text{otherwise} \end{cases}$$

$$h_i^j(x,w) = \begin{cases} E_{i,j} h_j(x,w) & \text{if } E_{i,j} \geq 0 \\ -E_{i,j} g_j(x,w) & \text{otherwise} \end{cases}$$

Lemma 2 *If E_i denotes the i -th row of matrix E , then $E_i \mathbf{f}(x,w) = g_i^E(x,w) - h_i^E(x,w)$, $i = 1, \dots, n$. Moreover, the functions $g_i^E(x,w)$ and $h_i^E(x,w)$ with $i = 1, \dots, n$ are convex.*

Proof: It is easy to see that $g_i^j(x,w) - h_i^j(x,w) = E_{i,j}(g_j(x,w) - h_j(x,w)) = E_{i,j} f_j(x,w)$. Therefore,

$$E_i \mathbf{f}(x,w) = \sum_{j=1}^n E_{i,j} f_j(x,w) =$$

$$\sum_{j=1}^n g_i^j(x, w) - h_i^j(x, w) = g_i^E(x, w) - h_i^E(x, w)$$

To finish the proof, note that by construction, $g_i^j(x, w)$ and $h_i^j(x, w)$ are convex. Thus, $g_i^E(x, w)$ and $h_i^E(x, w)$ are also convex. ■

Now, a second bounding operator to improve the results obtained by the operator presented in theorem 1 is enunciated. So, it is assumed that a zonotope Z such that $f(X, W) \subseteq Z$ has been computed. This new theorem uses the operator presented in appendix A to bound the zonotope Z by a parallelotope \tilde{P} .

Theorem 2 *Suppose that $f(X, W)$ is included in the zonotope: $Z = p_z \oplus H_z \mathbf{B}^{m+r_w+n}$. Suppose also that the parallelotope $\hat{P} = \{ x : \|\hat{E}x - \hat{q}\|_\infty \leq 1 \}$ is an outer approximation of Z ($Z \subseteq \hat{P}$) and the parallelotope P bounds the set X ($X \subseteq P$). Under this assumption, obtain:*

$$\bar{g}_i^{\hat{E}}(x, w) = g_i^{\hat{E}}(p_z, c_w) + \hat{u}_{g_i}^\top \begin{bmatrix} x - p_z \\ w - c_w \end{bmatrix}, i = 1, \dots, n$$

$$\bar{h}_i^{\hat{E}}(x, w) = h_i^{\hat{E}}(p_z, c_w) + \hat{u}_{h_i}^\top \begin{bmatrix} x - p_z \\ w - c_w \end{bmatrix}, i = 1, \dots, n$$

where \hat{u}_{g_i} , \hat{u}_{h_i} are subgradients at $(x, w) = (p_z, c_w)$ of $g_i^{\hat{E}}(x, w)$ and $h_i^{\hat{E}}(x, w)$ respectively. Compute now:

$$\gamma_i^+ = \max_{x, w \in \text{vert}(P, W)} g_i^{\hat{E}}(x, w) - \bar{h}_i^{\hat{E}}(x, w)$$

$$\gamma_i^- = \min_{x, w \in \text{vert}(P, W)} \bar{g}_i^{\hat{E}}(x, w) - h_i^{\hat{E}}(x, w)$$

where $i = 1, \dots, n$. Then

$$f(X, W) \subseteq Z \cap \tilde{P}$$

where $\tilde{P} = \{ x : \gamma_i^- \leq \hat{E}x \leq \gamma_i^+, i = 1, \dots, n \}$.

Proof: If \hat{E}_i is the i -th row of matrix \hat{E} , then $\hat{E}_i f(x, w) = g_i^{\hat{E}}(x, w) - h_i^{\hat{E}}(x, w)$ where $g_i^{\hat{E}}(x, w)$ and $h_i^{\hat{E}}(x, w)$ are convex functions by lemma 2. Bearing in mind that $\bar{g}_i^{\hat{E}}(x, w)$ and $\bar{h}_i^{\hat{E}}(x, w)$ are linear minorants of $g_i^{\hat{E}}(x, w)$ and $h_i^{\hat{E}}(x, w)$ it is clear that: $\gamma_i^- \leq \bar{g}_i^{\hat{E}}(x, w) - \bar{h}_i^{\hat{E}}(x, w) \leq \hat{E}_i f(x, w) \leq g_i^{\hat{E}}(x, w) - \bar{h}_i^{\hat{E}}(x, w) \leq \gamma_i^+ \quad \forall x, w \in X, W$. Then it is inferred that $f(X, W) \subseteq \tilde{P}$. ■

Remark 2 *Note that the parallelotope \tilde{P} obtained in theorem 2 can be used to improve the bound Z obtained by*

theorem 1. Parallelotope \tilde{P} is defined by the intersection of n strips. The operation $Z \cap \tilde{P}$ can be implemented by the intersection of Z with n strips. A new efficient operator to bound the intersection of a zonotope and a strip has been presented by the authors in (Bravo, Alamo & Camacho 2006). Given a strip and a zonotope of order r , the operator allows one to obtain a new zonotope, of order r , containing the intersection. The cited intersection operator can be used in this paper to obtain bounds of the intersection of a zonotope with a strip.

The next section boards the computation of a strip that bounds the set of states that are consistent with a given measurement.

5 Bound on the consistent state set

In this section, a bound of the consistent state set is provided. Given a measure $y_k \in \mathbb{R}^p$, the consistent state set was defined in section 2 as:

$$\mathcal{X}_{y_k} = \{ x \in \mathbb{R}^n : y_k \in \mathbf{d}(x, V) \}$$

where $V = c_v \oplus M_w \mathbf{B}^{p_v}$. Define now sets $\mathcal{X}_{y_k}(i)$, $i = 1, \dots, p$ as the region of the state space consistent with the i -th component of output y_k :

$$\mathcal{X}_{y_k}(i) = \{ x \in \mathbb{R}^n : y_k(i) \in d_i(x, V) \}$$

where $d_i(x, v)$ denotes the i -th component of $\mathbf{d}(x, v) \in \mathbb{R}^p$. With this definition it is clear that:

$$\mathcal{X}_{y_k} \subseteq \bigcap_{i=1}^p \mathcal{X}_{y_k}(i).$$

In the following it will be shown how to bound $\mathcal{X}_{y_k}(i)$ by means of a strip in the state space. If x_k belongs to the zonotope $\bar{\mathcal{X}}_k$ then the i -th component of the measured output y_k can be used to obtain a sharper bound of the state as $x_k \in \bar{\mathcal{X}}_k \cap \mathcal{X}_{y_k}(i)$. The following property shows that it is possible to bound $\bar{\mathcal{X}}_k \cap \mathcal{X}_{y_k}(i)$ by means of the intersection of $\bar{\mathcal{X}}_k$ and a strip in the state space.

Property 1 *Given the zonotope $\bar{\mathcal{X}}_k$, the measured output $y_k(i)$, and vector $c_i \in \mathbb{R}^n$, obtain a parallelotope P such that $\bar{\mathcal{X}}_k \subseteq P$, and the scalars $s_i, \sigma_i \in \mathbb{R}$ such that:*

$$s_i = \frac{\rho_i^+ + \rho_i^-}{2} \quad \sigma_i = \rho_i^+ - s_i$$

$$\rho_i^+ = \max_{x, v \in \text{vert}(P, V)} c_i^\top x - (\bar{a}_i(x, v) - b_i(x, v))$$

$$\rho_i^- = \min_{x, v \in \text{vert}(P, V)} c_i^\top x - (a_i(x, v) - \bar{b}_i(x, v))$$

Then, defining the strip $\bar{\mathcal{X}}_{y_k}(i) = \{ x : |c_i^\top x - y_k(i) - s_i| \leq \sigma_i \}$, it results that:

$$\bar{\mathcal{X}}_k \cap \mathcal{X}_{y_k}(i) \subseteq \bar{\mathcal{X}}_k \cap \bar{\mathcal{X}}_{y_k}(i)$$

Note that the convex functions $a_i(x, v)$ and $b_i(x, v)$ with $i = 1, \dots, p$ are the i -th components of $\mathbf{a}(x, v)$ and $\mathbf{b}(x, v)$ and the functions $\bar{a}_i(x, v)$ and $\bar{b}_i(x, v)$ are their linear minorants.

Proof: If $x \in \bar{\mathcal{X}}_k \cap \mathcal{X}_{y_k}(i)$ then there exists $v \in V$ such that $y_k(i) = d_i(x, v)$. Multiplying the equality by -1 and adding $c_i^\top x$:

$$c_i^\top x - y_k(i) = c_i^\top x - d_i(x, v) \subseteq [\rho_i^-, \rho_i^+] = [s_i - \sigma_i, s_i + \sigma_i]$$

Therefore, $|c_i^\top x - y_k(i) - s_i| \leq \sigma_i$ for every $x \in \bar{\mathcal{X}}_k \cap \mathcal{X}_{y_k}(i)$. ■

Note that if $d_i(\cdot, \cdot)$ is differentiable, an appropriate choice of c_i is $c_i = \frac{\partial}{\partial x} d_i(p_k, c_v)$. If not, vector c_i , along with a constant τ_i could be obtained in such a way that $c_i^\top x + \tau_i$ constitutes an affine approximation of function $d_i(\cdot, \cdot)$ in (P, V) .

In the next section, a detailed version of the new state estimation algorithm is presented.

6 Guaranteed state estimation algorithm

Suppose that an outer bound of the exact uncertain state set is available at time $k - 1$ (this bound will be denoted $\hat{\mathcal{X}}_{k-1}$). Suppose also that a measured output y_k is obtained at sample time k . Under these assumptions the following algorithm estimates an outer bound of the exact uncertain state set.

Algorithm 2

Step 1: Using theorem 1, compute a zonotope $\bar{\mathcal{X}}_k$ such that $f(\hat{\mathcal{X}}_{k-1}, W) \subseteq \bar{\mathcal{X}}_k$

Step 2: Using theorem 2, obtain a parallelotope \tilde{P} such that $f(\hat{\mathcal{X}}_{k-1}, W) \subseteq \tilde{P}$.

Step 3: Using property 1, compute an outer bound of the consistent state set \mathcal{X}_{y_k} . Denote it as $\bar{\mathcal{X}}_{y_k}$.

Step 4: Compute a zonotope $\hat{\mathcal{X}}_k \supseteq \bar{\mathcal{X}}_k \cap (\tilde{P} \cap \bar{\mathcal{X}}_{y_k})$ (see remark 2).

End of algorithm 2

The algorithm starts (first and second steps) computing the sets $\bar{\mathcal{X}}_k$ and \tilde{P} . These sets are outer bounds of the evolution of the system and they are computed using DC programming. An outer bound of the set of states

that are consistent with the new measurement y_k is obtained in step three using DC programming. Finally, an intersection operator of zonotope and strip (Bravo et al. 2006) is used in step four to obtain the outer bound $\hat{\mathcal{X}}_k$.

7 Example

A non-linear estimation example is presented here. Given the functions:

$$f_1(x_1, x_2) = -0.7x_2 + 0.1x_2^2 + 0.1x_1x_2 + 0.1exp(x_1)$$

$$f_2(x_1, x_2) = x_1 + x_2 - 0.1x_1^2 + 0.2x_1x_2$$

The system is described by the expression:

$$x_1(k+1) = f_1(x_1(k), x_2(k)) + w_1(k)$$

$$x_2(k+1) = f_2(x_1(k), x_2(k)) + w_2(k)$$

where $|w_1(k)| \leq 0.1$ and $|w_2(k)| \leq 0.1$. The measurements are:

$$y_k = x_1(k) + x_2(k) + v(k)$$

The error is bounded by $|v(k)| \leq 0.2$, $k \geq 0$. The initial state belongs to the box $3\mathbf{I}\mathbf{B}^2$ where \mathbf{I} is the identity matrix. The signal to be estimated is $z_k = [1 \ 0]x_k$. Knowing that for each $n \times n$ matrix Q , there exist two positive semidefinite $n \times n$ matrices A, B such that $x^\top Qx = x^\top Ax - x^\top Bx$ (Horst & Thoai 1999) and considering that $0.1exp(x_1)$ is a convex term, it is easy to obtain a DC representation of the considered system.

Figure 1 presents with a solid line, the evolution of the volume of the guaranteed bound of the state obtained with the new proposed method. The dashed line shows the volume obtained with the method presented in (Alamo et al. 2005). In this case, interval arithmetic is used to bound the evolution of the systems. The new proposed method improves the estimations obtained with the results of (Alamo et al. 2005). Figure 2 compares the obtained bounds on x_1 with the one corresponding to the exact uncertain state sets. Note that the exact uncertain state sets are obtained from the min and max values resulting from the uncertain evolution of a sufficiently dense cloud of points. Figure 3 shows a succession of sets $\bar{\mathcal{X}}_k$ and how the proposed algorithm reduces their volumes by intersection, obtaining sets $\hat{\mathcal{X}}_k$. Figure 4 shows a succession of sets $\hat{\mathcal{X}}_k$ obtained by the proposed algorithm.

8 Conclusions

A new approach to guaranteed state estimation for non-linear discrete-time systems with a bounded description

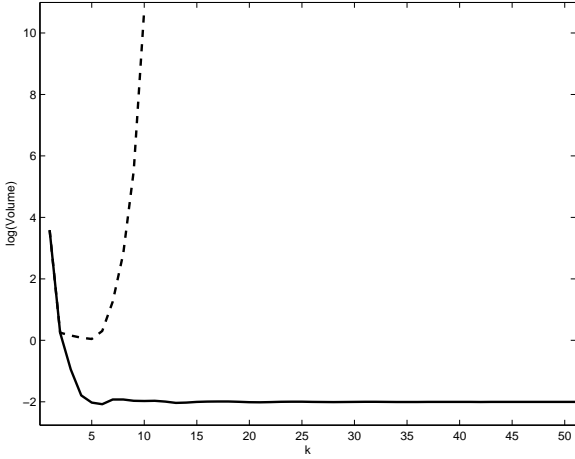


Fig. 1. Evolution of the volume of the guaranteed bound of the state.

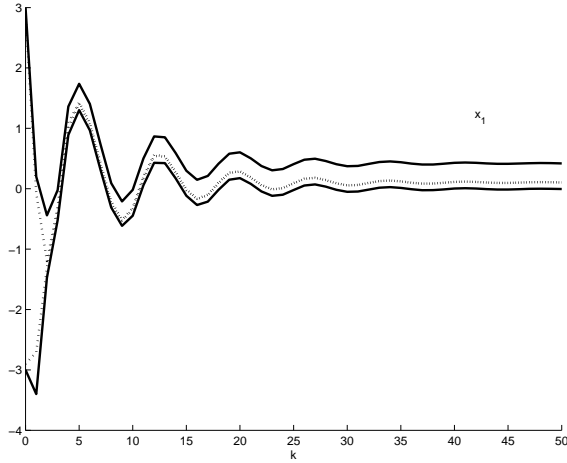


Fig. 2. Solid lines represent the guaranteed bounds of the state x_1 obtained by the presented algorithm. Dotted lines represent the bounds of x_1 obtained from the exact uncertain sets

of noise and parameters has been proposed. The algorithm bounds the set of all the states that are consistent with the measured output and the given noise and parameters. The evolution of the system is captured by zonotopes and DC programming is used to compute these zonotopes. The obtained measurements are used to intersect the computed zonotopes with strips of consistent states. Finally, an example has been provided to clarify the algorithm.

Acknowledgements

This research has been supported by CICYT DPI2006-15476-C02-01 and DPI2007-66718-C04-01.

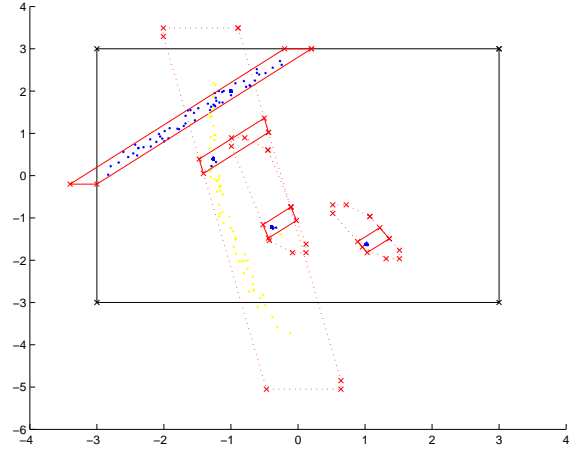


Fig. 3. Dotted lines show the sets $\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2$ and $\bar{\mathcal{X}}_3$. Solid lines represent the sets $\hat{\mathcal{X}}_0, \hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2$ and $\hat{\mathcal{X}}_3$. Dark clouds of points show the exact uncertain sets $\mathcal{X}_0, \mathcal{X}_1, \mathcal{X}_2$ and \mathcal{X}_3 . Sets $f(\mathcal{X}_0, W), f(\mathcal{X}_1, W)$ and $f(\mathcal{X}_2, W)$ are displayed as a light grey clouds of points.

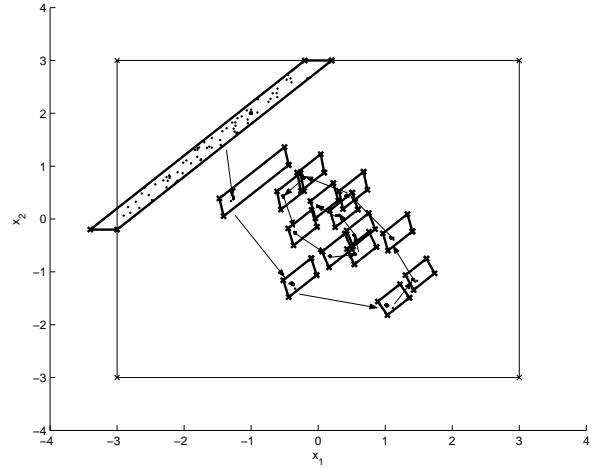


Fig. 4. Solid lines represent the sets $\hat{\mathcal{X}}_0, \hat{\mathcal{X}}_1, \dots, \hat{\mathcal{X}}_{15}$. Clouds of points show the exact uncertain sets $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_{15}$. Thin arrows represent the actual evolution of the system.

A Bounding a zonotope with a parallelotope

Lemma 3 Consider the zonotope $Z = p \oplus MB^m$ with $M \in \mathbb{R}^{n \times m}$, where $n \leq m$ and $\text{rank}(M) = n$. Consider also the singular value decomposition $M = U\Sigma V^T$, where $\Sigma = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\}$. Denote now D the diagonal matrix with components $D_{ii} = \|\sigma_i V_i\|_1, i = 1, \dots, n$, where V_i is the i -th column of matrix V . Under these assumptions it results that $Z \subseteq P = p \oplus UDB^n$.

Proof.

$$\begin{aligned} MB^m &= U\Sigma V^T B^m \\ &= U \left[\sigma_1 V_1 \ \sigma_2 V_2 \ \dots \ \sigma_n V_n \right]^T B^m \subseteq UDB^n \end{aligned}$$

Note that the last inclusion relies on the fact that $\sigma_i V_i^\top \mathbf{B}^m \subseteq \|\sigma_i V_i\|_1 \mathbf{B}^1 = D_{ii} \mathbf{B}^1$, where $\|\cdot\|_1$ denotes the vectorial norm equal to the sum of the absolute values of the components of a given vector. ■

As it will be justified in what follows, the assumption $n \leq m$ and $\text{rank}(M) = n$ is not restrictive. Consider the zonotope $\tilde{Z}(\epsilon) = p \oplus \begin{bmatrix} M & \epsilon \mathbf{I} \end{bmatrix} \mathbf{B}^{m+n} = p \oplus \tilde{M} \mathbf{B}^{m+n}$. It is clear that $Z = \tilde{Z}(0)$ and $Z \subseteq \tilde{Z}(\epsilon)$, $\forall \epsilon$. Moreover, \tilde{M} satisfies the assumptions of the lemma for every $\epsilon \neq 0$. Therefore, using lemma 3, it is possible to obtain for a given $\epsilon \neq 0$ a parallelotope $\tilde{P}(\epsilon)$ such that $Z \subseteq \tilde{Z}(\epsilon) \subseteq \tilde{P}(\epsilon)$. Choosing ϵ such that it is different from zero but arbitrarily small, an appropriate parallelotope that bounds Z can be obtained.

References

- Adjiman, C. & Floudas, C. (1996), ‘Rigorous convex underestimators for general twice-differentiable problems’, *J. Global Optimization* **9**, 23–40.
- Alamo, T., Bravo, J. & Camacho, E. (2005), ‘Guaranteed state estimation by zonotopes’, *Automatica* **41**(6), 1035–1043.
- Boyd, S. & Vandenberghe, L. (2004), *Convex Optimization*, Cambridge University Press.
- Bravo, J., Alamo, T. & Camacho, E. (2006), ‘Bounded error identification of systems with time-varying parameters’, *IEEE Transactions on Automatic Control* **51**(7), 1144–1150.
- Calafiore, G. (2005), ‘Reliable localization using set-valued nonlinear filters’, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* **35**(2), 189–197.
- Chisci, L., Garulli, A., Vicino, A. & Zappa, G. (1998), ‘Block recursive parallelotopic bounding in set membership identification’, *Automatica* **34**, 15–22.
- Chisci, L., Garulli, A. & Zappa, G. (1996), ‘Recursive state bounding by parallelotopes’, *Automatica* **32**, 1049–1056.
- Combastel, C. (2003), A state bounding observer based on zonotopes, in ‘Proceedings of 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.
- El Ghaoui & Calafiore, G. (2001), ‘Robust filtering for discrete-time system with bounded noise and parametric uncertainty’, *IEEE Transactions on Automatic Control* **46**(7), 1084–1089.
- Garulli, A., Tesi, A. & Vicino, A. (1999), *Robustness in Identification and Control*, Springer Verlag, Berlin Heidelberg, Germany.
- Horst, R. & Thoai, N. (1999), ‘Dc programming: Overview’, *Journal of Optimization Theory and Applications* **103**(1), 1–43.
- Jaulin, L. (2002), ‘Nonlinear bounded-error state estimation of continuous-time system’, *Automatica* **36**(7), 1079–1082.
- Kieffer, M., Jaulin, L. & Walter, E. (2002), ‘Guaranteed recursive nonlinear state estimation using interval analysis’, *International journal of adaptive control and signal processing* **16**, 193–218.
- Kuntsevich, V. & Lychak, M. (1985), *Synthesis of Optimal and Adaptive Control Systems: The Game Approach [in Russian]*, Naukova Dumka, Kiev.
- Kurzhanski, A. & Valyi, I. (1996), *Ellipsoidal Calculus for Estimation and Control*, Birkhäuser, Boston, Massachusetts.
- Milanese, M., Norton, J., Piet-Lahanier, H. & Walter, E. (1996), *Bounding Approaches to System Identification*, Plenum Press, New York.
- Montgomery, H. (1989), ‘Computing the volume of a zonotope’, *Amer. Math. Monthly* **97**, 431.
- Moore, R. (1966), *Interval Analysis*, Prentice-Hall, Englewood Cliffs, NJ.
- Pinter, J. (1996), *Global Optimization in Action*, Kluwer Academic Publisher, Dordrecht.
- Puig, V., Cugueró, P. & Quevedo, J. (2001), Worst-case estimation and simulation of uncertain discrete-time systems using zonotopes, in ‘Proceedings of European Control Conference’, Portugal.
- Raković, S. V. & Mayne, D. Q. (2004), State estimation for piecewise affine, discrete time systems with bounded disturbances, in ‘Submitted to the 43rd IEEE Conference on Decision and Control’, Atlantis, Paradise Island, Bahamas.
- Rockafellar, R. (1970), *Convex Analysis*, Princeton University.
- Savkin, A. & Petersen, I. (1998), ‘Robust state estimation and model validation for discrete-time uncertain system with a deterministic description of noise and uncertainty’, *Automatica* **34**(2), 271–274.
- Schweppe, F. (1968), ‘Recursive state estimation: Unknown but bounded errors and system inputs’, *IEEE Transactions on Automatic Control* **13**, 22–28.
- Shephard, G. (1974), ‘Combinatorial properties of associated zonotopes’, *Canadian Journal of Mathematics* **26**, 302–321.
- Tuy, H. (1995), *Handbook of Global Optimization*, Kluwer Academic Publishers, chapter D.C. Optimization: Theory, Methods and Algorithms, pp. 149–216.
- Tuy, H. (1998), *Convex Analysis and Global Optimization*, Kluwer Academic Publisher, Dordrecht.