2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

# Game-Based Inner Approximation of Maximal Output Admissible Sets under References Unknown in Advance

Kiminao Kogiso, Koji Kita and Kenji Sugimoto

*Abstract*— **This paper proposes a game-based inner approximation method that efficiently arranges multiple ellipsoids approximating a parameterization of maximal output admissible sets in terms of a reference vector. The approximation is contributive when considering a tracking control problem for constrained systems in which a reference value is not given in advance. Our idea for the approximation is to characterize search points of the ellipsoids in a state based game and a log-linear learning algorithm by dynamics of the constrained system. Consequently, an efficient arrangement of the ellipsoids can be achieved at a lower computation cost. A numerical example is finally illustrated to make sure of the consequence, compared to our previous works.**

## I. Introduction

Constrained control methods are required to enforce a state and/or control constraint, and the methods such as reference governor [1]–[3], model predictive control [4], [5], and multimode switching control [6], work for preventing degradation of control performance to an extent possible. Most of them use a mutual tool that plays a key role in constraint satisfaction for an infinite horizon, which is called a maximal output admissible set [7], [8].

The maximal output admissible set is very beneficial for describing the constraint satisfaction as a necessary and sufficient condition, while we have to take care of two requirements about how to use. One is that in a control problem setting a reference is constant and its value is known in advance. This is because the set is basically defined by the reference value, where it is called *known reference*. On the other hand, when considering a situation where the reference information is not given in advance, *unknown reference*, an alternative definition and its calculation way are needed. The other requirement is about time that it takes to calculate the set. In the case of the unknown reference, it is natural to promptly perform online calculation as soon as the reference value is confirmed by measurement. However, this is not necessarily followed, especially in mechanical control systems, because the calculation time by the methods [7], [8] is generally much more than a millisecond-order sampling period. The relaxation of the requirements, therefore, can lead to success in broadening control problems that we can handle, with not only the known reference but also the unknown one.

Straightforwardly extending a definition of the output admissible set to the unknown reference case is possible, but the resulting is not so much useful. The set may not exist for a relatively large range of allowable reference [8], or, even if it existed, the resulting output admissible set might be too small and lead to conservativeness in the constraint satisfaction, since the constraint must be fulfilled for any enormous reference changes. In other words, this is caused by requiring being invariant regarding the reference changes as well. There are some related works proposed so far. In [3], [9], limiting a variety of intended constraints achieves realtime calculation of the maximal output admissible set in the unknown reference case. In [10], an ellipsoidal level set of an appropriate Lyapunov function is proposed. This approach considers the reference as a bounded external disturbance, and the resulting level set is invariant regarding state evolutions as well as reference changes. When considering the large range of the reference, accordingly, it will produce the conservative consequence.

As for different approaches to the relaxation, our previous work [11] has proposed a method that generates multiple ellipsoids approximating a parameterization of the maximal output admissible sets in terms of a reference vector. The generated ellipsoids are arranged inside the parameterization. To improve an approximation accuracy, moreover, our another work [12] has proposed a game-based learning algorithm that finds the better efficient arrangement of the multiple ellipsoids. The work also has clarified that it is too difficult to reach the optimum arrangement. These methods, however, succeed to extend the methods [3], [9] in point of requiring no limitations on the intended constraints. The previous works [11], [12] do not take care about calculation time that it takes to acquire the approximation. From a point of view of practical uses or applications, consideration of reducing the computation cost is worth conducting, in preparation for a situation where a system's size increases.

This paper, then, proposes an inner approximation method that efficiently arranges the multiple ellipsoids at a lower computation cost. An union of the arranged ellipsoids is a set that approximates the parameterization of the maximal output admissible sets in terms of the reference. An essential difficulty of this problem is that although crossovers among the ellipsoids are allowed, they must be inside the parameterization, where the crossovers are difficult to compute. Our idea to overcome is to construct an approximation algorithm based on a state-based game [13], [14] and a log-linear learning algorithm [15]. This idea supports specification of search points preferable for the ellipsoid's arrangement in the game and also does efficient learning with dynamics of the constrained control system. Consequently, the number of

This work was partially supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 25709014.

The authors is with Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayamacho, Nara, Ikoma 6300192, Japan kogiso@is.naist.jp

the search points can be reduced and at the same time it can be easy to reach points near facets or vertices of the parameterization (polytope) where larger ellipsoids contributive to approximation accuracy improvement can be made. In other words, this consequence helps reduce calculation time that it takes to obtain the approximation at the similar accuracy as one of the previous work [12]. A numerical example is finally illustrated to make sure that the efficient ellipsoid arrangement is achieved and simultaneously the computation cost is lower, compared to the results obtained by [11], [12].

## II. Parameterization of Maximal Output Admissible Sets in Terms of Reference

This section starts with an already-designed tracking control system with state and control constraints under a constant but unknown reference. With these information, this section explains a set that we want to obtain and a difficulty to obtain.

### A. Tracking Control System with Constraints

Let us consider the linear closed-loop system

$$\tilde{x}(k+1) = A\tilde{x}(k) + Bw(k), \tag{1a}$$

$$\tilde{z}_1(k) = C_1\tilde{x}(k), \tag{1b}$$

$$\tilde{z}_0(k) = C_0\tilde{x}(k) + D_0 w(k), \tag{1c}$$

with a reference $w \in \Re^{p_1}$ and a pointwise-in-time constraint: $\tilde{z}_0(k) \in \tilde{Z}\ \forall k \in \mathcal{Z}^+$, where $k \in \mathcal{Z}^+ := \{0, 1, 2, \cdots\}$, $\tilde{x} \in \Re^n$ is a state vector, the initial state $\tilde{x}(0) = \tilde{x}_0$ is given, the reference $w(k) = \bar{w}\ \forall k \in \mathcal{Z}^+$ is constant but unknown in advance, $\tilde{z}_1 \in \Re^{p_1}$ is a controlled output, and $\tilde{z}_0 \in \Re^{p_0}$ is a vector to be constrained within the prescribed convex set $\tilde{Z} \subset \Re^{p_0}$; we also have $M_{\tilde{z}} \in \Re^{s_{\tilde{z}} \times p_0}$ and $m_{\tilde{z}} \in \Re^{s_{\tilde{z}}}$, i.e., $\tilde{Z} := \{\tilde{z}_0 \in \Re^{p_0} \mid M_{\tilde{z}}\tilde{z}_0 \le m_{\tilde{z}},\ m_{\tilde{z}} > 0\}$. Assumed that $A$ is invertible and that the control system from $w$ to $\tilde{z}_1$ has an integral property, which is to achieve tracking, i.e., $\tilde{z}_1 \to \bar{w}$ as $k \to \infty$. Note that in the absence of the constraint, the constrained control system (1) is supposed to be stabilized by an appropriate control design, so that the equilibrium is asymptotically stable.

Introducing the error variables $x := \tilde{x} - \tilde{x}_e(\bar{w})$, $z_1 := \tilde{z}_1 - \tilde{z}_{1e}(\bar{w})$, and $z_0 := \tilde{z}_0 - \tilde{z}_{0e}(\bar{w})$, system (1) is transformed into the following autonomous system:

$$x(k+1) = Ax(k), \tag{2a}$$

$$z_1(k) = C_1 x(k), \tag{2b}$$

$$z_0(k) = C_0 x(k), \tag{2c}$$

with the constraint: $z_0(k) \in Z(\bar{w})\ \forall k \in \mathcal{Z}^+$, where the definitions of each matrix and vector are $\tilde{x}_e(\bar{w}) := (I - A)^{-1}B\bar{w}$, $\tilde{z}_{1e}(\bar{w}) := C_1\tilde{x}_e(\bar{w})$, $\tilde{z}_{0e}(\bar{w}) := C_0\tilde{x}_e(\bar{w}) + D_0\bar{w}$, and

$$Z(\bar{w}) := \{z_0 \in \Re^{p_0} \mid M_{\tilde{z}} z_0 \le m_{\tilde{z}} - M_{\tilde{z}}\Phi\bar{w}\},$$
$$\Phi := C_0(I - A)^{-1}B + D_0. \tag{3}$$

Note that $Z(\bar{w}) = \tilde{Z}$ holds if $\Phi$ in (3) is zero.

For a particular value of the constant reference, the constraint satisfaction in the steady state depends on the prescribed constraint condition, $Z(\bar{w})$ (or $\tilde{Z}$). Technically, a reference setting of the control system that leads to a situation being out of the satisfaction in the steady state is impracticable. Then, we have the following assumption derived from $\tilde{z}_{0e}(\bar{w}) = M_{\tilde{z}}\Phi\bar{w} \in \tilde{Z}$.

*Assumption 1:* The constant $\bar{w}$ satisfies the following: $\bar{w} \in W := \{\bar{w} \in \Re^{p_1} \mid M_{\tilde{z}}\Phi\bar{w} \le m_{\tilde{z}}\}$, where $W \subset \Re^{p_1}$ is a convex set of allowable references.

### B. Maximal Output Admissible Set

The maximal output admissible set is known to be a powerful tool that can be used to analyze the satisfaction of the pointwise-in-time constraint. Its definition is given below.

*Definition 1:* **(Maximal Output Admissible Set)** Let $z_0(t; x_0, \bar{w})$ denote an output $z_0$ of (2c) for an initial state $x(0) = x_0$[1] and the constant reference $\bar{w} \in W$. Define a $\bar{w}$-dependent maximal output admissible set [7], [8] as

$$O_\infty(\bar{w}) = \{x_0 \in \Re^n \mid z_0(t; x_0, \bar{w}) \in Z(\bar{w})\ \ \forall t \in \mathbb{Z}^+\}.$$

The maximal output admissible set is convex polyhedral and invariant, and can be calculated by using linear-programming-based computational methods. As a result, it can be written as $O_\infty(\bar{w}) = \{x \in \Re^n \mid Mx \le m\}$ with appropriate $M \in \Re^{s \times n}$ and $m \in \Re^s$. Additionally, $0 \in O_\infty(\bar{w})$. This set has the invariance, which provides the following powerful theorem [7], [8].

*Theorem 1:* $\forall \bar{w} \in W,\ z_0(k) \in Z(\bar{w})\ \forall k \in \mathcal{Z}^+ \Leftrightarrow x_0 \in O_\infty(\bar{w})$.

This theorem is effective in the case where information on the reference value is given before starting control operations. Not in the case, real-time calculation of the set is needed, while it is not always possible due to its computation cost. Accordingly, we have to develop a new scheme for quickly calculating the set as soon as the reference value is obtained. This motivates to parameterize the maximal output admissible sets in terms of a reference vector. Actually, some of papers that have the similar motivation [3], [9], have considered a specific case $\Phi \equiv 0$, in which the parameterization can be realized by parallel shift every time the reference changes. This paper also avoids direct computation but considers a general case in $\Phi$.

### C. Parameterization in Terms of Reference

The philosophy that the maximal output admissible set is parameterized in terms of the reference vector must be useful for real-time acquisition of the parameterized set. The parameterized set $\mathcal{O}_\infty \subset \Re^{n+p_1}$ is then defined in an augmented space:

$$\mathcal{O}_\infty := \bigcup_{w \in W} (O_\infty(w), w). \tag{4}$$

This expresses a set that consists of an infinite number of the maximal output admissible sets along the $w$-axis, and the set is convex [11], where it is called an *ideal*

[1] Since $x_0 := \tilde{x}_0 - \tilde{x}_e(\bar{w})$, it is preferable to say "for an initial state $\tilde{x}_0$." However, we use $x_0$ for simplifying the notation.

*parameterization* in this paper. Importantly, since the set $W$ is infinite, $\mathcal{O}_\infty$ is generally difficult to compute. We then contrive to make an efficient approximation by multiple ellipsoids, also considering a good approximation accuracy in the sense that the individual ellipsoids are made as big as possible.

## III. Policy for Inner Approximation with Multiple Ellipsoids

This section defines an finitely-sampled version of the ideal parameterization and explains an inner-approximation technique of a convex polytope by a single ellipsoid. A technical procedure of the approximation by the multiple ellipsoids will appear in the next section.

With $l$ references sampled, the following set is defined:

$$\mathcal{O}_l := \bigcup_{i=1}^{l} (O_\infty(\bar{w}_i), \bar{w}_i) \quad \bar{w}_i \in W,$$

where $l$ is a sufficiently large non-negative integer and the references take different values each other. This set is unconnected in the augmented space, but is numerically tractable differently from the ideal parameterization because $l$ is finite. Now, letting $\mathcal{O}_{\rm CH}$ denote a convex hull of $\mathcal{O}_l$, from a convexity of $\mathcal{O}_\infty$, the following inclusion holds [11].

*Proposition 1:* $\mathcal{O}_l \subset \mathcal{O}_{\rm CH} \subset \mathcal{O}_\infty$.

As $l$ increases, $\mathcal{O}_l$ converges to $\mathcal{O}_\infty$ in the sense of Hausdorff norm, and at the same time $\mathcal{O}_{\rm CH}$ also converges to $\mathcal{O}_\infty$. The large number of $l$, however, causes explosive growth of matrix and vector sizes to express $\mathcal{O}_{\rm CH}$, which is not adequate in implementation and real-time operations. Moreover, to keep the invariance of the maximal output admissible set, the ellipsoids to be generated must be placed inside the ideal parameterization, at least inside the set $\mathcal{O}_{\rm CH}$. This is motivated to consider inner approximation of the ideal parameterization by the multiple ellipsoids.

*Remark 1:* The set $\mathcal{O}_l$ is influenced by how and how many references, $\bar{w}_i \in W$, are sampled. Under sufficiently large number of randomly-sampled references, $\mathcal{O}_l$ gets close to $\mathcal{O}_\infty$ as $l$ increases. This paper, therefore, assumes that $l$ is so large that the resulting convex hull $\mathcal{O}_{\rm CH}$ is quite close to the ideal parameterization. A detailed discussion from an analytical point of view will be made in a different chance.

The inner approximation procedure that this paper proposes, consists of two parts: the first one employs an optimization to obtain a maximum-volume ellipsoid inside the given polytope [16], [17],

$$\max_{P_0,\ c_0} \ \{\det(P_0)\}^{\frac{1}{n+p_1}} \ \ \text{s.t.} \ \ E(P_0, c_0) \subset \mathcal{O}_{\rm CH}, \qquad (5)$$

where an ellipsoid $E \subset \Re^{n+p_1}$ is defined as follows,

$$E(P_0, c_0) := \{(x, w) \in \Re^{n+p_1} \mid \begin{bmatrix} x \\ w \end{bmatrix} = P_0\Psi + c_0,\ ||\Psi||_2 \le 1\},$$

with a positive definite symmetric matrix $P \in \Re^{(n+p_1)\times(n+p_1)}$, a vector $c \in \Re^{n+p_1}$, and $\Psi \in \Re^{n+p_1}$. The resulting optimizers $P_0^*$ and $c_0^*$ provide the maximum-volume ellipsoid $E(P_0^*, c_0^*)$, which holds the following inclusion: $E(P_0^*, c_0^*) \subset \mathcal{O}_\infty$, from *Proposition 1*. In the first part, therefore, the ideal parameterization is approximated by the single maximum-volume ellipsoid, which is labeled as $E_0^*$.

*Remark 2:* It is ideal to acquire such an ellipsoid, solving an optimization problem replaced with $\mathcal{O}_\infty$: $\max_{P_0,\ c_0}\{\det(P_0)\}^{\frac{1}{n+p_1}}$ subject to $E(P_0, c_0) \subset \mathcal{O}_\infty$. This formulation with $\mathcal{O}_\infty$ is unfortunately difficult to numerically solve. This is why $\mathcal{O}_{\rm CH}$ is used in (5).

The second part is to attempt improving an approximation accuracy by generating $N$ ellipsoids plus the maximum-volume one. The accuracy considered in this approach is a filling rate of how much the ideal parameterization is filled with the multiple ellipsoids. Given ellipsoids' centers $c := (c_1, c_2, \cdots, c_N)$, an union of the ellipsoids is defined below as an approximation of the ideal parameterization,

$$E_s(c) := E_0^* \cup \bigcup_{i=1}^{N} E_i(P_i, c_i).$$

Given $c$, we can obtain the corresponding maximum-volume ellipsoids $E_i(P_i^*, c_i)$ $i \in \{1, 2, \cdots, N\}$ via the optimization (5) under fixed $c_i$. What matters is how the centers are updated to reach a (locally) optimum set $E_s^* \subset \mathcal{O}_\infty$,

$$E_s^* := E_s(c^*) = E_0^* \cup \bigcup_{i=1}^{N} E_i(P_i^*, c_i^*),$$

regarding the approximation accuracy.

The above statement is our policy for how to approximate the ideal parameterization using the multiple ellipsoids. In the next section, we will focus on how to make a decision of the center using a game theory framework.

## IV. The Inner Approximation via Game Theory

The approximation accuracy of the filling rate can be written as a function $\phi$,

$$\phi(c) := \frac{\mathrm{vol}(E_s(c))}{\mathrm{vol}(\mathcal{O}_\infty)} \quad (0 \le \phi \le 1), \qquad (6)$$

where $\mathrm{vol}(\mathcal{O}_\infty)$ denotes a volume of the set $\mathcal{O}_\infty$. Then, the inner-approximation problem of our interest can be described in the following optimization problem:

$$\max_{c} \phi(c) \ \text{s.t.}\ E_s(c) \subset \mathcal{O}_{\rm CH}, \qquad (7)$$

where $\mathcal{O}_{\rm CH}$ is supposed to be very close to $\mathcal{O}_\infty$ under the sufficiently large $l$, as stated in *Proposition 1* and *Remark 1*. A class of this problem is an optimal inner approximation of a convex polytope by multiple ellipsoids. As far as the authors know, unfortunately, it seems to be impossible to solve the problem in one-shot manner because the matrix $P_i^*$ depends on $c_i$, i.e., $P_i^* = \arg\max_{P_i}\{\det(P_i)\}^{\frac{1}{n+p_1}}$ subject to $E(P_i, c_i) \subset \mathcal{O}_{\rm CH}$. This paper, then, employs a game-based approach to obtain the efficient inner approximation $E_s^*$ via the respective players'actions (the center $c$) and the players' utility (maximizing an ellipsoid's volume). The game theoretic formulation is rational in case of handling the dependency between $c_i$ and $P_i$ and of searching $c_i$ that has a restriction.

*A. State Based Game with Restricted Action Set*

An interesting feature of the game-based approach is that the players make respective best actions to maximize the respective players' utility functions so that it attempts to maximize the function $\phi$ in (7), where assumed the players do not know the function and their actions are rationally made. This distributed process is quite useful when the function is directly incomputable. On the other hand, this approach must take care of how to design the utility functions, which is important for the achievement.

A state based game [14], an extension of a finite strategic form game, is used to formulate an algorithm for trying to solve the original problem of (7). The game $\mathcal{G}$ with $N$ players is denoted as a tuple $\mathcal{G} = (\mathcal{P}, \mathcal{X}, \mathcal{A}, \{f_i\}_{i\in\mathcal{P}}, \{U_i\}_{i\in\mathcal{P}}, \{R_i\}_{i\in\mathcal{P}})$. Its definition is as follows.

**Player.** $c_i$ is the player expressing the center of the ellipsoid to be generated. A set of the players is denoted as $\mathcal{P} := \{1, 2, \cdots, N\}$. In this game, increasing and decreasing the players are not allowed through plays repeated, i.e., $N$ is given and constant, and the players are assumed to be rational.

**State Space.** $\mathcal{X}_i \subset \Re^{n+p_1}$ corresponds to the augmented state space where the player moves, i.e., $c_i = [x_i^{\mathrm{T}} \;\; w_i^{\mathrm{T}}]^{\mathrm{T}} \in \mathcal{X}_i$, where $x_i$ and $w_i$ denote the state variable and the reference value of the $i$th player. For $c$, the profile other than player $i$ is denoted by $c_{-i} = (c_1, \cdots, c_{i-1}, c_{i+1}, \cdots, c_N)$ and we use the notation $c = (c_i, c_{-i}) \in \mathcal{X} := \Pi_{i\in\mathcal{P}}\mathcal{X}_i$.

**Action Set.** $\mathcal{A}_i \subset \mathcal{Z}^{p_1+1}$ is a finite set of an action $a_i = (l_i, d_i)$ with integer $l_i \in [-L_i, L_i]$ $(L_i > 0)$ and $d_i \in \{-1, 0, 1\}^{p_1}$. Similarly, for $a \in \mathcal{A} := \Pi_{i\in\mathcal{P}}\mathcal{A}_i$, let $a_{-i}$ denote the profile of player other than player $i$, i.e., $a_{-i} = (a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_N)$, and $a = (a_i, a_{-i})$ is also used.

**State Dynamics.** $f_i : \mathcal{X}_i \times \mathcal{A}_i \to \mathcal{X}_i$ expresses a stochastic update process for the centers,

$$c_i^+ = f_i(c_i, a_i) = \begin{bmatrix} A^{l_i} x_i \\ w_i + \Delta w d_i \end{bmatrix} \quad \forall i \in \mathcal{P},$$

where the matrix $A$ is the same as one of (2b) and $\Delta w > 0$ is a constant length of a grid in the reference. Its dynamics is related to a learning algorithm shown in the next section.

**Utility.** $U_i : \mathcal{X} \times \mathcal{A} \to \Re$ is defined as follows,

$$U_i(c,a) := \begin{cases} \max\limits_{P_i} \det\{P_i\}^{\frac{1}{n+p_1}} & \text{s.t. } E_i(P_i, c_i^+) \subset \mathcal{O}_{\mathrm{CH}} \\ & \text{if } \;\; c_i^+ \notin \Theta(c_{-i}, a_{-i}), \\ 0 & \text{if } \;\; c_i^+ \in \Theta(c_{-i}, a_{-i}), \end{cases}$$

where $\Theta(c_{-i}, a_{-i}) := E_0^* \cup \bigcup_{j\neq i} E_j(P_j^*, c_j^+)$ $\forall i, j \in \mathcal{P}$. Why the utility function is introduced is to make an effort to improve the approximation accuracy $\phi$ in an effective way of excluding a situation in which the ellipsoid's center $c_i^+$ is placed inside the other ellipsoids, i.e.,

$$c_i^+ \in \Theta(c_{-i}, a_{-i}) \subset \mathcal{O}_\infty \quad \forall i \in \mathcal{P}.$$

This situation is possibly not contributive to the accuracy improvement because a common region among the ellipsoids must get large. The rational player is not motivated to select an action $a_i'$ that yields $c_i^+$ giving $U_i(c, (a_i', a_{-i})) = 0$ so that the definition of the utility is workable for excluding the situation not conductive to the improvement.

**Restricted Action Set.** $R_i : \mathcal{X} \to \mathcal{A}_i$ is defined as follows,

$$R_i(c) := \{a_i' \in \mathcal{A}_i \mid f_i(c_i, a_i') \in \{\mathcal{O}_{\mathrm{CH}} \setminus \Theta(c_{-i}, a_{-i})\}\},$$

where it is a finite set because $\mathcal{A}_i$ is also finite. In repeated game plays, the next center is always chosen from among its candidates produced by the restricted action set. What its definition says is that the candidates are restricted onto a state trajectory of the dynamical system (2) in discrete time under an allowable reference value. This restriction brings two merits; one is to reduce the whole number of search points more than by defining a cubic lattice [18], and the other is that there are opportunities that the player can pass over the other ellipsoids so that he/she can easily reach points near facets or vertices of the parameterization (convex polytope) where larger ellipsoids contributive to approximation accuracy improvement can be made. As a result, we can expect that the definition helps reducing the computation time. This point will be checked in a numerical example section.

An equilibrium concept in the state based game is used to acquire the efficient approximation. First, introducing a Markovian state transition function $\Phi : \mathcal{X} \times \mathcal{A} \to \Delta(\mathcal{X})$, the state transitions can be written, i.e., $c^+ \sim \Phi(c, a)$, which means that the state is selected randomly according to the probability distribution $\Phi(c, a)$. Next, a reachable set $X(a^0|c^0)$ is also introduced, which denotes all states that may eventually emerge along the state trajectory starting from the pair $(c^0, a^0)$. Here, a state $c \in X(a^0|c^0)$ if and only if there exists a time $k > 0$ such that

$$\begin{aligned} &\mathtt{Pr}\left[c(k) = c \,|\, c(0) = c^0, \right. \\ &\qquad \left. c^+ \sim \Phi(c(t), a) \;\forall t \in \{0, 1, \cdots, k-1\}\right] > 0. \end{aligned}$$

And then, we can make a general definition of a recurrent state equilibrium in the state based game $\mathcal{G}$ and show its existence.

*Definition 2:* [14] The action state pair $(c^*, a^*)$ is a recurrent state equilibrium with respect to the state transition process $\Phi(\cdot)$ if the following two conditions are satisfied: (i) the state $x^*$ is recurrent according to the process $\Phi(c, a = a^*)$ with initial state $c(0) = c^*$. In terms of reachable states this implies that $c^* \in X(a^*|c)$ for every state $c \in X(a^*|c^*)$. (ii) For every $i \in \mathcal{P}$ and ever state $c \in X(a^*|c^*)$, $U_i(c, (a_i^*, a_{-i}^*)) = \max_{a_i\in\mathcal{A}_i} U_i(c, (a_i, a_{-i}^*))$.

*Theorem 2:* A recurrent state equilibrium exists in the state based game $\mathcal{G}$.

*Proof:* An evolution of the state $c_i$ follows the dynamics $f_i$ with the action $a_i$ inside an area $\mathcal{O}_{\rm CH} \setminus \Theta(c_{-i}, a_{-i})$, which is not void because $\mathcal{O}_{\rm CH}$ is a convex polytope and $\Theta(c_{-i}, a_{-i})$ is a sum of the ellipsoids. The number of the allowable pairs $(c, a)$ is finite because $c$ is restricted on the trajectory of $f_i$ and $\mathcal{A}_i$ is finite. A value of the utility is then bounded, i.e., for each $i \in \mathcal{P}$, $0 \leq U_i(c, a) < \det\{P_0^*\}^{\frac{1}{n+p_1}}$. We, therefore, can choose a certain pair $(c^*, a^*)$ and the process $P(c^*, a^*)$ generates $c^{**}$. Consequently, the pair $(c^{**}, a^*)$ is a recurrent state equilibrium. ■

### *B. Log-Linear Learning in the Game*

The game-theoretic learning algorithm is constructed for obtaining the efficient approximation $E_s^*$, where $\hat{k}$ is an iteration step by repeated play of the state based game. The algorithm is adequately customized by a restricted spatial adaptive play, keeping an essence of it. A terminal condition of the algorithm uses a best response for any pair $(c, a) \in \mathcal{X} \times \mathcal{A}$ as $BR_i(a; c) := \{a_i' \in \mathcal{A}_i \,|\, U_i(c, (a_i', a_{-i})) > U_i(c, a)\}$. If $BR_i(a; c)$ is void, it is a sign that all players cannot improve their ellipsoids' volume anymore.

---

**Game-Theoretical Learning Algorithm**

---

**Step 0.** Set iteration counter $\hat{k} = 0$ and generate $E_0^*$.
**Step 1.** Seed $c$ randomly so as to satisfy $c_i \in \{\mathcal{O}_\infty \setminus \Theta(c_{-i}, a_{-i})\}$ for all $i \in \mathcal{P}$ and $\hat{k} = 1$.
**Step 2.** $i \in \mathcal{P}$ is randomly chosen with equal probability.
**Step 3.** At $i$, a trial action $a_i'$ is randomly selected from its allowable set $R_i(c)$ with the following probabilities.

$$\texttt{Pr}[a_i' = a_i] = 1/q, \quad a_i \in \{R_i(c(\hat{k}-1)) \setminus a_i^z\},$$
$$\texttt{Pr}[a_i' = a_i^z] = 1 - (|R_i(c)| - 1)/q,$$

where $q = |\{c_i^+ \,|\exists c_i,\, c_i^+ = f_i(c_i, a_i)\ \forall a_i \in \mathcal{A}_i\}|$ and $a_i^z = (0, 0)$.
**Step 4.** $\hat{c}_i^+$ is determined via choosing an action with the following probabilities,

$$\texttt{Pr}[a_i(\hat{k}) = a_i'] = \frac{\exp\{\beta U_i(c_i(\hat{k}-1), (a_i', a_{-i}(\hat{k}-1)))\}}{D},$$

$$\texttt{Pr}[a_i(\hat{k}) = a_i^z] = \frac{\exp\{\beta U_i(c(\hat{k}-1), (a_i^z, a_{-i}(\hat{k}-1)))\}}{D},$$

where $D = \exp\{\beta U_i(c_i(\hat{k}-1), (a_i', a_{-i}(\hat{k}-1)))\} + \exp\{\beta U_i(c(\hat{k}-1), (a_i^z, a_{-i}(\hat{k}-1)))\}$ and $\beta$ is a positive scalar exploration parameter.
**Step 5.** If $BR_i(a(\hat{k}), c(\hat{k}-1))$ are void for all $i \in \mathcal{P}$, then this algorithm terminates. Otherwise, go to **Step 2** with $\hat{k} = \hat{k}+1$.

---

The obtained pair $(c^*, a^*)$ for all players that yields the ellipsoids, which is an efficient approximation (locally sub optimal) $E_s^*$ satisfying the following feature.

*Theorem 3:* Given the constrained control system under *Assumption 1*, the presented algorithm provides the equilibrium $(c^*, a^*)$ and the resulting approximate set $E_s^*$ holds $E_s^* \subset \mathcal{O}_\infty$. Then, $x \in E_s^*$ under $\bar{w}$ implies the constraint satisfaction $z_0(k) \in Z(\bar{w})\, \forall k \in \mathcal{Z}^+$.

The next section will confirm the proposed approach can provide the approximation $E_s^*$ and will discuss the computation time to get the approximation.

## V. Numerical Example

A numerical example is used to illustrate effectiveness of the proposed game-based approach, comparing to our previous works [11], [12].

### *A. Description of Constrained Control System*

The example is a position servomechanism of a DC motor with two constraints. The motor dynamics from an input voltage to a rotation angle is experimentally identified, to be $P(s) = \frac{49}{s(s+1.5)}$. One of the constraints pertains to the input voltage to the motor and the other is related to the allowable range of its rotation angle; they can be expressed as follows,

$$|u| \leq 5 \text{ (V)} \qquad |\theta| \leq 5/6\pi \text{ (rad)}.$$

By discretizing the dynamics with a zero-order hold of sampling period 5.0 (ms), we can obtain information on the plant dynamics in discrete time and on the constraint condition,

$$\tilde{x}(t+1) = \begin{bmatrix} 1 & 0.0099 \\ 0 & -0.9851 \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} 0.0024 \\ 0.4863 \end{bmatrix} u(t),$$
$$\tilde{z}_1(t) = \tilde{x}(t),$$
$$\tilde{z}_0(t) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t),$$
$$M_{\tilde{z}} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 1 \end{bmatrix},\ m_{\tilde{z}} = \begin{bmatrix} 5 \\ 5 \\ 5/6\pi \\ 5/6\pi \end{bmatrix},$$
$$\tilde{x} = [\tilde{x}_1\ \tilde{x}_2]' \in \Re^2,\ \ \tilde{x}_1 = \theta,\ \ \tilde{x}_2 = \dot{\theta},$$
$$\tilde{z}_0 = [\tilde{z}_{01}\ \tilde{z}_{02}]' \in \Re^2,\ \ \tilde{z}_{01} = u,\ \ \tilde{z}_{02} = \theta,$$

where $\theta$ and $\dot{\theta}$ denote the DC motor's rotation angle (rad) and velocity (rad/s), respectively, and the control law used is basically a state feedback law: $u = F(\bar{w} - \tilde{x})$. The reference is given, just before starting control operations, by $\bar{w} = [\bar{w}_\theta\ \bar{w}_{\dot{\theta}}]' \in \Re^2$ with $\bar{w}_\theta = \theta_i$ (rad) and $\bar{w}_{\dot{\theta}} \equiv 0$ (rad/s). We designed the gain by using the discrete LQR technique with appropriate weights: $F = [0.9578\ 0.1723]$.

### *B. Application of the Proposed Approach*

By setting $l = 19$, the set $\mathcal{O}_l$ is constructed; it is illustrated in Fig. 1(a), and a list of selected references is shown in TABLE I. The closed regions bounded by a frame in the augmented space denote the corresponding maximal output admissible sets and we can see that the set $\mathcal{O}_l$ is unconnected in that space. A quick hull algorithm [19] is used to obtain a convex set. The first maximum-volumed single ellipsoid $E_0^*$ is determined from the optimization (5) by YALMIP of [20]. The ellipsoid is shown in Fig. 1(b) and is given below,

$$E_0^* = \left\{ (x, \bar{w}_\theta) \in \Re^3 \;\middle|\; \begin{bmatrix} x \\ \bar{w}_\theta \end{bmatrix} = \begin{bmatrix} 2.9944 & -1.5333 & -0.9544 \\ -1.5333 & 32.8364 & 0.5203 \\ -0.9544 & 0.5203 & 2.2477 \end{bmatrix} \Psi + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ ||\Psi||_2 \leq 1 \right\}.$$

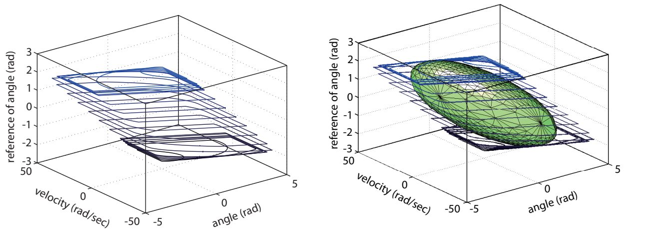
Next, define $\beta = 10^4$ and set $N$ and $L_i\ \forall i \in \mathcal{P}$ to 10 and 4 respectively. The result obtained by the proposed algorithm

TABLE I

LIST OF REFERENCES USED FOR THE APPROXIMATION OF $\mathcal{O}_\infty$ ($l$=19)

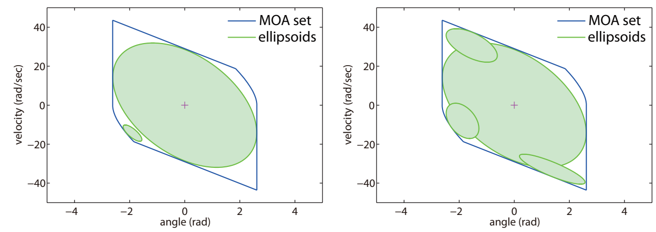
| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $\theta_i$ (deg) | $-4.9/6\pi$ | $-4.7/6\pi$ | $-4.5/6\pi$ | $-4.3/6\pi$ | $-4.1/6\pi$ | $-4/6\pi$ | $-3/6\pi$ | $-2/6\pi$ | $-1/6\pi$ |

| $i$ | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\theta_i$ (deg) | 0 | $1/6\pi$ | $2/6\pi$ | $3/6\pi$ | $4/6\pi$ | $4.1/6\pi$ | $4.3/6\pi$ | $4.5/6\pi$ | $4.7/6\pi$ | $4.9/6\pi$ |



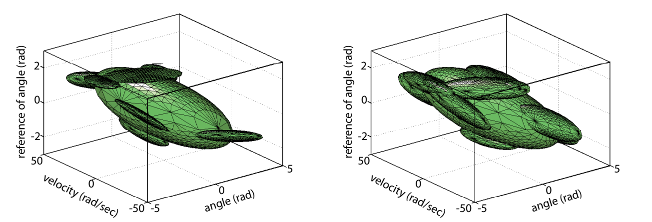

(a) The unconnected set $\mathcal{O}_l$. (b) The single maximum-volumed ellipsoid $E_0^*$ obtained by (5).

Fig. 1. Overview of $\mathcal{O}_l$ and $E_0^*$.



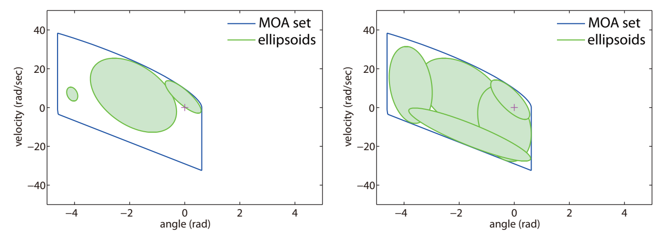

(a) $E_s(c)$ at the **Step 1** in the proposed algorithm. (b) $E_s^*$ obtained after applying the proposed algorithm.

Fig. 3. The cross sections of $E_s(c)$ and $E_s^*$ at $w = 0$, obtained before and after applying the game.




(a) $E_s(c)$ at the **Step 1** in the proposed algorithm. (b) $E_s^*$ obtained after applying the proposed algorithm.

Fig. 2. The approximations of $E_s(c)$ and $E_s^*$ consisting of 11 ellipsoids before and after applying the game.




(a) $E_s(c)$ at the **Step 1** in the proposed algorithm. (b) $E_s^*$ obtained after applying the proposed algorithm.

Fig. 4. The cross sections of $E_s(c)$ and $E_s^*$ at $w = 2$, obtained before and after applying the game.

is shown in Figs. 2-4. Fig. 2(a) shows the ellipsoids arranged before applying the game, and Fig. 2(b) does them aligned after applying the proposed game-based algorithm. From these figures, it can be seen that the individual ellipsoids' position changes. To check the position change more clearly, two cross sections at $\bar{w}_\theta = 0$ and 2 are shown in Fig. 3 and Fig. 4, respectively. Similarly, Fig. 3(a) and Fig. 4(a) are at the time just before starting the game, and Fig. 3(b) and Fig. 4(b) are the results in the approximation. Apparently, it looks to gain the better filling rate that means that the approximation accuracy is improved.

The filling rate[2] is numerically checked using 500 results since in the proposed game-based approach the action of $f_i$ is chosen stochastically by the learning algorithm. We then compare the result of the proposed game with ones of the two methods: a random algorithm [11] and a conventional game [12], and the comparison is made in Fig. 5. Figs. 5(a) and (b), respectively, show the maximum and minimum values and standard deviations among the 500 results data for each, where their averages are commonly plotted in the both figures. Their data values is summarized in TABLE II. The green segment shows the filling rate of ellipsoids arranged at **Step 1** of the learning algorithm, and the blue segment does the one of ellipsoids after the learning. Furthermore, from Fig. 5(a), it can be seen that a champion data of the proposed approach is better than the conventional one. Although the proposed approach restricts the candidate centers onto the each trajectory, it works as much as the conventional method [12], which never consider such a restriction.

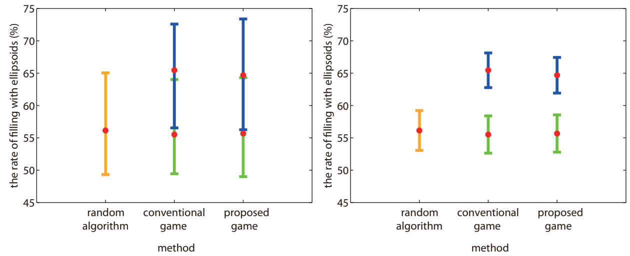
As for the computation time that it takes to get the approximation, the standard deviation in the times for each approach is shown in Fig. 6. From this figure, because the

[2]In the volume measurement, Monte Carlo method is used, that is, many points $(x, \bar{w})$ are randomly seeded in the augmented space, and the filling rate is calculated by the ratio of the number of the points in $E_s^*$ to the number of points $x$ in $O_\infty(\bar{w})$.

TABLE II

THE LIST OF FILLING RATE $\phi \times 100$ (%) ($N$=10).

| approaches | min | max | average | std. devi. |
|---|---|---|---|---|
| random algorithm [11] | 49.32 | 65.05 | 56.13 | 3.08 |
| conventional game [12] | 60.31 | 69.64 | 65.40 | 2.68 |
| proposed game | 56.25 | 73.38 | 64.68 | 2.88 |




(a) Maximum and mininum.

(b) Standard deviation.

Fig. 5. Filling rate for each algorithm. The averages are common.

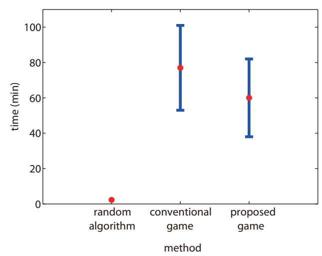



Fig. 6. The standard deviation in the calculation time for each algorithm.

random algorithm just arranges the centers randomly, its time is so fast, but its filling rate is not so good. On the other hand, the proposed approach is faster than the conventional one. This is because the grid points to be searched by the learning algorithm is smaller than the conventional one.

Moreover, the proposed approach has another merit that the player can pass over the ellipsoids the other players yielded. Actually, a situation in which the player pass over the other ellipsoid to get player's better utility value occurs, which is illustrated in Fig. 7 and whose detail explanation can be seen in its caption. The corresponding utility values are plotted in Fig. 8. Such a situation is an unique advantage of the proposed approach and is contributive to efficiently reach the similarly good filling rate to reduce the computation time.

Consequently, from the obtained results and the comparison with the conventional approaches, we can conclude that the proposed game-based inner approximation method of the ideal parameterization by the multiple ellipsoids is effective.

*Remark 3:* To improve the approximation accuracy more, the larger number of ellipsoids should be provided. Actually, the relationship between the filling rate and the number of ellipsoids is discussed in [12], which illustrates that the accuracy improves as the number of ellipsoids increases.

## VI. Conclusion

This paper has presented the game-based inner approximation method by the multiple ellipsoids. The restriction in updating the centers of the ellipsoids to be generated onto the state trajectory enables to efficiently reach the locally (sub-) optimal arrangement of the multiple ellipsoids. The advantage that the computation time can be reduced keeping the good approximation accuracy, has been confirmed through the numerical example. From the discussion, we can see that the game-based method proposed in this paper is effective

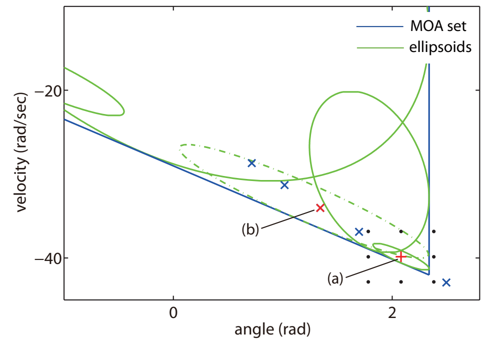



Fig. 7. The cross sections of $E_s^*$ at $w = 0.28$. The current player is located on the point (a). The 8 black points around (a) is a set of the candidates where the player can move by the conventional [12], and the blue $\times$ is a set of ones given by the proposed game. This shows a turn where the player passes over the other ellipsoid to get a better utility value at the point (b).

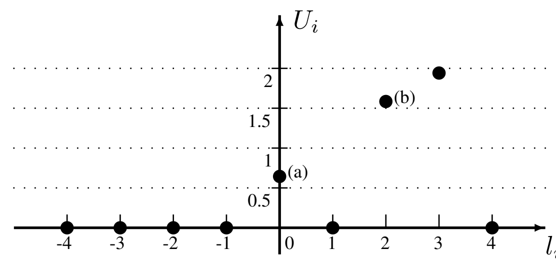



Fig. 8. The utility values $U_i$ of where the player located on the point (a) in Fig. 7 can reach with $l_i \in [-4, 4]$.

for calculating in realtime the efficient approximation of the maximal output admissible set under a situation in which the reference value is not known until starting controls.

In the future works, we will consider how to change the trajectory where the candidates are restricted, such as adding a control input term to the dynamics $f_i$, in order to improve the algorithm to be faster.

## Acknowledgments

The authors would like to acknowledge Prof. Jeff S. Shamma for useful suggestions and comments.

## Appendix

Introduce the following three sets,

$$\begin{aligned} A_1 &:= \{a' \in \mathcal{A} \,|\, \exists c,\ \phi(f_i(c_i, a_i')) - \phi(c) > 0\ \forall i \in \mathcal{P}\}, \\ A_2 &:= \{a' \in \mathcal{A} \,|\, \exists c,\ \phi(f_i(c_i, a_i')) - \phi(c) > 0, \\ &\qquad\qquad U_i(c, a') - U_i(c, a^z) > 0\ \forall i \in \mathcal{P}\}, \\ A_3 &:= \{a' \in \mathcal{A} \,|\, \exists c,\ U_i(c, a') - U_i(c, a^z) > 0\ \forall i \in \mathcal{P}\}, \end{aligned}$$

where $a^z = (a_1^z, \cdots, a_N^z)$. The relation among these sets is

$$A_1 \supset A_2 \subset A_3 \Leftrightarrow (A_1 \cap A_3) \supseteq A_2$$

Therefore, for the considered state game to become a potential game, $A_3 \setminus A_1 = \{\}$ must hold. However, through the numerical example, a situation was confirmed such that $A_3 \setminus A_1$ is not void. That is, although the utility value goes up, the value of $\phi$ goes down. This is why we cannot employ the state potential game approach for the approximation problem by the multiple ellipsoids. In other words, the game we have considered in this paper is in a class of the games that it is difficult to solve and to find the optimum Nash equilibrium.

## REFERENCES


[1] A. Bemporad, "Reference governor for constrained nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 3, pp. 415–419, 1998.

[2] I. Kolmanovsky and J. Sun, "Parameter governors for discrete-time nonlinear systems with pointwise-in-time state and control constraints," *Automatica*, vol. 42, no. 5, pp. 841–848, 2006.

[3] K. Kogiso and K. Hirata, "Reference governor for constrained linear systems with time-varying references," *Robotics and Autonomous Systems*, vol. 57, no. 3, pp. 289–295, 2009.

[4] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[5] D. Angeli, A. Casavola, G. Franze, and E. Mosca, "An ellipsoidal off-line MPC scheme for uncertain polytopic discrete-time systems," *Automatica*, vol. 44, no. 12, pp. 3113–3119, 2008.

[6] K. Hirata and M. Fujita, "Multimode switching state feedback control of constrained linear discrete-time systems," in *Proceedings of the 1998 IEEE International Conference on Control Applications*, Trieste, 1998, pp. 148–152.

[7] E. G. Gilbert and K. T. Tan, "Linear systems with state and control constraints: The theory and application of maximal output admissible sets," *IEEE Transactions on Automatic Control*, vol. 36, no. 9, pp. 1008–1030, 1991.

[8] K. Hirata and Y. Ohta, "Exact determinations of the maximal output admissible set for a class of nonlinear systems," *Automatica*, vol. 44, no. 2, pp. 526–533, 2008.

[9] S. Oh-Hara and K. Hirata, "Experimental evaluation of on-line reference governors for constrained systems," in *Proceedings of SICE Annual Conference*, 2003, pp. 1458–1463.

[10] E. Gilbert and I. Kolmanovsky, "Nonlinear tracking control in the presence of state and control constraints: A generalized reference governor," *Automatica*, vol. 38, no. 12, pp. 2063–2073, 2002.

[11] K. Kogiso, Y. Doi, K. Hirata, and N. Hirade, "Parametric approximation of maximal output admissible sets and its application to multi-mode switching control under unknown constant reference," in *Proceedings of IEEE Multi-conference on Systems and Control*, 2010, pp. 2105–2111.

[12] K. Kita and K. Kogiso, "Game-theoretic consideration of efficient approximation of maximal output admissible sets by multiple ellipsoids," *SICE Annual Conference 2011*, pp. 1727–1733, 2011.

[13] N. Li and J. R. Marden, "Designing games for distributed optimization," *Proceedings of the 50th IEEE Conference on Decision and Control*, 2011.

[14] J. R. Marden, "State based potential games," *Automatica*, vol. 48, no. 12, pp. 3075–3088, 2012.

[15] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and a payoff-based implementation," *Games and Economic Behavior*, 2011, to appear.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

[17] F. Dabbene, P. Gay, and B. Polyak, "Recursive algorithm for inner ellipsoidal approximation of convex polytopes," *Automatica*, vol. 39, no. 10, pp. 1773–1781, 2003.

[18] J. R. Marden, "Learning in large-scale games and cooperative control," Ph.D. dissertation, UCLA.

[19] C. Barber and D. Dobkin, "The quick algorithm for convex hulls," *ACM Transactions on Mathematical Software*, vol. 22, no. 4, pp. 469–483, 1996.

[20] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proceedings of the 2004 IEEE International Symposium on Computer-Aided Control System Design*, Taipei, 2004, pp. 284–289, http://control.ee.ethz.ch/˜ joloef/yalmip.php.