Preamble. This is a reprint of the article:

M. Schulze Darup and M. Mönnigmann. Accurate approximation of the largest null-controllable set for single-input bilinear system. In *Proc. of the 52th IEEE Conference on Decision and Control*, pp. 3951–3956, 2013.

The digital object identifier (DOI) of the original article is:
10.1109/CDC.2013.6760493

Accurate approximation of the largest null-controllable set for single-input bilinear system

Moritz Schulze Darup† *and* M. Mönnigmann†

Abstract

We present a method for the accurate approximation of the largest null-controllable set $\mathcal{N}_\infty$ for constrained bilinear systems. It is central to the presented approach that a simple quantitative measure of the accuracy of approximation can be determined. This measure can be used as a termination criterion for an iterative approximation of $\mathcal{N}_\infty$ with step sets. If the termination criterion is met, the proposed method results in an inner approximation of $\mathcal{N}_\infty$ that covers a requested percentage of $\mathcal{N}_\infty$.

1. Introduction

Consider a nonlinear discrete time system of the form

$$x(k+1) = f(x(k), u(k)), \quad x(0) = x_0, \tag{1}$$

with input and state constraints

$$u(k) \in \mathcal{U} \subset \mathbb{R}^m, \qquad x(k) \in \mathcal{X} \subset \mathbb{R}^n, \qquad \forall\, k \in \mathbb{N}, \tag{2}$$

where $\mathcal{U}$ and $\mathcal{X}$ are solid convex polytopes that contain the origin. Assume that the origin is an equilibrium point of the system, i.e., $f(0,0) = 0$. We call a sequence of inputs *admissible* if all its elements and the resulting trajectory $x(k)$ respect the constraints $u(k) \in \mathcal{U}$ and $x(k) \in \mathcal{X}$, respectively.

It is a recurring and important problem to calculate or approximate the *largest null-controllable set* $\mathcal{N}_\infty$, i.e., the set of all states $x_0 \in \mathcal{X}$ for which there exists an admissible input sequence that steers the system to the origin in a finite number of steps (see, e.g., [5,6] or [9]). The set $\mathcal{N}_\infty$ can be approximated by the set of all states $x_0 \in \mathcal{X}$ that can be

† M. Schulze Darup and M. Mönnigmann are with Automatic Control and Systems Theory, Department of Mechanical Engineering, Ruhr-Universität Bochum, 44801 Bochum, Germany. E-mail: moritz.schulzedarup@rub.de.

steered to the origin with an admissible input sequence with at most i steps. We call this set the *i-step null-controllable set* and denote it by $\mathcal{N}_i$. More precisely, $\mathcal{N}_i$ is iteratively defined by

$$\mathcal{N}_{i+1} = \mathcal{Q}(\mathcal{N}_i) \quad \text{with} \quad \mathcal{N}_0 = \{0\}, \tag{3}$$

where $\mathcal{Q}(\mathcal{T})$ refers to the so-called *one-step-set*

$$\mathcal{Q}(\mathcal{T}) := \{x \in \mathcal{X} \,|\, \exists u \in \mathcal{U} : f(x,u) \in \mathcal{T}\}. \tag{4}$$

The sequence $\{\mathcal{N}_i\}_{i=0}^{\infty}$ is known to tend towards the largest null-controllable set, i.e., $\mathcal{N}_i \to \mathcal{N}_\infty$ as $i \to \infty$. Moreover, $\{\mathcal{N}_i\}_{i=0}^{\infty}$ is non-decreasing, i.e., $\mathcal{N}_i \subseteq \mathcal{N}_{i+1} \subseteq \mathcal{N}_\infty$ for all $i \in \mathbb{N}$ [7], and therefore approaches $\mathcal{N}_\infty$ from its interior. The volume fraction

$$\eta_i := \frac{\text{vol}(\mathcal{N}_i)}{\text{vol}(\mathcal{N}_\infty)} \in [0,1] \tag{5}$$

indicates how well $\mathcal{N}_i$ approximates $\mathcal{N}_\infty$. We call $\mathcal{N}_i$ *accurate approximation* of the largest null-controllable set $\mathcal{N}_\infty$, if $\eta_i = \text{vol}(\mathcal{N}_i)/\text{vol}(\mathcal{N}_\infty) \geq \eta^*$ for a given accuracy η^*. Unfortunately, η_i can in general not be determined, since $\text{vol}(\mathcal{N}_\infty)$ is unknown.

As a remedy we can derive and work with an underestimator $\hat{\eta}_i \leq \eta_i$. Assuming such an $\hat{\eta}_i$ can be computed, we can terminate the iterative approximation (4) of $\mathcal{N}_\infty$ once $\hat{\eta}_i \geq \eta^*$, since this implies

$$\frac{\text{vol}(\mathcal{N}_i)}{\text{vol}(\mathcal{N}_\infty)} = \eta_i \geq \hat{\eta}_i \geq \eta^* \tag{6}$$

and therefore $\mathcal{N}_i$ is accurate in the sense defined above.

For linear systems, it is more or less simple to calculate the sets $\mathcal{N}_i$ and the underestimators $\hat{\eta}_i$. For nonlinear systems, in contrast, this is not the case. We present a method for the accurate approximation of the largest null-controllable set for *single-input bilinear systems*, where

$$f(x,u) = A\,x + (b + N\,x)\,u \tag{7}$$

with $A, N \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$ (see [4] for details on bilinear systems). The approach builds on [**?**], where we showed how to evaluate (3) for bilinear systems. It is the purpose of the present paper to derive an underestimator $\hat{\eta}_i$, and to show how $\hat{\eta}_i$ can be used to control the accuracy of the approximation of $\mathcal{N}_\infty$ for the system class (7). For this purpose we adapt methods for the computation of $\hat{\eta}_i$ for linear systems. Specifically, we use outer approximations of the largest controlled invariant set $\mathcal{C}_\infty$ (see Sect. 1.1 for terminology) to overestimate the set $\mathcal{N}_\infty$. While this is trivial for linear systems, the extension to bilinear systems is challenging.

The paper is organized as follows. We briefly recall how to compute underestimators for linear systems and one-step-sets for bilinear systems in Sect. 2. The main results of the paper, i.e., the derivation of the underestimator $\hat{\eta}_i$ and the formulation of an algorithm for the accurate approximation of $\mathcal{N}_\infty$, are treated in Sect. 3. Finally, Sects. 4 and 5 present two illustrative examples and state conclusions, respectively.

1.1. Notation and Preliminaries

We denote matrices by capital letters, vectors and scalars with lowercase letters and sets with calligraphic letters. Let $I_n \in \mathbb{R}^{n\times n}$ and $e_i \in \mathbb{R}^n$ refer to the identity matrix and the i-th canonical basis vector, respectively. Define $\mathbb{N}_i^k := \{j \in \mathbb{N} \,|\, i \leq j \leq k\}$. Let $\lambda\,\mathcal{T} := \{\lambda\,x \,|\, x \in \mathcal{T}\}$ and $\mathcal{T} + \mu := \{x + \mu \,|\, x \in \mathcal{T}\}$ for any $\mathcal{T} \subset \mathbb{R}^n$, $\lambda \in \mathbb{R}$ and $\mu \in \mathbb{R}^n$. By $P\,\mathcal{Z}$ and $P^{-1}\,\mathcal{T}$ denote the sets $P\,\mathcal{Z} := \{P\,z \,|\, z \in \mathcal{Z}\}$ and $P^{-1}\,\mathcal{T} := \{z \in \mathbb{R}^p \,|\, P\,z \in \mathcal{T}\}$ for any $\mathcal{Z} \subset \mathbb{R}^p$ and $P \in \mathbb{R}^{n\times p}$. Note that $P^{-1}\mathcal{T}$ is well defined even if P is not invertible. Recall that a set $\mathcal{T} \subseteq \mathcal{X}$ with $0 \in \mathcal{T}$ is called *controlled invariant*, if, for every $x \in \mathcal{T}$, there exists a $u \in \mathcal{U}$ such that $f(x,u) \in \mathcal{T}$. Some basic facts about the relation between controlled invariant and null-controllable sets are summarized in Sect. 2.1.

2. Existing Methods

Section 2.1 applies to the general system class (1). The techniques summarized in Sects. 2.2 and 2.3 are restricted to linear and bilinear systems, respectively.

2.1. The Largest Controlled Invariant Set

We introduced (3) to determine the sequence of null-controllable sets $\{\mathcal{N}_i\}_0^\infty$. In order to approximate the largest controlled invariant set $\mathcal{C}_\infty$, a related sequence $\{\mathcal{C}_i\}_0^\infty$ can be constructed with

$$\mathcal{C}_{i+1} = \mathcal{Q}(\mathcal{C}_i) \quad \text{with} \quad \mathcal{C}_0 = \mathcal{X}, \tag{8}$$

which only differs from (3) with respect to the initial set $\mathcal{C}_0$ [3]. The sequence $\{\mathcal{C}_i\}_0^\infty$ defined by (8) tends to $\mathcal{C}_\infty$ for $i \to \infty$ [2]. Some other basic properties, which follow from the definitions (3) and (8), are as follows. While $\{\mathcal{N}_i\}_0^\infty$ is non-decreasing, i.e. $\mathcal{N}_i \subseteq \mathcal{N}_{i+1} \subseteq \mathcal{N}_\infty$, $\{\mathcal{C}_i\}_0^\infty$ is non-increasing, i.e. $\mathcal{C}_i \supseteq \mathcal{C}_{i+1} \supseteq \mathcal{C}_\infty$. By definition, the largest null-controllable set $\mathcal{N}_\infty$ is a controlled invariant set, which implies $\mathcal{N}_\infty \subseteq \mathcal{C}_\infty$. By collecting all inclusion properties we obtain

$$\mathcal{N}_i \subseteq \mathcal{N}_\infty \subseteq \mathcal{C}_\infty \subseteq \mathcal{C}_i \tag{9}$$

from which we infer $\mathrm{vol}(\mathcal{N}_\infty) \leq \mathrm{vol}(\mathcal{C}_i)$ for all $i \in \mathbb{N}$. Thus,

$$\hat{\eta}_i = \frac{\mathrm{vol}(\mathcal{N}_i)}{\mathrm{vol}(\mathcal{C}_i)} \leq \frac{\mathrm{vol}(\mathcal{N}_i)}{\mathrm{vol}(\mathcal{N}_\infty)} = \eta_i \tag{10}$$

for all $i \in \mathbb{N}$, which implies that $\hat{\eta}_i$ constitutes an underestimator for the accuracy η_i of $\mathcal{N}_i$ defined in (5).

2.2. A Tailored Underestimator for Linear Systems

The sets $\mathcal{N}_i$ and $\mathcal{C}_i$ need to be known to calculate $\hat{\eta}_i$ according to (10). Obviously, $\mathcal{N}_i$ and $\mathcal{C}_i$ can be computed by (3) and (8), respectively, if the one-step-set $\mathcal{Q}(\mathcal{T})$ defined in (4) can be evaluated. For linear systems[1], where $f(x,u) = A\,x + b\,u$ for some $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, Keerthi and Gilbert [6] showed how to evaluate (4). It is convenient to introduce the following extended state z and the associated constraints $\mathcal{Z}$ to summarize the results due to Keerthi and Gilbert [6]:

$$z := \begin{bmatrix} x \\ u \end{bmatrix} \quad \text{and} \quad \mathcal{Z} := \{z \in \mathbb{R}^{n+1} \,|\, x \in \mathcal{X},\, u \in \mathcal{U}\}$$

Furthermore, define the matrices $P := [\ I_n \quad 0\ ]$ and $S := [\ A \quad b\ ]$ with $P,\, S \in \mathbb{R}^{n\times(n+1)}$. Then, according to [6], the one-step-set can be determined with

$$\mathcal{Q}(\mathcal{T}) = P\left(S^{-1}\mathcal{T} \cap \mathcal{Z}\right) \tag{11}$$

[1] The approach in [6] is not restricted to single-input linear systems.

for an arbitrary set $\mathcal{T} \subset \mathbb{R}^n$. Note that (11) results in a convex polytope, if both $\mathcal{T}$ and $\mathcal{Z}$ are convex polytopes (see, e.g., Prop. 3.2 in [7]). Since $\mathcal{Z}$, $\mathcal{N}_0$ and $\mathcal{C}_0$ are convex polytopes for linear systems, the sets $\mathcal{N}_i$ and $\mathcal{C}_i$ are convex polytopes for all $i \in \mathbb{N}$ in this case.

Hence, in the linear case, the evaluation of (10) requires the computation of the volume of two polytopes. Since the volume computation for polytopes is computationally expensive, it is advisable to use another, more efficient underestimator. By $\lambda_i \in [0,1]$ denote the largest scaling factor such that $\lambda_i \, \mathcal{C}_i \subseteq \mathcal{N}_i$, i.e.,

$$\lambda_i := \max_{\lambda} \lambda \quad \text{s.t.} \quad \lambda \, \mathcal{C}_i \subseteq \mathcal{N}_i \tag{12}$$

and note that (12) is a linear optimization problem. Then, an underestimator $\hat{\eta}_i \leq \eta_i$ is given by

$$\hat{\eta}_i = \lambda^n, \tag{13}$$

since $\lambda_i \, \mathcal{C}_i \subseteq \mathcal{N}_i$ implies $\mathrm{vol}(\lambda_i \, \mathcal{C}_i) \leq \mathrm{vol}(\mathcal{N}_i)$ and since

$$\mathrm{vol}(\lambda_i \, \mathcal{C}_i) = \mathrm{vol}(\lambda_i \, I_n \, \mathcal{C}_i) = \det(\lambda_i \, I_n) \, \mathrm{vol}(\mathcal{C}_i) = \lambda_i^n \, \mathrm{vol}(\mathcal{C}_i).$$

In summary we have

$$\lambda_i^n \leq \frac{\mathrm{vol}(\mathcal{N}_i)}{\mathrm{vol}(\mathcal{C}_i)} \leq \frac{\mathrm{vol}(\mathcal{N}_i)}{\mathrm{vol}(\mathcal{N}_\infty)} = \eta_i$$

for λ_i from (12).

2.3. One-Step-Set Computation for Bilinear Systems

We briefly summarize how to compute exact null-controllable sets for bilinear systems (7) subject to polytopic constraints (2). The approach summarized here was introduced in [?]. Lemma 1 states conditions under which the bilinear system (7) can be transformed into a linear one.

Lemma 1 (Lem. 2 in [?]): *Let $c \in \mathbb{R}^n$ be such that*

$$c^T A^k (b + N\,x) = 0 \;\; \textit{for all} \, x \in \mathbb{R}^n \;\; \textit{and} \;\; k \in \mathbb{N}_0^{n-2}, \tag{14}$$

$$c^T A^{n-1} (b + N\,x^*) \neq 0 \;\; \textit{for all} \, x^* \in \{x \,|\, b + N\,x \neq 0\}, \tag{15}$$

and let $\hat{A} := M^{-1} \tilde{A}\, M$ and $\hat{b} := M^{-1} \tilde{b}$, where

$$\tilde{A} := \begin{bmatrix} 0 & I_{n-1} \\ 0 & 0 \end{bmatrix}, \;\; \tilde{b} := e_n \;\; \textit{and} \;\; M := \begin{bmatrix} c^T A^0 \\ \vdots \\ c^T A^{n-1} \end{bmatrix}.$$

Then, the relation

$$A\,x + (b + N\,x)\,u = \hat{A}\,x + \hat{b}\,\varphi(x,u), \tag{16}$$

holds for all $x \in \mathbb{R}^n$ and $u \in \mathbb{R}$, where

$$\varphi(x,u) := c^T A^{n-1} (b + N\,x)\,u + c^T A^n x. \tag{17}$$

See [?] for details on the geometrical meaning of the vector c. Lemma 1 implies that one-step-sets for the bilinear system (7) are equal to those of the linear system $f(x,\hat{u}) = \hat{A}\,x + \hat{b}\,\hat{u}$ subject to the constraints $\hat{z} \in \hat{\mathcal{Z}}$, where

$$\hat{z} := \begin{bmatrix} x \\ \hat{u} \end{bmatrix} \text{and } \hat{\mathcal{Z}} := \{\hat{z} \,|\, \hat{u} = \varphi(x,u), \; x \in \mathcal{X}, \; u \in \mathcal{U}\}. \tag{18}$$

Specifically, we have $\mathcal{Q}(\mathcal{T}) = P(\hat{S}^{-1}\mathcal{T} \cap \hat{\mathcal{Z}})$, where $\hat{S} := [\ \hat{A}\ \ \hat{b}\ ]$ and $\hat{\mathcal{Z}}$ is as in (18). Obviously, the system dynamics are simplified by applying Lem. 1, but $\mathcal{Z}$ must be replaced by the more complicated set $\hat{\mathcal{Z}}$. Essentially, the bilinearity was removed from the system dynamics, but now appears in $\hat{\mathcal{Z}}$. In particular, $\hat{\mathcal{Z}}$ is in general not convex, while $\mathcal{Z}$ is. However, according to the following lemma, the set $\hat{\mathcal{Z}}$ can be written as the union of two convex polytopes.

Lemma 2 (Lem. 4 in [?])**:** *Let $\mathcal{U}$ and $\mathcal{X}$ be convex polytopes of the form*

$$\mathcal{U} = \{u \in \mathbb{R}\,|h_u u \leq d_u\} \text{ and } \mathcal{X} = \{x \in \mathbb{R}^n | H_x\, x \leq d_x\},$$

where $h_u,\, d_u \in \mathbb{R}^2$ and $H_x \in \mathbb{R}^{q\times n}$, $d_x \in \mathbb{R}^q$ with $q \in \mathbb{N}$. Then, the set $\hat{\mathcal{Z}}$ as defined in (18) can be expressed as the union $\hat{\mathcal{Z}} = \hat{\mathcal{Z}}^1 \cup \hat{\mathcal{Z}}^2$ of the two convex polytopes

$$\hat{\mathcal{Z}}^1 = \{\hat{z} \in \mathbb{R}^{n+1}|\, H_{\hat{z}}^{(+)}\, \hat{z} \leq d_{\hat{z}}^{(+)}\} \quad \text{and} \tag{19}$$

$$\hat{\mathcal{Z}}^2 = \{\hat{z} \in \mathbb{R}^{n+1}|\, H_{\hat{z}}^{(-)}\, \hat{z} \leq d_{\hat{z}}^{(-)}\}, \tag{20}$$

where

$$H_{\hat{z}}^{(\pm)} = \begin{bmatrix} H_x & 0 \\ \mp c^T A^{n-1} N & 0 \\ \mp h_u\, c^T A^n \mp d_u\, c^T A^{n-1} N & \pm h_u \end{bmatrix}, \tag{21}$$

$$d_{\hat{z}}^{(\pm)} = \begin{bmatrix} d_x \\ \pm c^T A^{n-1} b \\ \pm d_u\, c^T A^{n-1} b \end{bmatrix}. \tag{22}$$

Using the decomposition from Lem. 2, null-controllable sets $\mathcal{N}_i$ of bilinear systems can be calculated as follows. Essentially, Lemma 3 implements the iteration (3) and exploits the special structure $\hat{\mathcal{Z}}$ of the transformed bilinear system.

Lemma 3 (Lem. 5 in [?])**:** *Let $\hat{\mathcal{Z}}^1$ and $\hat{\mathcal{Z}}^2$ be defined as in Eqs. (19)–(22). Assume there exist convex sets $\mathcal{N}_i^1, \ldots, \mathcal{N}_i^l$ such that $\mathcal{N}_i = \bigcup_{j=1}^l \mathcal{N}_i^j$. Define the sets*

$$\mathcal{N}_{i+1}^{2\,j-1} = P(\hat{S}^{-1}\mathcal{N}_i^j \cap \hat{\mathcal{Z}}^1),\;\; \mathcal{N}_{i+1}^{2\,j} = P(\hat{S}^{-1}\mathcal{N}_i^j \cap \hat{\mathcal{Z}}^2) \tag{23}$$

for every $j \in \mathbb{N}_1^l$. Then, for every $j \in \mathbb{N}_1^l$, $\mathcal{N}_{i+1}^{2\,j-1}$ as well as $\mathcal{N}_{i+1}^{2\,j}$ is a convex set and

$$\mathcal{N}_{i+1} = \textstyle\bigcup_{j=1}^l \mathcal{N}_{i+1}^{2\,j-1} \cup \mathcal{N}_{i+1}^{2\,j} \tag{24}$$

for $\mathcal{N}_{i+1}$ as specified in (3)*.*

The set $\mathcal{C}_{i+1}$ can be calculated analogously to $\mathcal{N}_{i+1}$. Assume $\mathcal{C}_i$ is given as the union $\mathcal{C}_i = \bigcup_{j=1}^l \mathcal{C}_i^j$ of l convex sets $\mathcal{C}_i^j$. The expressions (23) and (24) can be replaced by

$$\mathcal{C}_{i+1}^{2\,j-1} = P(\hat{S}^{-1}\mathcal{C}_i^j \cap \hat{\mathcal{Z}}^1),\;\; \mathcal{C}_{i+1}^{2\,j} = P(\hat{S}^{-1}\mathcal{C}_i^j \cap \hat{\mathcal{Z}}^2) \tag{25}$$

for every $j \in \mathbb{N}_1^l$ and

$$\mathcal{C}_{i+1} = \textstyle\bigcup_{j=1}^l \mathcal{C}_{i+1}^{2\,j-1} \cup \mathcal{C}_{i+1}^{2\,j} \tag{26}$$

to calculate $\mathcal{C}_{i+1}$ specified in (8).

Note that the union of convex regions may be convex or non-convex. In general, we obtain non-convex null-controllable sets for bilinear systems (see Ex. 2 in Sec. 4).

3. Accurate Approximation of the Largest Null-Controllable Set for Bilinear Systems

This section contains the main results of the paper. Section 3.1 explains why a naive extension of the underestimators (10) and (13) from linear to bilinear systems is *not* appropriate. This motivates the approach explained in the remainder of Sect. 3, which introduces a convenient, tree-based representation of $\mathcal{N}_i = \bigcup_j \mathcal{N}_i^j$ and $\mathcal{C}_i = \bigcup_j \mathcal{C}_i^j$ from Lem. 3 (Sect. 3.2), the actual calculation of the understimator $\hat{\eta}_i$ (Sect. 3.3), and an algorithm for its computation (Sect. 3.4).

3.1. Naive extension of (10), (13) to the bilinear case fails

Assume $\mathcal{N}_i = \cup_{j=1}^{l} \mathcal{N}_i^j$ and $\mathcal{C}_i = \cup_{j=1}^{l} \mathcal{C}_i^j$ have been determined. In order to evaluate the underestimator $\hat{\eta}_i$ from (10), it remains to calculate the volumes $\text{vol}(\mathcal{N}_i)$ and $\text{vol}(\mathcal{C}_i)$. Unfortunately, this is computationally expensive, since the sets $\mathcal{N}_i^j$ (resp. $\mathcal{C}_i^j$) are in general not pairwise disjoint. Loosely speaking, we must not just sum over the volumes $\text{vol}(\mathcal{N}_i^j)$ (resp. $\text{vol}(\mathcal{C}_i^j)$), but need to subtract the volumes of nonempty intersections. More precisely,

$$\text{vol}(\mathcal{N}_i) = \sum_{\mathcal{J} \subseteq \mathbb{N}_1^l,\, \mathcal{J} \neq \emptyset} (-1)^{|\mathcal{J}|+1} \text{vol}\left(\bigcap\nolimits_{j \in \mathcal{J}} \mathcal{N}_i^j\right) \tag{27}$$

results with the inclusion-exclusion principle [1, p. 61]. In order to evaluate (27) the volume of $\sum_{j=1}^{l} \frac{l!}{j!\,(l-j)!} = 2^l - 1$ polytopes must be calculated. If the set $\mathcal{N}_i$ consists of $l = 16$ subsets, for example, $2^l - 1 = 65535$ polytopes result. Since the computational effort is high for a single polytope, this extension of (10) from the linear to the bilinear case is not attractive from a computational point of view.

Similarly, it is not straight forward to extend the underestimator (13) to the bilinear case. In contrast to the linear case, the sets $\mathcal{C}_i$ may be non-convex (see Ex. 2 in Sect. 4). It is easy to prove and illustrated in Fig. 1b that $\lambda_i \mathcal{C}_i \nsubseteq \mathcal{C}_i$ may result for some or all $\lambda_i \in [0, 1]$ if $\mathcal{C}_i$ is not convex. This implies that $\lambda_i \mathcal{C}_i \subseteq \mathcal{N}_i$ cannot hold, since $\mathcal{N}_i \subseteq \mathcal{C}_i$ according to (9). Consequently, the optimization problem (12) is in general not meaningful in the bilinear case.

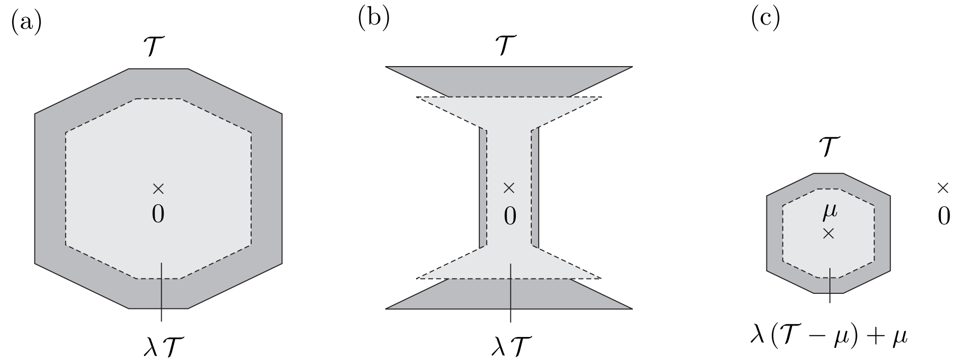

Figure 1: Let $\lambda < 1$. (a) Scaled convex set $\mathcal{T}$ with $\lambda\mathcal{T} \subseteq \mathcal{T}$. (b) Scaled non-convex set $\mathcal{T}$ with $\lambda\mathcal{T} \nsubseteq \mathcal{T}$. (c) Scaled convex set $\mathcal{T}$ with $0 \notin \mathcal{T}$, $\mu \in \mathcal{T}$ and $\lambda(\mathcal{T}-\mu)+\mu \subseteq \mathcal{T}$.

3.2. Null-controllable set representation with binary tree

It proves to be convenient to describe $\mathcal{N}_i$ (resp. $\mathcal{C}_i$) with a binary tree, where each node corresponds to exactly one of the subsets $\mathcal{N}_i^j$ of the union $\mathcal{N}_i = \bigcup_j \mathcal{N}_i^j$. A node is uniquely characterized by the tuple (i, j) of its subset $\mathcal{N}_i^j$, where i and j are the depth of the node and the number of the branch counted from left to right, respectively (see Fig. 2). The root node $(0, 1)$ corresponds to $\mathcal{N}_0 = \mathcal{N}_0^1 = \{0\}$, which obviously is a convex set. Recall $\mathcal{N}_i^j$ associated with the node (i, j) (where $j \in \mathbb{N}_1^{2^i}$) is recursively defined by (28). The following lemma essentially states that the set $\mathcal{N}_i$ is given by the union of all subsets $\mathcal{N}_i^j$ associated with the nodes on the level i of the binary tree (cf. Fig. 2). Sets $\mathcal{C}_i = \bigcup_j \mathcal{C}_i^j$ can be represented by a binary tree correspondingly.

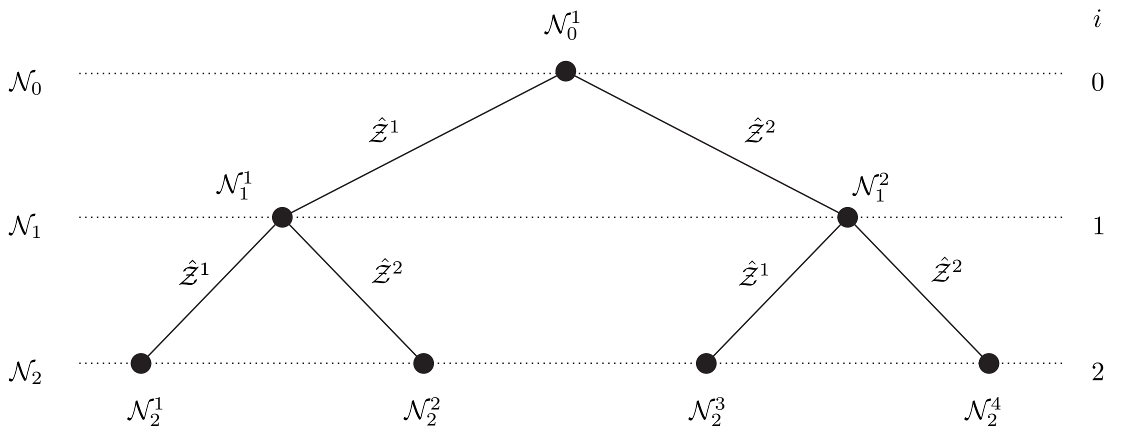

Figure 2: Binary tree associated with the computation of the null-controllable sets $\mathcal{N}_i$ up to level $i = 2$. Nodes of the tree correspond to convex subsets $\mathcal{N}_i^j$ defined in (28). Left and right child of a node (i, j) result from evaluating $P\big(\hat{S}^{-1}\mathcal{N}_i^j \cap \hat{\mathcal{Z}}^1\big)$ and $P\big(\hat{S}^{-1}\mathcal{N}_i^j \cap \hat{\mathcal{Z}}^2\big)$, respectively. Thus, left and right outgoing edges are associated with the convex constraints $\hat{\mathcal{Z}}^1$ or $\hat{\mathcal{Z}}^2$, respectively.

Lemma 4: *Let $i \in \mathbb{N}$ and define the sets*

$$\mathcal{N}_i^j = \begin{cases} \{0\} & \text{if } i = 0, \\ P\big(\hat{S}^{-1}\mathcal{N}_{i-1}^{\frac{j+1}{2}} \cap \hat{\mathcal{Z}}^1\big) & \text{if } i > 0,\ j \text{ odd}, \\ P\big(\hat{S}^{-1}\mathcal{N}_{i-1}^{\frac{j}{2}} \cap \hat{\mathcal{Z}}^2\big) & \text{if } i > 0,\ j \text{ even} \end{cases} \tag{28}$$

and

$$\mathcal{C}_i^j = \begin{cases} \mathcal{X} & \text{if } i = 0, \\ P\big(\hat{S}^{-1}\mathcal{C}_{i-1}^{\frac{j+1}{2}} \cap \hat{\mathcal{Z}}^1\big) & \text{if } i > 0,\ j \text{ odd}, \\ P\big(\hat{S}^{-1}\mathcal{C}_{i-1}^{\frac{j}{2}} \cap \hat{\mathcal{Z}}^2\big) & \text{if } i > 0,\ j \text{ even}. \end{cases} \tag{29}$$

for every $j \in \mathbb{N}_1^{2^i}$. Then, we have

$$\mathcal{N}_i = \bigcup_{j=1}^{2^i} \mathcal{N}_i^j \quad \text{and} \quad \mathcal{C}_i = \bigcup_{j=1}^{2^i} \mathcal{C}_i^j, \tag{30}$$

where $\mathcal{N}_i$ and $\mathcal{C}_i$ are defined as in (3) *and* (8)*, respectively. Moreover, the relation $\mathcal{N}_i^j \subseteq \mathcal{C}_i^j$ holds for every $j \in \mathbb{N}_1^{2^i}$.*

The proof of Lem. 4 is given in the appendix.

3.3. An underestimator for the current accuracy

The following proposition shows how to calculate $\hat{\eta}_i$ for $\mathcal{N}_i$ from the subsets $\mathcal{N}_i^j$ and $\mathcal{C}_i^j$.

Proposition 1: *Let $i \in \mathbb{N}$ be arbitrary and, for all $j \in \mathbb{N}_1^{2^i}$, define $\mathcal{N}_i^j$ and $\mathcal{C}_i^j$ as in Lemma 4. Let $\mathcal{J}_i \subseteq \mathbb{N}_1^{2^i}$ be arbitrary such that $\mathcal{N}_i^j \neq \emptyset$ for all $j \in \mathcal{J}$ and assume*

$$\bigcup_{j=1}^{2^i} \mathcal{N}_i^j = \bigcup_{j\in\mathcal{J}_i} \mathcal{N}_i^j \quad \text{and} \quad \mathcal{Q}(\mathcal{B}) \subseteq \mathcal{B}, \tag{31}$$

where $\mathcal{B} := \bigcup_{j\in\mathcal{J}_i} \mathcal{C}_i^j$. Then there exist, for every $j \in \mathcal{J}_i$, $\mu_{ij} \in \mathcal{N}_i^j$ and $\lambda_{ij} \in [0,1]$ such that

$$\lambda_{ij}\,(\mathcal{C}_i^j - \mu_{ij}) + \mu_{ij} \subseteq \mathcal{N}_i^j \ \text{and} \tag{32}$$

$$\hat{\eta}_i \leq \frac{\mathrm{vol}(\mathcal{N}_i)}{\mathrm{vol}(\mathcal{N}_\infty)}, \ \text{where} \ \hat{\eta}_i^{-1} = 1 + \sum_{j\in\mathcal{J}_i}(\lambda_{ij}^{-n} - 1). \tag{33}$$

We prepare the proof of Prop. 1 with some remarks and lemmas.

Remark 1: *A subset $\mathcal{J}_i \subseteq \mathbb{N}_1^{2^i}$ is introduced to simplify $\mathcal{N}_i = \cup_{j=0}^{2^i}\mathcal{N}_i^j$ by removing empty $\mathcal{N}_i^j$ (see Ex. 1 in Sect. 4, for an example with some $\mathcal{N}_i^j = \emptyset$). In fact, nonempty $\mathcal{N}_i^j$ may sometimes also be removed by exploiting that the $\mathcal{N}_i^j$ are not pairwise disjoint and may cover each other. The first condition in (31) guarantees that the description of $\mathcal{N}_i$ remains exact despite this simplification.*

For the sake of a concise description the best choice of $\mathcal{J}_i$ can be formalized as

$$\mathcal{J}_i := \arg\min_{\mathcal{J}\subseteq\tilde{\mathcal{J}}_i} |\mathcal{J}| \quad \text{s.t.} \quad \bigcup_{j\in\mathcal{J}} \mathcal{N}_i^j = \bigcup_{j\in\tilde{\mathcal{J}}_i} \mathcal{N}_i^j, \tag{34}$$

where $\tilde{\mathcal{J}}_i = \mathbb{N}_1^{2^i}$. While (34) can in general not be solved exactly, there exist algorithms[2] that provide an approximate solution. It is, for example, easy to exclude empty sets $\mathcal{N}_i^j$ and those which are completly contained in other subsets $\mathcal{N}_i^k$.

The following lemma relates the underestimator

$$\hat{\eta}_\cdot \leq \mathrm{vol}(\mathcal{T})/\mathrm{vol}(\mathcal{B})$$

for the unions $\mathcal{T} = \bigcup_j \mathcal{T}_j$, $\mathcal{B} = \bigcup_j \mathcal{B}_j$ to the underestimators

$$\hat{\eta}_{\cdot j} \leq \mathrm{vol}(\mathcal{T}^j)/\mathrm{vol}(\mathcal{B}^j) \tag{35}$$

of the elements of the unions. Note that $\hat{\eta}_\cdot$ and $\hat{\eta}_{\cdot j}$ as introduced above (and in Lem. 5) refer to $\hat{\eta}_i$ and $\hat{\eta}_{ij} = \lambda_{ij}^n$ as specified in Prop. 1.

Lemma 5: *Let $l \in \mathbb{N}$ and let $\mathcal{T} := \bigcup_{j=1}^l \mathcal{T}^j$ and $\mathcal{B} := \bigcup_{j=1}^l \mathcal{B}^j$, where $\mathcal{T}^j \subset \mathbb{R}^n$ and $\mathcal{B}^j \subset \mathbb{R}^n$ are such that $\mathcal{T}^j \subseteq \mathcal{B}^j$ for every $j \in \mathbb{N}_1^l$. Assume there exists, for every $j \in \mathbb{N}_1^l$, a $\hat{\eta}_{\cdot j} \in [0,1]$ such that (35) holds. Then*

$$\hat{\eta}_\cdot \leq \frac{\mathrm{vol}(\mathcal{T})}{\mathrm{vol}(\mathcal{B})} \quad \text{with} \quad \hat{\eta}_\cdot^{-1} = 1 + \sum_{j=1}^l(\hat{\eta}_{\cdot j}^{-1} - 1). \tag{36}$$

---

[2] In Alg. 1, the function `reduceunion` within the multiparametric toolbox (MPT) [8] is used to solve (34).

The proof of Lem. 5 is given in the appendix. Lemma 5 is applied to $\mathcal{T} = \bigcup_{j\in\mathcal{J}} \mathcal{N}_i^j = \mathcal{N}_i$ below. The choice of $\mathcal{B}$ is slightly more complicated in that we can in general not set $\mathcal{B} = \mathcal{N}_\infty$ but $\mathcal{N}_\infty \subseteq \mathcal{B}$. We need Lemma 6 to show that the second condition in (31) ensures $\mathcal{N}_\infty \subseteq \mathcal{B}$.

Lemma 6: *Let $\mathcal{B} \subseteq \mathcal{X}$ with $0 \in \mathcal{B}$. If $\mathcal{Q}(\mathcal{B}) \subseteq \mathcal{B}$, then $\mathcal{N}_\infty \subseteq \mathcal{B}$.*

The proof of Lemma 6 is given in the appendix. Note that it is computationally demanding to check if $\mathcal{Q}(\mathcal{B}) \subseteq \mathcal{B}$. However, there exists appropriate algorithms[3]. Proposition 1 can now be proved as follows.

Proof of Prop. 1. According to Rem. 1 $\mathcal{N}_i^j \neq \emptyset$ for all $j \in \mathcal{J}_i$. Since the $\mathcal{N}_i^j$ are nonempty convex polytopes, we can choose μ_{ij} to be their Chebyshev centers. To see that $\lambda_{ij} \in [0,1]$ exists for all $j \in \mathcal{J}_i$ consider the linear optimization problem

$$\lambda_{ij} := \max_{\lambda} \lambda \quad \text{s.t.} \quad \lambda\,(\mathcal{C}_i^j - \mu_{ij}) + \mu_{ij} \subseteq \mathcal{N}_i^j \tag{37}$$

which yields $\lambda_{ij} \in [0,1]$, since

$$0 \cdot (\mathcal{C}_i^j - \mu_{ij}) + \mu_{ij} = \mu_{ij} \subseteq \mathcal{N}_i^j \subseteq 1 \cdot (\mathcal{C}_i^j - \mu_{ij}) + \mu_{ij} = \mathcal{C}_i^j.$$

It remains to prove (33). Assumption (32) implies

$$\mathrm{vol}(\lambda_{ij}\,(\mathcal{C}_i^j - \mu_{ij}) + \mu_{ij}) \leq \mathrm{vol}(\mathcal{N}_i^j).$$

Since $\mathrm{vol}(\lambda_{ij}\,(\mathcal{C}_i^j - \mu_{ij}) + \mu_{ij}) = \mathrm{vol}(\lambda_{ij}\,(\mathcal{C}_i^j - \mu_{ij})) = \lambda_{ij}^n \,\mathrm{vol}(\mathcal{C}_i^j - \mu_{ij}) = \lambda_{ij}^n \,\mathrm{vol}(\mathcal{C}_i^j)$, we have $\lambda_{ij}^n \,\mathrm{vol}(\mathcal{C}_i^j) \leq \mathrm{vol}(\mathcal{N}_i^j)$ for every $j \in \mathcal{J}_i$. Now, according to Lem. 5,

$$\hat{\eta}_i \leq \frac{\mathrm{vol}(\mathcal{T})}{\mathrm{vol}(\mathcal{B})} \quad \text{with} \quad \mathcal{T} := \bigcup\nolimits_{j\in\mathcal{J}_i} \mathcal{N}_i^j,$$

where $\hat{\eta}_i$ is as defined in (33). We obviously have $\mathcal{T} = \bigcup_{j\in\mathcal{J}_i} \mathcal{N}_i^j = \bigcup_{j=1}^{2^i} \mathcal{N}_i^j = \mathcal{N}_i$, due to the first condition in (31) and according to Lem. 4. Moreover, the second condition in (31) in combination with Lem. 6 guarantees that $\mathcal{N}_\infty \subseteq \mathcal{B}$. Thus, we have $\mathrm{vol}(\mathcal{N}_\infty) \leq \mathrm{vol}(\mathcal{B})$ and finally

$$\hat{\eta}_i \leq \frac{\mathrm{vol}(\mathcal{T})}{\mathrm{vol}(\mathcal{B})} = \frac{\mathrm{vol}(\mathcal{N}_i)}{\mathrm{vol}(\mathcal{B})} \leq \frac{\mathrm{vol}(\mathcal{N}_i)}{\mathrm{vol}(\mathcal{N}_\infty)},$$

which completes the proof. ■

3.4. An algorithm for the accurate approximation of $\mathcal{N}_\infty$

Algorithm 1 implements the iterative computation of the null-controllable sets $\mathcal{N}_i$ according to (23)−(24). In order to provide a lower bound for the current accuracy of the approximation of $\mathcal{N}_\infty$, the underestimator $\hat{\eta}_i$ as introduced in Prop. 1 is considered. Obviously, if the termination criterion $\hat{\eta}_i \geq \eta^*$ is met, Alg. 1 returns an accurate approximation of the largest null-controllable set in terms of $\mathcal{N}_i$. Otherwise, the algorithm stops unsuccessfully after a finite, user-defined number of steps $i^* \in \mathbb{N}$. Nevertheless, in the last case, a measure of the achieved accuracy $\hat{\eta}_{i^*}$ is still returned.

[3] In Alg. 1, the function `regiondiff` within the multiparametric toolbox (MPT) [8] is used to detect $\mathcal{Q}(\mathcal{B}) \subseteq \mathcal{B}$.

Algorithm 1: *Accurate approximation of the largest null-controllable set for bilinear systems.*

set $\mathcal{N}_0^1 = \{0\}$, $\mathcal{C}_0^1 = \mathcal{X}$, $i = 0$, $\mathcal{J}_0 = \{1\}$ and $\hat{\eta}_0 = 0$.
while $\eta^* > \hat{\eta}_i$ and $i^* > i$ **do**
  foreach $j \in \mathcal{J}_i$ **do**
    compute $\mathcal{N}_{i+1}^{2j-1}$ and $\mathcal{N}_{i+1}^{2j}$ according to (23).
    compute $\mathcal{C}_{i+1}^{2j-1}$ and $\mathcal{C}_{i+1}^{2j}$ according to (25).
  set $\tilde{\mathcal{J}}_{i+1} \leftarrow \{2j-1 \,|\, j \in \mathcal{J}_i\} \cup \{2j \,|\, j \in \mathcal{J}_i\}$.
  set $i \leftarrow i+1$ and compute $\mathcal{J}_i$ according to (34).
  set $\mathcal{N}_i \leftarrow \bigcup_{j \in \mathcal{J}_i} \mathcal{N}_i^j$ and $\mathcal{B} \leftarrow \bigcup_{j \in \mathcal{J}_i} \mathcal{C}_i^j$.
  if $\mathcal{Q}(\mathcal{B}) \subseteq \mathcal{B}$ **then**
    foreach $j \in \mathcal{J}_i$ **do**
      choose $\mu_{ij} \in \mathcal{N}_i^j$ and compute λ_{ij} according to (37).
    set $\hat{\eta}_i \leftarrow \max(\hat{\eta}_{i-1}, (1 + \sum_{j \in \mathcal{J}_i} (\lambda_{ij}^{-n} - 1))^{-1})$.
  else $\hat{\eta}_i \leftarrow \hat{\eta}_{i-1}$.
return set $\mathcal{N}_i$, accuracy $\hat{\eta}_i$ and terminate.

4. Numerical examples

We first demonstrate the stepwise approximation of $\mathcal{N}_\infty$ with a one-dimensional example. We then present a two-dimensional example and illustrate the non-convexity of $\mathcal{N}_\infty$. For both examples, we try to achieve the accuracy $\eta^* = 0.99$ and consider the (polytopic) constraints

$$\mathcal{X} = \{x \in \mathbb{R}^n \,|\, \|x\|_\infty \leq 2\} \text{ and } \mathcal{U} = \{u \in \mathbb{R} \,|\, \|u\|_\infty \leq 1\}.$$

Note that $\eta^* = 0.99$ essentially means that the resulting approximation covers 99% of the exact largest null-controllable set.

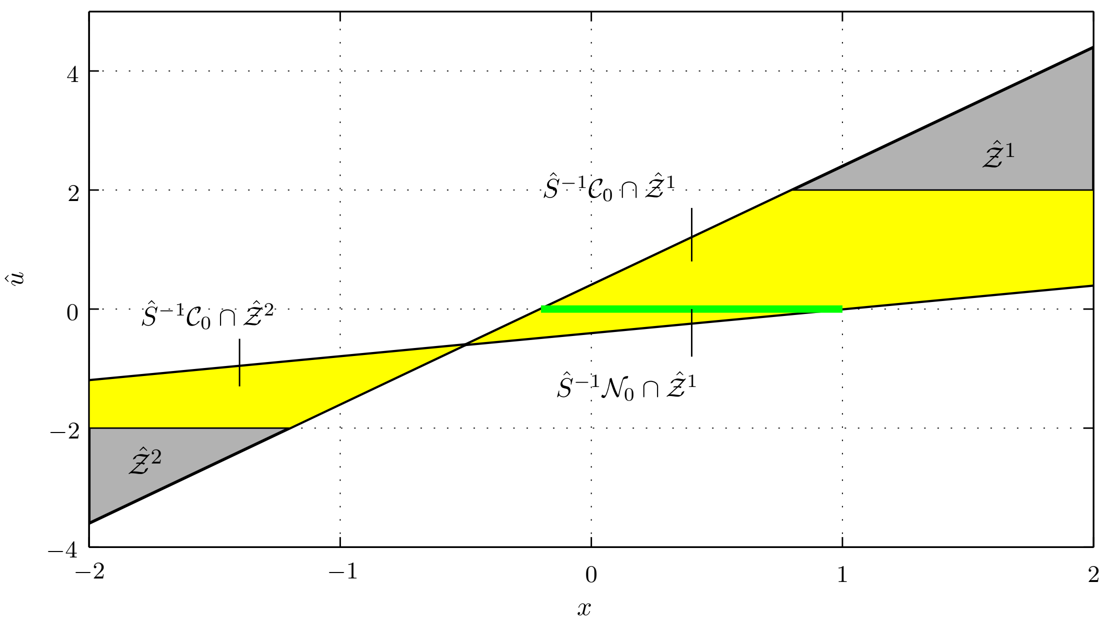

Figure 3: Sets $\hat{\mathcal{Z}}^1$ and $\hat{\mathcal{Z}}^2$ for Ex. 1. Sets $\hat{S}^{-1}\mathcal{N}_0 \cap \hat{\mathcal{Z}}^1$, $\hat{S}^{-1}\mathcal{C}_0 \cap \hat{\mathcal{Z}}^1$ and $\hat{S}^{-1}\mathcal{C}_0 \cap \hat{\mathcal{Z}}^2$, which are required in step $i = 0$ of Alg. 1, are marked in green and yellow, respectively.

Example 1: Consider the bilinear system with

$$A = 1.2, \quad b = 0.4 \quad \text{and} \quad N = 0.8.$$

Without giving details, we claim that $c = 1.0$ fulfills the conditions (14)−(15) and that the exact linearization reads

$$\hat{A} = 0.0, \quad \hat{b} = 1.0 \quad \text{with} \quad M = c^T = 1.0.$$

The sets $\hat{\mathcal{Z}}^1$ and $\hat{\mathcal{Z}}^2$ defined in (19)−(22) are visualized in Fig. 3. We refer to [?] for more details on the representation of $\hat{\mathcal{Z}}$.

Table 1: Numerical results for Ex. 1.

i	$\mathcal{N}_i = \mathcal{N}_i^1$	$\mathcal{C}_i^1$	$\hat{\eta}_i$
0	$[\ 0.0000000, 0.0]$	$[-2.0000000, 2.0]$	0.0000000
1	$[-0.2000000, 1.0]$	$[-0.5000000, 2.0]$	0.3750012
2	$[-0.3000000, 2.0]$	$[-0.4500000, 2.0]$	0.8846156
3	$[-0.3500000, 2.0]$	$[-0.4250000, 2.0]$	0.9400001
5	$[-0.3875000, 2.0]$	$[-0.4062500, 2.0]$	0.9845361
6	$[-0.3937500, 2.0]$	$[-0.4031250, 2.0]$	0.9922279
10	$[-0.3996094, 2.0]$	$[-0.4001953, 2.0]$	0.9995118

We analyze the first step of Alg. 1 in detail. Lines 3 and 4 are carried out with the sets

$$\begin{aligned} \hat{S}^{-1}\mathcal{N}_0^1 &= [\,0.0\ \ 1.0\,]^{-1}\{0\} = \{\hat{z} \in \mathbb{R}^2 \mid \hat{z}_2 = \hat{u} = 0\} \text{ and} \\ \hat{S}^{-1}\mathcal{C}_0^1 &= [\,0.0\ \ 1.0\,]^{-1}\mathcal{X} \ \ = \{\hat{z} \in \mathbb{R}^2 \mid -2 \le \hat{z}_2 \le 2\}, \end{aligned}$$

respectively. Subsequently, the intersections of $\hat{S}^{-1}\mathcal{N}_0^1$ and $\hat{S}^{-1}\mathcal{C}_0^1$ with the sets $\hat{\mathcal{Z}}^1$ and $\hat{\mathcal{Z}}^2$ are calculated. The resulting sets are visualized in Fig. 3. Obviously, the set $\hat{S}^{-1}\mathcal{N}_0^1 \cap \hat{\mathcal{Z}}^2$ is empty. Finally, evaluating the projections in (23) and (25) yields

$$\begin{aligned} \mathcal{N}_1^1 &= [-0.2, 1.0], & \mathcal{N}_1^2 &= \emptyset \text{ and} \\ \mathcal{C}_1^1 &= [-0.5, 2.0], & \mathcal{C}_1^2 &= [-2.0, -0.5]. \end{aligned}$$

We make two interesting observations. First, the set $\mathcal{N}_1^2$ is empty. Thus, we have $\mathcal{N}_1^2 \subset \mathcal{N}_1^1$ and the evaluation of the index set $\mathcal{J}_i$ in line 7 of Alg. 1 will result in $\mathcal{J}_1 = \{1\}$. Second, the set $\mathcal{C}_1$ reads $\mathcal{C}_1 = \mathcal{C}_1^1 \cup \mathcal{C}_1^2 = [-2.0, 2.0] = \mathcal{C}_0 = \mathcal{X}$. Thus, the largest controlled invariant set equals the state constraints, i.e., $\mathcal{C}_\infty = \mathcal{X}$. An analysis of the following steps reveals that $\mathcal{N}_i = \mathcal{N}_i^1$, thus $\mathcal{J}_i = \{1\}$, for every $i \in \mathbb{N}$. Moreover, we find that the set $\mathcal{C}_i^1$ tends towards $\mathcal{N}_i^1$ for $i \to \infty$ (see Tab. 1). In fact, we meet the termination criterion $\hat{\eta}_i \ge \eta^*$ of Alg. 1 for $i = 6$. Thus, the set $\mathcal{N}_6 = [-0.393750, 2.0]$ approximates $\mathcal{N}_\infty$ within the chosen accuracy $\eta^* = 0.99$. Note that it is easy to prove that the largest null-controllable set reads $\mathcal{N}_\infty = (-0.4, 2.0]$. The proper termination of Alg. 1 is remarkable, since $\mathcal{C}_\infty = [-2.0, 2.0]$. Thus, the direct evaluation of the underestimator (10) (as proposed for linear systems) yields

$$\hat{\eta}_i = \frac{\text{vol}(\mathcal{N}_i)}{\text{vol}(\mathcal{C}_i)} \le \frac{\text{vol}(\mathcal{N}_\infty)}{\text{vol}(\mathcal{C}_\infty)} = \frac{2.4}{4.0} = 0.6 \ll \eta^* = 0.99.$$

Obviously, it is not possible to meet the termination criterion $\hat{\eta}_i \ge \eta^*$ by calculating $\frac{\text{vol}(\mathcal{N}_i)}{\text{vol}(\mathcal{C}_i)}$ for this example.

Example 2: Consider the bilinear system with

$$A = \begin{bmatrix} 1.12 & 0.54 \\ 0.76 & 0.92 \end{bmatrix},\ b = \begin{bmatrix} 0.5 \\ -1.0 \end{bmatrix},\ N = \begin{bmatrix} 0.4 & -0.6 \\ -0.8 & 1.2 \end{bmatrix}.$$

We claim without proof that $c^T = \begin{bmatrix} 2.0 & 1.0 \end{bmatrix}$ fulfills (14)–(15) and that the exact linearization is given in terms of

$$\hat{A} = \begin{bmatrix} 6.0 & 4.0 \\ -9.0 & -6.0 \end{bmatrix},\ \hat{b} = \begin{bmatrix} -1.0 \\ 2.0 \end{bmatrix},\ M = \begin{bmatrix} 2.0 & 1.0 \\ 3.0 & 2.0 \end{bmatrix}.$$

Algorithm 1 terminates after the 27th step (see Tab. 2). The set $\mathcal{N}_{27}$ as well as some intermediate results for the steps $i = \{2, 5, 10\}$ are visualized in Fig. 4. Obviously, the largest null-controllable set is not convex.

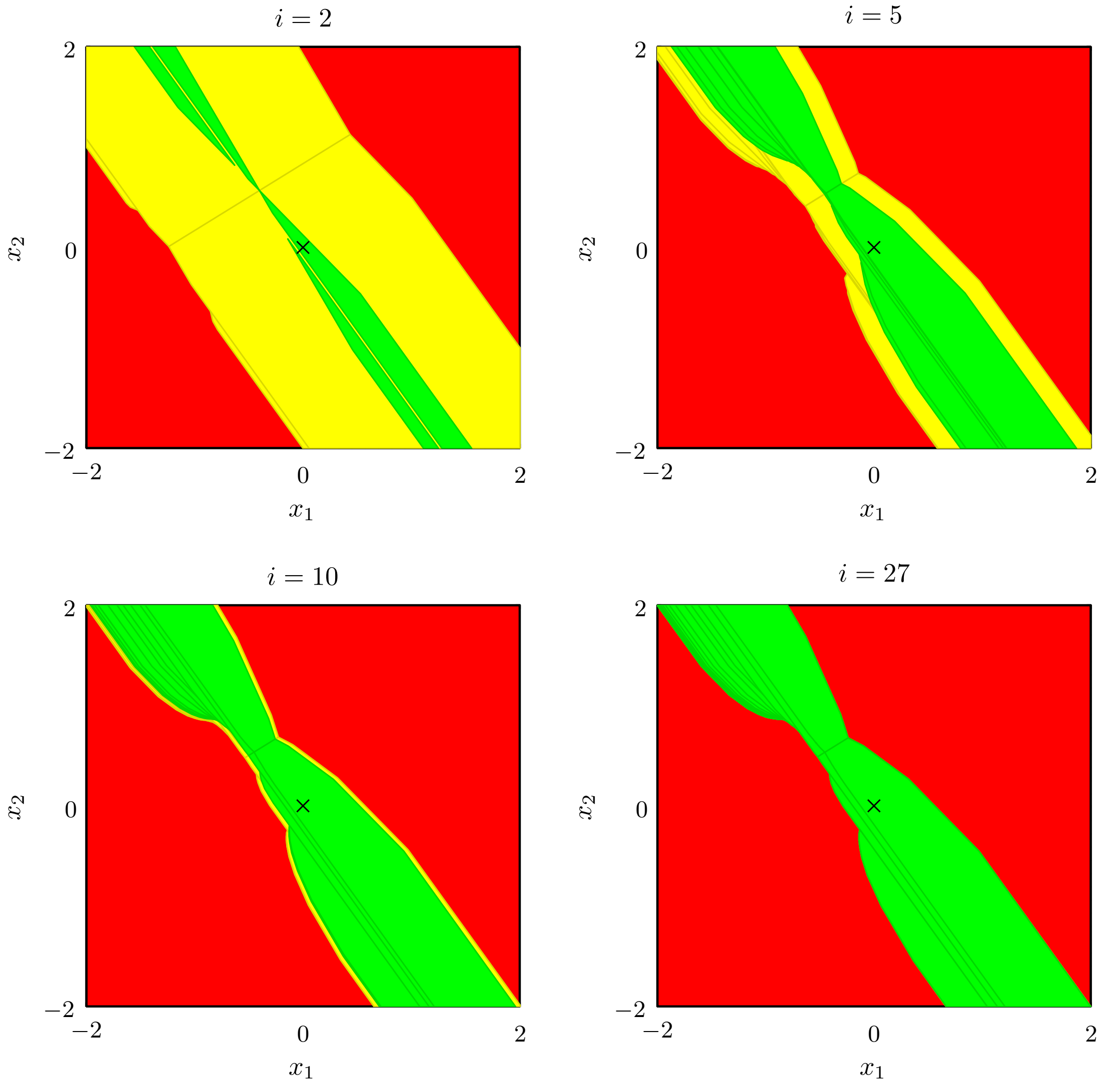

Figure 4: Null-controllable sets $\mathcal{N}_i$ (green) and corresponding overestimations $\mathcal{B}$ (yellow) evaluated for Ex. 2. Red areas visualize $\mathcal{X} \setminus \mathcal{B}$. Sets $\mathcal{N}_i$ and $\mathcal{B}$ are intermediate results of Alg. 1 for the steps $i \in \{2, 5, 10, 27\}$.

It is apparent from Fig. 4 that the proposed underestimator is conservative. While the approximations obtained after 10 steps is found to be very close to that after 27 steps by visual inspection, the value of the underestimator after 10 steps, $\hat{\eta}_{10} = 0.0485$, suggests that the approximation is still far from accurate. Table 2 provides another interesting result. Assume that the set $\mathcal{J}_i$ (as evaluated in line 7 of Alg. 1) reads $\mathcal{J}_i = \mathbb{N}_1^{2^i}$ for each step. Then, the number of subsets describing $\mathcal{N}_i$ would evaluate to $l = 2^i$. Thus, the number of subsets would increase exponential with the number of steps (see fourth column in Tab. 2). Fortunately, Ex. 2 (as well as Ex. 1) illustrates that this dramatic increase does not necessarily occur (see third column in Tab. 2). In contrast, the number of subsets seem to stagnate for $i \to \infty$ for both examples.

Table 2: Numerical results for example 2.

i	$\hat{\eta}_i$	$\lvert\mathcal{J}_i\rvert$	$l = 2^i$
0	0.0000000	1	1
2	0.0010133	4	4
5	0.0085739	15	32
10	0.0484568	18	1024
20	0.8650666	25	1048576
27	0.9919346	24	134217728
30	0.9964801	25	1073741824

5. Conclusions

We presented a method for the accurate approximation of the largest null-controllable set $\mathcal{N}_\infty$ for bilinear systems with input and state constraints. The proposed approach builds on the computation of the i-step null-controllable set $\mathcal{N}_i$ as introduced in [10]. It is the main contribution of the present paper to derive a measure for the accuracy of the approximation $\mathcal{N}_i$ of the largest null-controllable set $\mathcal{N}_\infty$. This measure can be used to state a termination criterion for the iterative approximation of $\mathcal{N}_\infty$ with $\mathcal{N}_i$.

We illustrated the resulting method with two examples. For both examples, the proposed algorithm returned an accurate approximation of the largest null-controllable set. In fact, in both cases, the approximation includes more than 99% of all null-controllable states. Future work has to address the extension to multi-input systems.

References

[1] R. B. J. T. Allenby and A. Slomson. *How to Count: An Introduction to Combinatorics.* CRC Press, 2010.

[2] D. P. Bertsekas. Infinite-time reachability of state-space regions by using feedback control. *IEEE Trans. Autom. Control*, 17:604–613, 1972.

[3] F. Blanchini and S. Miani. *Set-Theoretic Methods in Control.* Birkhäuser, 2008.

[4] C. Bruni, G. DiPillo, and G. Koch. Bilinear systems: An appealing class of "nearly linear" systems in theory and applications. *IEEE Trans. Autom. Control*, 19(4):334–348, 1974.

[5] P. O. Gutman and M. Cwikel. An algorithm to find maximal state constraint sets for discrete-time linear dynamical systems with bounded control and states. *IEEE Trans. Autom. Control*, 32(3):251–253, 1987.

[6] S. S. Keerthi and E. G. Gilbert. Computation of minimum-time feedback control laws for discrete-time systems with state-control constraints. *IEEE Trans. Autom. Control*, 32(5):432–435, 1987.

[7] E. C. Kerrigan. *Robust Constraint Satisfaction: Invariant Sets and Predictive Control.* Ph.D. thesis, University of Cambridge, 2000.

[8] M. Kvasnica, P. Grieder, M. Baotić, and M. Morari. Multi-parametric toolbox (MPT). In *Proc. of 7th International Workshop on Hybrid Systems - Computation and Control*, pp. 448–462, 2004.

[9] M. Schulze Darup and M. Mönnigmann. Low complexity suboptimal explicit nmpc. In *Proc. of 4th IFAC Nonlinear Model Predictive Control Conference*, pp. 406–411, 2012.

[10] M. Schulze Darup and M. Mönnigmann. Null-controllable set computation for a class of constrained bilinear systems. In *Proc. of 2013 European Control Conference*, pp. 2758–2763, 2013.

A. Supplementary proofs

Proof of Lem. 4 by induction. We start with the first relation in (30). The claim holds for $i=0$, since $\mathcal{N}_0 = \{0\} = \mathcal{N}_0^1$ by definition. We have to prove

$$\mathcal{N}_{i+1} = \bigcup_{j=1}^{2^{i+1}} \mathcal{N}_{i+1}^j. \tag{38}$$

under the assumption $\mathcal{N}_i = \bigcup_{j=1}^{2^i} \mathcal{N}_i^j$ and for $\mathcal{N}_i^j$ in (28). The r.h.s. in (38) can be written as

$$\bigcup_{j=1}^{2^{i+1}} \mathcal{N}_{i+1}^j = \bigcup_{j\in\mathcal{J}^1} \mathcal{N}_{i+1}^j \cup \bigcup_{j\in\mathcal{J}^2} \mathcal{N}_{i+1}^j, \tag{39}$$

where $\mathcal{J}^1 = \{j \in \mathbb{N}_1^{2^{i+1}} \,|\, j \,\text{odd}\}$ and $\mathcal{J}^2 = \{j \in \mathbb{N}_1^{2^{i+1}} \,|\, j \,\text{even}\}$. According to (28) and due to $i+1>0$, the r.h.s. in (39) is equal to

$$\bigcup_{j\in\mathcal{J}^1} P\big(\hat{S}^{-1}\mathcal{N}_i^{\frac{j+1}{2}} \cap \hat{\mathcal{Z}}^1\big) \cup \bigcup_{j\in\mathcal{J}^2} P\big(\hat{S}^{-1}\mathcal{N}_i^{\frac{j}{2}} \cap \hat{\mathcal{Z}}^2\big). \tag{40}$$

Since $\mathcal{J}^1 = \{2\,k-1 \,|\, k \in \mathbb{N}_1^{2^i}\}$ and $\mathcal{J}^2 = \{2\,k \,|\, k \in \mathbb{N}_1^{2^i}\}$, (40) can be written as

$$\begin{aligned} &\bigcup_{k=1}^{2^i} P\big(\hat{S}^{-1}\mathcal{N}_i^{\frac{2\,k-1+1}{2}} \cap \hat{\mathcal{Z}}^1\big) \cup \bigcup_{k=1}^{2^i} P\big(\hat{S}^{-1}\mathcal{N}_i^{\frac{2\,k}{2}} \cap \hat{\mathcal{Z}}^2\big) \\ &= \bigcup_{k=1}^{2^i} P\big(\hat{S}^{-1}\mathcal{N}_i^k \cap \hat{\mathcal{Z}}^1\big) \cup P\big(\hat{S}^{-1}\mathcal{N}_i^k \cap \hat{\mathcal{Z}}^2\big). \end{aligned} \tag{41}$$

which holds according to Lemma 3 (specifically (24) with (23) substituted). This proves (38). The second relation in (30), i.e., $\mathcal{C}_i = \bigcup_{j=1}^{2^i} \mathcal{C}_i^j$ can be shown accordingly. It remains to prove $\mathcal{N}_i^j \subseteq \mathcal{C}_i^j$ for every $j \in \mathbb{N}_1^{2^i}$. Since we have $0 \in \mathcal{X}$ by definition, the claim is trivially fulfilled for $i=0$ in that $\mathcal{N}_0^1 = \{0\} \subseteq \mathcal{X} = \mathcal{C}_0^1$. We have to show

$$\mathcal{N}_{i+1}^j \subseteq \mathcal{C}_{i+1}^j \quad \text{for every} \quad j \in \mathbb{N}_1^{2^{i+1}} = \mathcal{J}^1 \cup \mathcal{J}^2 \tag{42}$$

to complete the induction. We consider an arbitrary $j \in \mathcal{J}^1$; the case can be treated $j \in \mathcal{J}^2$ accordingly. As a preparation note that

$$\mathcal{T} \subseteq \mathcal{B} \quad \Longrightarrow \quad P\big(\hat{S}^{-1}\mathcal{T} \cap \hat{\mathcal{Z}}^1\big) \subseteq P\big(\hat{S}^{-1}\mathcal{B} \cap \hat{\mathcal{Z}}^1\big) \tag{43}$$

for arbitrary sets $\mathcal{T}, \mathcal{B} \subset \mathbb{R}^n$. According to (28) and (29), we have to show

$$P\big(\hat{S}^{-1}\mathcal{N}_i^{\frac{j+1}{2}} \cap \hat{\mathcal{Z}}^1\big) \subseteq P\big(\hat{S}^{-1}\mathcal{C}_i^{\frac{j+1}{2}} \cap \hat{\mathcal{Z}}^1\big) \text{ for every } j \in \mathcal{J}^1. \tag{44}$$

With the same argumentation as above, we may rewrite (44) as

$$P\big(\hat{S}^{-1}\mathcal{N}_i^k \cap \hat{\mathcal{Z}}^1\big) \subseteq P\big(\hat{S}^{-1}\mathcal{C}_i^k \cap \hat{\mathcal{Z}}^1\big) \text{ for every } k \in \mathbb{N}_1^{2^i} \tag{45}$$

Relation (45) follows from applying (43) to $\mathcal{N}_i^k \subseteq \mathcal{C}_i^k$, which holds for every $k \in \mathbb{N}_1^{2^i}$ by induction hypothesis. ■

Proof of Lem. 5. Since $\mathcal{T}^j \subseteq \mathcal{B}^j$ for every $j \in \mathbb{N}_1^l$, we have $\mathcal{T} \subseteq \mathcal{B}$. Thus, $\mathrm{vol}(\mathcal{B})$ can be written as

$$\mathrm{vol}(\mathcal{B}) = \mathrm{vol}(\mathcal{T}) + \mathrm{vol}(\mathcal{B} \setminus \mathcal{T}). \tag{46}$$

The term $\mathrm{vol}(\mathcal{B} \setminus \mathcal{T})$ can be bounded above as follows

$$\mathrm{vol}(\mathcal{B} \setminus \mathcal{T}) = \mathrm{vol}(\textstyle\bigcup_{j=1}^l \mathcal{B}^j \setminus \mathcal{T}^j) \leq \sum_{j=1}^l \mathrm{vol}(\mathcal{B}^j \setminus \mathcal{T}^j). \tag{47}$$

With the same argumentation as above, we have

$$\mathrm{vol}(\mathcal{B}^j \setminus \mathcal{T}^j) = \mathrm{vol}(\mathcal{B}^j) - \mathrm{vol}(\mathcal{T}^j), \tag{48}$$

since $\mathcal{T}^j \subseteq \mathcal{B}^j$. Rearranging (35) yields

$$\mathrm{vol}(\mathcal{B}^j) \leq \hat{\eta}_{\cdot j}^{-1}\mathrm{vol}(\mathcal{T}^j). \tag{49}$$

Substituting (47), (48) and (49) into (46) yields

$$\begin{aligned} \mathrm{vol}(\mathcal{B}) \quad &\leq \quad \mathrm{vol}(\mathcal{T}) + \textstyle\sum_{j=1}^l (\hat{\eta}_{\cdot j}^{-1} - 1)\,\mathrm{vol}(\mathcal{T}^j), \\ &\leq \quad \big(1 + \textstyle\sum_{j=1}^l (\hat{\eta}_{\cdot j}^{-1} - 1)\big)\,\mathrm{vol}(\mathcal{T}), \end{aligned} \tag{50}$$

where the second inequality holds because of $\mathcal{T}^j \subseteq \mathcal{T}$. Rearranging (50) results in (36). ■

Proof of Lem. 6 by contradiction. Assume $\mathcal{Q}(\mathcal{B}) \subseteq \mathcal{B}$ but $\mathcal{N}_\infty \not\subseteq \mathcal{B}$. If $\mathcal{N}_\infty \not\subseteq \mathcal{B}$, there exists an $x_0 \in \mathcal{N}_\infty$ such that $x_0 \notin \mathcal{B}$. Since $0 \in \mathcal{B}$ by assumption, we obviously have $x_0 \neq 0$. By definition of $\mathcal{N}_\infty$, there exist $i \in \mathbb{N}$ and $u(0), \ldots, u(i-1) \in \mathcal{U}$ such that $x(k) \in \mathcal{X}$ for all $k \in \mathbb{N}_0^i$ and $x(i) = 0$. Note that $i > 0$, since we need at least one step to steer $x_0 \neq 0$ into the origin. Combining $x_0 \notin \mathcal{B}$, which holds by construction, and $x(i) \in \mathcal{B}$, which follows from $x(i) = 0$ and $0 \in \mathcal{B}$, we infer that there must be a step in which $x(k)$ enters $\mathcal{B}$, i.e., there must be a $k \in \mathbb{N}_0^{i-1}$ such that

$$x(k) \notin \mathcal{B} \quad \text{and} \quad x(k+1) = f(x(k), u(k)) \in \mathcal{B}.$$

According to the definition of $\mathcal{Q}(\mathcal{B})$ from (4), $f(x(k), u(k)) \in \mathcal{B}$ (along with $u(k) \in \mathcal{U}$ and $x(k) \in \mathcal{X}$) implies $x(k) \in \mathcal{Q}(\mathcal{B})$. Obviously, $x(k) \in \mathcal{Q}(\mathcal{B})$ with $x(k) \notin \mathcal{B}$ contradicts $\mathcal{Q}(\mathcal{B}) \subseteq \mathcal{B}$. ■