A branch-and-bound algorithm for non-convex Nash equilibrium problems

Peter Kirst* Stefan Schwarze# Oliver Stein‡

December 19, 2022

Abstract

This paper introduces a spatial branch-and-bound method for the approximate computation of the set of all ε-Nash equilibria of continuous box-constrained non-convex Nash equilibrium problems. We explain appropriate discarding and fathoming techniques, provide a termination proof for a prescribed approximation tolerance, and report our computational experience.

1 Introduction

The basic solution concept for a collection of finitely many coupled optimization problems was introduced by Nash [18,19] and is known as the Nash equilibrium. For historical reasons the owners of the optimization problems are termed players, and their feasible sets are called strategy sets.

The present paper considers a finite number of players $\nu = 1, \ldots, N$, whose strategy sets are assumed to be boxes $\Omega_\nu \subset \mathbb{R}^{n_\nu}$. With $n = \sum_{\nu=1}^{N} n_\nu$ and $\Omega := \Omega_1 \times \ldots \times \Omega_N \subset \mathbb{R}^n$, each player ν's objective function $\theta_\nu : \Omega \to \mathbb{R}$ is assumed to be continuous but, in contrast to the vast majority of literature, we do not impose any convexity assumptions on θ_ν.

*Operations Research and Logistics (ORL), Wageningen University & Research (WUR), The Netherlands, peter.kirst@wur.nl

#Institute for Operations Research (IOR), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, stefan.schwarze@kit.edu

‡Institute for Operations Research (IOR), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, stein@kit.edu

A Nash equilibrium is a point $\bar{x} = (\bar{x}^1, \ldots, \bar{x}^N) \in \Omega$ such that

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \quad \text{for all} \quad x^\nu \in \Omega_\nu$$

holds for all players $\nu = 1, \ldots, N$. This means that for each $\nu = 1, \ldots, N$ the point $\bar{x}^\nu$ solves the problem

$$P_\nu(\bar{x}^{-\nu}): \quad \min_{x^\nu} \ \theta_\nu(x^\nu, \bar{x}^{-\nu}) \qquad \text{s. t.} \qquad x^\nu \in \Omega_\nu.$$

The collection of the problems $P_\nu(x^{-\nu})$, $\nu = 1, \ldots, N$, with $x \in \Omega$ is called Nash equilibrium problem. The notation $x = (x^\nu, x^{-\nu})$ with $x^{-\nu} = \left(x^1, \ldots, x^{\nu-1}, x^{\nu+1}, \ldots, x^N\right) \in \mathbb{R}^{n-n_\nu}$ emphasizes the ν-th player's decision variables, but does not reorder the entries of the vector $\left(x^1, \ldots, x^\nu, \ldots, x^N\right)$.

Subsequently it will be useful to consider, for each $\nu = 1, \ldots, N$, the set $S_\nu(x^{-\nu})$ of minimal points of $P_\nu(x^{-\nu})$ as well as the graph $\operatorname{gph} S_\nu = \{(x^\nu, x^{-\nu}) \mid x^\nu \in S_\nu(x^{-\nu})\}$ of the corresponding set-valued mapping S_ν. The set of Nash equilibria E then coincides with the set $\bigcap_{\nu=1}^N \operatorname{gph} S_\nu$.

For a survey of theory and methods for continuous Nash equilibrium problems we refer to [11], where also more general problem classes with not necessarily box-constrained strategy sets are considered, which may even depend on the other players' decsions (so-called generalized Nash equilibrium problems). Prominent solution techniques comprise the Karush-Kuhn-Tucker approach, which aims at solving the joint Karush-Kuhn-Tucker systems of all players [10], and the variational inequality approach [11], which both rely on the convexity of the players' problems or even stronger convexity assumptions. The Nikaido-Isoda approach from [22], on the other hand, does not need convexity assumptions, but it depends on an intrinsically nonsmooth auxiliary problem.

A common feature of all these methods is that they only compute a single Nash equilibrium while, even under convexity assumptions, the set of Nash equilibria does neither need to be a singleton nor convex, but is typically disconnected (see Section 6.2 for an illustration of this fact). We are only aware of few algorithms which compute the whole set of equilibria for continuous Nash equilibrium problems, namely the ones from [6] for convex quadratic player problems with one-dimensional strategy sets, from [7] for problems satisfying polyhedrality assumptions, and from [12] for jointly convex problems. For (mixed-)integer Nash equilibrium problems a branch-and-prune approach for the computation of all equilibria was introduced in [23, 25] and generalized in [27].

So far, branch-and-bound approaches for continuous Nash equilibrium problems have not been proposed in the literature. The present paper intends

to fill this gap where, as a first step, we consider box constrained problems. Since our spatial branch-and-bound approach is heavily based on techniques from global optimization, it is not only applicable to non-convex Nash equilibrium problems, but it also approximately computes the whole set of Nash equilibria.

For an introduction to global optimization we refer to the monographs [13,15, 26]. Spatial branch-and-bound methods rely on subsequently refined boxes and suitable lower and upper bounding procedures that are applied over these domains. In fact, there are several possibilities for bounding procedures available in the literature, which vary with regard to their assumptions. Hence, although our method itself can be stated based on very mild assumptions regarding the defining functions, depending on the concrete choice of bounding procedure one might want to impose additional requirements such as, e.g., differentiability. In principle, however, this is not needed and there are many lower bounding procedures commonly used in global optimization that work under our mild assumptions. In particular, the widely used interval arithmetic as described, for instance, in the monograph [21] does not require differentiability.

Well-known alternatives that might impose additional requirements are, for example, the so-called αBB relaxation as described in [1–3], centered forms as described in [5, 16], lower bounds based on duality concepts as examined in [8, 9] or lower bounds based on linearization techniques (see [20, 26, 28, 29]). Basically all these approaches require that the defining functions are factorable, i.e. composed of a finite number of operations such as $+$, $-$, sin, cos, exp, etc.

This paper is structured as follows. In Section 2 we provide preliminaries that are crucial for the development of the proposed branch-and-bound method, like convergent lower bounding procedures and ε-Nash equilibria. Based on these concepts, in Section 3 we provide a suitable discarding test and a first, conceptual algorithm for computing all Nash equilibria of possibly non-convex NEPs with box constraints. Section 4 addresses a fathoming step which improves the performance of the branch-and-bound method. A more suitable termination criterion than the one in the conceptual method from Section 3 is presented in Section 5, where the appropriate modifications lead to the second and final version of the proposed algorithm. After all these theoretical considerations are established, the proposed method is illustrated along some numerical examples in Section 6. Eventually, Section 7 closes this paper with some final remarks.

Our notation and terminology is standard. In particular, we denote a box

$X \subseteq \mathbb{R}^n$ by $[\underline{x}, \overline{x}]$ with $\underline{x}, \overline{x} \in \mathbb{R}^n$, $\underline{x} \leq \overline{x}$. By the width of a box X we mean its diagonal length $\operatorname{diag}(X) = \|\overline{x} - \underline{x}\|_2$ with respect to the Euclidean distance, and we denote its midpoint by $\operatorname{mid}(X) = \frac{1}{2}(\underline{x} + \overline{x})$.

2 Preliminaries

In this section we recall two concepts that will be important for our algorithm. Section 2.1 deals with the notion of M-independent convergent lower bounding procedures, which we will use to show convergence properties of the two proposed methods. The first, conceptual version of the method only includes a basic termination criterion. In order to improve it, we employ the concept of ε-Nash equilibria (Section 2.2). It will turn out that this definition leads to a natural termination criterion for our new method (cf. Section 5).

2.1 Convergent lower bounding procedures

In the following it will be important that lower bounding procedures as the ones mentioned in Section 1 are M-independent and convergent [14]. For convenience, we briefly review this concept here. In fact, a function ℓ_f from the set of all sub-boxes X of Ω to $\mathbb{R}$ is called *M-independent lower bounding procedure* for a function f, if it satisfies $\ell_f(X) \leq \min_{x \in X} f(x)$ for all sub-boxes $X \subseteq \Omega$ and any choice of f. Furthermore, we call an M-independent lower bounding procedure *convergent* if

$$\lim_{k \to \infty} \ell_f(X^k) = \lim_{k \to \infty} \min_{x \in X^k} f(x)$$

holds for any exhaustive sequence of boxes $(X^k)_{k \in \mathbb{N}}$ and any choice of the function f, respectively. A sequence of boxes $(X^k)_{k \in \mathbb{N}}$ is called exhaustive (see, e.g., [15]) if $X^{k+1} \subseteq X^k$ as well as $\lim_{k \to \infty} \operatorname{diag}(X^k) = 0$ hold. The concept of M-independent upper bounds is defined analogously.

We remark that for any exhaustive sequence of boxes (X^k) the set $\bigcap_{k \in \mathbb{N}} X^k$ is a singleton, say $\{\widetilde{x}\}$. If we have $x^k \in X^k$ for all k, then in view of $\lim_{k \to \infty} \operatorname{diag}(X^k) = 0$ the sequence (x^k) converges to $\widetilde{x}$, and the convergence of ℓ_f and continuity yield

$$\lim_{k \to \infty} \ell_f(X^k) = \lim_{k \to \infty} \min_{x \in X^k} f(x) = f(\widetilde{x}) = \lim_{k \to \infty} f(x^k).$$

2.2 The notion of an ε-Nash equilibrium

A crucial concept for the termination of global optimization methods is the notion of an ε-optimal point. In order to extend spatial branch-and-bound algorithms from nonlinear problems to NEPs, it is important to have an analogy here as well. In the following let us briefly recall the notion of an ε-optimal point for nonlinear problems. To this end, we consider an optimization problem of the form

$$\min_{x \in \mathbb{R}^n} \; f(x) \qquad \text{s. t.} \qquad x \in M$$

where f is a possibly non-convex function and M is a feasible set. Then, for some $\varepsilon > 0$ a feasible point x^* with $f(x^*) \leq f(x) + \varepsilon$ for all $x \in M$ is said to be an *ε-optimal point*. That means, although x^* might not be a globally optimal point, other points can be at most slightly better. This idea possesses a natural extension to *ε-approximate Nash equilibria* or, briefly, *ε-Nash equilibria* for NEPs (cf., e.g., [30]). In the following, note that we also use a strict inequality in addition to the standard definition from the literature. While for the mere existence of some $\varepsilon > 0$ this difference is irrelevant, this will be helpful to describe certain inner and outer approximations in the remainder of this article.

Definition 2.1. *Given a Nash equilibrium problem and some $\varepsilon > 0$, we say that a point $(\bar{x}^1, \ldots, \bar{x}^N) \in \Omega$ is an* ε-Nash equilibrium*, if for every player $\nu = 1, \ldots, N$ it holds*

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) + \varepsilon \quad \text{for all} \quad x^\nu \in \Omega_\nu. \tag{1}$$

We denote the set of all ε-Nash equilibria by E_ε.

If inequality (1) *holds strictly, we refer to $(\bar{x}^1, \ldots, \bar{x}^N) \in \Omega$ as a* strict ε-Nash equilibrium*. Moreover, the set of all strict ε-Nash equilibria is denoted by $E_\varepsilon^<$.*

Although there might be an incentive for some players to deviate from a (strict or non-strict) ε-Nash equilibrium, this incentive is rather small, depending on the choice of ε.

3 A conceptual algorithm for non-convex NEPs with box constraints

In this section we first introduce the main idea of our new method, which is based on a so-called discarding criterion that helps to recognize sets that do

not contain any Nash equilibria. This is described in Section 3.1 and allows us to establish the conceptual version of our new algorithm in Section 3.2. While its finite termination is easily seen, a convergence property with regard to the set of Nash equilibria is studied in Section 3.3.

This first version of the method is preliminary in the sense that we do not yet include fathoming steps here, so that the described approach might not be efficient. Moreover, the termination criterion is very simple and is not based on the notion of ε-Nash equilibria yet. These features will be added in the remaining sections, as improvements of the basic version of the method described here.

3.1 Basic idea of a discarding criterion

A main idea of this article is to establish and apply a criterion that helps to exclude boxes that cannot contain any Nash equilibria. In order to derive this, let us start by considering a single point $(x^1, \dots, x^N) \in \Omega$. Moreover, we consider a player ν and a point $(y^1, \dots, y^N) \in \Omega$ with $y^{-\nu} = x^{-\nu}$ and $\theta_\nu(y^\nu, x^{-\nu}) < \theta_\nu(x^\nu, x^{-\nu})$. Then the point $x = (x^1, \dots, x^N)$ cannot be a Nash equilibrium, and we say that y is *ν-dominated* by x.

Let us now extend this concept to an entire box X. Hence, we assume that for some ν and every point $x \in X$ there is a point $y(x) \in \Omega$ that ν-dominates x, that is, $y^{-\nu}(x) = x^{-\nu}$ and $\theta_\nu(y^\nu(x), x^{-\nu}) < \theta_\nu(x^\nu, x^{-\nu})$ hold. Then the box X cannot contain any Nash equilibria. These considerations motivate the following result.

Proposition 3.1. *Let a box $X \subseteq \Omega$ as well as a set $Z \subseteq \Omega$ be given. Furthermore, let us assume that there is at least one player ν such that the following two conditions hold:*

- *For the projections $\mathrm{pr}_{x^{-\nu}} X$ of the box X and $\mathrm{pr}_{x^{-\nu}} Z$ of the set Z onto the $x^{-\nu}$-space we have the relation $\mathrm{pr}_{x^{-\nu}} X \subseteq \mathrm{pr}_{x^{-\nu}} Z$.*
- *There is a lower bound $\ell_{\theta_\nu}(X)$ at the function θ_ν over the box X and an upper bound $u_{\theta_\nu}(Z)$ at the function θ_ν over the set Z such that the inequality $\ell_{\theta_\nu}(X) > u_{\theta_\nu}(Z)$ holds.*

Then, the box X does not contain any Nash equilibria.

Proof. Due to $\mathrm{pr}_{x^{-\nu}} X \subseteq \mathrm{pr}_{x^{-\nu}} Z$, for every point $x \in X$ there is a point $z(x) \in Z$ with $z^{-\nu}(x) = x^{-\nu}$. Moreover, it holds

$$\theta_\nu(x) \quad \geq \quad \ell_{\theta_\nu}(X) \quad > \quad u_{\theta_\nu}(Z) \quad \geq \quad \theta_\nu(z(x))$$

and, thus, the point $z(x)$ is ν-dominated by x. Since x is chosen arbitrarily in X there cannot be a Nash equilibrium in this box and, thus, the assertion holds. □

The role of the set Z in Proposition 3.1 is reminiscent of the incumbent in classical branch-and-bound algorithms. There, the incumbent is some feasible point, and a box is discarded if a lower bound over this box exceeds the objective value at the incumbent (because the incumbent possesses a better objective value than all points from the box). Since in NEPs different players may have different incumbents, and these may even depend on the other players' decisions, the generalization of this bounding concept is the main challenge in the design of a branch-and-bound method for NEPs. The next section will introduce our approach to this problem, where the single incumbent from the classical setting is replaced by a list of boxes which may be used for discarding purposes.

3.2 Formal description of the algorithm

The main idea of the branch-and-bound method for finding Nash equilibria is to iteratively refine a list $\mathcal{N}$ of boxes that might contain Nash equilibria. The list is initialized to $\mathcal{N} := \{\Omega\}$ and, since Ω contains all feasible points, clearly, this is a superset of all Nash equilibria. During the algorithm boxes from this list are chosen and divided into sub-boxes along a longest edge. For all these sub-boxes we check whether they can be excluded from further consideration using a discarding criterion based on Proposition 3.1.

It remains to be clarified how a set Z in Proposition 3.1 may be suitably chosen. Whereas, in classical branch-and-bound algorithms for solving non-convex problems, as incumbents we can use feasible points with good objective values to discard boxes from the list, in the setting of NEPs this is not possible, as motivated above. The reason is that, in order to rule out a box, we need at least one single player ν who has an incentive to unilaterally deviate from all points within this box, and this fact forms the challenge for the design of a branch-and-bound method for NEPs. In our new algorithm this is reflected by the existence of a second list $\mathcal{L}$. In this list we keep boxes that can be used to exclude other boxes from the list of possible Nash equilibria $\mathcal{N}$, where points contained in boxes of the list $\mathcal{L}$ may or may not be Nash equilibria.

In order to improve the list $\mathcal{N}$ and drive it closer to an approximation of the set of all Nash equilibria, as common in spatial branch-and-bound algorithms

we choose boxes from this list and divide those in order to compute improved bounds at functions over these boxes. In addition to this, boxes in the list $\mathcal{L}$ must be refined as well, in order to obtain improved bounds for excluding boxes in $\mathcal{N}$. Moreover, since every player might have a reason to deviate and, thus, every player might provide a reason to exclude boxes, for every player $\nu = 1, \ldots, N$ we may choose a different box from the list $\mathcal{L}$ and divide all these boxes. In the following, the selection of boxes is described more precisely.

In fact, in every iteration we start by choosing a box $X = X_1 \times \ldots \times X_N$ with a largest diagonal length in the list $\mathcal{N}$, divide it into two sub-boxes and compute improved bounds at the functions θ_ν. Using this, we try to discard these sub-boxes of X by applying a criterion based on Proposition 3.1. That means that we search for possibilities to deviate for every player $\nu = 1, \ldots, N$. At least in theory we could do so by using a second branch-and-bound algorithm and try to compute a kind of globally optimal deviations for every player $\nu = 1, \ldots, N$. However, from the computational point of view this is prohibitive. For this reason, the idea is to use the boxes that are already available in the list $\mathcal{L}$.

More precisely, in every iteration we consider the sub-lists

$$\mathcal{L}_\nu \;=\; \mathcal{L}_\nu(X) \;:=\; \Big\{ Y \in \mathcal{L} \;\Big|\; \Big(\operatorname{pr}_{x^{-\nu}} Y \Big) \cap \Big(\operatorname{pr}_{x^{-\nu}} X \Big) \neq \emptyset \Big\}, \quad \nu = 1, \ldots, N, \tag{2}$$

of $\mathcal{L}$ containing boxes that are of interest for player ν, since they may contain points which unilaterally improve points $x \in X$ in the player variable x^ν.

Now, for every player ν we search for a box Y^ν in $\mathcal{L}_\nu$ that provides a smallest lower bound $\ell_{\theta_\nu}(Y^\nu)$ over all boxes $Y \in \mathcal{L}_\nu$. After doing so we treat the ν-part $\widehat{y}^\nu$of the midpoint $\widehat{y}$ of this box as a possible deviation for player ν, depending on the objective value of this point. However, a single point in the box Y^ν cannot be used to apply a discarding criterion as described in Proposition 3.1. Thus, based on this current approximation, we construct yet another box Z^ν which results from X by replacing its ν-part with the singleton $[\widehat{y}^\nu, \widehat{y}^\nu]$, that is, we put

$$Z^\nu := X_1 \times \ldots \times X_{\nu-1} \times [\widehat{y}^\nu, \widehat{y}^\nu] \times X_{\nu+1} \times \ldots \times X_N.$$

In view of $\operatorname{pr}_{x^{-\nu}} X = \operatorname{pr}_{x^{-\nu}} Z^\nu$ the lower-dimensional box Z^ν is formally suitable for a discarding test based on Proposition 3.1. Moreover, its choice is promising for the calculation of a small upper bound.

In addition to this, it is also important to refine boxes in the list $\mathcal{L}$ in order to obtain improved approximations at globally optimal points for the player

problems $P_\nu(x^{-\nu})$ later on in the algorithm. To this end, in the same iteration for each player $\nu = 1, \dots, N$ we divide the boxes Y^ν from the previous construction. We divide these boxes into sub-boxes along the midpoint of a longest edge, which enables us to compute improved bounds at functions in subsequent iterations. After doing so, these sub-boxes are added to the list $\mathcal{L}$ again.

An important step during classical branch-and-bound methods is the fathoming step which removes boxes from the maintained lists that have become superfluous after the improvement of bounds, and that we have not considered up to now. While we try to discard boxes from $\mathcal{N}$ as often as possible, so far we never exclude boxes from $\mathcal{L}$. For now, it is sufficient to understand that it is, indeed, possible to rule out certain boxes in $\mathcal{L}$, as long as we ensure that there are sufficiently good points contained in boxes of this list. Clearly, globally optimal points of problems $P_\nu(x^{-\nu})$ can serve this purpose. Formally, this can be required as follows.

Assumption 3.2. *In every iteration of the algorithm the union of the boxes in the list $\mathcal{L}$ contains all pairs $(x^\nu, x^{-\nu})$ for which x^ν is a globally minimal point of $P_\nu(x^{-\nu})$, $\nu = 1, \dots, N$, that is,*

$$\bigcup_{\nu=1}^{N} \operatorname{gph} S_\nu \subseteq \bigcup_{Y \in \mathcal{L}} Y$$

holds.

In our conceptual algorithm Assumption 3.2 is simply guaranteed by never removing a box from the list $\mathcal{L}$ at all. We will address this topic in more detail in Section 4. Under Assumption 3.2 it is clear, however, that the list $\mathcal{L}$ may be not interpreted as a list of incumbents which approximates the set of Nash equilibria $\bigcap_{\nu=1}^{N} \operatorname{gph} S_\nu$ upon termination, which would be the analogy the classical branch-and-bound methods. Instead, this task will be achieved by the list $\mathcal{N}$, as we will see in the next section.

Formally, the algorithm as described so far is stated in Algorithm 1.

3.3 A convergence property

For a termination tolerance $\tau > 0$ Algorithm 1 terminates after a finite number of iterations, since we always choose a largest box from the list $\mathcal{N}$, and dividing boxes along longest edges drives their width to zero. Moreover, using Proposition 3.1 it is ensured that only boxes are excluded from $\mathcal{N}$ that

Algorithm 1: Branch-and-bound algorithm for non-convex, box-constrained NEPs

Input: Non-convex NEP with box constraints, termination tolerance $\tau > 0$

Output: List $\mathcal{N}$ of boxes with width below τ, whose union forms a superset of the set of all Nash equilibria

Initialization: Put list $\mathcal{L} := \{\Omega\}$, list $\mathcal{N} := \{\Omega\}$;

while $\exists X' \in \mathcal{N}$ *with* $\operatorname{diag}(X') > \tau$ **do**

- *Step 1:* Choose box: Choose some $X \in \mathcal{N}$ with $\operatorname{diag}(X) \geq \operatorname{diag}(X')$ for all $X' \in \mathcal{N}$;
- *Step 2:* Remove X from the list $\mathcal{N}$;
- *Step 3:* Divide X along the midpoint of a longest edge into X^1 and X^2;
- *Step 4:* Add X^1 and X^2 to the list $\mathcal{N}$;
- *Step 5a:* Try to discard box X^1:
- **for** $\nu = 1, \ldots, N$ **do**
  - Find a box $Y^\nu \in \mathcal{L}_\nu(X^1)$ with $\ell_{\theta_\nu}(Y^\nu) \leq \ell_{\theta_\nu}(Y)$ for all $Y \in \mathcal{L}_\nu(X^1)$;
  - Denote the corresponding midpoint $\operatorname{mid}(Y^\nu)$ by $(\widehat{y}^1, \ldots, \widehat{y}^N)$;
  - Put $Z^\nu := X_1^1 \times \ldots \times [\widehat{y}^\nu, \widehat{y}^\nu] \times \ldots \times X_N^1$;
  - **if** $\ell_{\theta_\nu}(X^1) > u_{\theta_\nu}(Z^\nu)$ **then**
    - Remove X^1 from the list $\mathcal{N}$;
  - **end**
- **end**
- *Step 5b:* Proceed analogously for box X^2;
- *Step 6:* Improve boxes for discarding criterion:
- **for** $\nu = 1, \ldots, N$ **do**
  - Remove Y^ν from $\mathcal{L}$;
  - Divide Y^ν along a longest edge into sub-boxes;
  - Add both sub-boxes to the list $\mathcal{L}$;
- **end**
- *Step 7:* Fathoming step for $\mathcal{L}$ (according to Assumption 3.2);

end

do not contain any Nash equilibria. For that reason it is straightforward to see that in any iteration, and thus also upon termination, the union of all boxes in the list $\mathcal{N}$ forms a superset of the set of all Nash equilibria. While this shows the finite termination of Algorithm 1, so far it is not clear whether for termination tolerances τ tending to zero the set $\bigcup_{X\in\mathcal{N}} X$ actually converges to the set E of all Nash equilibria in some sense. This is indeed the case and will be discussed in the current subsection.

The following lemma will be helpful.

Lemma 3.3. *We consider the infinite branch-and-bound procedure of Algorithm 1 corresponding to $\tau = 0$ with an M-independent and convergent lower bounding procedure. Moreover, we assume that the fathoming step fulfills Assumption 3.2. Then, there is an exhaustive sequence of boxes $(X^s)_{s\in\mathbb{N}}$ created by the algorithm with $X^s \in \mathcal{N}$ for all $s \in \mathbb{N}$ as well as a point $(x^1, \ldots, x^N) \in X^s$ for all $s \in \mathbb{N}$.*

Furthermore, for any such sequence $(X^s)_{s\in\mathbb{N}}$ with corresponding $(x^1, \ldots, x^N)$ and for every player $\nu \in \{1, \ldots, N\}$ there are also exhaustive sequences of boxes $(X^{\nu,t})_{t\in\mathbb{N}}$ and $(Y^{\nu,t})_{t\in\mathbb{N}}$ created by the algorithm with $(y^\nu, x^{-\nu}) \in Y^{\nu,t}$ for all $t \in \mathbb{N}$ and $y^\nu \in S_\nu(x^{-\nu})$ and where $(X^{\nu,t})_{t\in\mathbb{N}}$ is a sub-sequence of $(X^s)_{s\in\mathbb{N}}$.

Proof. Since the algorithm does not terminate for $\tau = 0$, there are at least one exhaustive sequence of boxes $(X^s)_{s\in\mathbb{N}}$ and a corresponding unique point $(x^1, \ldots, x^N) \in \bigcap_{s\in\mathbb{N}} X^s$. We continue to establish the existence of the asserted exhaustive sequences of boxes $(Y^{\nu,t})_{t\in\mathbb{N}}$ for $\nu = 1, \ldots, N$. Based on this, we shall prove that there are minimal points y^ν of $P_\nu(x^{-\nu})$ with the asserted property.

Let us consider an exhaustive sequence of boxes $(X^s)_{s\in\mathbb{N}}$ as well as a fixed player ν. Moreover, without loss of generality we limit our consideration to iterations where a box X^s from the exhaustive sequence of boxes is chosen from the list $\mathcal{N}_s$, which denotes the list $\mathcal{N}$ in iteration s. Then, in the same iteration we choose boxes from the list $\mathcal{L}_\nu(X^s)$. Furthermore, we have

$$\bigcup_{Y\in\mathcal{L}_\nu(X^s)} Y \subseteq \bigcup_{Y\in\mathcal{L}_\nu(X^{\widehat{s}})} Y \qquad \text{for} \qquad \widehat{s} < s.$$

Since the list $\mathcal{L}_\nu(X^s)$ contains only a finite number of elements, there is indeed such an exhaustive sequence of boxes $(Y^{\nu,t})_{t\in\mathbb{N}}$. This can be seen as follows. For a fixed iteration $\widehat{s}$ there is a finite number of boxes in $\mathcal{L}_\nu(X^{\widehat{s}})$ and every such box can be interpreted as a node of the branch-and-bound tree. Further

subdividing these boxes leads to a new but still finite number of boxes that partitions the set $\bigcup_{Y \in \mathcal{L}_\nu(X^s)} Y$. Thus, in view of the aforementioned inclusion property on every level of the branch-and-bound tree there is a finite number of boxes. Since in the infinite branch-and-bound procedure an infinite number of boxes is created, there must be a nested sequence of boxes $(Y^{\nu,t})_{t\in\mathbb{N}}$ and in view of the box division rule we also have $\lim_{t\to\infty} \operatorname{diag}(Y^{\nu,t}) = 0$. This line of arguments is standard in the theory of global optimization. For a more formal explanation we refer to the literature (see, e.g., the proof of Theorem IV.1. as well as Corollary IV.1. in [15]).

Next, as for any exhaustive sequence of boxes, there is a point $(y^1, \dots, y^N) \in Y^{\nu,t}$ for all $t \in \mathbb{N}$. We complete the proof by showing $y^{-\nu} = x^{-\nu}$ and the minimality of y^ν for $P_\nu(x^{-\nu})$. Let us assume that there is some $\mu \neq \nu$ such that $y^\mu \neq x^\mu$. Then, due to $\lim_{s\to\infty} \operatorname{diag}(X^s) = 0$ and $\lim_{t\to\infty} \operatorname{diag}(Y^{\nu,t}) = 0$, for s and t sufficiently large we have

$$\left(\operatorname{pr}_{x^{-\nu}} X^s\right) \cap \left(\operatorname{pr}_{x^{-\nu}} Y^{\nu,t}\right) = \emptyset$$

and, by (2), it follows $Y^{\nu,t} \notin \mathcal{L}_\nu^s$. This, however, contradicts the choice of $Y^{\nu,t}$ and proves $y^{-\nu} = x^{-\nu}$.

It remains to be shown that $y^\nu \in S_\nu(x^{-\nu})$ holds. In order to see this, assume that there is a minimal point $\widehat{y}^\nu$ of $P_\nu(x^{-\nu})$ with $\theta_\nu(y^\nu, x^{-\nu}) > \theta_\nu(\widehat{y}^\nu, x^{-\nu})$. Due to the convergence of the lower bounding procedure we have $\lim_{t\to\infty} \ell_{\theta_\nu}(Y^{\nu,t}) = \theta_\nu(y^\nu, y^{-\nu}) = \theta_\nu(y^\nu, x^{-\nu})$ and therefore $\ell_{\theta_\nu}(Y^{\nu,t}) > \theta_\nu(\widehat{y}^\nu, x^{-\nu})$ for some sufficiently large t. On the other hand, by $(\widehat{y}^\nu, x^{-\nu}) \in \operatorname{gph} S_\nu$ and since in view of Assumption 3.2 we never exclude a box from $\mathcal{L}$ that contains an element of $\operatorname{gph} S_\nu$, there is a box $\widehat{Y} \in \mathcal{L}_\nu^t$ that contains the point $(\widehat{y}^\nu, x^{-\nu})$, yielding $\ell_{\theta_\nu}(Y^{\nu,t}) > \theta_\nu(\widehat{y}^\nu, x^{-\nu}) \geq \ell_{\theta_\nu}(\widehat{Y})$. Thus, the box $\widehat{Y} \in \mathcal{L}_\nu^t$ would have been chosen for further subdivision in Step 5 and 6 of the algorithm and, thus, the sequence $(Y^{\nu,t})$ is not created at all, which contradicts our assumption. For every player ν the asserted sub-sequence $(X^{\nu,t})_{t\in\mathbb{N}}$ is now obtained by omitting those elements of the sequence $(X^s)_{s\in\mathbb{N}}$ where no corresponding element $(Y^{\nu,t})_{t\in\mathbb{N}}$ is chosen. □

We can now prove the announced approximation property for the set E of Nash equilibria.

Theorem 3.4. *We consider the infinite branch-and-bound sequence of Algorithm 1 corresponding to $\tau = 0$ with an M-independent and convergent lower bounding procedure, where we assume that the fathoming step fulfills*

Assumption 3.2. Moreover, we put

$$E^k := \bigcup_{X \in \mathcal{N}_k} X,$$

where $\mathcal{N}_k$ denotes the list $\mathcal{N}$ in iteration k of the algorithm. Then with the Hausdorff metric δ we have $\lim_{k\to\infty} \delta(E^k, E) = 0$.

Proof. Assume that there is a point $\widetilde{x} = (\widetilde{x}^1, \ldots, \widetilde{x}^N)$ outside of E (i.e., $\widetilde{x}$ is not a Nash equilibrium) which is never discarded by Algorithm 1 after a finite number of iterations, even for the infinite branch-and-bound sequence corresponding to $\tau = 0$. In view of the box selection in Step 1 of Algorithm 1 there must then be an exhaustive sequence of boxes $(X^s)_{s\in\mathbb{N}}$ created by the algorithm with $\widetilde{x} \in X^s$ for all $s \in \mathbb{N}$.

By Lemma 3.3 there is also an exhaustive sequence of boxes $(Y^{\nu,t})_{t\in\mathbb{N}}$ created by the algorithm with a point $(y^\nu, \widetilde{x}^{-\nu}) \in Y^{\nu,t}$ for all $t \in \mathbb{N}$ which satisfies $y^\nu \in S_\nu(\widetilde{x}^{-\nu})$. Thus we have $\lim_{t\to\infty} \operatorname{mid}(Y^{\nu,t}) = (y^\nu, \widetilde{x}^{-\nu})$. Moreover, there is a sub-sequence $(X^{\nu,t})_{t\in\mathbb{N}}$ of $X^s_{s\in\mathbb{N}}$ and for the convergent lower bounding procedure we obtain $\lim_{s\to\infty} \ell_{\theta_\nu}(X^s) = \lim_{t\to\infty} \ell_{\theta_\nu}(X^{\nu,t}) = \theta_\nu(\widetilde{x})$. Since $\widetilde{x}$ is not a Nash equilibrium, we obtain

$$\theta_\nu(y^\nu, \widetilde{x}^{-\nu}) \quad < \quad \theta_\nu(\widetilde{x}^\nu, \widetilde{x}^{-\nu}). \tag{3}$$

Similarly, for every box $X^{\nu,t}$ there is a box $Z^{\nu,t}$ created by the algorithm such that $\operatorname{pr}_{x^{-\nu}} X^{\nu,t} = \operatorname{pr}_{x^{-\nu}} Z^{\nu,t}$. From $0 \le \operatorname{diag}(Z^{\nu,t}) \le \operatorname{diag}(X^{\nu,t})$ and $\lim_{t\to\infty} \operatorname{diag}(X^{\nu,t}) = 0$ we obtain $\lim_{t\to\infty} \operatorname{diag}(Z^{\nu,t}) = 0$. Moreover, we have $\lim_{t\to\infty} \operatorname{mid}(Z^{\nu,t}) = (y^\nu, \widetilde{x}^{-\nu})$, since $\lim_{t\to\infty} \operatorname{mid}(Y^{\nu,t}) = (y^\nu, \widetilde{x}^{-\nu})$ and $\lim_{t\to\infty} \operatorname{mid}(X^{\nu,t}) = (\widetilde{x}^\nu, \widetilde{x}^{-\nu})$).

We now consider the sequence of boxes $(\widehat{Z}^{\nu,t})_{t\in\mathbb{N}}$ that are defined as the smallest boxes that contain all points in $Z^{\nu,t}$ as well as the point $(y^\nu, \widetilde{x}^{-\nu})$, which is an exhaustive sequence of boxes with $(y^\nu, \widetilde{x}^{-\nu}) \in \widehat{Z}^{\nu,t}$ for all $t \in \mathbb{N}$. Thus, for every convergent upper bounding procedure we have $\lim_{t\to\infty} u_{\theta_\nu}(\widehat{Z}^{\nu,t}) = \theta_\nu(y^\nu, \widetilde{x}^{-\nu})$. Moreover, we have

$$u_{\theta_\nu}(\widehat{Z}^{\nu,t}) \quad \ge \quad u_{\theta_\nu}(Z^{\nu,t}) \quad \text{for all} \;\; t \in \mathbb{N}$$

and in view of inequality (3) we have for sufficiently large t (i.e., sufficiently small boxes) the inequality

$$\ell_{\theta_\nu}(X^{\nu,t}) \quad > \quad u_{\theta_\nu}(\widehat{Z}^{\nu,t}) \quad \ge \quad u_{\theta_\nu}(Z^{\nu,t}).$$

Thus, using Proposition 3.1 we see that the box $X^{\nu,t}$ cannot contain a Nash equilibrium for sufficiently large t and, hence, it is removed from the list $\mathcal{N}$ by Algorithm 1 in Step 5. Moreover, since $(X^{\nu,t})_{t\in\mathbb{N}}$ is a sub-sequence of $(X^s)_{s\in\mathbb{N}}$ this holds for these boxes analogously, in contradiction to our assumption.

Finally, we have $E \subseteq E^k$ for all $k \in \mathbb{N}$. Since the algorithm removes every point $\widetilde{x}$ which is not a Nash equilibrium after a finite number of iterations, we have $\lim_{k\to\infty} \delta(E^k, E) = 0$. □

4 Fathoming steps

An important step in classical branch-and-bound methods is the fathoming step, which is crucial in order to obtain suitable performance. So far, although we try to discard boxes from the list $\mathcal{N}$, we never exclude any boxes from the list $\mathcal{L}$. However, from the theory of global optimization we know that it is important to accelerate these methods and, thus, we also include such a fathoming step in our new algorithm for the computation of Nash equilibria.

So far, in our current version of the algorithm as stated in Algorithm 1 we always have $\bigcup_{Y\in\mathcal{L}} Y = \Omega$, because whenever we remove some box Y from $\mathcal{L}$ in order to divide it, we immediately add the resulting sub-boxes to $\mathcal{L}$ again within the same iteration. However, one may ask if it is really necessary to consider all these boxes, since from a computational point of view it would be beneficial to exclude some of them.

In fact, in order to make sure that a single point $x = (x^\nu, x^{-\nu})$ is not a Nash equilibrium, it is sufficient to show that there is at least *one single* globally minimal point $\bar{x}^\nu$ of $P_\nu(x^{-\nu})$ that yields an incentive for this player to deviate. Conversely, if a box $Y \in \mathcal{L}$ does not contain a point $(x^\nu, x^{-\nu}) \in \operatorname{gph} S_\nu$ for any player $\nu = 1, \ldots, N$, that is, $Y \subseteq \bigcap_{\nu=1}^N (\operatorname{gph} S_\nu)^c = \left(\bigcup_{\nu=1}^N \operatorname{gph} S_\nu\right)^c$ holds, we may safely delete this box from $\mathcal{L}$, as long as we make sure to always keep boxes that contain points that might solve at least one problem of a player ν. Note that by doing so we ensure Assumption 3.2. Since this is an important observation, we state it in the following remark.

Remark 4.1 (Fathoming criterion)**.** *Let $Y \in \mathcal{L}$ be a box such that for every $\nu = 1, \ldots, N$ and for every $(x^\nu, x^{-\nu}) \in Y$ we have $x^\nu \notin S_\nu(x^{-\nu})$. Then Y can be deleted from the list $\mathcal{L}$.*

The question is, how such a box $Y \in \mathcal{L}$ can be recognized. To this end, we may apply a similar idea as already used in the proof of Theorem 3.4. For

every $\nu \in \{1, \ldots, N\}$, in anology to (2) we may construct the list

$$\mathcal{L}_\nu(Y) \;=\; \Big\{ Y' \in \mathcal{L} \;\Big|\; \Big(\mathrm{pr}_{x^{-\nu}} Y' \Big) \cap \Big(\mathrm{pr}_{x^{-\nu}} Y \Big) \neq \emptyset \Big\}.$$

Next, we find a box $Y^\nu \in \mathcal{L}_\nu(Y)$ with $\ell_{\theta_\nu}(Y^\nu) \leq \ell_{\theta_\nu}(Y')$ for all $Y' \in \mathcal{L}_\nu(Y)$. Then, analogously to what is done in Algorithm 1 for a box X from the list $\mathcal{N}$, we may proceed for the box $Y = Y_1 \times \ldots \times Y_N$ from the list $\mathcal{L}$. Hence, denoting the midpoint of Y^ν by $(\widehat{y}^1, \ldots, \widehat{y}^N)$ we put $Z^\nu = Y_1 \times \ldots \times [\widehat{y}^\nu, \widehat{y}^\nu] \times \ldots \times Y_N$. If we have

$$u_{\theta_\nu}(Z^\nu) \quad < \quad \ell_{\theta_\nu}(Y)$$

for *all* $\nu = 1, \ldots, N$, then the box Y cannot contain a minimal point of $P_\nu(y^{-\nu})$ for any $(y^\nu, y^{-\nu}) \in Y$, since we have

$$\min_{x^\nu \in \Omega_\nu} \theta_\nu(x^\nu, y^{-\nu}) \quad \leq \quad u_{\theta_\nu}(Z^\nu) \quad < \quad \ell_{\theta_\nu}(Y) \quad \leq \quad \theta_\nu(y^\nu, y^{-\nu}).$$

Note, however, that this fathoming step is much more elaborate than classical fathoming steps for nonlinear problems in global optimization. This holds especially in view of the fact that this has to be checked for all players $\nu = 1, \ldots, N$. Most likely it takes a substantial amount of time to proceed like this, and so there might be a trade-off between whether it takes more time to search in the complete list of boxes or to use this fathoming step that might help to shorten the list. Moreover, this can be expected to be highly dependent on implementation details. It might be beneficial to not use this fathoming step in every iteration.

Formally, the fathoming step is described in Algorithm 2, which may be used in Algorithm 1 in Step 7. Note that the proof of convergence is not affected by this, since Assumption 3.2 continues to hold, and so this situation has already been taken into consideration.

We remark that in principle also a fathoming step for the list $\mathcal{N}$ could be added to Algorithm 1, in analogy to Algorithm 2. We expect, however, that the additional computational effort of checking all boxes X in $\mathcal{N}$, including the computation of the sublists $\mathcal{N}_\nu(X)$ in analogy to (2), would be prohibitive, so that we refrain from this in the present study.

5 A suitable termination criterion

So far, our conceptual method as presented in Algorithm 1 does not include a convincing termination criterion. Instead, we simply stop as soon as all boxes have become sufficiently small. This, however, does not provide any

Algorithm 2: Fathoming step for branch-and-bound algorithm for non-convex, box-constrained NEPs

Input: Non-convex NEP with box constraints, current list $\mathcal{L}$ of Algorithm 1

Output: Possibly fathomed list $\mathcal{L}$

Every box in $\mathcal{L}$ is checked:

for $Y \in \mathcal{L}$ **do**

  Try to exclude box Y:

  Set remove_from_$\mathcal{L}$:=true;

  Check for all player specific problems:

  for $\nu = 1, \ldots, N$ **do**

    Find a box $Y^\nu \in \mathcal{L}_\nu(Y)$ with $\ell_{\theta_\nu}(Y^\nu) \leq \ell_{\theta_\nu}(Y')$ for all $Y' \in \mathcal{L}_\nu(Y)$;

    Denote the corresponding midpoint $\mathrm{mid}(Y^\nu)$ by $(\widehat{y}^1, \ldots, \widehat{y}^N)$;

    Put $Z^\nu := Y_1 \times \ldots \times [\widehat{y}^\nu, \widehat{y}^\nu] \times \ldots \times Y_N$;

    if $\ell_{\theta_\nu}(Y) \leq u_{\theta_\nu}(Z^\nu)$ **then**

      remove_from_$\mathcal{L}$:=false;

    end

  end

  if *remove_from_* $\mathcal{L}$ **then**

    Remove Y from the list $\mathcal{L}$;

  end

end

information regarding the approximation quality for the set of Nash equilibria, except that we know that we compute supersets. In the current section we improve this behavior of the algorithm, which, however, comes at the cost of computing ε-Nash equilibria instead of exact Nash equilibria.

In the next subsection we present the main idea to achieve this. Based on this, we finally describe our complete branch-and-bound algorithm including the improved termination criterion formally in Section 5.2. In Section 5.3 we show that our improved termination criterion is fulfilled after a finite number of iterations so that the algorithm terminates, provided that some mild assumptions hold.

5.1 Inner and outer approximations

As already indicated at the beginning of this article, we now aim at the computation of ε-Nash equilibria. To this end, let us consider a box $X \in \mathcal{N}$ as well as upper bounds $u_{\theta_\nu}(X)$ for all players $\nu = 1, \ldots, N$. Moreover, we consider the player specific lists

$$\mathcal{L}_\nu(X) = \Big\{ Y \in \mathcal{L} \;\Big|\; \Big(\operatorname{pr}_{x^{-\nu}} Y \Big) \cap \Big(\operatorname{pr}_{x^{-\nu}} X \Big) \neq \emptyset \Big\}, \quad \nu = 1, \ldots, N$$

from (2). Recall that $\mathcal{L}_\nu(X)$ contains the boxes from $\mathcal{L}$ which are of interest for player ν, since they may contain points which unilaterally improve points $x \in X$ in the player variable x^ν.

In particular, in view of Assumption 3.2 the list $\mathcal{L}$ is maintained in such a way that for every player ν and for every $(x^\nu, x^{-\nu}) \in X$, each minimal point y^ν of $P_\nu(x^{-\nu})$ satisfies $(y^\nu, x^{-\nu}) \in Y$ for some $Y \in \mathcal{L}_\nu(X)$. Thus, we have

$$\min_{Y \in \mathcal{L}_\nu(X)} \ell_{\theta_\nu}(Y) \quad \leq \quad \theta_\nu(y^\nu, x^{-\nu}) \quad \leq \quad \theta_\nu(x^\nu, x^{-\nu}) \quad \leq \quad u_{\theta_\nu}(X) \qquad (4)$$

for all $(x^\nu, x^{-\nu}) \in X$ and all $\nu = 1, \ldots, N$. If, in addition, the difference between the upper bound $u_{\theta_\nu}(X)$ and $\min_{Y \in \mathcal{L}_\nu(X)} \ell_{\theta_\nu}(Y)$ becomes sufficiently small, then we know that there is at most a small incentive for player ν to deviate from any point $(x^\nu, x^{-\nu}) \in X$. In fact, if we have

$$u_{\theta_\nu}(X) - \min_{Y \in \mathcal{L}_\nu(X)} \ell_{\theta_\nu}(Y) \quad < \quad \varepsilon$$

for all players $\nu = 1, \ldots, N$, then all points in X are strict ε-Nash equilibria, since we may then rearrange (4) to

$$\theta_\nu(x^\nu, x^{-\nu}) \;\leq\; u_{\theta_\nu}(X) \;<\; \min_{Y \in \mathcal{L}_\nu(X)} \ell_{\theta_\nu}(Y) + \varepsilon \;\leq\; \theta_\nu(y^\nu, x^{-\nu}) + \varepsilon$$

for any $x \in X$ and $\nu = 1, \ldots, N$.

Analogously to spatial branch-and-bound algorithms in global optimization, where one aims at ε-minimal points of nonlinear problems, here we are satisfied with the computation of (strict) ε-Nash equilibria as well. For this reason we copy these boxes into another list $\widetilde{\mathcal{N}}$. Boxes in this list are considered to be valid solutions of our branch-and-bound algorithm and will not be further sub-divided in the course of our method, but only those from $\mathcal{N} \setminus \widetilde{\mathcal{N}}$.

Unfortunately, we cannot expect that sooner or later all boxes from $\mathcal{N}$ are either excluded or added to the list $\widetilde{\mathcal{N}}$. In fact, it is possible that for a given $\varepsilon > 0$ even very small boxes contain, both, strict ε-Nash equilibria as well as points that are not ε-Nash equilibria. For that reason we consider the union of the boxes in the list $\widetilde{\mathcal{N}}$ as an inner approximation of the set $E_\varepsilon^<$ of all strict ε-Nash equilibria and the union of all boxes in $\mathcal{N}$ as an outer approximation of the set E_ε of all ε-Nash equilibria, that is, we have

$$\bigcup_{X \in \widetilde{\mathcal{N}}} X \;\subseteq\; E_\varepsilon^< \;\subseteq\; E_\varepsilon \;\subseteq\; \bigcup_{X \in \mathcal{N}} X.$$

We may terminate if the Hausdorff distance

$$\begin{aligned} \delta\left(\bigcup_{X \in \mathcal{N}} X, \bigcup_{\widetilde{X} \in \widetilde{\mathcal{N}}} \widetilde{X}\right) &= \max_{x \in \bigcup_{X \in \mathcal{N}} X} \min_{\widetilde{x} \in \bigcup_{\widetilde{X} \in \widetilde{\mathcal{N}}} \widetilde{X}} \|x - \widetilde{x}\|_2 \\ &= \max_{X \in \mathcal{N}} \max_{x \in X} \min_{\widetilde{X} \in \widetilde{\mathcal{N}}} \min_{\widetilde{x} \in \widetilde{X}} \|x - \widetilde{x}\|_2 \end{aligned}$$

between the outer and inner approximation of E_ε and $E_\varepsilon^<$ falls below a prescribed termination tolerance $\tau > 0$. Unfortunately, this formula itself is too costly to be evaluated in every iteration of the branch-and-bound method. However, the special structure of this formula allows the construction of an upper bound that is suited as a termination criterion for our algorithm. This is shown in the following result where, following a usual convention, a minimum over the empty set is considered the extended real value $+\infty$.

Lemma 5.1. *For $\mathcal{N} \neq \emptyset$ the Hausdorff distance between the inner and the outer approximation is bounded above in the form*

$$\delta\left(\bigcup_{X \in \mathcal{N}} X, \bigcup_{\widetilde{X} \in \widetilde{\mathcal{N}}} \widetilde{X}\right) \;\leq\; \max_{X \in \mathcal{N}} \min_{\widetilde{X} \in \widetilde{\mathcal{N}}} \|\Delta(X, \widetilde{X})\|_2,$$

where for boxes $A = [\underline{a}, \overline{a}]$ and $B = [\underline{b}, \overline{b}]$ the entries of the vector $\Delta(A, B)$ are

$$\Delta_i(A, B) \;=\; \max\{0, \underline{b}_i - \underline{a}_i, \overline{a}_i - \overline{b}_i\}, \quad i = 1, \ldots, n.$$

Proof. In the case $\widetilde{\mathcal{N}} = \emptyset$ the bound formally holds, since both terms coincide with the extended real value $+\infty$. Thus, in the following we may assume $\widetilde{\mathcal{N}} \neq \emptyset$. Due to

$$\begin{aligned}\delta\left(\bigcup_{X\in\mathcal{N}} X, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}} \widetilde{X}\right) &= \max_{X\in\mathcal{N}} \max_{x\in X} \min_{\widetilde{X}\in\widetilde{\mathcal{N}}} \min_{\widetilde{x}\in\widetilde{X}} \|x-\widetilde{x}\|_2 \\ &\leq \max_{X\in\mathcal{N}} \min_{\widetilde{X}\in\widetilde{\mathcal{N}}} \max_{x\in X} \min_{\widetilde{x}\in\widetilde{X}} \|x-\widetilde{x}\|_2\end{aligned}$$

it remains to explicitly compute the term $\max_{a\in A} \min_{b\in B} \|a-b\|_2$ for two boxes $A, B \subseteq \mathbb{R}^n$. We have

$$\begin{aligned}\left(\max_{a\in A} \min_{b\in B} \|a-b\|_2\right)^2 &= \max_{a\in A} \min_{b\in B} \sum_{i=1}^n (a_i-b_i)^2 \\ &= \sum_{i=1}^n \max_{a_i\in A_i} \min_{b_\in B_i} (a_i-b_i)^2 \\ &= \sum_{i=1}^n \left(\max_{a_i\in A_i} \min_{b_\in B_i} |a_i-b_i|\right)^2,\end{aligned}$$

where for each i the term $\max_{a_i\in A_i} \min_{b_\in B_i} |a_i-b_i|$ is the maximal distance of points in A_i from B_i. Since the latter is easily seen to coincide with $\Delta_i(A,B)$, we arrive at $\max_{a\in A} \min_{b\in B} \|a-b\|_2 = \|\Delta(A,B)\|_2$ and, thus, the assertion. □

5.2 Formal description of the branch-and-bound algorithm

In this section we discuss some more details of our new branch-and-bound procedure and describe the entire algorithm formally.

We start with a brief consideration of the box selection rule. In our conceptual version of the branch-and-bound in Algorithm 1 we always choose a largest box from the list $\mathcal{N}$ for further subdivision. This is in contrast to classical box selection strategies for spatial branch-and-bound algorithms in global optimization where it is common, for instance, to choose a box with a smallest lower bound, i.e. a box where the termination criterion is most violated. Other box selection rules are possible as well and also for our new method to compute ε-Nash equilibria we allow for some more flexibility in the following.

In fact, at least two aspects are required for a suitable box selection rule. First, if a box $X \in \mathcal{N}$ contains a point which is no ε-Nash equilibrium, then we need to ensure that this box is selected sooner or later in order to make sure that our outer approximation is improved. Second, if a box contains a strict ε-Nash equilibrium but is not contained in $\widetilde{\mathcal{N}}$ yet, then it must be selected after a finite number of iterations as well, because otherwise the inner approximation cannot be improved. In the sequel we shall prove that this is not only necessary but also sufficient for the convergence of our branch-and-bound procedure. Formally, it is stated in the following assumption.

Assumption 5.2. *Given a box $X \in \mathcal{N}$ such that one of the two conditions*

- *X contains a point which is no ε-Nash equilibrium,*
- *X contains a strict ε-Nash equilibrium and X is not contained in $\widetilde{\mathcal{N}}$,*

holds, then this box must be subdivided after finitely many iterations.

Note that the simple box selection rule from Algorithm 1, where we always choose a largest box in $\mathcal{N}$, fulfills this requirement.

The final version of our new branch-and-bound algorithm for the computation of all ε-Nash equilibria is described formally in Algorithm 3. This includes the improved termination criterion, possible fathoming steps as well as the more general box selection rule.

With regard to the usage of the two tolerances ε and τ in Algorithm 3 we remark that in classical branch-and-bound algorithms the termination tolerance ε on the gap between upper and lower bounds yields a measure for the approximation quality of the terminal incumbent, in the sense that it is an ε-minimal point. If there one aimed at computing even the whole set of ε-minimal points, one would need to fix ε, introduce an approximation tolerance τ for this set and modify the branch-and-bound algorithm accordingly. In particular, ε served as an approximation tolerance in the image space, and τ as an independent approximation tolerance in the decision space.

The same idea is implemented in Algorithm 3 for the computation of a τ-approximation of the set of ε-Nash equilibria. As opposed to classical branch-and-bound methods, here the computation of only a single ε-Nash equilibrium as the algorithmic output would not be satisfactory. This explains the presence of the two tolerances ε and τ.

Algorithm 3: Branch-and-bound algorithm for non-convex, box-constrained NEPs

Input: Non-convex NEP with box constraints, $\varepsilon > 0$, $\tau > 0$

Output: Outer approximation $\bigcup_{X \in \mathcal{N}} X$ at the set of all ε-Nash equilibria and inner approximation $\bigcup_{\widetilde{X} \in \widetilde{\mathcal{N}}} \widetilde{X}$ at the set of all strict ε-Nash equilibria with approximation quality τ, or certificate $\mathcal{N} = \emptyset$ that there are no ε-Nash equilibria

Initialization: Put list $\mathcal{L} := \{\Omega\}$, list $\mathcal{N} := \{\Omega\}$, list $\widetilde{\mathcal{N}} = \emptyset$;

while $\mathcal{N} \neq \emptyset$ **and** $\max_{X \in \mathcal{N}} \min_{\widetilde{X} \in \widetilde{\mathcal{N}}} \|\Delta(X, \widetilde{X})\|_2 > \tau$ **do**

- *Step 1:* Choose box: Select $X \in \mathcal{N}$ according to some rule that fulfills Assumption 5.2;
- *Step 2:* Remove X from the list $\mathcal{N}$;
- *Step 3:* Divide X along a longest edge into X^1 and X^2;
- *Step 4:* Add X^1 and X^2 to the list $\mathcal{N}$;
- *Step 5a:* Process box X^1:
- Set add_to_$\widetilde{\mathcal{N}}$:=true;
- **for** $\nu = 1, \ldots, N$ **do**
  - Find a box $Y^\nu \in \mathcal{L}_\nu(X^1)$ with $\ell_{\theta_\nu}(Y^\nu) \leq \ell_{\theta_\nu}(Y)$ for all $Y \in \mathcal{L}_\nu(X^1)$;
  - Denote the corresponding midpoint $\mathrm{mid}(Y^\nu)$ by $(\widehat{y}^1, \ldots, \widehat{y}^N)$;
  - Put $Z^\nu := X_1^1 \times \ldots \times [\widehat{y}^\nu, \widehat{y}^\nu] \times \ldots \times X_N^1$;
  - **if** $\ell_{\theta_\nu}(X^1) > u_{\theta_\nu}(Z^\nu)$ **then**
    - Remove X^1 from the list $\mathcal{N}$;
    - add_to_$\widetilde{\mathcal{N}}$:=false;
  - **else if** $u_{\theta_\nu}(X^1) - \min_{Y \in \mathcal{L}_\nu(X)} \ell_{\theta_\nu}(Y) \geq \varepsilon$ **then**
    - add_to_$\widetilde{\mathcal{N}}$:=false;
  - **end**
- **end**
- **if** *add_to_*$\widetilde{\mathcal{N}}$ **then**
  - Add X^1 to the list $\widetilde{\mathcal{N}}$;
- **end**
- *Step 5b:* Proceed analogously for box X^2;
- *Step 6:* Improve boxes for discarding criterion:
- **for** $\nu = 1, \ldots, N$ **do**
  - Remove Y^ν from $\mathcal{L}$;
  - Divide Y^ν along a longest edge into sub-boxes;
  - Add both sub-boxes to the list $\mathcal{L}$;
- **end**
- *Step 7:* Fathoming step for $\mathcal{L}$ (according to Assumption 3.2);

end

5.3 Proof of convergence for the branch-and-bound algorithm

In this section we prove that the inner and outer approximations converge to each other under mild assumptions and, thus, Algorithm 3 terminates after finitely many steps. To this end, we start with the convergence of the inner as well as convergence of the outer approximation. Finally, we combine these results in order to show termination of the proposed branch-and-bound algorithm under mild assumptions.

By construction, in every iteration k of the algorithm the union of the boxes in the list $\widetilde{\mathcal{N}}_k$ is an inner approximation of the set of strict ε-Nash equilibria $E_\varepsilon^<$. We now show that this approximation is not only improved in the course of the algorithm but also converges to the set of strict ε-Nash equilibria.

Proposition 5.3. *We consider a Nash equilibrium problem with $E_\varepsilon^< \neq \emptyset$ for some $\varepsilon > 0$ and the possibly infinite branch-and-bound procedure of Algorithm 3 corresponding to $\tau = 0$, where we assume that a convergent and M-independent lower bounding procedure as well as some box selection rule fulfilling Assumption 5.2 are used. Then, if the algorithm does not terminate, we have*

$$\lim_{k\to\infty} \delta\left(E_\varepsilon^<, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}\right) = 0$$

where $\widetilde{\mathcal{N}}_k$ is the list $\widetilde{\mathcal{N}}$ of the algorithm in iteration k.

Proof. We assume that the assertion does not hold and derive a contradiction. Hence, let us assume that there is a strict ε-Nash equilibrium $\bar{x} = (\bar{x}^1, \ldots, \bar{x}^N)$ and some $\sigma > 0$ such that

$$\operatorname{dist}\left(\bar{x}, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}\right) := \min_{x\in\bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k}\widetilde{X}} \|x - \bar{x}\|_2 > \sigma$$

holds in every iteration k of the algorithm. Then, in each iteration s, the point $\bar{x}$ is contained in some box $X^s \in \mathcal{N}$ which is neither discarded nor copied to $\widetilde{\mathcal{N}}$. In view of the box selection rule and Assumption 5.2 this box is branched in some later iteration, so that there is an exhaustive sequence of boxes $(X^s)_{s\in\mathbb{N}}$ with $\bar{x} \in X^s$ for all s.

By Lemma 3.3 for every player ν there are also exhaustive sequences of boxes $(X^{\nu,t})_{t\in\mathbb{N}}$ and $(Y^{\nu,t})_{t\in\mathbb{N}}$ created by the algorithm such that the following conditions holds:

- The sequence $(X^{\nu,t})_{t\in\mathbb{N}}$ is a sub-sequence of $(X^s)_{s\in\mathbb{N}}$ and for this reason we also have $\bar{x} \in X^{\nu,t}$ for all t.
- There is a point $(y^\nu, \bar{x}^{-\nu}) \in Y^{\nu,t}$ for all $t \in \mathbb{N}$, where y^ν is a globally minimal point of the problem $P_\nu(\bar{x}^{-\nu})$.

Since $\bar{x}$ is assumed to be a strict ε-Nash equilibrium, we also have

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(y^\nu, \bar{x}^{-\nu}) \quad < \quad \varepsilon.$$

For this reason, due to convergence of the bounding operations we obtain for every ν and for t sufficiently large the inequality

$$u_{\theta_\nu}(X^{\nu,t}) - \ell_{\theta_\nu}(Y^{\nu,t}) \quad < \quad \varepsilon. \tag{5}$$

Moreover, the boxes $Y^{\nu,t}$ are chosen such that $\ell_{\theta_\nu}(Y^{\nu,t}) = \min_{Y\in\mathcal{L}_\nu(X^{\nu,t})} \ell_{\theta_\nu}(Y)$ holds and, thus, inequality (5) represents exactly the criterion that is used to add the box $X^{\nu,t}$ to the list $\widetilde{\mathcal{N}}$. For this reason, this box will not be divided any further and, hence, neither the exhaustive sequences of boxes $(X^s)_{s\in\mathbb{N}}$ nor the sub-sequence $(X^{\nu,t})_{t\in\mathbb{N}}$ are created by the algorithm, which contradicts our assumption. □

We now consider the outer approximation of the set of ε-Nash equilibria E_ε. By construction, this may be obtained by taking the union of the boxes in the list $\mathcal{N}$. However, it still needs to be shown that this approximation eventually converges as well, which is done in the following.

Proposition 5.4. *We consider a Nash equilibrium problem with $E_\varepsilon \neq \emptyset$ for some $\varepsilon > 0$ and we consider the possibly infinite branch-and-bound procedure of Algorithm 3 corresponding to $\tau = 0$, where we assume that a convergent and M-independent lower bounding procedure as well as some box selection rule fulfilling Assumption 5.2 are used. Then, if the algorithm does not terminate, we have*

$$\lim_{k\to\infty} \delta\left(\bigcup_{X\in\mathcal{N}_k} X, E_\varepsilon\right) = 0$$

where $\mathcal{N}_k$ is the list $\mathcal{N}$ of the algorithm in iteration k.

Proof. We assume that the assertion does not hold and derive a contradiction. Hence, let us assume that there exists a point $\bar{x} = (\bar{x}^1, \ldots, \bar{x}^N) \in \bigcup_{X\in\mathcal{N}_k} X$ for all iterations k that is not an ε-Nash equilibrium. Then, in view of Assumption 5.2, every box that contains $\bar{x}$ must be divided sooner or later

and, thus, there is an exhaustive sequence of boxes $(X^s)_{s\in\mathbb{N}}$ with $\bar{x} \in X^s$ for all s.

Again, similar to the proof of Proposition 5.3, by Lemma 3.3 for every player ν there are also exhaustive sequences of boxes $(X^{\nu,t})_{t\in\mathbb{N}}$ and $(Y^{\nu,t})_{t\in\mathbb{N}}$ created by the algorithm such that the following conditions holds:

- The sequence $(X^{\nu,t})_{t\in\mathbb{N}}$ is a sub-sequence of $(X^s)_{s\in\mathbb{N}}$ and for this reason we also have $\bar{x} \in X^{\nu,t}$ for all t.
- There is a point $(y^\nu, \bar{x}^{-\nu}) \in Y^{\nu,t}$ for all $t \in \mathbb{N}$, where y^ν is a globally minimal point of the problem $P_\nu(\bar{x}^{-\nu})$.

By the minimality of y^μ for $P_\mu(\bar{x}^{-\nu})$, and since $\bar{x}$ is assumed to not be an ε-Nash equilibrium, there are at least one player μ and some $\widetilde{y}^\mu \in \Omega_\mu$ with

$$\theta_\mu(y^\mu, \bar{x}^{-\mu}) \quad \leq \quad \theta_\mu(\widetilde{y}^\mu, \bar{x}^{-\mu}) \quad < \quad \theta_\mu(\bar{x}^\mu, \bar{x}^{-\mu}) - \varepsilon.$$

Due to convergence of the lower and upper bounding procedures, for t sufficiently large we obtain

$$u_{\theta_\mu}(Y^{\mu,t}) \quad < \quad \ell_{\theta_\mu}(X^{\mu,t}) - \varepsilon.$$

With the same line of arguments as in the proof of Theorem 3.4 we may construct the corresponding boxes $Z^{\mu,t}$ and, again, analogously to this proof we obtain

$$u_{\theta_\mu}(Z^{\mu,t}) \quad < \quad \ell_{\theta_\mu}(X^{\mu,t}) - \varepsilon.$$

Therefore, for t sufficiently large a box $X^{\mu,t}$ is excluded by the algorithm and, thus, neither the exhaustive sequences of boxes $(X^s)_{s\in\mathbb{N}}$ nor the sub-sequence $(X^{\mu,t})_{t\in\mathbb{N}}$ is created by the algorithm, which contradicts our assumption. □

Hence, from Proposition 5.3 we obtain convergent inner approximations of the set $E_\varepsilon^<$ and from Proposition 5.4 we also have convergent outer approximations of the set E_ε. For this reason, it is also ensured that both approximations converge to each other, provided that the Hausdorff distance $\delta(E_\varepsilon, E_\varepsilon^<)$ is zero. Formally, this is stated in the next result.

Theorem 5.5. *We consider a Nash equilibrium problem with $E_\varepsilon^< \neq \emptyset$ for some $\varepsilon > 0$ and we consider the possibly infinite branch-and-bound procedure of Algorithm 3 corresponding to $\tau = 0$, where we assume that a convergent and M-independent lower bounding procedure as well as some box selection*

rule fulfilling Assumption 5.2 are used. Furthermore, we assume $\delta(E_\varepsilon, E_\varepsilon^<) = 0$. *Then, if the algorithm does not terminate, we have*

$$\lim_{k\to\infty} \delta\left(\bigcup_{X\in\mathcal{N}_k} X, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}\right) = 0$$

where $\mathcal{N}_k$ *is the list* $\mathcal{N}$ *of the algorithm in iteration* k *and* $\widetilde{\mathcal{N}}_k$ *denotes the list* $\widetilde{\mathcal{N}}$ *in the same iteration.*

Proof. Due to $E_\varepsilon^< \subseteq E_\varepsilon$ we know that $E_\varepsilon \neq \emptyset$ and, therefore, the assumptions of Proposition 5.3 and Proposition 5.4 are fulfilled. Hence, for every $\tau > 0$ and sufficiently large k we have the inequalities

$$\delta\left(\bigcup_{X\in\mathcal{N}_k} X, E_\varepsilon\right) \leq \tau \quad \text{and} \quad \delta\left(E_\varepsilon^<, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}\right) \leq \tau,$$

as well as $\delta(E_\varepsilon, E_\varepsilon^<) = 0$ by assumption. In a next step we intend to use the triangle inequality for the Hausdorff metric which holds, however, only on the set of nonempty compact subsets of $\mathbb{R}^n$. Indeed, the sets $\bigcup_{X\in\mathcal{N}_k} X$ and $\bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}$ are compact as finite unions of compact sets, and E_ε is compact as a closed subset of the compact set Ω, while the set $E_\varepsilon^<$ is in general not compact. However, the Hausdorff distance only depends on the closure of the involved sets, and the closure $\operatorname{cl} E_\varepsilon^<$ is compact.

Thus, the triangle inequality for the Hausdorff metric yields

$$\begin{aligned}
\delta\left(\bigcup_{X\in\mathcal{N}_k} X, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}\right) &\leq \delta\left(\bigcup_{X\in\mathcal{N}_k} X, E_\varepsilon\right) + \delta(E_\varepsilon, \operatorname{cl} E_\varepsilon^<) + \delta\left(\operatorname{cl} E_\varepsilon^<, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}\right) \\
&= \delta\left(\bigcup_{X\in\mathcal{N}_k} X, E_\varepsilon\right) + \delta(E_\varepsilon, E_\varepsilon^<) + \delta\left(E_\varepsilon^<, \bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k} \widetilde{X}\right) \\
&\leq 2\tau.
\end{aligned}$$

Since $\tau > 0$ can be chosen arbitrarily close to zero, the assertion must hold.

□

So far, from Theorem 5.5 we know that the inner and outer approximations converge to each other in the Hausdorff metric under mild assumptions. However, since in the termination criterion of Algorithm 3 not the Hausdorff

distance between the two sets is used, but an algorithmically more accessible upper bound, it still remains to be shown that also this coarser termination criterion is fulfilled after a finite number of iterations, and the algorithm eventually terminates.

Corollary 5.6. *We consider a Nash equilibrium problem with $E_\varepsilon^< \neq \emptyset$ for some $\varepsilon > 0$ and we consider the branch-and-bound procedure of Algorithm 3 with $\tau > 0$, where we assume that a convergent and M-independent lower bounding procedure as well as some box selection rule fulfilling Assumption 5.2 are used. Furthermore, we assume $\delta(E_\varepsilon, E_\varepsilon^<) = 0$. Then, the algorithm terminates after a finite number of iterations.*

Proof. Assume that the algorithm does not terminate. Then in each iteration k the termination criterion is violated, that is, the lists $\mathcal{N}_k$ and $\widetilde{\mathcal{N}}_k$ satisfy $\mathcal{N}_k \neq \emptyset$ and $\max_{X \in \mathcal{N}_k} \min_{\widetilde{X} \in \widetilde{\mathcal{N}}_k} \|\Delta(X, \widetilde{X})\|_2 > \tau$. While the first of the two conditions must hold in view of

$$\emptyset \;\neq\; E_\varepsilon^< \;\subseteq\; \bigcup_{X \in \mathcal{N}_k} X,$$

the second condition is ruled out by the property

$$\lim_{k \to \infty} \max_{X \in \mathcal{N}_k} \min_{\widetilde{X} \in \widetilde{\mathcal{N}}_k} \|\Delta(X, \widetilde{X})\|_2 \;=\; 0.$$

To see the latter property, recall from the proof of Lemma 5.1 that for each $i = 1, \ldots, n$ we have $\Delta_i(A, B) = \max_{a_i \in A_i} \min_{b_i \in B_i} |a_i - b_i|$. This yields

$$\Delta_i(A, B) \;\leq\; \min_{a_i \in A_i} \min_{b_i \in B_i} |a_i - b_i| + (\overline{a}_i - \underline{a}_i),$$

so that the definition $\Delta_i'(A, B) := \min_{a_i \in A_i} \min_{b_i \in B_i} |a_i - b_i|$ leads to $\Delta(A, B) \leq \Delta'(A, B) + (\overline{a} - \underline{a})$ and, by the absolute monotonicity of the Euclidean norm,

$$\|\Delta(A, B)\|_2 \;\leq\; \|\Delta'(A, B) + (\overline{a} - \underline{a})\|_2 \;\leq\; \|\Delta'(A, B)\|_2 + \operatorname{diag}(A).$$

Since with analogous arguments as in the proof of Lemma 5.1 one obtains

$$\|\Delta'(A, B)\|_2 \;=\; \min_{a \in A} \min_{b \in B} \|a - b\|_2,$$

we arrive at

$$\begin{aligned}
\max_{X\in\mathcal{N}_k}\min_{\widetilde{X}\in\widetilde{\mathcal{N}}_k}\|\Delta(X,\widetilde{X})\|_2 &\leq \max_{X\in\mathcal{N}_k}\min_{\widetilde{X}\in\widetilde{\mathcal{N}}_k}\left(\min_{x\in X}\min_{\widetilde{x}\in\widetilde{X}}\|x-\widetilde{x}\|_2+\operatorname{diag}(X)\right)\\
&\leq \max_{X\in\mathcal{N}_k}\min_{\widetilde{X}\in\widetilde{\mathcal{N}}_k}\min_{x\in X}\min_{\widetilde{x}\in\widetilde{X}}\|x-\widetilde{x}\|_2+\max_{X\in\mathcal{N}_k}\operatorname{diag}(X)\\
&= \max_{X\in\mathcal{N}_k}\min_{x\in X}\min_{\widetilde{X}\in\widetilde{\mathcal{N}}_k}\min_{\widetilde{x}\in\widetilde{X}}\|x-\widetilde{x}\|_2+\max_{X\in\mathcal{N}_k}\operatorname{diag}(X)\\
&\leq \max_{X\in\mathcal{N}_k}\max_{x\in X}\min_{\widetilde{X}\in\widetilde{\mathcal{N}}_k}\min_{\widetilde{x}\in\widetilde{X}}\|x-\widetilde{x}\|_2+\max_{X\in\mathcal{N}_k}\operatorname{diag}(X)\\
&= \delta\left(\bigcup_{X\in\mathcal{N}_k}X,\bigcup_{\widetilde{X}\in\widetilde{\mathcal{N}}_k}\widetilde{X}\right)+\max_{X\in\mathcal{N}_k}\operatorname{diag}(X).
\end{aligned}$$

The first of the two summands converges to zero by Theorem 5.5. Moreover, in view of the box selection rule of the branch-and-bound algorithm, the diagonal length of all boxes in $\mathcal{N}$ converges to zero as well, which can be seen as follows. Let us assume for a moment that a box $X \in \mathcal{N}$ is not subdivided any further. Then, in view of Assumption 5.2 this box only consists of non-strict ε-Nash equilibria. This, however, violates $\delta(E_\varepsilon, E_\varepsilon^<) = 0$ and, hence, concludes the proof. □

In spatial branch-and-bound algorithms in standard global optimization typically not only convergence to ε-optimal points for solvable problems is ensured, but also termination of algorithms with an empty list of boxes in case there are no optimal points at all. We now derive an analogous result for our branch-and-bound algorithm.

The next result follows with arguments similar to the ones in the proof of Proposition 5.4.

Corollary 5.7. *We consider the branch-and-bound procedure of Algorithm 3 with $\tau > 0$. Moreover, we assume that a convergent and M-independent lower bounding procedure is used and, furthermore, we assume $E_\varepsilon = \emptyset$. Then, the algorithm terminates after a finite number of iterations with $\mathcal{N} = \emptyset$.*

Proof. Let us assume that the assertion does not hold and that the algorithm does not terminate. Then, there is an exhaustive sequence of boxes $(X^s)_{s\in\mathbb{N}}$ with $X^s \in \mathcal{N}_s$ created by the algorithm. By assumption none of the boxes X^s contains an ε-Nash equilibrium. Then, analogously to the proof of Proposition 5.4 we can show that sooner or later such a box must be excluded by the algorithm and, thus, such a sequence of boxes cannot be created at all. Hence, the list $\mathcal{N}$ becomes empty after a finite number of iterations. □

By Corollary 5.7, Algorithm 3 is able to detect the case $E_\varepsilon = \emptyset$ and, by Corollary 5.6, to otherwise approximate E_ε with precision τ. We point out that the assumption $\delta(E_\varepsilon, E_\varepsilon^<) = 0$ prevents us from studying how the algorithm behaves in the case $\delta(E_\varepsilon, E_\varepsilon^<) > 0$. This situation, however, is ruled out in presence of an appropriate constraint qualification in the set E_ε.

6 Illustrative examples

In this section, we present the first concrete implementation of the branch-and-bound method as a proof-of-concept. We illustrate the output of our method and discuss its properties with three numerical examples from the literature. At first, in Section 6.1, we look at a two-player potential NEP with a unique equilibrium and use an implementation of the branch-and-bound method without the termination criterion introduced in Section 5. In Section 6.2 we consider an instance with multiple Nash equilibria and look at the advantages of the suitable termination criterion from Section 5. In Section 6.3 we illustrate the results for a three-dimensional instance. Previously, however, we will mention implementational details that are identical in all of the following examples.

The algorithm was implemented in Python 3.10.8 and the experiments ran on an Intel i7 processor with 3.60 GHz and 32 GB of RAM. All figures were produced with Matplotlib 3.6.2.

There are some degrees of freedom in the implementation of Algorithm 3 which we will now address. In *Step 5a*, the lists $\mathcal{L}_\nu(X^1)$ are computed from scratch in every iteration. Furthermore, the lower and upper bounds within this step are computed by simple centered forms as described in [16]. For practical reasons, we remove a box from $\mathcal{N}$, if we add it to $\widetilde{\mathcal{N}}$ after *Step 5a* because it does not need to be further subdivided given Assumption 5.2. Finally, we refrain from fathoming in *Step 7* and fulfill Assumption 3.2 by keeping all boxes in $\mathcal{L}$. Explanations of *Step 1* and the termination criterion will be given in the following subsections.

6.1 A two-player potential NEP

As a first experiment, we execute Algorithm 3 without the stopping criterion based on the distance between $\mathcal{N}$ and $\widetilde{\mathcal{N}}$ and use the while-statement from Algorithm 1 instead. In *Step 1*, we always take the first element from the list $\mathcal{N}$ and in *Step 4* we append X^1 and X^2 to the end of the list. Consequently,

the selected box in *Step 1* is always one of the boxes with largest diagonal length. The method terminates when this length falls below a specified value τ. By finiteness of the list $\mathcal{N}$, each box that is added to the end of the list will be split again after finitely many iterations, given that it is not discarded or added to $\widetilde{\mathcal{N}}$. Hence, Assumption 5.2 is fulfilled.

We test this implementation on a potential Nash equilibrium problem taken from [17]. In this problem class, the players unknowingly minimize the same objective function in their respective variables. We have two players, each controlling one variable. Player $i = 1, 2$ aims to minimize the function

$$\theta_i(x^1, x^2) = \frac{(x^1 + x^2)^2}{4} + \frac{(x^1 - x^2)^2}{9}$$

over the interval $\Omega_i = [-10, 10]$. It is easy to verify that we have a player-convex setting, meaning that θ_i is convex with respect to x^i and Ω_i is convex for $i = 1, 2$. Thus we can illustrate the sets $S_i(x^{-i})$ using first order optimality conditions. Figure 1 illustrates the best response of player one ($\operatorname{gph} S_1$) in orange and $\operatorname{gph} S_2$ in blue as well as some level lines of the common objective function. The unique Nash equilibrium is the intersection point of both lines at $(0, 0)$.

With our procedure, we aim to inner approximate the set of all strict ε-Nash equilibria $E_\varepsilon^<$ and outer approximate the set of all ε-Nash equilibria E_ε. To illustrate these approximations, we plotted the slim orange and blue lines, which enclose all ε-best responses for both players with $\varepsilon = 0.05$. Indeed, E_ε is the set where both players give ε-best responses (cf. (1)). The lower two images in Figure 1 show the output, when we execute the described method with different termination criteria, namely $\tau = 0.1$ on the left and $\tau = 0.05$ on the right. The boxes contained in $\widetilde{\mathcal{N}}$ are pictured in green and the boxes in $\mathcal{N} \setminus \widetilde{\mathcal{N}}$ are black. We can observe that the desired approximation properties hold in this example and that the approximations become more accurate with smaller τ.

6.2 A two-player NEP with multiple equilibria

In this example, we consider a slight modification of Example 1.1 taken from [4]. The objective functions are defined as

$$\theta_1(x^1, x^2) = \frac{(x^1)^2}{2} - q(x^2) \cdot x^1,$$
$$\theta_2(x^1, x^2) = \frac{(x^2)^2}{2} - q(x^1) \cdot x^2,$$

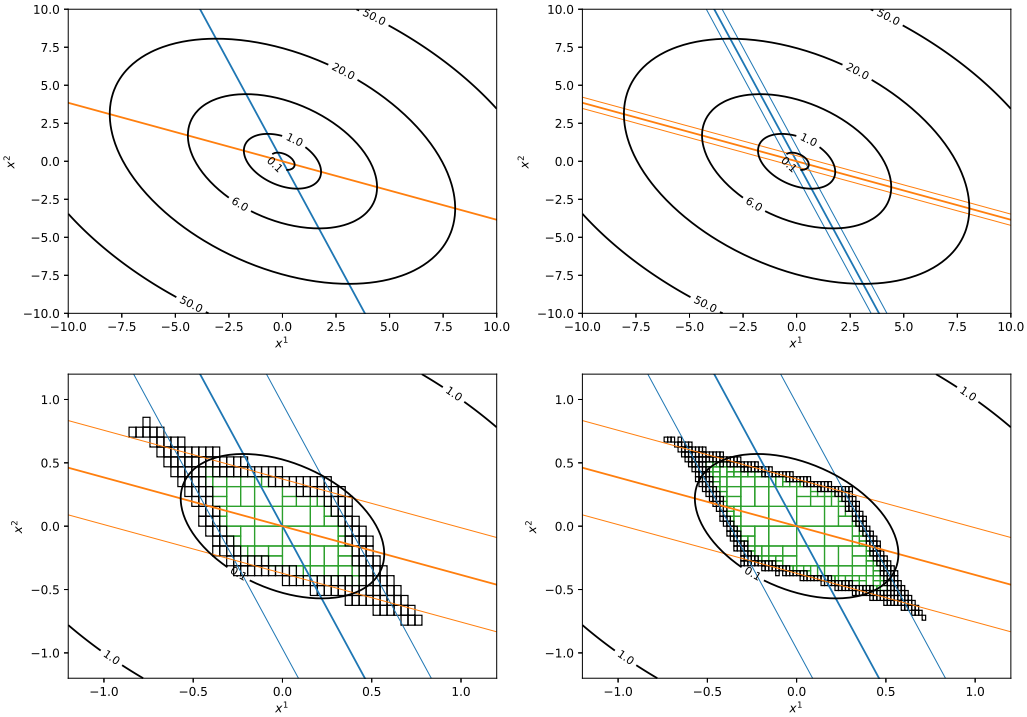

Figure 1: Approximation of $E^{<}_{0.05}$ and $E_{0.05}$ by uniformly refining $\mathcal{N} \setminus \widetilde{\mathcal{N}}$.

τ	k	$\lvert\mathcal{N} \setminus \widetilde{\mathcal{N}}\rvert$	$\lvert\widetilde{\mathcal{N}}\rvert$	$\lvert\mathcal{L}\rvert$
0.05	7145	3355	0	22837
0.02	22872	11155	0	73201
0.01	52782	28847	709	174794
0.005	129132	71376	11810	444998

Table 1: Approximation of $E^{<}_{0.05}$ and $E_{0.05}$ by uniformly refining $\mathcal{N} \setminus \widetilde{\mathcal{N}}$.

with $q(x) = (x-4)^2 + 2$ and the strategy sets are $\Omega_1 = \Omega_2 = [0, 10]$. Figure 2 again illustrates the best response of player one ($\operatorname{gph} S_1$) in orange and $\operatorname{gph} S_2$ in blue. We can see that there are five Nash equilibria, i.e. points in $E = \operatorname{gph} S_1 \cap \operatorname{gph} S_2$. In the following, we approximate the strict and non-strict ε-Nash equilibria for $\varepsilon = 0.05$ in two different ways.

Approximation of $E^{<}_{0.05}$ and $E_{0.05}$ by uniformly refining $\mathcal{N} \setminus \widetilde{\mathcal{N}}$.

At first, we solve this instance as outlined in the previous subsection and terminate when the largest diagonal length of the boxes in $\mathcal{N}$ falls below a given value. Table 1 shows the number of iterations k and the length of all lists for different length parameters τ. Notably, the list $\widetilde{\mathcal{N}}$ is empty for $\tau \in \{0.05, 0.02\}$, so we are only receiving an outer approximation of $E_{0.05}$ as output for these parameters. For smaller τ-values the uniform box refinement leads to a special behavior of the approximations $E^{<}_{0.05}$ and $E_{0.05}$. The results for $\tau = 0.005$ are plotted in Figure 2. Again, the boxes contained in $\widetilde{\mathcal{N}}$ are pictured in green and the boxes in $\mathcal{N} \setminus \widetilde{\mathcal{N}}$ are black. We can see that for some points in E we have plenty of green boxes which inner approximate $E^{<}_{0.05}$ in this area (upper right image). In contrast, there is not a single box in $\widetilde{\mathcal{N}}$ for two other Nash equilibria as we can see in the lower two images of Figure 2.

Approximation of $E^{<}_{0.05}$ and $E_{0.05}$ by Algorithm 3.

As a next step, we approximate Nash equilibria with Algorithm 3 and use the therein specified termination criterion from Section 5 based on the Hausdorff distance between the inner approximation of $E^{<}_{\varepsilon}$ and the outer approximation of E_{ε}.

In *Step 1*, as long as $\widetilde{\mathcal{N}}$ is empty, we take the first box X from the list $\mathcal{N}$ and append X^1 and X^2 in *Step 4* to the end of the list. In this case, there is no termination because the distance to an empty set is defined as the extended real value $+\infty > \tau$. As soon as $\widetilde{\mathcal{N}}$ is nonempty, we take the first box X from

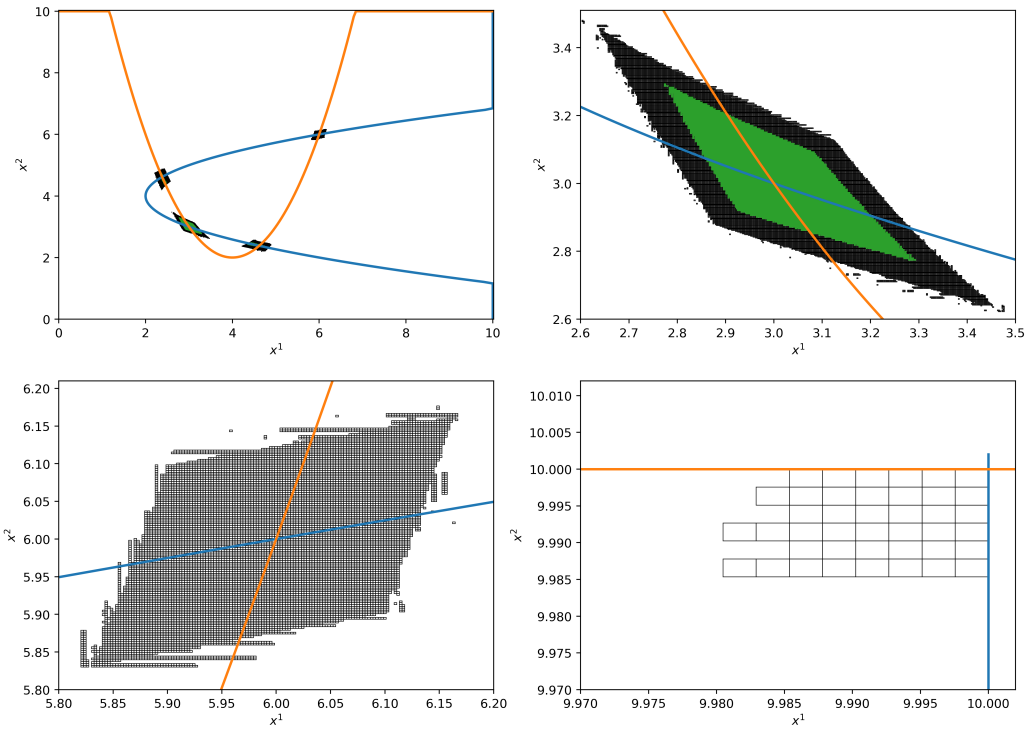

Figure 2: Approximation of $E_{0.05}^{<}$ and $E_{0.05}$ by uniformly refining $\mathcal{N} \setminus \widetilde{\mathcal{N}}$.

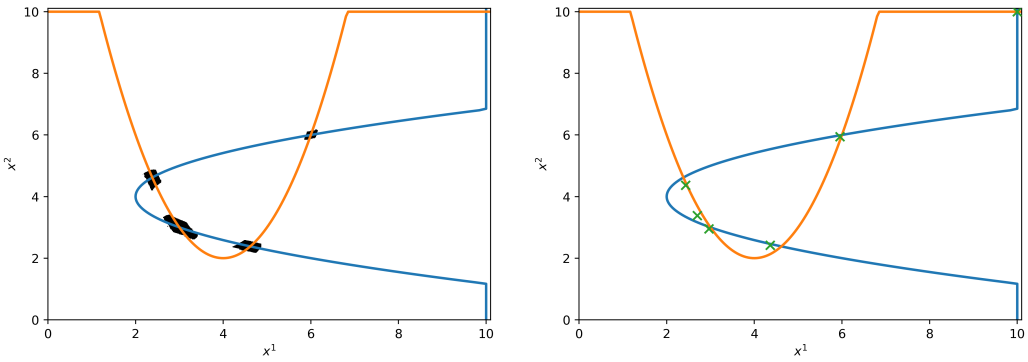

Figure 3: Approximation of $E^{<}_{0.05}$ and $E_{0.05}$ by Algorithm 3 with $\tau = 0.5$

$\mathcal{N}$ for which

$$\min_{\widetilde{X}\in\widetilde{\mathcal{N}}} \|\Delta(X, \widetilde{X})\|_2 > \tau \tag{6}$$

holds in *Step 1*. After dividing X, we still append X^1 and X^2 in *Step 4* to the end of the list. If there is no box for which (6) holds, Algorithm 3 terminates. It remains to be shown that the termination occurs after finitely many steps. We can see in the while statement, that boxes X' for which $\min_{\widetilde{X}\in\widetilde{\mathcal{N}}} \|\Delta(X', \widetilde{X})\|_2 \leq \tau$ holds are not preventing termination. This naturally also holds for all boxes in $\widetilde{\mathcal{N}}$. Due to Assumption 5.2 and Corollary 5.6, the termination is ensured when we subdivide each box from $\mathcal{N}$ which satisfies (6) after finitely many iterations. This is the case as we employ a first in first out scheme for these boxes and the list $\mathcal{N}$ is finite. We remark that various other strategies would also be possible, for example always selecting a box which maximizes the left-hand side of (6).

For $\varepsilon = 0.05$ and $\tau = 0.5$, the method terminates after 121520 iterations with six boxes in $\widetilde{\mathcal{N}}$ and 64194 boxes in $\mathcal{N} \setminus \widetilde{\mathcal{N}}$. In Figure 3, the boxes from $\mathcal{N}$ are plotted in black (left), the green crosses are markers for the six boxes from $\widetilde{\mathcal{N}}$ which would not be visible in original size (right). In contrast to all results obtained by uniformly refining $\mathcal{N} \setminus \widetilde{\mathcal{N}}$, we have at least one box in $\widetilde{\mathcal{N}}$ near each Nash equilibrium.

6.3 A three-dimensional example

In this example, we consider a two-player oligopoly model. The first player controls two variables and the second player can only set one variable. Each variable represents the production quantities of a specific good. The objective

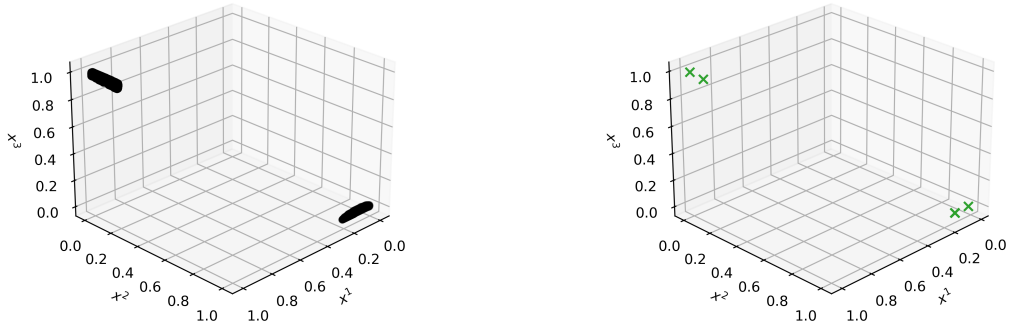

Figure 4: Approximation of $E^{<}_{0.01}$ and $E_{0.01}$ by Algorithm 3 with $\tau = 0.1$.

functions are defined as

$$\theta_1(x^1, x^2) = \frac{(x_1^1)^2}{2} + \frac{(x_2^1)^2}{2} + x_1^1 x_2^1 - x_1^1 x_1^2 - x_1^1 - x_2^1$$
$$\theta_2(x^1, x^2) = \frac{(x_1^2)^2}{2} + x_2^1 x_1^2 - x_1^2$$

and the strategy sets are $\Omega_1 = [0,1]^2$ and $\Omega_2 = [0,1]$. The economic interpretation of this instance is that the first and third good are complements to each other while the second good is a substitute of the other goods and vice versa. For more details, see Example 1 and 2 in [24]. The parameters are slightly changed, so that this instance is not player convex. We apply Algorithm 3 to this instance, in the same form as described in the previous subsection. For $\varepsilon = 0.01$ and $\tau = 0.1$, we receive 4377 boxes in $\mathcal{N} \setminus \widetilde{\mathcal{N}}$ and five boxes in $\widetilde{\mathcal{N}}$ as output. The results are illustrated in Figure 4, where we plotted black dots as markers for boxes in $\mathcal{N} \setminus \widetilde{\mathcal{N}}$ (left) and green crosses as markers for boxes in $\widetilde{\mathcal{N}}$ (right). We can see that both approximations cluster around the points $(1,0,1)^\top$ and $(0,1,0)^\top$. This indicates that, in Nash equilibria, either both players produce the goods which complement one another or the first player produces only the substitute for both goods.

7 Conclusions

In this paper we proposed the first branch-and-bound algorithm for the computation of all Nash equilibria of continuous non-convex Nash equilibrium problems. A proof of convergence is given and numerical results on some

illustrative examples are provided. However, there are a few issues we would like to mention.

First we point out that our implementation is just meant as a proof-of-concept. It is written in Python with a very simple lower bounding procedure. No additional acceleration steps such as bound tightening are considered so far. Furthermore, we think that the presented method would be ideally suited to be run in parallel. For this reason, we think that the performance of the proposed algorithm can be significantly improved by developing a more sophisticated implementation. However, this is left for future research.

Moreover, in our current version of the branch-and-bound algorithm for non-convex Nash equilibrium games we have not taken into consideration any other constraints than box constraints. However, we expect that this is possible and that even non-convex generalized Nash equilibrium problems may be tackled by an appropriately adapted method. Again, these questions are left for future research.

References

[1] C.S. Adjiman, S. Dallwig, C.A. Floudas, A. Neumaier, *A global optimization method, αBB, for general twice-differentiable constrained NLPs – I. Theoretical advances*, Computers & Chemical Engineering, 22 (1998), 1137-1158

[2] C.S. Adjiman, I.P. Androulakis, C.A. Floudas, *A global optimization method, αBB, for general twice-differentiable constrained NLPs – II. Implementation and computational results*, Computers & Chemical Engineering, 22 (1998), 1159-1179

[3] I.P. Androulakis, C.D. Maranas, C.A. Floudas, *αBB: A global optimization method for general constrained nonconvex problems*, Journal of Global Optimization, 7 (1995), 337-363

[4] M. Beck, O. Stein, *Semi-infinite models for equilibrium selection,* Minimax Theory and its Applications, to appear.

[5] E. Baumann, *Optimal centered forms*, BIT Numerical Mathematics, 28 (1988), 80-87

[6] A. Dreves, *Finding all solutions of affine generalized Nash equilibrium problems with one-dimensional strategy sets*, Math Meth Oper Res, 80 (2014), 139–159

[7] A. Dreves, *Computing all solutions of linear generalized Nash equilibrium problems*, Math Meth Oper Res, 85 (2017), 207–221

[8] M. Dür, *Dual bounding procedures lead to convergent branch-and-bound algorithms*, Mathematical Programming, 91 (2001), 117-125

[9] M. Dür, *A class of problems where dual bounds beat underestimation bounds*, Journal of Global Optimization, 22 (2002), 49-57

[10] F. Facchinei, A. Fischer, V. Piccialli, *Generalized Nash equilibrium problems and Newton methods*, Mathematical Programming, 117 (2009), 163–194

[11] F. Facchinei, C. Kanzow, *Generalized Nash equilibrium problems*, 4OR, 5 (2007), 173–210

[12] F. Facchinei, S. Sagratella, *On the computation of all solutions of jointly convex generalized Nash equilibrium problems*, Optim Lett, 5 (2011), 531–547

[13] C.A. Floudas. *Deterministic Global Optimization. Theory, Methods, and Applications.* Kluwer, Dordrecht (2000)

[14] P. Kirst, O. Stein, P. Steuermann, *Deterministic upper bounds for spatial branch-and-bound methods in global minimization with nonconvex constraints*, TOP, 23 (2015), 591-616

[15] R. Horst, H. Tuy, *Global Optimization: Deterministic Approaches*, Springer, Berlin (1996)

[16] R. Krawczyk, K. Nickel, *Die zentrische Form in der Intervallarithmetik, ihre quadratische Konvergenz und ihre Inklusionsisotonie*, Computing, 28 (1982), 117-137

[17] J.B. Krawczyk, S. Uryasev, *Relaxation algorithms to find Nash equilibria with economic applications*, Environmental Modeling & Assessment, 5 (2000), 63-73

[18] J.F. Nash, *Equilibrium points in n-person games*, Proceedings of the National Academy of Sciences, 36 (1950), 48–49

[19] J.F. Nash, *Non-Cooperative Games*, The Annals of Mathematics, 54 (1951), 286–295

[20] G.P. McCormick, *Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems*, Mathematical Programming, 10 (1976), 145-175

[21] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge (1990)

[22] H. Nikaido, K. Isoda, *Note on noncooperative convex games*, Pacific Journal of Mathematics, 5 (1955), 807–815

[23] S. Sagratella, *Computing all solutions of Nash equilibrium problems with discrete strategy sets*, SIAM Journal on Optimization, 26 (2016), 2190–2218

[24] S. Sagratella, *Computing equilibria of Cournot oligopoly models with mixed-integer quantities* Mathematical Methods of Operations Research, 86 (2017), 549–565

[25] S. Sagratella, *On generalized Nash equilibrium problems with linear coupling constraints and mixed-integer variables*, Optimization, 68 (2019), 197–226

[26] M. Tawarmalani, N.V. Sahinidis, *Convexification and Global Optimization in Continous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software and Applications*, Springer Science and Business Media, (2002)

[27] S. Schwarze, O. Stein, *A branch-and-prune algorithm for discrete Nash equilibrium problems*, Optimization Online, Preprint ID 2022-03-8836, 2022

[28] E.M. Smith, C.C. Pantelides, *Global optimisation of nonconvex MINLPs*, Computers & Chemical Engineering, 21 (1997), S791-S796

[29] E.M. Smith, C.C. Pantelides, *A symbolic reformulation/spatial branch-and-bound algorithm for the global optimization of nonconvex MINLPs*, Computers & Chemical Engineering, 23 (1999), 457-478

[30] S.H. Tijs, *Nash equilibria for noncooperative n-person games in normal form*, SIAM Review, 23 (1981), 225-237