Convergence-order analysis of branch-and-bound algorithms for constrained problems

Citation: Kannan, Rohit, and Paul I. Barton. "Convergence-Order Analysis of Branch-and-Bound Algorithms for Constrained Problems." Journal of Global Optimization 71, no. 4 (June 1, 2017): 753–813.

As Published: https://doi.org/10.1007/s10898-017-0532-y

Publisher: Springer-Verlag

Persistent URL: http://hdl.handle.net/1721.1/117047

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Journal of Global Optimization manuscript No.
(will be inserted by the editor)

Convergence-Order Analysis of Branch-and-Bound Algorithms for Constrained Problems

Rohit Kannan · **Paul I. Barton**

Received: date / Accepted: date

Abstract The performance of branch-and-bound algorithms for deterministic global optimization is strongly dependent on the ability to construct tight and rapidly convergent schemes of lower bounds. One metric of the efficiency of a branch-and-bound algorithm is the convergence order of its bounding scheme. This article develops a notion of convergence order for lower bounding schemes for constrained problems, and defines the convergence order of convex relaxation-based and Lagrangian dual-based lower bounding schemes. It is shown that full-space convex relaxation-based lower bounding schemes can achieve first-order convergence under mild assumptions. Furthermore, such schemes can achieve second-order convergence at KKT points, at Slater points, and at infeasible points when second-order pointwise convergent schemes of relaxations are used. Lagrangian dual-based full-space lower bounding schemes are shown to have at least as high a convergence order as convex relaxation-based full-space lower bounding schemes. Additionally, it is shown that Lagrangian dual-based full-space lower bounding schemes achieve first-order convergence even when the dual problem is not solved to optimality. The convergence order of some widely-applicable reduced-space lower bounding schemes is also analyzed, and it is shown that such schemes can achieve first-order convergence under suitable assumptions. Furthermore, such schemes can achieve second-order convergence at KKT points, at unconstrained points in the reduced-space, and at infeasible points under suitable assumptions when the problem exhibits a specific separable structure. The importance of constraint propagation techniques in boosting the convergence order of reduced-space lower bounding schemes (and helping mitigate clustering in the process) for problems which do not possess such a structure is demonstrated.

Keywords Global optimization · Constrained optimization · Convergence order · Convex relaxation · Lagrangian dual · Branch-and-bound · Lower bounding scheme · Reduced-space

Mathematics Subject Classification (2010) 49M20 · 49M29 · 49M37 · 49N15 · 65K05 · 68Q25 · 90C26

The authors gratefully acknowledge financial support from BP. This work was conducted as a part of the BP-MIT conversion research program.

Rohit Kannan · Paul I. Barton (✉)
Process Systems Engineering Laboratory, Department of Chemical Engineering,
Massachusetts Institute of Technology, Cambridge, MA, USA

Paul I. Barton
E-mail: pib@mit.edu

Rohit Kannan
E-mail: rohitk@mit.edu

1 Introduction

Global optimization has found widespread applications in various areas of engineering and the sciences [11]. Deterministic global optimization algorithms attempt to determine an approximate optimal solution within a specified tolerance and terminate with a certificate of its optimality in finite time [13]. While efficient algorithms are known for classes of convex optimization problems [4], no such algorithms are currently known for most classes of nonconvex problems. Deterministic global optimization algorithms for nonconvex problems usually involve the concept of partitioning the domain of ('branching on') the decision variables [13]. The performance of branch-and-bound algorithms for deterministic global optimization is strongly dependent on the ability to construct tight and rapidly convergent relaxations of nonconvex functions.

Since the worst-case running time of all known branch-and-bound algorithms is exponential in the dimension of the variables partitioned, it may be advantageous to utilize 'reduced-space' algorithms which only require branching on a subset of the variables (as opposed to 'full-space' branch-and-bound algorithms which may branch on all of the variables) to guarantee convergence. Despite the potential advantages of reduced-space algorithms for nonconvex problems [3, 9, 10, 38], such methods have not been widely adopted in the literature and in commercial software. One potential reason is that most widely-applicable reduced-space branch-and-bound algorithms often do not seem to exhibit favorable convergence rates compared to their full-space counterparts [22]. The convergence properties of reduced-space branch-and-bound algorithms have not been thoroughly investigated, although some progress has been made in this direction [8, 38]. The reader is directed to the work of Epperly and Pistikopoulos [10] for a survey of reduced-space branch-and-bound algorithms.

One metric of the efficiency of a deterministic branch-and-bound algorithm is the order of convergence of its bounding scheme, which, for the case of unconstrained optimization, compares the rate of convergence of an estimated range of a function to its true range [23]. Recently, Bompadre and coworkers [5, 6] developed the notions of Hausdorff and pointwise convergence orders of bounding schemes and established sharp rules for the propagation of convergence orders of bounding schemes constructed using McCormick [20], Taylor [25], and McCormick-Taylor [28] models. In addition, they showed that if a function is twice continuously differentiable, the scheme of relaxations corresponding to its envelopes is at least second-order pointwise convergent which, in turn, implies Hausdorff convergence of at least second-order. Najman and Mitsos [24] used the framework developed in [5, 6] to establish sharp rules for the propagation of convergence orders of multivariate McCormick relaxations [37]. Khan and coworkers [17] developed a continuously differentiable variant of McCormick relaxations [20, 37], and established second-order pointwise convergence of schemes of the differentiable McCormick relaxations for twice continuously differentiable functions. Also note the definition of rate of convergence of bounding schemes for geometric branch-and-bound methods proposed by Schöbel and Scholz [30], and the proof of second-order Hausdorff convergence of centered forms in [18, 31]. Establishing that a scheme of relaxations is at least second-order Hausdorff convergent is important from many viewpoints, notably in mitigating the so-called cluster effect in unconstrained global optimization [7, 39]. Recently, the authors of this work have analyzed the cluster problem for constrained global optimization and determined that, under certain conditions, first-order convergence of the lower bounding scheme may be sufficient to avoid the cluster problem at constrained minima [15]. However, an analysis of convergence order for constrained problems is currently lacking.

In this work, we investigate the convergence orders of some full-space and reduced-space deterministic branch-and-bound algorithms by extending the convergence analysis of Bompadre and coworkers to constrained problems. Specifically, we propose a definition of convergence order for lower bounding schemes, analyze the convergence orders of commonly-used full-space lower bounding schemes, and analyze the convergence orders of some widely-applicable reduced-space lower bounding schemes in the literature. Throughout this work, we tacitly assume that a branch-and-bound algorithm utilizes efficient heuristics for finding feasible points which determine a global optimal solution early on in the branch-and-bound tree.

This paper is organized as follows. Section 2 formulates the problem of interest, and provides some background definitions. Section 3 develops the notion of convergence order of a lower bounding scheme, and Section 4 provides some results on the convergence orders of commonly-used full-space lower bounding schemes. Section 5 lists some widely-applicable reduced-space lower bounding schemes in the literature, provides some

results on their convergence orders, and highlights the importance of constraint propagation in reduced-space branch-and-bound algorithms. Finally, Section 6 lists the conclusions and some avenues for future work.

2 Problem Formulation and Background

Consider the problem

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \ & f(\mathbf{x},\mathbf{y}) \\ \text{s.t. } & \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}, \\ & \mathbf{x} \in X, \ \mathbf{y} \in Y, \end{aligned} \tag{P}$$

where $X \subset \mathbb{R}^{n_x}$ and $Y \subset \mathbb{R}^{n_y}$ are nonempty convex sets, $f : X \times Y \to \mathbb{R}$ and $\mathbf{g} : X \times Y \to \mathbb{R}^{m_I}$ are partially convex with respect to $\mathbf{x}$, i.e., $f(\cdot,\mathbf{y})$ and $\mathbf{g}(\cdot,\mathbf{y})$ are convex on X for each $\mathbf{y} \in Y$, $\mathbf{h} : X \times Y \to \mathbb{R}^{m_E}$ is affine with respect to $\mathbf{x}$, i.e., $\mathbf{h}(\cdot,\mathbf{y})$ is affine on X for each $\mathbf{y} \in Y$, and $\mathbf{0}$ denotes a vector of zeros of appropriate dimension. The following assumption will be made throughout this work.

Assumption 1 The sets X and Y are compact, and the functions f, $\mathbf{g}$, and $\mathbf{h}$ are continuous on $X \times Y$.

When the dimension n_y of the Y-space corresponding to the nonconvexities in the functions in Problem (P) is significantly smaller than the dimension n_x of the X-space, it may be computationally advantageous to partition only the Y-space during the course of a branch-and-bound algorithm (assuming, of course, that the reduced-space algorithm is guaranteed to converge). However, the convergence rate of a reduced-space branch-and-bound algorithm may be different compared to a similar full-space algorithm, which makes it difficult to judge a priori whether using a reduced-space branch-and-bound approach would be advantageous. Before we analyze the convergence orders of some full-space and reduced-space lower bounding schemes in the literature, we need to define formally the notion of convergence order for constrained problems. For this purpose, we review some relevant definitions [5, 6].

Throughout this work, we use $\mathbb{I}Z$ to denote the set of nonempty, closed, and bounded interval subsets of $Z \subset \mathbb{R}^n$, $\mathbb{R}_+$ and $\mathbb{R}_-$ to respectively denote the sets of nonnegative and nonpositive reals, z_j to denote the j^{th} component of a vector $\mathbf{z}$, $(z_1, z_2, \cdots, z_n)$ to denote a vector $\mathbf{z} \in \mathbb{R}^n$ with components $z_1, z_2, \cdots, z_n \in \mathbb{R}$, $(\mathbf{v},\mathbf{w})$ to denote the column vector $[\mathbf{v}^{\mathrm{T}}\ \mathbf{w}^{\mathrm{T}}]^{\mathrm{T}}$ corresponding to (column) vectors $\mathbf{v}$ and $\mathbf{w}$, $\|\mathbf{z}\|$ to denote the Euclidean norm of a vector $\mathbf{z} \in \mathbb{R}^n$, $\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}$ to denote a vector-valued function with domain Z and codomain $\mathbb{R}^{m+n}$ corresponding to vector-valued functions $\mathbf{g} : Z \to \mathbb{R}^m$ and $\mathbf{h} : Z \to \mathbb{R}^n$, $\mathrm{conv}(S)$ to denote the convex hull of a set $S \subset \mathbb{R}^n$, and $\mathrm{int}(S)$ to denote the interior of a set $S \subset \mathbb{R}^n$.

Definition 1 (Width of an Interval) Let $Z = [z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] \times \cdots \times [z_n^{\mathrm{L}}, z_n^{\mathrm{U}}]$ be an element of $\mathbb{I}\mathbb{R}^n$. The width of Z, denoted by $w(Z)$, is given by

$$w(Z) := \max_{i=1,\cdots,n} (z_i^{\mathrm{U}} - z_i^{\mathrm{L}}).$$

Definition 2 (Distance Between Two Sets) Let $Z, V \subset \mathbb{R}^n$. The distance between Z and V, denoted by $d(Z,V)$, is given by

$$d(Z,V) := \inf_{\substack{\mathbf{z}\in Z,\\ \mathbf{v}\in V}} \|\mathbf{z}-\mathbf{v}\|.$$

Note that the above definition of distance does not define a metric; however, it will prove useful in defining a measure of infeasibility for points in $X \times Y$ for Problem (P). The following result holds.

Lemma 1 *Let $\mathbf{z}, \mathbf{v} \in \mathbb{R}^n$, and let $K \subset \mathbb{R}^n$ be a convex cone. Then*

$$d(\{\mathbf{z}\}, K) - d(\{\mathbf{v}\}, K) \leq d(\{\mathbf{z}-\mathbf{v}\}, K).$$

Proof See [32]. □

Corollary 1 *Let $\mathbf{z}, \mathbf{v} \in \mathbb{R}^{m+n}$. Then*

$$d(\{\mathbf{z}\}, \mathbb{R}^m_- \times \{\mathbf{0}\}) - d(\{\mathbf{v}\}, \mathbb{R}^m_- \times \{\mathbf{0}\}) \leq d(\{\mathbf{z}-\mathbf{v}\}, \mathbb{R}^m_- \times \{\mathbf{0}\}).$$

Proof This result is a direct consequence of Lemma 1. □

Lemma 2 *All norms on $\mathbb{R}^n$ are equivalent. Specifically, if $\|\cdot\|_p$ and $\|\cdot\|_q$ are two norms in $\mathbb{R}^n$ for any $p, q \in \mathbb{N} \cup \{+\infty\}$ with $p \neq q$, then there exist constants $c_1, c_2 \in \mathbb{R}_+$ such that $c_1\|\mathbf{z}\|_p \leq \|\mathbf{z}\|_q \leq c_2\|\mathbf{z}\|_p$, $\forall \mathbf{z} \in \mathbb{R}^n$. Furthermore, for $(p,q) = (1,2)$, $c_2 = 1$ provides a valid upper bound and for $(p,q) = (+\infty, 2)$, $c_2 = \sqrt{n}$ provides a valid upper bound.*

Proof For the first part of the lemma, see, for instance, [29, Theorem 4.2]. The second part of the lemma follows from the inequalities

$$\|\mathbf{z}\|_2^2 = \sum_{i=1}^n z_i^2 \leq \sum_{i=1}^n z_i^2 + \sum_{i=1}^n \sum_{j=i+1}^n 2|z_i||z_j| = \|\mathbf{z}\|_1^2$$

and

$$\|\mathbf{z}\|_2^2 = \sum_{i=1}^n z_i^2 \leq n \max_{i=1,\cdots,n} z_i^2 = n\|\mathbf{z}\|_\infty^2$$

for any $\mathbf{z} \in \mathbb{R}^n$. □

Definition 3 (Lipschitz Continuous Function) Let $Z \subset \mathbb{R}^n$. A function $f : Z \to \mathbb{R}$ is said to be Lipschitz continuous with Lipschitz constant $M \geq 0$ if $|f(\mathbf{z}_1) - f(\mathbf{z}_2)| \leq M\|\mathbf{z}_1 - \mathbf{z}_2\|$, $\forall \mathbf{z}_1, \mathbf{z}_2 \in Z$.

Remark 1 Locally Lipschitz continuous functions are Lipschitz continuous on compact subsets of their domains. Therefore, the assumption that the functions f, $\mathbf{g}$, and $\mathbf{h}$ in Problem (P) are Lipschitz continuous on $X \times Y$ is not particularly strong when Assumption 1 is made.

Definition 4 (Hausdorff Metric) Let $X = [x^L, x^U]$ and $Y = [y^L, y^U]$ be two intervals in $\mathbb{IR}$. The Hausdorff metric between X and Y, denoted by $d_H(X,Y)$, is given by

$$d_H(X,Y) = \max\{|x^L - y^L|, |x^U - y^U|\} = \max\left\{\max_{x\in X}\min_{y\in Y}|x-y|, \max_{y\in Y}\min_{x\in X}|x-y|\right\}.$$

Definition 5 (Inclusion Function) Let $V \subset \mathbb{R}^n$ and suppose $\mathbf{f} : V \to \mathbb{R}^m$ is continuous. For any $Z \subset V$, let $\overline{\mathbf{f}}(Z)$ denote the image of Z under $\mathbf{f}$. A mapping $F : \mathbb{I}V \to \mathbb{IR}^m$ is called an inclusion function for $\mathbf{f}$ on $\mathbb{I}V$ if, for every $Z \in \mathbb{I}V$, we have $\overline{\mathbf{f}}(Z) \subset F(Z)$.

Definition 6 (Range Order) Let $V \subset \mathbb{R}^n$ be a bounded set. Let $f : V \to \mathbb{R}$ be continuous, and let F be an inclusion function for f on $\mathbb{I}V$. The inclusion function F is said to have range of order $\alpha > 0$ at a point $\mathbf{v} \in V$ if there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}V$ with $\mathbf{v} \in Z$,

$$w(F(Z)) \leq \tau w(Z)^\alpha.$$

The function f itself is said to have a range of order $\alpha > 0$ at $\mathbf{v} \in V$ if its image $\overline{f}$ has range of order α at $\mathbf{v}$. The functions F and f are said to have ranges of order $\alpha > 0$ on V if they have ranges of order (at least) α at each $\mathbf{v} \in V$, with the constant τ independent of $\mathbf{v}$.

The reader is directed to Remark 3 for a discussion on the assumption that the set V in the above definition be bounded. Since the convergence order analysis in this work is asymptotic in nature (see Remark 4 and Lemma 5), we will need the following asymptotic notations.

Definition 7 (Big O and Little o Notations) Let $Z \subset \mathbb{R}$, $f : Z \to \mathbb{R}$, and $g : Z \to \mathbb{R}$. We say that $f(z) = O(g(z))$ as $z \to \bar{z} \in Z$ if and only if there exist $\delta, M > 0$ such that

$$|f(z)| \leq M|g(z)|, \quad \forall z \in Z \text{ with } |z - \bar{z}| < \delta.$$

Similarly, we say that $f(z) = o(g(z))$ as $z \to \bar{z} \in Z$ if and only if for all $M' > 0$ there exists $\delta' > 0$ such that

$$|f(z)| \leq M'|g(z)|, \quad \forall z \in Z \text{ with } |z - \bar{z}| < \delta'.$$

Note that unless otherwise specified, we consider $\bar{z} = 0$ in this work.

The following lemma is from Proposition 11.7 in [14].

Lemma 3 *Let $Z \subset \mathbb{R}^n$ be nonempty, and $f : Z \to \mathbb{R}$ and $g : Z \to \mathbb{R}$ be bounded on Z. Then*

$$\left|\sup_{z\in Z} f(z) - \sup_{z\in Z} g(z)\right| \leq \sup_{z\in Z} |f(z) - g(z)|,$$
$$\left|\inf_{z\in Z} f(z) - \inf_{z\in Z} g(z)\right| \leq \sup_{z\in Z} |f(z) - g(z)|.$$

Definition 8 (Convex and Concave Relaxations) Given a convex set $Z \subset \mathbb{R}^n$ and a function $f : Z \to \mathbb{R}$, a convex function $f_Z^{\text{cv}} : Z \to \mathbb{R}$ is called a convex relaxation of f on Z if $f_Z^{\text{cv}}(\mathbf{z}) \leq f(\mathbf{z})$, $\forall \mathbf{z} \in Z$. Similarly, a concave function $f_Z^{\text{cc}} : Z \to \mathbb{R}$ is called a concave relaxation of f on Z if $f_Z^{\text{cc}}(\mathbf{z}) \geq f(\mathbf{z})$, $\forall \mathbf{z} \in Z$.

Definition 9 (Convex and Concave Envelopes) Given a convex set $Z \subset \mathbb{R}^n$ and a function $f : Z \to \mathbb{R}$, a convex function $f_Z^{\text{cv,env}} : Z \to \mathbb{R}$ is called the convex envelope of f on Z if $f_Z^{\text{cv,env}}$ is a convex relaxation of f on Z and for every convex relaxation $f_Z^{\text{cv}} : Z \to \mathbb{R}$, we have $f_Z^{\text{cv,env}}(\mathbf{z}) \geq f_Z^{\text{cv}}(\mathbf{z})$, $\forall \mathbf{z} \in Z$. Similarly, a concave function $f_Z^{\text{cc,env}} : Z \to \mathbb{R}$ is called the concave envelope of f on Z if $f_Z^{\text{cc,env}}$ is a concave relaxation of f on Z and for every concave relaxation $f_Z^{\text{cc}} : Z \to \mathbb{R}$, we have $f_Z^{\text{cc,env}}(\mathbf{z}) \leq f_Z^{\text{cc}}(\mathbf{z})$, $\forall \mathbf{z} \in Z$.

The following result establishes sufficient conditions for lower semicontinuity of the convex envelope. Note that a weaker version of this result is presented in [26, Corollary 17.2.1], and stronger versions of this result are stated without proof in [9, Page 349] (where the assumption that the function f is bounded above is relaxed) and in [34, Page 253] (where the assumptions that the function f is bounded above and the set W is bounded are relaxed).

Lemma 4 *Let $W \subset \mathbb{R}^{n_w}$ be a nonempty compact convex set and $f : W \to \mathbb{R}$ be a lower semicontinuous function on W bounded above by M. Let $f_W^{\text{cv,env}}$ denote the convex envelope of f on W. Then $f_W^{\text{cv,env}}$ is lower semicontinuous on W.*

Proof The function f is lower semicontinuous on the compact set W iff its epigraph $\{(\mathbf{x}, r) : \mathbf{x} \in W, r \geq f(\mathbf{x})\}$ is closed. Consequently, the set $S := \{(\mathbf{x}, r) : \mathbf{x} \in W, r \geq f(\mathbf{x}), r \leq M\}$ is compact. Theorem 17.2 in [26] implies that $\text{conv}(S)$ is a compact convex set. Therefore, the set $\text{conv}(S) \cup \{(\mathbf{x}, r) : \mathbf{x} \in W, r \geq f(\mathbf{x})\}$ is closed, which implies that $\{(\mathbf{x}, r) : \mathbf{x} \in W, r \geq f_W^{\text{cv,env}}(\mathbf{x})\}$ is closed, which in turn implies that $f_W^{\text{cv,env}}$ is lower semicontinuous on W. □

Remark 2 Although convex and concave relaxations of classes of functions can be constructed on general convex sets, the typical application requires construction of relaxations on bounded intervals. Therefore, we will implicitly assume that the sets X and Y are intervals and that relaxations are constructed on intervals in subsequent sections. The assumption that X and Y are intervals is not restrictive since general convex constraints defining X and Y that are available in factorable form can be equivalently reformulated to appear as part of the constraints $\mathbf{g}$ and $\mathbf{h}$. The proposed definitions of convergence order in the next section will be based on schemes of relaxations constructed on intervals. Note that similar notions of convergence order can be developed for schemes of relaxations constructed, for instance, on simplices.

3 Definitions of Convergence Order

This section reviews the definitions of convergence orders of schemes of relaxations [5, 6] and defines the convergence order of a (reduced-space) lower bounding scheme. It is also shown that the convergence order of a convergent scheme of relaxations at a point is governed by the tiny intervals around that point. We begin with the following definition, adapted from [5, Definition 6], that defines schemes of relaxations in a reduced-space.

Definition 10 (Schemes of Convex and Concave Relaxations) Let $V \subset \mathbb{R}^{n_v}$ and $W \subset \mathbb{R}^{n_w}$ be nonempty convex sets, and let $f : V \times W \to \mathbb{R}$. Suppose, for every $Z \in \mathbb{I}W$, we can construct functions $f^{\text{cv}}_{V\times Z} : V \times Z \to \mathbb{R}$ and $f^{\text{cc}}_{V\times Z} : V \times Z \to \mathbb{R}$ that are convex and concave relaxations, respectively, of f on $V \times Z$. The sets of functions $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ define schemes of convex and concave relaxations of f in W, respectively, and the set of pairs of functions $(f^{\text{cv}}_{V\times Z}, f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ defines a scheme of relaxations of f in W. The schemes of relaxations are said to be continuous when $f^{\text{cv}}_{V\times Z}$ and $f^{\text{cc}}_{V\times Z}$ are continuous on $V \times Z$ for each $Z \in \mathbb{I}W$.

Bompadre and coworkers [5, 6] define Hausdorff convergence of inclusion functions. Note that an inclusion function can be associated with schemes of relaxations in a natural way (see [5, Definition 7]).

Definition 11 (Hausdorff Convergence Order of an Inclusion Function) Let $V \in \mathbb{IR}^{n_v}$ and $W \subset \mathbb{R}^{n_w}$ be nonempty sets, $h : V \times W \to \mathbb{R}$ be a continuous function, and H be an inclusion function of h on $\mathbb{I}(V \times W)$.

The inclusion function H is said to have Hausdorff convergence of order $\beta > 0$ at a point $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau \geq 0$ such that

$$d_H(\overline{h}(V \times Z), H(V \times Z)) \leq \tau w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

Moreover, H is said to have Hausdorff convergence of order $\beta > 0$ on W if it has Hausdorff convergence of order (at least) β at each $\mathbf{w} \in W$, with the constant τ independent of $\mathbf{w}$.

In the context of (constrained) global optimization, the following definition of convergence of schemes of convex and concave relaxations is more pertinent.

Definition 12 (Convergence Order of Schemes of Convex and Concave Relaxations) Let $V \subset \mathbb{R}^{n_v}$, $W \subset \mathbb{R}^{n_w}$ be nonempty convex sets, and $f : V \times W \to \mathbb{R}$ be a continuous function. Let $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ respectively denote schemes of convex and concave relaxations of f in W.

The scheme of convex relaxations $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have convergence of order $\beta > 0$ at $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau^{\text{cv}} \geq 0$ such that

$$\inf_{(\mathbf{v},\mathbf{z})\in V\times Z} f(\mathbf{v},\mathbf{z}) - \inf_{(\mathbf{v},\mathbf{z})\in V\times Z} f^{\text{cv}}_{V\times Z}(\mathbf{v},\mathbf{z}) \leq \tau^{\text{cv}} w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

Similarly, the scheme of concave relaxations $(f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have convergence of order $\beta > 0$ at $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau^{\text{cc}} \geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} f^{\text{cc}}_{V\times Z}(\mathbf{v},\mathbf{z}) - \sup_{(\mathbf{v},\mathbf{z})\in V\times Z} f(\mathbf{v},\mathbf{z}) \leq \tau^{\text{cc}} w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

The scheme of relaxations $(f^{\text{cv}}_{V\times Z}, f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have (Hausdorff) convergence of order $\beta > 0$ at $\mathbf{w} \in W$ if the corresponding schemes of convex and concave relaxations have convergence of orders (at least) β at $\mathbf{w}$. The schemes $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ are said to have convergence of order $\beta > 0$ on W if they have convergence of order (at least) β at each $\mathbf{w} \in W$, with constants τ^{cv} and τ^{cc} independent of $\mathbf{w}$.

Definition 13 (Pointwise Convergence Order of Schemes of Convex and Concave Relaxations) Let $V \subset \mathbb{R}^{n_v}$, $W \subset \mathbb{R}^{n_w}$ be nonempty convex sets, and $f : V \times W \to \mathbb{R}$ be a continuous function. Let $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ respectively denote schemes of convex and concave relaxations of f in W. The scheme of convex

relaxations $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have pointwise convergence of order $\gamma>0$ at $\mathbf{w}\in W$ if for each bounded $Q\subset W$ with $\mathbf{w}\in Q$, there exists $\tau^{\mathrm{cv}}\geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f(\mathbf{v},\mathbf{z})-f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z})|\leq \tau^{\mathrm{cv}} w(Z)^{\gamma},\quad \forall Z\in\mathbb{I}Q \text{ with } \mathbf{w}\in Z.$$

Similarly, the scheme of concave relaxations $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have pointwise convergence of order $\gamma>0$ at $\mathbf{w}\in W$ if for each bounded $Q\subset W$ with $\mathbf{w}\in Q$, there exists $\tau^{\mathrm{cc}}\geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f^{\mathrm{cc}}_{V\times Z}(\mathbf{v},\mathbf{z})-f(\mathbf{v},\mathbf{z})|\leq \tau^{\mathrm{cc}} w(Z)^{\gamma},\quad \forall Z\in\mathbb{I}Q \text{ with } \mathbf{w}\in Z.$$

The scheme of relaxations $(f^{\mathrm{cv}}_{V\times Z}, f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have pointwise convergence of order $\gamma>0$ at $\mathbf{w}\in W$ if the corresponding schemes of convex and concave relaxations have pointwise convergence of orders (at least) γ at $\mathbf{w}$. Furthermore, the schemes of relaxations are said to have pointwise convergence of order $\gamma>0$ on W if they have pointwise convergence of order at least γ at each $\mathbf{w}\in W$, with constants τ^{cv} and τ^{cc} independent of $\mathbf{w}$.

Note that we simply say that a scheme of relaxations, $(f^{\mathrm{cv}}_{V\times Z}, f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$, of a function $f:V\times W\to\mathbb{R}$ in W has (pointwise) convergence order of $\gamma>0$ if it has (pointwise) convergence of order γ on W.

Remark 3 Definitions 11, 12, and 13 are based on a modification (see [16, Definition 9.2.35]) of the definitions of convergence order proposed in [5, 6], which incorporates the set Q. Note that the use of the set Q is necessary when the schemes of relaxations are constructed on unbounded sets, but may be omitted (set to W) when the schemes of relaxations are constructed over bounded sets (which is the typical application). Henceforth, the use of Q shall be omitted for brevity since we are only interested in compact sets V and W (see Assumption 1).

Remark 4 The pointwise convergence order of a convergent scheme of convex and concave relaxations on W is governed by the strength of the relaxations over small intervals in W. This observation is made precise in Lemma 5. Also note that the pointwise convergence order of schemes of either convex, or concave relaxations (as per Definition 13) can be arbitrarily high for nonlinear functions in contrast to the pointwise convergence order of schemes of convex *and* concave relaxations (see Theorem 2 in [5]). For instance, consider the function $f:[0,1]\times[0,1]=:V\times W\to\mathbb{R}$ with $f(v,w)=v^2-\sqrt{w}$ and a corresponding scheme of convex relaxations $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ defined by $f^{\mathrm{cv}}_{V\times Z}(v,z)=v^2-\sqrt{w}$ on $[w^{\mathrm{L}},w^{\mathrm{U}}]\subset[0,1]$. The scheme of convex relaxations $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ has arbitrarily high pointwise convergence order on W.

Remark 5 Unlike the pointwise convergence order of a scheme of relaxations, the convergence order of a scheme of convex and concave relaxations can be arbitrarily high for any function. For instance, consider the scheme of constant relaxations of the function $f:[0,1]\times[0,1]=:V\times W\to\mathbb{R}$ with $f(v,w)=w-\sqrt{v}$ defined by $f^{\mathrm{cv}}_{V\times Z}(v,z)=w^{\mathrm{L}}-1$, $f^{\mathrm{cc}}_{V\times Z}(v,z)=w^{\mathrm{U}}$ on $[w^{\mathrm{L}},w^{\mathrm{U}}]\subset[0,1]$. The scheme of constant relaxations $(f^{\mathrm{cv}}_{V\times Z}, f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ has arbitrarily high convergence order on W, but is not pointwise convergent of any order on W.

Lemma 5 *Let $V\subset\mathbb{R}^{n_v}, W\subset\mathbb{R}^{n_w}$ be nonempty compact convex sets and $f:V\times W\to\mathbb{R}$. Let $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ denote a scheme of convex relaxations of f in W with pointwise convergence order $\gamma^{\mathrm{cv}}>0$ and corresponding prefactor $\tau^{\mathrm{cv}}\geq 0$ (on W). If there exist constants $\gamma\geq\gamma^{\mathrm{cv}}$, $\tau\geq 0$, and $\delta>0$ such that for every $Z\in\mathbb{I}W$ with $w(Z)\leq\delta$,*

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f(\mathbf{v},\mathbf{z})-f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z})|\leq \tau w(Z)^{\gamma},$$

then $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ converges pointwise with order γ to f on W.

Proof Since $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ converges pointwise with order γ^{cv} to f on W which is compact, there exists $M\geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f(\mathbf{v},\mathbf{z})-f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z})|\leq \tau^{\mathrm{cv}} w(Z)^{\gamma^{\mathrm{cv}}}\leq M,\quad \forall Z\in\mathbb{I}W.$$

The desired result then follows from the fact that for every $Z \in \mathbb{I}W$,

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f(\mathbf{v},\mathbf{z}) - f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z})| \leq \left(\tau + \frac{M}{\delta^{\gamma}}\right) w(Z)^{\gamma}.$$

□

Results similar to Lemma 5 are applicable to other notions of convergence order presented in this work. Note that if the constant δ in Lemma 5 is relatively small, then the bound on the prefactor obtained can be relatively large making the result weak on intervals with $w(Z) \gg \delta$.

The next result shows that for schemes of relaxations, the notion of pointwise convergence is stronger than the notion of convergence in Definition 12 (also see [5, Theorem 1]).

Lemma 6 *Let $V \subset \mathbb{R}^{n_v}$, $W \subset \mathbb{R}^{n_w}$ be nonempty compact convex sets, and $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ respectively denote schemes of convex and concave relaxations of a bounded function $f : V \times W \to \mathbb{R}$ in W. If either scheme has pointwise convergence of order $\gamma > 0$, it has convergence of order $\beta \geq \gamma$.*

Proof By noting from Definition 13 that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f(\mathbf{v},\mathbf{z}) - f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z})| \leq \tau^{\mathrm{cv}} w(W)^{\gamma}, \quad \forall Z \in \mathbb{I}W,$$
$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f^{\mathrm{cc}}_{V\times Z}(\mathbf{v},\mathbf{z}) - f(\mathbf{v},\mathbf{z})| \leq \tau^{\mathrm{cc}} w(W)^{\gamma}, \quad \forall Z \in \mathbb{I}W,$$

the result follows from Lemma 3 via

$$\inf_{(\mathbf{v},\mathbf{z})\in V\times Z} f(\mathbf{v},\mathbf{z}) - \inf_{(\mathbf{v},\mathbf{z})\in V\times Z} f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z}) \leq \sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f(\mathbf{v},\mathbf{z}) - f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z})|, \quad \forall Z \in \mathbb{I}W,$$

and

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} f^{\mathrm{cc}}_{V\times Z}(\mathbf{v},\mathbf{z}) - \sup_{(\mathbf{v},\mathbf{z})\in V\times Z} f(\mathbf{v},\mathbf{z}) \leq \sup_{(\mathbf{v},\mathbf{z})\in V\times Z} |f^{\mathrm{cc}}_{V\times Z}(\mathbf{v},\mathbf{z}) - f(\mathbf{v},\mathbf{z})|, \quad \forall Z \in \mathbb{I}W.$$

□

The following lemma establishes mild sufficient conditions under which the scheme of envelopes of a function is first-order pointwise convergent.

Lemma 7 *Let $W \subset \mathbb{R}^{n_w}$ be a nonempty compact convex set and $f : W \to \mathbb{R}$ be Lipschitz continuous on W. Let $(f^{\mathrm{cv,env}}_Z, f^{\mathrm{cc,env}}_Z)_{Z\in\mathbb{I}W}$ denote the scheme of envelopes of f in W. Then the scheme $(f^{\mathrm{cv,env}}_Z, f^{\mathrm{cc,env}}_Z)_{Z\in\mathbb{I}W}$ is at least first-order pointwise convergent on W.*

Proof We wish to show that there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}W$,

$$\sup_{\mathbf{z}\in Z} |f(\mathbf{z}) - f^{\mathrm{cv,env}}_Z(\mathbf{z})| \leq \tau w(Z),$$
$$\sup_{\mathbf{z}\in Z} |f(\mathbf{z}) - f^{\mathrm{cc,env}}_Z(\mathbf{z})| \leq \tau w(Z).$$

Consider the scheme of relaxations $(f^{\mathrm{cv}}_Z, f^{\mathrm{cc}}_Z)_{Z\in\mathbb{I}W}$ defined by

$$f^{\mathrm{cv}}_Z(\mathbf{z}) = \min_{\mathbf{w}\in Z} f(\mathbf{w}), \; f^{\mathrm{cc}}_Z(\mathbf{z}) = \max_{\mathbf{w}\in Z} f(\mathbf{w}), \quad \forall Z \in \mathbb{I}W.$$

From the fact that f^{cv}_Z and f^{cc}_Z are convex and concave relaxations of f in Z and the assumption that f is Lipschitz continuous, we have that $(f^{\mathrm{cv}}_Z, f^{\mathrm{cc}}_Z)_{Z\in\mathbb{I}W}$ is at least first-order pointwise convergent on W. The desired result then follows from the definition of $(f^{\mathrm{cv,env}}_Z, f^{\mathrm{cc,env}}_Z)_{Z\in\mathbb{I}W}$. □

The definitions provided thus far facilitate a theoretical analysis of the (reduced-space) convergence order of a scheme of relaxations to a corresponding scalar function, or, in the context of global optimization, provide a way to analyze theoretically the (reduced-space) convergence order of a (lower) bounding scheme for an unconstrained problem. The subsequent definitions seek to extend naturally the analysis of convergence order to constrained problems.

Definition 14 (Convergence Order of a Lower Bounding Scheme) Consider Problem (P) (satisfying Assumption 1). For any $Z \in \mathbb{I}Y$, let $\mathscr{F}(Z) = \{(\mathbf{x},\mathbf{y}) \in X \times Z : \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}\}$ denote the feasible set of Problem (P) with $\mathbf{y}$ restricted to Z.

Consider a scheme of lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ for Problem (P). We associate with the scheme $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ a scheme of pairs $(\mathscr{O}(Z),\mathscr{I}_C(Z))_{Z\in\mathbb{I}Y}$, where $(\mathscr{O}(Z))_{Z\in\mathbb{I}Y}$ is a scheme of lower bounds on the scheme of problems $\left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}Y}$ and $(\mathscr{I}_C(Z))_{Z\in\mathbb{I}Y}$ is a scheme of subsets of $\mathbb{R}^{m_I+m_E}$ that indicate the feasibility of the lower bounding scheme $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$. The schemes $(\mathscr{O}(Z))_{Z\in\mathbb{I}Y}$ and $(\mathscr{I}_C(Z))_{Z\in\mathbb{I}Y}$ (are required to) satisfy

$$\begin{aligned}
&\mathscr{O}(Z) \leq \min_{(\mathbf{x},\mathbf{z})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{z}), \quad \forall Z \in \mathbb{I}Y,\\
&d(\mathscr{I}_C(Z),\mathbb{R}^{m_I}_- \times \{\mathbf{0}\}) \leq d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X\times Z),\mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right), \quad \forall Z \in \mathbb{I}Y,\\
&\mathscr{O}(Z) = +\infty \iff d(\mathscr{I}_C(Z),\mathbb{R}^{m_I}_- \times \{\mathbf{0}\}) > 0, \quad \forall Z \in \mathbb{I}Y,
\end{aligned}$$

where $\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X\times Z)$ denotes the image of $X \times Z$ under the vector-valued function $\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}$. The scheme of lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ is said to have convergence of order $\beta > 0$ at

1. a feasible point $\mathbf{y} \in Y$ if there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$\min_{(\mathbf{x},\mathbf{z})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{z}) \; - \mathscr{O}(Z) \leq \tau w(Z)^\beta.$$

2. an infeasible point $\mathbf{y} \in Y$ if there exists $\overline{\tau} \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X\times Z),\mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) - d\left(\mathscr{I}_C(Z),\mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) \leq \overline{\tau} w(Z)^\beta.$$

The scheme of lower bounding problems is said to have convergence of order $\beta > 0$ on Y if it has convergence of order (at least) β at each $\mathbf{y} \in Y$, with constants τ and $\overline{\tau}$ independent of $\mathbf{y}$.

Remark 6 Definition 14 is motivated by the requirements of a lower bounding scheme to fathom feasible and infeasible regions in a branch-and-bound procedure [13]. The first condition requires that the sequence of lower bounds converges rapidly to the corresponding sequence of minimum objective values on nested sequences of intervals converging to a feasible point of Problem (P). On nested sequences of intervals converging to an infeasible point of Problem (P), the second condition requires that the sequence of lower bounding problems rapidly detect the (eventual) infeasibility of the corresponding sequences of intervals for Problem (P). In simple terms, the first condition can be used to require that feasible points with 'good objective values' are fathomed rather easily, while the second condition can be used to require that infeasible points that are 'close to the feasible region', as determined by the distance function d, are fathomed with relatively less effort [15]. Note that Definition 14 reduces to the definition of convergence order for unconstrained minimization in [39, Definition 1] when n_x, m_I, and m_E are all set to zero.

Definition 14 can be readily applied to analyze the convergence order of a convex relaxation-based lower bounding scheme as follows.

Suppose, for each $Z \in \mathbb{I}Y$, we associate a convex set $X(Z)$ such that $X \supset X(Z) \supset \mathscr{F}_X(Z)$, where $\mathscr{F}_X(Z) := \{\mathbf{x} \in X : \exists \mathbf{y} \in Z \text{ s.t. } \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}\}$ denotes the projection of $\mathscr{F}(Z)$ on X. The set $X(Z)$ could, for instance, correspond to an interval subset of X that is obtained using bounds tightening techniques [2] when $\mathbf{y}$ is restricted to Z (the motivation for considering the set $X(Z)$ in the definition of convergence order below will become clear in Section 5). Note that the restriction $X(Z) \supset \mathscr{F}_X(Z)$ can be relaxed when optimality-based bounds tightening techniques are employed. Also note that unless otherwise specified, we simply use $X(Z) = X, \forall Z \in \mathbb{I}Y$.

By an abuse of Definition 10, let $(f^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ and $(\mathbf{g}^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ denote continuous schemes of convex relaxations of f and $\mathbf{g}$, respectively, in Y, and let $(\mathbf{h}^{\text{cv}}_{X(Z)\times Z}, \mathbf{h}^{\text{cc}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ denote a continuous scheme of relaxations of $\mathbf{h}$ in Y. For any $Z \in \mathbb{I}Y$, let

$$\mathscr{F}^{\text{cv}}(Z) = \left\{(\mathbf{x},\mathbf{y}) \in X(Z)\times Z : \mathbf{g}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\text{cc}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \geq \mathbf{0}\right\}$$

denote the feasible set of the scheme of convex relaxations. The lower bounding scheme $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ with

$$\begin{aligned}(\mathscr{O}(Z))_{Z\in\mathbb{I}Y} &:= \left(\min_{(\mathbf{x},\mathbf{z})\in\mathscr{F}^{\text{cv}}(Z)} f^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z})\right)_{Z\in\mathbb{I}Y},\\ (\mathscr{I}_C(Z))_{Z\in\mathbb{I}Y} &:= \left(\left\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I}\times\mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z}), \mathbf{h}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z}) \leq \mathbf{w} \leq \mathbf{h}^{\text{cc}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z})\right.\right.\\ &\qquad\qquad \left.\left.\text{for some } (\mathbf{x},\mathbf{z}) \in X(Z)\times Z\right\}\right)_{Z\in\mathbb{I}Y} \qquad (1)\end{aligned}$$

is said to have convergence of order $\beta > 0$ at

1. a feasible point $\mathbf{y} \in Y$ if there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$\min_{(\mathbf{x},\mathbf{z})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{z}) - \min_{(\mathbf{x},\mathbf{z})\in\mathscr{F}^{\text{cv}}(Z)} f^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z}) \leq \tau w(Z)^{\beta}.$$

2. an infeasible point $\mathbf{y} \in Y$ if there exists $\overline{\tau} \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X(Z)\times Z), \mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right) - d\left(\mathscr{I}_C(Z), \mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right) \leq \overline{\tau} w(Z)^{\beta},$$

where $\mathscr{I}_C(Z)$ is defined by Equation (1).

Definition 14 can also be used to analyze the convergence orders of alternative lower bounding schemes such as those based on Lagrangian duality (see Section 4.2).

4 Full-Space Branch-and-Bound Algorithms

In this section, we present some results on the convergence order of lower bounding schemes for Problem (P) when both the X and Y sets may be partitioned during the course of the branch-and-bound algorithm (we consider schemes of relaxations in $X \times Y$ instead of schemes of relaxations in Y as was considered in Section 3). This section is divided into two parts. First, we look at the convergence order of lower bounding schemes which utilize convex and concave relaxations (see, for instance, [1, 17, 20, 36, 37] for techniques to construct relaxations) of the objective and the constraints for the lower bounding scheme. Next, the convergence order of duality-based lower bounding schemes (see, for instance, [9]) is investigated.

4.1 Convex relaxation-based branch-and-bound

This section derives bounds on the convergence order of convex relaxation-based lower bounding schemes by making assumptions on the convergence orders of the schemes of relaxations used by the lower bounding schemes. The reader is directed to [5], [6], [24], and [17] for details on how to construct schemes of (convex) relaxations that have the requisite convergence orders.

The following result establishes a lower bound on the convergence order of the lower bounding scheme at infeasible points based on the convergence orders of schemes of convex relaxations of the inequality constraints and schemes of relaxations of the equality constraints. Note that this is the primary result that is used to derive a lower bound on the convergence order of such relaxation-based lower bounding schemes at infeasible points.

Lemma 8 *Let $(g^{\mathrm{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j=1,\cdots,m_I$, denote continuous schemes of convex relaxations of $g_1,\cdots,g_{m_I}$, respectively, in $X\times Y$ with pointwise convergence orders $\gamma^{\mathrm{cv}}_{g,1}>0,\cdots,\gamma^{\mathrm{cv}}_{g,m_I}>0$ and corresponding constants $\tau^{\mathrm{cv}}_{g,1},\cdots,\tau^{\mathrm{cv}}_{g,m_I}$, and $(h^{\mathrm{cv}}_{k,Z},h^{\mathrm{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k=1,\cdots,m_E$, denote continuous schemes of relaxations of $h_1,\cdots,h_{m_E}$, respectively, in $X\times Y$ with pointwise convergence orders $\gamma_{h,1}>0,\cdots,\gamma_{h,m_E}>0$ and corresponding constants $\tau_{h,1},\cdots,\tau_{h,m_E}$. Then, there exists $\overline{\tau}\geq 0$ such that for every $Z\in\mathbb{I}(X\times Y)$*

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)-d\left(\mathscr{I}_C(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)\leq\overline{\tau}w(Z)^{\beta},$$

where $\mathscr{I}_C(Z)$ is defined as

$$\mathscr{I}_C(Z):=\{(\mathbf{v},\mathbf{w})\in\mathbb{R}^{m_I}\times\mathbb{R}^{m_E}:\mathbf{v}=\mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}),\mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})\leq\mathbf{w}\leq\mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y})\ \textit{for some}\ (\mathbf{x},\mathbf{y})\in Z\},$$

and β is defined as

$$\beta:=\min\left\{\min_{j\in\{1,\cdots,m_I\}}\gamma^{\mathrm{cv}}_{g,j},\min_{k\in\{1,\cdots,m_E\}}\gamma_{h,k}\right\}.$$

Proof Suppose $Z\in\mathbb{I}(X\times Y)$. Then for each $j\in\{1,\cdots,m_I\}$, $k\in\{1,\cdots,m_E\}$, we have from Definition 13 that

$$\begin{aligned}\max_{(\mathbf{x},\mathbf{y})\in Z}|g_j(\mathbf{x},\mathbf{y})-g^{\mathrm{cv}}_{j,Z}(\mathbf{x},\mathbf{y})|&\leq\tau^{\mathrm{cv}}_{g,j}w(Z)^{\gamma^{\mathrm{cv}}_{g,j}},\\ \max_{(\mathbf{x},\mathbf{y})\in Z}|h_k(\mathbf{x},\mathbf{y})-h^{\mathrm{cv}}_{k,Z}(\mathbf{x},\mathbf{y})|&\leq\tau_{h,k}w(Z)^{\gamma_{h,k}},\\ \max_{(\mathbf{x},\mathbf{y})\in Z}|h_k(\mathbf{x},\mathbf{y})-h^{\mathrm{cc}}_{k,Z}(\mathbf{x},\mathbf{y})|&\leq\tau_{h,k}w(Z)^{\gamma_{h,k}},\end{aligned}$$

since $(g^{\mathrm{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$ and $(h^{\mathrm{cv}}_{k,Z},h^{\mathrm{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$ converge pointwise to g_j and h_k, respectively, on $X\times Y$ with orders $\gamma^{\mathrm{cv}}_{g,j}$ and $\gamma_{h,k}$. Let $(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)\in Z$ and $(\mathbf{v}^{\mathrm{cv}}_Z,\mathbf{w}^{\mathrm{cv}}_Z)\in\mathscr{I}_C(Z)$ such that $\mathbf{v}^{\mathrm{cv}}_Z=\mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z),\mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)\leq\mathbf{w}^{\mathrm{cv}}_Z\leq\mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)$, and $d(\{(\mathbf{v}^{\mathrm{cv}}_Z,\mathbf{w}^{\mathrm{cv}}_Z)\},\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\})=d\left(\mathscr{I}_C(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)$. The existence of $(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)$ and

$(\mathbf{v}_Z^{\mathrm{cv}}, \mathbf{w}_Z^{\mathrm{cv}})$ follows from the continuity of $\mathbf{g}_Z^{\mathrm{cv}}$, $\mathbf{h}_Z^{\mathrm{cv}}$, and $\mathbf{h}_Z^{\mathrm{cc}}$ and the compactness of Z. We have

$$
\begin{aligned}
& d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) - d\left(\mathscr{I}_C(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) \\
& \leq d\left(\left\{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\right\}, \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) - d\left(\{(\mathbf{v}_Z^{\mathrm{cv}}, \mathbf{w}_Z^{\mathrm{cv}})\}, \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) \\
& \leq d\left(\left\{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - (\mathbf{v}_Z^{\mathrm{cv}}, \mathbf{w}_Z^{\mathrm{cv}})\right\}, \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) \\
& \leq \|\mathbf{g}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{v}_Z^{\mathrm{cv}}\| + \|\mathbf{h}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{w}_Z^{\mathrm{cv}}\| \\
& \leq \|\mathbf{g}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\| + \max\{\|\mathbf{h}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\|, \|\mathbf{h}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\|\} \\
& \leq \max_{(\mathbf{x},\mathbf{y})\in Z} \|\mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\| + \max\left\{\max_{(\mathbf{x},\mathbf{y})\in Z} \|\mathbf{h}(\mathbf{x},\mathbf{y}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\|, \max_{(\mathbf{x},\mathbf{y})\in Z} \|\mathbf{h}(\mathbf{x},\mathbf{y}) - \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y})\|\right\} \\
& \leq \sum_{j=1}^{m_I} \max_{(\mathbf{x},\mathbf{y})\in Z} |g_j(\mathbf{x},\mathbf{y}) - g_{j,Z}^{\mathrm{cv}}(\mathbf{x},\mathbf{y})| + \max\left\{\sum_{k=1}^{m_E} \max_{(\mathbf{x},\mathbf{y})\in Z} |h_k(\mathbf{x},\mathbf{y}) - h_{k,Z}^{\mathrm{cv}}(\mathbf{x},\mathbf{y})|, \sum_{k=1}^{m_E} \max_{(\mathbf{x},\mathbf{y})\in Z} |h_k(\mathbf{x},\mathbf{y}) - h_{k,Z}^{\mathrm{cc}}(\mathbf{x},\mathbf{y})|\right\} \\
& \leq \sum_{j=1}^{m_I} \tau_{g,j}^{\mathrm{cv}} w(Z)^{\gamma_{g,j}^{\mathrm{cv}}} + \sum_{k=1}^{m_E} \tau_{h,k} w(Z)^{\gamma_{h,k}} \\
& \leq \left(\sum_{j=1}^{m_I} \tau_{g,j}^{\mathrm{cv}} w(X\times Y)^{\gamma_{g,j}^{\mathrm{cv}}-\beta} + \sum_{k=1}^{m_E} \tau_{h,k} w(X\times Y)^{\gamma_{h,k}-\beta}\right) w(Z)^{\beta},
\end{aligned}
$$

where Corollary 1 is used to derive Step 2, Step 3 follows from the triangle inequality, and Lemma 2 is used to derive Step 6. The desired result follows by choosing $\overline{\tau}$ as

$$
\overline{\tau} = \left(\sum_{j=1}^{m_I} \tau_{g,j}^{\mathrm{cv}} w(X\times Y)^{\gamma_{g,j}^{\mathrm{cv}}-\beta} + \sum_{k=1}^{m_E} \tau_{h,k} w(X\times Y)^{\gamma_{h,k}-\beta}\right).
$$

□

Note that the conclusions of Lemma 8 hold even when the schemes of convex relaxations $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $\forall j \in \{1,\cdots,m_I\}$, and $(h_{k,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $\forall k \in \{1,\cdots,m_E\}$, are merely lower semicontinuous, and the schemes of concave relaxations $(h_{k,Z}^{\mathrm{cc}})_{Z\in\mathbb{I}(X\times Y)}$, $\forall k \in \{1,\cdots,m_E\}$, are merely upper semicontinuous.

Remark 7 The analysis in Lemma 8 can be refined under the following assumptions. Let $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1,\cdots,m_I$, denote schemes of convex relaxations of $g_1,\cdots,g_{m_I}$, respectively, in $X\times Y$ with convergence orders $\beta_{g,1}^{\mathrm{cv}} > 0,\cdots,\beta_{g,m_I}^{\mathrm{cv}} > 0$ and corresponding constants $\tau_{g,1}^{\mathrm{cv}},\cdots,\tau_{g,m_I}^{\mathrm{cv}}$, and let $(h_{k,Z}^{\mathrm{cv}}, h_{k,Z}^{\mathrm{cc}})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1,\cdots,m_E$, denote schemes of relaxations of $h_1,\cdots,h_{m_E}$, respectively, in $X\times Y$ with convergence orders $\beta_{h,1} > 0,\cdots,\beta_{h,m_E} > 0$ and corresponding constants $\tau_{h,1},\cdots,\tau_{h,m_E}$. Suppose for each interval $Z \in \mathbb{I}(X\times Y)$, there exists $(\mathbf{x}_Z, \mathbf{y}_Z) \in Z$ such that

$$
\begin{aligned}
& d\left(\{(\mathbf{x}_Z,\mathbf{y}_Z)\}, \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) = d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right), \\
& (\mathbf{x}_Z,\mathbf{y}_Z) \in \operatorname*{arg\,min}_{(\mathbf{x},\mathbf{y})\in Z} g_j(\mathbf{x},\mathbf{y}), \quad \forall j \in \{1,\cdots,m_I\}, \\
& \text{either } (\mathbf{x}_Z,\mathbf{y}_Z) \in \operatorname*{arg\,min}_{(\mathbf{x},\mathbf{y})\in Z} h_k(\mathbf{x},\mathbf{y}), \text{ or } (\mathbf{x}_Z,\mathbf{y}_Z) \in \operatorname*{arg\,max}_{(\mathbf{x},\mathbf{y})\in Z} h_k(\mathbf{x},\mathbf{y}), \quad \forall k \in \{1,\cdots,m_E\}.
\end{aligned}
$$

Then, there exists $\overline{\tau} \geq 0$ such that for every $Z \in \mathbb{I}(X\times Y)$

$$
d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) - d\left(\mathscr{I}_C(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) \leq \overline{\tau} w(Z)^{\beta},
$$

where β is defined as

$$\beta := \min\left\{\min_{j\in\{1,\cdots,m_I\}} \beta^{\mathrm{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}} \beta_{h,k}\right\}.$$

Note that the above assumptions are trivially satisfied when Problem (P) only has one constraint (cf. Example 1).

The following example demonstrates the importance of a sufficiently high convergence order at nearly-feasible points (also see [15, Example 4]).

Example 1 Let $X = [0,0], Y = [-1,1]$, $m_I = 1$, and $m_E = 0$ with $f(x,y) = y$ and $g(x,y) = -y$. For any $[0,0] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$, let $f^{\mathrm{cv}}_Z(x,y) = y$, $g^{\mathrm{cv}}_Z(x,y) = -y^{\mathrm{U}} - (y^{\mathrm{U}} - y^{\mathrm{L}})^{\alpha}$ for some constant $\alpha > 0$. Note that $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order and arbitrarily high convergence order on $X \times Y$, whereas $(g^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has $\min\{\alpha, 1\}$-order pointwise convergence and α-order convergence on $X \times Y$.

Pick $\delta \in (0,1)$ and let $\varepsilon \in (0,\delta)$. Let $y^{\mathrm{L}} = -\delta - \varepsilon$, $y^{\mathrm{U}} = -\delta + \varepsilon$. The width of Z is $w(Z) = 2\varepsilon$. We have $\overline{g}(Z) = [\delta - \varepsilon, \delta + \varepsilon]$, which yields $d(\overline{g}(Z), \mathbb{R}^{m_I}_-) = \delta - \varepsilon$ (this confirms that g is infeasible at each $(x,y) \in Z$). Furthermore, $\overline{g}^{\mathrm{cv}}_Z(Z) = [\delta - \varepsilon - (2\varepsilon)^{\alpha}, \delta - \varepsilon - (2\varepsilon)^{\alpha}]$, which yields $d(\overline{g}^{\mathrm{cv}}_Z(Z), \mathbb{R}^{m_I}_-) = \max\{0, \delta - \varepsilon - (2\varepsilon)^{\alpha}\}$. Therefore, for ε sufficiently small, the lower bounding problem detects the infeasibility of Z and we have $d(\overline{g}(Z), \mathbb{R}^{m_I}_-) - d(\overline{g}^{\mathrm{cv}}_Z(Z), \mathbb{R}^{m_I}_-) = (2\varepsilon)^{\alpha}$, which implies that convergence of the lower bounding scheme at the infeasible point $(0, -\delta)$ is at most of order α.

For $\alpha = 1$, the maximum value of ε for which the interval Z can be fathomed by infeasibility by the lower bounding scheme is $\varepsilon = \frac{\delta}{3}$, whereas for $\alpha = 0.5$, the maximum value of ε for which the interval Z can be fathomed by infeasibility by the lower bounding scheme is $\varepsilon = \left(\frac{-\sqrt{2}+\sqrt{2}\sqrt{1+2\delta}}{2}\right)^2$, which is $O(\delta^2)$ for $\delta \ll 1$.

Therefore, a lower bounding scheme with a low convergence order at infeasible points may result in a large number of partitions on nearly-feasible points before they are fathomed, thereby resulting in the cluster problem.

The next result shows that under mild assumptions on the objective, the constraints, and the schemes of relaxations, first-order convergence to a global minimum is guaranteed.

Theorem 1 *Consider Problem* (P). *Suppose* f, g_j, $j = 1,\cdots,m_I$, *and* h_k, $k = 1,\cdots,m_E$, *are Lipschitz continuous on* $X \times Y$ *with Lipschitz constants* M_f, $M_{g,1},\cdots,M_{g,m_I}$, $M_{h,1},\cdots,M_{h,m_E}$, *respectively. Let* $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$, $(g^{\mathrm{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1,\cdots,m_I$, *denote continuous schemes of convex relaxations of* f, $g_1,\cdots,g_{m_I}$, *respectively, in* $X \times Y$ *with pointwise convergence orders* $\gamma^{\mathrm{cv}}_f \geq 1$, $\gamma^{\mathrm{cv}}_{g,1} \geq 1,\cdots,\gamma^{\mathrm{cv}}_{g,m_I} \geq 1$ *and corresponding constants* τ^{cv}_f, $\tau^{\mathrm{cv}}_{g,1},\cdots,\tau^{\mathrm{cv}}_{g,m_I}$. *Let* $(h^{\mathrm{cv}}_{k,Z}, h^{\mathrm{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1,\cdots,m_E$, *denote continuous schemes of relaxations of* $h_1,\cdots,h_{m_E}$, *respectively, in* $X \times Y$ *with pointwise convergence orders* $\gamma_{h,1} \geq 1,\cdots,\gamma_{h,m_E} \geq 1$ *and corresponding constants* $\tau_{h,1},\cdots,\tau_{h,m_E}$. *The scheme of lower bounding problems* $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ *with*

$$(\mathscr{O}(Z))_{Z\in\mathbb{I}(X\times Y)} := \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}(X\times Y)},$$

$$(\mathscr{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)} := \left(\left\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I}\times\mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}), \mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y}) \right.\right.$$
$$\left.\left. \text{for some } (\mathbf{x},\mathbf{y}) \in Z\right\}\right)_{Z\in\mathbb{I}(X\times Y)}$$

is at least first-order convergent on $X \times Y$.

Proof Lemma 8 establishes first-order convergence at infeasible points $(\mathbf{x},\mathbf{y}) \in X \times Y$ with the prefactor $\overline{\tau}$ independent of $(\mathbf{x},\mathbf{y})$; therefore, it suffices to prove first-order convergence at feasible points $(\mathbf{x},\mathbf{y}) \in X \times Y$ with a prefactor independent of $(\mathbf{x},\mathbf{y})$.

In order to do so, suppose $\mathscr{F}(X \times Y) \neq \emptyset$ and consider $Z \in \mathbb{I}(X \times Y)$ such that $Z \cap \mathscr{F}(X \times Y) \neq \emptyset$. Let

$$\mathscr{F}^{\mathrm{cv}}(Z) := \{(\mathbf{x},\mathbf{y}) \in Z : \mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y}) \geq \mathbf{0}\}$$

denote the feasible set of the convex relaxation-based lower bounding scheme. Then

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \\
&= \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y})\right) + \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\right) \\
&\leq \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y})\right) \; + \max_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} |f(\mathbf{x},\mathbf{y}) - f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})|, \qquad (2)
\end{aligned}$$

where the above inequality follows from Lemma 3. The second term in Equation (2) can be bounded from above as

$$\max_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} |f(\mathbf{x},\mathbf{y}) - f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})| \leq \tau_f^{\mathrm{cv}} w(Z)^{\gamma_f^{\mathrm{cv}}},$$

since $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ converges pointwise to f on $X\times Y$ with order $\gamma_f^{\mathrm{cv}} \geq 1$.

Let $(\mathbf{x}_Z^*,\mathbf{y}_Z^*) \in \operatorname*{arg\,min}_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y})$ and $(\mathbf{x}_Z^{\mathrm{cv}},\mathbf{y}_Z^{\mathrm{cv}}) \in \operatorname*{arg\,min}_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y})$. The first term in Equation (2) can be bounded from above as

$$\begin{aligned}
\left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y})\right) &= f(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - f(\mathbf{x}_Z^{\mathrm{cv}},\mathbf{y}_Z^{\mathrm{cv}}) \\
&\leq M_f\sqrt{n_x+n_y}\, w(Z),
\end{aligned}$$

where the last step follows from the Lipschitz continuity of f on $X\times Y$ and Lemma 2.

Plugging in the above bounds in Equation (2), we get

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \leq \left(M_f\sqrt{n_x+n_y} + \tau_f^{\mathrm{cv}} w(X\times Y)^{\gamma_f^{\mathrm{cv}}-1}\right) w(Z),$$

which establishes first-order convergence of $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ at feasible points $(\mathbf{x},\mathbf{y}) \in X\times Y$ with the prefactor independent of $(\mathbf{x},\mathbf{y})$. □

The following examples show that the convergence order of the lower bounding scheme may be as low as one under the assumptions of Theorem 1.

Example 2 Let $X=[-1,1], Y=[-1,1]$, $m_I=1$, and $m_E=0$ with $f(x,y)=2x+2y$ and $g(x,y)=-x-y$. For any $[x^{\mathrm{L}},x^{\mathrm{U}}]\times[y^{\mathrm{L}},y^{\mathrm{U}}] =: Z \in \mathbb{I}(X\times Y)$, let $f_Z^{\mathrm{cv}}(x,y)=2x+2y$ and $g_Z^{\mathrm{cv}}(x,y)=-x^{\mathrm{U}}-y^{\mathrm{U}}$. The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$ and the scheme $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has first-order pointwise convergence on $X\times Y$. Note that $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high convergence order on $X\times Y$.

Let $x^{\mathrm{L}}=y^{\mathrm{L}}=-\varepsilon$, $x^{\mathrm{U}}=y^{\mathrm{U}}=\varepsilon$ with $0<\varepsilon\leq 1$. The width of Z is $w(Z)=2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is -4ε. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Example 3 Let $X=[-1,1], Y=[-1,1]$, $m_I=1$, and $m_E=0$ with $f(x,y)=2x+2y$ and $g(x,y)=-x-y$. For any $[x^{\mathrm{L}},x^{\mathrm{U}}]\times[y^{\mathrm{L}},y^{\mathrm{U}}] =: Z \in \mathbb{I}(X\times Y)$, let $f_Z^{\mathrm{cv}}(x,y)=2x^{\mathrm{L}}+2y^{\mathrm{L}}$ and $g_Z^{\mathrm{cv}}(x,y)=-x-y$. The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has first-order pointwise convergence on $X\times Y$ and the scheme $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$. Note that $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high convergence order on $X\times Y$.

Let $x^{\mathrm{L}}=y^{\mathrm{L}}=-\varepsilon$, $x^{\mathrm{U}}=y^{\mathrm{U}}=\varepsilon$ with $0<\varepsilon\leq 1$. The width of Z is $w(Z)=2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is -4ε. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Example 4 Let $X=[0,0], Y=[-1,1]$, $m_I=1$, and $m_E=0$ with $f(x,y)=y$ and $g(x,y)=\min\{-0.5y,-y\}$. For any $[0,0]\times[y^{\mathrm{L}},y^{\mathrm{U}}]=:Z\in\mathbb{I}(X\times Y)$ with $y^{\mathrm{L}}<0<y^{\mathrm{U}}$, let

$$f_Z^{\mathrm{cv}}(x,y)=y,\quad g_Z^{\mathrm{cv}}(x,y)=-\frac{y^{\mathrm{U}}-0.5y^{\mathrm{L}}}{y^{\mathrm{U}}-y^{\mathrm{L}}}y+\frac{0.5y^{\mathrm{L}}y^{\mathrm{U}}}{y^{\mathrm{U}}-y^{\mathrm{L}}}.$$

Note that g_Z^{cv} corresponds to the convex envelope of g on Z. The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$ and the scheme $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has first-order pointwise convergence on $X\times Y$. Note that $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high convergence order on $X\times Y$.

Let $y^{\mathrm{L}}=-\varepsilon$, $y^{\mathrm{U}}=\varepsilon$ with $0<\varepsilon\leq 1$. The width of Z is $w(Z)=2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-\frac{\varepsilon}{3}$. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Example 5 Let $X=[0,0], Y=[-1,1]$, $m_I=0$, and $m_E=1$ with $f(x,y)=y$ and $h(x,y)=\min\{-0.5y,-y\}$. For any $[0,0]\times[y^{\mathrm{L}},y^{\mathrm{U}}]=:Z\in\mathbb{I}(X\times Y)$ with $y^{\mathrm{L}}<0<y^{\mathrm{U}}$, let

$$f_Z^{\mathrm{cv}}(x,y)=y,\quad h_Z^{\mathrm{cv}}(x,y)=-\frac{y^{\mathrm{U}}-0.5y^{\mathrm{L}}}{y^{\mathrm{U}}-y^{\mathrm{L}}}y+\frac{0.5y^{\mathrm{L}}y^{\mathrm{U}}}{y^{\mathrm{U}}-y^{\mathrm{L}}},\quad h_Z^{\mathrm{cc}}(x,y)=\min\{-0.5y,-y\}.$$

Note that h_Z^{cv} and h_Z^{cc} correspond to the convex and concave envelopes of h on Z, respectively. The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$ and the scheme $(h_Z^{\mathrm{cv}},h_Z^{\mathrm{cc}})_{Z\in\mathbb{I}(X\times Y)}$ has first-order pointwise convergence on $X\times Y$. Note that $(h_Z^{\mathrm{cv}},h_Z^{\mathrm{cc}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high convergence order on $X\times Y$.

Let $y^{\mathrm{L}}=-\varepsilon$, $y^{\mathrm{U}}=\varepsilon$ with $0<\varepsilon\leq 1$. The width of Z is $w(Z)=2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-\frac{\varepsilon}{3}$. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Despite the fact that the schemes of relaxations used in Examples 4 and 5 correspond to the envelopes of the functions involved (unlike those used in Examples 2 and 3), we only have first-order convergence at the global minimizer $(0,0)$. However, the reader can verify that first-order convergent lower bounding schemes may be sufficient to mitigate the cluster problem in Examples 4 and 5, whereas at least second-order convergent lower bounding schemes are required to mitigate the cluster problem in Examples 2 and 3 [15]. Furthermore, Examples 2 to 5 illustrate that high convergence orders of schemes of relaxations of the objective and constraints do not guarantee a high convergence order of the lower bounding scheme (cf. Remark 7) at constrained minima (which may be required to mitigate clustering). This is because a high convergence order of a scheme of relaxations of the objective function may only place a restriction on the gap between the minimum value of the relaxation and the minimum value of the objective function without taking the feasible region into account; this restriction may not be sufficient in a constrained setting because the gap between the minimum value of the relaxed problem and the minimum value of the original problem may be relatively large when their respective feasible regions are taken into consideration (see Example 3 for an extreme case). Similarly, a high convergence order of a scheme of relaxations of the constraints may not exclude infeasible regions of the search space in which the objective function value is less than the optimal (constrained) objective value (Example 2 provides an extreme case), potentially leading to relatively large underestimation gaps.

The following result proves second-order convergence at certain points in $X\times Y$.

Theorem 2 *Consider Problem* (P). *Suppose* f *is Lipschitz continuous on* $X\times Y$ *with Lipschitz constant* M_f. *Let* $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ *denote a continuous scheme of convex relaxations of* f *with pointwise convergence order* $\gamma_f^{\mathrm{cv}}\geq 2$ *and corresponding constant* τ_f^{cv}.

Suppose there exists a feasible point $(\mathbf{x}^{\mathrm{f}},\mathbf{y}^{\mathrm{f}})\in\mathscr{F}(X\times Y)$, *continuous schemes of convex relaxations* $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $j=1,\cdots,m_I$, *of* $g_1,\cdots,g_{m_I}$, *respectively, in* $X\times Y$, *continuous schemes of relaxations*

$(h^{\mathrm{cv}}_{k,Z}, h^{\mathrm{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1,\cdots,m_E$, of $h_1,\cdots,h_{m_E}$, respectively, in $X\times Y$, and a constant $\delta > 0$ such that for each $Z \in \mathbb{I}(X\times Y)$ with $(\mathbf{x}^{\mathrm{f}},\mathbf{y}^{\mathrm{f}}) \in Z$ and $w(Z) \le \delta$, we have

$$d\left(\mathscr{F}(Z), \underset{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)}{\arg\min} f(\mathbf{x},\mathbf{y})\right) \le \hat{\tau} w(Z)^{\gamma} \tag{3}$$

for constants $\gamma \ge 2$ and $\hat{\tau} \ge 0$. Then, the scheme of lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$\begin{aligned}(\mathscr{O}(Z))_{Z\in\mathbb{I}(X\times Y)} &:= \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}(X\times Y)},\\ (\mathscr{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)} &:= \left(\left\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I}\times\mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}), \mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) \le \mathbf{w} \le \mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y})\right.\right.\\ &\qquad\qquad \left.\left. \text{for some } (\mathbf{x},\mathbf{y}) \in Z\right\}\right)_{Z\in\mathbb{I}(X\times Y)}\end{aligned}$$

is at least $\min\{\gamma^{\mathrm{cv}}_f, \gamma\}$-order convergent at $(\mathbf{x}^{\mathrm{f}},\mathbf{y}^{\mathrm{f}})$.

Proof Suppose $Z \in \mathbb{I}(X\times Y)$ such that $(\mathbf{x}^{\mathrm{f}},\mathbf{y}^{\mathrm{f}}) \in Z$ and $w(Z) \le \delta$. From the proof of Theorem 1, we have

$$\begin{aligned}&\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})\\ &\le \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y})\right) + \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})\right)\\ &\le \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y})\right) + \tau^{\mathrm{cv}}_f w(Z)^{\gamma^{\mathrm{cv}}_f}.\end{aligned} \tag{4}$$

Consider $(\mathbf{x}^*_Z,\mathbf{y}^*_Z) \in \underset{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)}{\arg\min} f(\mathbf{x},\mathbf{y})$. Choose $(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z) \in \mathscr{F}(Z)$ and $(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z) \in \underset{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)}{\arg\min} f(\mathbf{x},\mathbf{y})$ such that $d(\{(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z)\},\{(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)\}) \le \hat{\tau} w(Z)^{\gamma}$ (note that $(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z)$ and $(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)$ exist by assumption). The first term in Equation (4) can be bounded from above as

$$\begin{aligned}\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y}) &= f(\mathbf{x}^*_Z,\mathbf{y}^*_Z) - f(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)\\ &\le f(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z) - f(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)\\ &\le M_f\, d(\{(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z)\},\{(\mathbf{x}^{\mathrm{cv}}_Z,\mathbf{y}^{\mathrm{cv}}_Z)\})\\ &\le M_f \hat{\tau} w(Z)^{\gamma},\end{aligned}$$

where Step 3 above follows from the Lipschitz continuity of f. Therefore, from Equation (4),

$$\begin{aligned}&\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})\\ &\le \left(M_f \hat{\tau} w(X\times Y)^{\gamma-\min\{\gamma^{\mathrm{cv}}_f,\gamma\}} + \tau^{\mathrm{cv}}_f w(X\times Y)^{\gamma^{\mathrm{cv}}_f-\min\{\gamma^{\mathrm{cv}}_f,\gamma\}}\right) w(Z)^{\min\{\gamma^{\mathrm{cv}}_f,\gamma\}}.\end{aligned}$$

The desired result follows by analogy to Lemma 5 by noting that the lower bounding scheme $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ has convergence of order at least one at $(\mathbf{x}^{\mathrm{f}},\mathbf{y}^{\mathrm{f}})$ from Theorem 1. □

The key assumption of Theorem 2, Equation (3), is rather unwieldy since verifying it involves the solution of the optimization problem $\underset{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)}{\min} f(\mathbf{x},\mathbf{y})$ for each $Z \in \mathbb{I}(X\times Y)$ with $(\mathbf{x}^{\mathrm{f}},\mathbf{y}^{\mathrm{f}}) \in Z$ and $w(Z) \le \delta$. The following more restrictive (but potentially more easily verifiable) condition implies Equation (3):

$$\exists \delta > 0, \hat{\tau} \ge 0, \gamma \ge 2 :\ d_H(\mathscr{F}(Z),\mathscr{F}^{\mathrm{cv}}(Z)) \le \hat{\tau} w(Z)^{\gamma}, \quad \forall Z \in \mathbb{I}(X\times Y) \text{ with } (\mathbf{x}^{\mathrm{f}},\mathbf{y}^{\mathrm{f}}) \in Z \text{ and } w(Z) \le \delta.$$

The following example shows that the convergence order may be as low as two under the assumptions of Theorem 2.

Example 6 Let $X=[-1,1], Y=[-1,1]$, $m_I=1$, and $m_E=0$ with $f(x,y)=-xy$ and $g(x,y)=x+y-1$. For any $[x^{\mathrm{L}},x^{\mathrm{U}}]\times[y^{\mathrm{L}},y^{\mathrm{U}}]=:Z\in\mathbb{I}(X\times Y)$, let

$$f_Z^{\mathrm{cv}}(x,y)=\max\{-x^{\mathrm{U}}y+(-x)y^{\mathrm{L}}-(-x^{\mathrm{U}})y^{\mathrm{L}},-x^{\mathrm{L}}y+(-x)y^{\mathrm{U}}-(-x^{\mathrm{L}})y^{\mathrm{U}}\},\quad g_Z^{\mathrm{cv}}(x,y)=x+y-1.$$

The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, which corresponds to the scheme of convex envelopes of f on $X\times Y$, has (at least) second-order pointwise convergence on $X\times Y$ (see [5, Theorem 10]) and the scheme $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$. Note that $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high convergence order on $X\times Y$.

Let $x^{\mathrm{L}}=y^{\mathrm{L}}=0.5-\varepsilon$, $x^{\mathrm{U}}=y^{\mathrm{U}}=0.5+\varepsilon$ with $0<\varepsilon\leq 0.5$. The width of Z is $w(Z)=2\varepsilon$. The optimal objective value of Problem (P) on Z is -0.25, while $f_Z^{\mathrm{cv}}(0.5,0.5)=-0.25-\varepsilon^2$ and $g_Z^{\mathrm{cv}}(0.5,0.5)=0$. Convergence at the point $(0.5,0.5)$ is, therefore, at most second-order.

Note that the use of feasibility-based bounds tightening techniques is ineffective in boosting the convergence order for the above example. This is in contrast to the similar Example 16 where the use of constraint propagation techniques improves the convergence order of reduced-space branch-and-bound algorithms (see Examples 17 and 18 in Section 5.2).

Remark 8 Theorem 2 requires, at a minimum, second-order pointwise convergence of the scheme of convex relaxations $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, which cannot be achieved in general by relaxations constructed purely using interval arithmetic [23]. Consequently, lower bounding schemes constructed using interval arithmetic can, in general, only be expected to possess first-order convergence (see Theorem 1). When the functions f, $\mathbf{g}$, and $\mathbf{h}$ are twice continuously differentiable, references [27] and [35] imply that polyhedral outer-approximation schemes of second-order pointwise convergent schemes of relaxations, that are employed by most state-of-the-art software for nonconvex problems (P) [2, 21, 36], also produce second-order pointwise convergent schemes of relaxations.

The following corollary of Theorem 2 shows that second-order convergence is guaranteed at points $(\mathbf{x},\mathbf{y})\in X\times Y$ such that $\mathbf{g}(\mathbf{x},\mathbf{y})<\mathbf{0}$, assuming Problem (P) contains no equality constraints (note the weaker assumption on the pointwise convergence order of the scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, and the slight abuse of notation in the description of $\mathscr{I}_C(Z)$ where we simply discard the components corresponding to $\mathbf{h}$ since $m_E=0$). A consequence of the corollary is that second-order convergence to unconstrained minima is guaranteed.

Corollary 2 *Consider Problem (P) with $m_E=0$. Suppose f is Lipschitz continuous on $X\times Y$. Let $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ denote a continuous scheme of convex relaxations of f in $X\times Y$ with pointwise convergence order $\gamma_f^{\mathrm{cv}}\geq 1$, and convergence order $\beta_f^{\mathrm{cv}}\geq 2$ with corresponding constant τ_f^{cv}. Furthermore, let $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $j=1,\cdots,m_I$, denote continuous schemes of convex relaxations of $g_1,\cdots,g_{m_I}$, respectively, in $X\times Y$ with pointwise convergence orders $\gamma_{g,1}^{\mathrm{cv}}>0,\cdots,\gamma_{g,m_I}^{\mathrm{cv}}>0$ and corresponding constants $\tau_{g,1}^{\mathrm{cv}},\cdots,\tau_{g,m_I}^{\mathrm{cv}}$.*

Suppose $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\in X\times Y$ is such that $\mathbf{g}(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})<\mathbf{0}$ (i.e. $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$ is a Slater point). Then, the scheme of lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$(\mathscr{O}(Z))_{Z\in\mathbb{I}(X\times Y)}:=\left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}(X\times Y)},$$
$$(\mathscr{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)}:=(\overline{\mathbf{g}}_Z^{\mathrm{cv}}(Z))_{Z\in\mathbb{I}(X\times Y)}$$

is at least β_f^{cv}-order convergent at $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$.

Proof Since we are interested in the convergence order at the feasible point $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$, it suffices to show that the assumptions of Theorem 2 hold.

Let $g_j(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})=-\varepsilon_j<0$, $j=1,\cdots,m_I$. Since g_j is continuous for each $j\in\{1,\cdots,m_I\}$ by virtue of Assumption 1, there exists $\delta_j>0$, $\forall j\in\{1,\cdots,m_I\}$, such that $\|(\mathbf{x},\mathbf{y})-(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\|_\infty<\delta_j$ implies $|g_j(\mathbf{x},\mathbf{y})-g_j(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})|<\frac{\varepsilon_j}{2}$ (see Lemma 2).

Define $\delta := \min_{j\in\{1,\cdots,m_I\}} \delta_j$, and note that $\delta > 0$. Consider $Z \in \mathbb{I}(X\times Y)$ such that $(\mathbf{x}^S,\mathbf{y}^S) \in Z$ and $w(Z) \leq \delta$. For each $(\mathbf{x},\mathbf{y}) \in Z$, $j \in \{1,\cdots,m_I\}$ we have $|g_j(\mathbf{x},\mathbf{y}) - g_j(\mathbf{x}^S,\mathbf{y}^S)| < \frac{\varepsilon_j}{2}$. Consequently, for each $j \in \{1,\cdots,m_I\}$, $g_j(\mathbf{x},\mathbf{y}) < -\frac{\varepsilon_j}{2}$, $\forall (\mathbf{x},\mathbf{y}) \in Z$. Since $\mathbf{g}_Z^{\text{cv}}(\mathbf{x},\mathbf{y}) \leq \mathbf{g}(\mathbf{x},\mathbf{y}) < -\frac{\varepsilon_j}{2}$, $\forall (\mathbf{x},\mathbf{y}) \in Z$, we have $\mathbf{g}_Z^{\text{cv}}(\mathbf{x},\mathbf{y}) < -\frac{\varepsilon_j}{2}$, $\forall (\mathbf{x},\mathbf{y}) \in Z$, i.e. every point in Z is feasible for Problem (P) and the lower bounding problem $\mathscr{L}(Z)$.

Therefore, $\delta := \min_{j\in\{1,\cdots,m_I\}} \delta_j$, any $(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z) \in \operatorname{arg\,min}_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\text{cv}}(Z)} f(\mathbf{x},\mathbf{y})$, $\gamma = \beta_f^{\text{cv}} + 1$, and $\hat{\tau} = 0$ satisfies the (necessary) assumptions of Theorem 2 which yield an upper bound on the first term in Equation (4). The second term in Equation (4) can be bounded from above as

$$\begin{aligned}\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\text{cv}}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\text{cv}}(Z)} f_Z^{\text{cv}}(\mathbf{x},\mathbf{y}) &= \min_{(\mathbf{x},\mathbf{y})\in Z} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in Z} f_Z^{\text{cv}}(\mathbf{x},\mathbf{y}) \\ &\leq \tau_f^{\text{cv}} w(Z)^{\beta_f^{\text{cv}}}\end{aligned}$$

since f_Z^{cv} converges with order β_f^{cv} to f on $X\times Y$, and $\mathscr{F}^{\text{cv}}(Z) = Z$. Substituting the above bounds in Equation (4), we obtain

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\text{cv}}(Z)} f_Z^{\text{cv}}(\mathbf{x},\mathbf{y}) \leq \tau_f^{\text{cv}} w(Z)^{\beta_f^{\text{cv}}}.$$

The desired result follows by analogy to Lemma 5 by noting that the lower bounding scheme $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ is at least first-order convergent at $(\mathbf{x}^S,\mathbf{y}^S)$ from Theorem 1. □

Note that the bound on the prefactor obtained from Corollary 2 for convergence at points where a constraint is 'nearly active' can be relatively large (also see the comment after Lemma 5).

Remark 9 Corollary 2 does not apply to Problem (P) with active constraints; however, Theorem 2 can be used to demonstrate second-order convergence when Problem (P) contains active convex constraints (note that this includes affine equality constraints) if the schemes of relaxations of the active constraints are the (convex) functions themselves and the scheme of convex relaxations of the objective function is second-order pointwise convergent. Examples 8 and 9 illustrate cases where the above modification of Corollary 2 does not apply when the schemes of relaxations of active convex constraints are not the constraints themselves (note that if the schemes of relaxations of active convex constraints used are the constraints themselves, then the convergence orders of the lower bounding schemes in these examples would be arbitrarily high at their minimizers), thereby highlighting the importance of convexity detection in boosting the convergence order.

The following example shows that the convergence order of the lower bounding scheme is dictated by the convergence order, β_f^{cv}, of the scheme $(f_Z^{\text{cv}})_{Z\in\mathbb{I}(X\times Y)}$ when the assumptions of Corollary 2 are satisfied.

Example 7 Let $X = [0,0], Y = [0,1]$, $m_I = 1$, and $m_E = 0$ with $f(x,y) = y^4 - y^2$ and $g(x,y) = 1 - 2y$. For any $[0,0]\times[y^L,y^U] =: Z \in \mathbb{I}(X\times Y)$, let $f_Z^{\text{cv}}(x,y) = y^4 - (y^L + y^U)y + y^L y^U$, $g_Z^{\text{cv}}(x,y) = 1 - 2y$. The scheme $(f_Z^{\text{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has second-order pointwise convergence and second-order convergence on $X\times Y$, while the scheme $(g_Z^{\text{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$.

Let $y^L = \frac{1}{\sqrt{2}} - \varepsilon, y^U = \frac{1}{\sqrt{2}} + \varepsilon$ with $0 < \varepsilon \leq 0.25$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is -0.25, while the optimal objective of the lower bounding problem on Z is $-0.25 - \varepsilon^2$. Convergence at the point $\left(0, \frac{1}{\sqrt{2}}\right)$ is, therefore, at most second-order.

Example 8 Let $X = [-3,3], Y = [-3,3]$, $m_I = 1$, and $m_E = 0$ with $f(x,y) = x + y$ and $g(x,y) = x^2 + y^2 - 8$. For any $[x^L,x^U]\times[y^L,y^U] =: Z \in \mathbb{I}(X\times Y)$, let $f_Z^{\text{cv}}(x,y) = x + y$, $g_Z^{\text{cv}}(x,y) = x^2 + y^2 - 8 - (w(Z))^2$. The scheme $(f_Z^{\text{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$, while the scheme $(g_Z^{\text{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has second-order pointwise convergence on $X\times Y$.

Let $x^L = y^L = -2 - \varepsilon$, $x^U = y^U = -2 + \varepsilon$ with $0 < \varepsilon \leq 1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is -4, while the optimal objective of the lower bounding problem on Z is $-\sqrt{16 + 8\varepsilon^2} = -4 - \varepsilon^2 + O(\varepsilon^4)$ for $\varepsilon \ll 1$. Convergence at the point $(-2,-2)$ is, therefore, at most second-order.

Example 9 Let $X=[0,1], Y=[0,1]$, $m_I=1$, and $m_E=0$ with $f(x,y)=-x-y$ and $g(x,y)=x^2+2xy+y^2-1$. For any $[x^{\mathrm{L}},x^{\mathrm{U}}]\times[y^{\mathrm{L}},y^{\mathrm{U}}]=:Z\in\mathbb{I}(X\times Y)$, let

$$f_Z^{\mathrm{cv}}(x,y)=-x-y,\quad g_Z^{\mathrm{cv}}(x,y)=x^2+2\max\left\{x^{\mathrm{L}}y+y^{\mathrm{L}}x-x^{\mathrm{L}}y^{\mathrm{L}},x^{\mathrm{U}}y+y^{\mathrm{U}}x-x^{\mathrm{U}}y^{\mathrm{U}}\right\}+y^2-1.$$

The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$, while the scheme $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has second-order pointwise convergence on $X\times Y$.

Let $x^{\mathrm{L}}=y^{\mathrm{L}}=0.5-\varepsilon$, $x^{\mathrm{U}}=y^{\mathrm{U}}=0.5+\varepsilon$ with $0<\varepsilon\leq 0.5$. The width of Z is $w(Z)=2\varepsilon$. The optimal objective value of Problem (P) on Z is -1, while the point $\left(0.5-0.25\varepsilon^2,0.5+0.5\varepsilon^2\right)$ is feasible for the lower bounding problem on Z with objective value $-1-0.25\varepsilon^2$. Convergence at the point $(0.5,0.5)$ is, therefore, at most second-order.

The next result provides a slight generalization of Corollary 2 by showing that under the assumptions of Corollary 2, the lower bounding scheme $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ in fact exhibits (at least) second-order convergence on a neighborhood of $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$ (this result is motivated by the assumptions on the convergence order of the lower bounding scheme in the analysis of the cluster problem in [15]).

Corollary 3 *Consider Problem (P) with $m_E=0$. Suppose f is Lipschitz continuous on $X\times Y$. Let $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ denote a continuous scheme of convex relaxations of f in $X\times Y$ with pointwise convergence order $\gamma_f^{\mathrm{cv}}\geq 1$, and convergence order $\beta_f^{\mathrm{cv}}\geq 1$ with corresponding constant τ_f^{cv}. Furthermore, let $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $j=1,\cdots,m_I$, denote continuous schemes of convex relaxations of $g_1,\cdots,g_{m_I}$, respectively, in $X\times Y$ with pointwise convergence orders $\gamma_{g,1}^{\mathrm{cv}}>0,\cdots,\gamma_{g,m_I}^{\mathrm{cv}}>0$ and corresponding constants $\tau_{g,1}^{\mathrm{cv}},\cdots,\tau_{g,m_I}^{\mathrm{cv}}$.*

Suppose $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\in X\times Y$ such that $\mathbf{g}(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})<\mathbf{0}$ (i.e. $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$ is a Slater point). Then, $\exists\delta>0$ such that the scheme of lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$(\mathscr{O}(Z))_{Z\in\mathbb{I}(X\times Y)}:=\left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)}f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}(X\times Y)},$$
$$(\mathscr{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)}:=(\overline{\mathbf{g}}_Z^{\mathrm{cv}}(Z))_{Z\in\mathbb{I}(X\times Y)}$$

is at least β_f^{cv}-order convergent on $\left\{(\mathbf{x},\mathbf{y}):\|(\mathbf{x},\mathbf{y})-(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\|_\infty<\delta\right\}$.

Proof Let $g_j(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})=-\varepsilon_j<0$, $j=1,\cdots,m_I$. Since g_j is continuous for each $j\in\{1,\cdots,m_I\}$, there exists $\delta_j>0$, $\forall j\in\{1,\cdots,m_I\}$, such that $\|(\mathbf{x},\mathbf{y})-(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\|_\infty\leq\delta_j$ implies $|g_j(\mathbf{x},\mathbf{y})-g_j(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})|<\varepsilon_j$ (see Lemma 2). Define $\bar{\delta}:=\min_{j\in\{1,\cdots,m_I\}}\delta_j$, note that $\bar{\delta}>0$, and let $\delta:=\frac{\bar{\delta}}{2}$.

Consider $Z\in\mathbb{I}(X\times Y)$ with $Z\cap\left\{(\mathbf{x},\mathbf{y}):\|(\mathbf{x},\mathbf{y})-(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\|_\infty<\delta\right\}\neq\emptyset$ and $w(Z)\leq\delta$. Similar to the proof of Corollary 2, it can be shown that

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)}f(\mathbf{x},\mathbf{y})\ -\ \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)}f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\leq\tau_f^{\mathrm{cv}}w(Z)^{\beta_f^{\mathrm{cv}}}.$$

The desired result follows by analogy to Lemma 5 by noting that the lower bounding scheme $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ has at least first-order convergence on $\left\{(\mathbf{x},\mathbf{y}):\|(\mathbf{x},\mathbf{y})-(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\|_\infty<\delta\right\}$ from Theorem 1. □

While it may appear that the neighborhood of a Slater point on which second-order convergence of the lower bounding scheme is guaranteed by Corollary 3 can be unnecessarily small, Example 10 shows that a stronger result cannot be deduced without additional assumptions.

A natural question is whether second-order convergence is guaranteed on $X\times Y$ when second-order pointwise convergent schemes of (convex) relaxations of $f, g_1,\cdots,g_{m_I}, h_1,\cdots,h_{m_E}$ are used by the lower bounding scheme. The following example shows that even when schemes of (convex) envelopes are used to underestimate smooth functions f, $\mathbf{g}$, and $\mathbf{h}$, at most first-order convergence can be guaranteed at certain points in $X\times Y$.

Example 10 Let $X = [0,0], Y = [-1,1]$, $m_I = 1$, and $m_E = 0$ with $f(x,y) = -y$ and $g(x,y) = y^3$. For any $[0,0] \times [-\varepsilon,\varepsilon] =: Z \in \mathbb{I}(X \times Y)$ with $\varepsilon > 0$, let

$$f_Z^{\text{cv}}(x,y) = -y, \quad g_Z^{\text{cv}}(x,y) = \begin{cases} -0.25\varepsilon^3 + 0.75\varepsilon^2 y, & \text{if } y < 0.5\varepsilon \\ y^3, & \text{if } y \geq 0.5\varepsilon \end{cases}.$$

Note that the scheme $(f_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$ and the scheme $(g_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$, which is the scheme of convex envelopes of g on $X \times Y$ [19], has (at least) second-order pointwise convergence on $X \times Y$. Also note that $(g_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high convergence order on $X \times Y$.

The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-\frac{\varepsilon}{3}$. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Despite the fact that we only have first-order convergence at the global minimizer in Example 10, the reader can verify that the natural interval extension-based lower bounding scheme along with the interval bisection branching rule and lowest lower bound node selection rule is sufficient to mitigate the cluster problem for this case [15].

The following result establishes second-order convergence of a convex relaxation-based lower bounding scheme at a feasible point $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}}) \in X \times Y$ when second-order pointwise convergent schemes of relaxations are used and the dual lower bounding scheme (see Section 4.2) is second-order convergent at $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$. This result will be used to prove second-order convergence of such convex relaxation-based lower bounding schemes at KKT points in Corollary 4.

Theorem 3 *Consider Problem* (P)*, and let* $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}}) \in X \times Y$ *be a feasible point. Suppose the dual lower bounding scheme has convergence of order* $\beta_d > 0$ *at* $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$ *with a corresponding scheme of bounded dual variables* $\left(\left(\boldsymbol{\mu}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}, \boldsymbol{\lambda}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}\right)\right)_{Z \in \mathbb{I}(X \times Y)}$ *(not necessarily optimal, but which yield* β_d*-order convergence at* $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$*) with* $\left(\boldsymbol{\mu}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}, \boldsymbol{\lambda}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}\right) \in \mathbb{R}_+^{m_I} \times \mathbb{R}^{m_E}$, $\left\|\boldsymbol{\mu}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}\right\|_\infty \leq \bar{\mu}$ *and* $\left\|\boldsymbol{\lambda}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}\right\|_\infty \leq \bar{\lambda}$, $\forall Z \in \mathbb{I}(X \times Y)$, *for some constants* $\bar{\mu}, \bar{\lambda} \geq 0$ *(see Section 4.2). Let* $(f_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$, $(g_{j,Z}^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, *denote continuous schemes of convex relaxations of* f, $g_1, \cdots, g_{m_I}$, *respectively, in* $X \times Y$ *with pointwise convergence orders* $\gamma_f^{\text{cv}} \geq 1$, $\gamma_{g,1}^{\text{cv}} \geq 1, \cdots, \gamma_{g,m_I}^{\text{cv}} \geq 1$ *and corresponding constants* τ_f^{cv}, $\tau_{g,1}^{\text{cv}}, \cdots, \tau_{g,m_I}^{\text{cv}}$. *Let* $(h_{k,Z}^{\text{cv}}, h_{k,Z}^{\text{cc}})_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, *denote continuous schemes of relaxations of* $h_1, \cdots, h_{m_E}$, *respectively, in* $X \times Y$ *with pointwise convergence orders* $\gamma_{h,1} \geq 1, \cdots, \gamma_{h,m_E} \geq 1$ *and corresponding constants* $\tau_{h,1}, \cdots, \tau_{h,m_E}$. *Then, the scheme of lower bounding problems* $(\mathscr{L}(Z))_{Z \in \mathbb{I}(X \times Y)}$ *with*

$$(\mathscr{O}(Z))_{Z \in \mathbb{I}(X \times Y)} := \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}^{\text{cv}}(Z)} f_Z^{\text{cv}}(\mathbf{x},\mathbf{y})\right)_{Z \in \mathbb{I}(X \times Y)},$$

$$(\mathscr{I}_C(Z))_{Z \in \mathbb{I}(X \times Y)} := \left(\left\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}_Z^{\text{cv}}(\mathbf{x},\mathbf{y}), \mathbf{h}_Z^{\text{cv}}(\mathbf{x},\mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}_Z^{\text{cc}}(\mathbf{x},\mathbf{y}) \right.\right.$$
$$\left.\left. \text{for some } (\mathbf{x},\mathbf{y}) \in Z\right\}\right)_{Z \in \mathbb{I}(X \times Y)}$$

is at least $\min\left\{\min\left\{\gamma_f^{\text{cv}}, \min_{j \in \{1,\cdots,m_I\}} \gamma_{g,j}^{\text{cv}}, \min_{k \in \{1,\cdots,m_E\}} \gamma_{h,k}\right\}, \beta_d\right\}$*-order convergent at* $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$.

Proof Let $\beta := \min\left\{\min\left\{\gamma_f^{\text{cv}}, \min_{j \in \{1,\cdots,m_I\}} \gamma_{g,j}^{\text{cv}}, \min_{k \in \{1,\cdots,m_E\}} \gamma_{h,k}\right\}, \beta_d\right\}$, and let $\boldsymbol{\mu}_Z := \boldsymbol{\mu}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}$, $\boldsymbol{\lambda}_Z := \boldsymbol{\lambda}_Z^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}$, $\forall Z \in \mathbb{I}(X \times Y)$, denote the scheme of dual variables corresponding to the dual lower bounding scheme (we omit the dependence of the dual variables on $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$ for ease of exposition). Since we are concerned about the convergence order at the feasible point $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$, it suffices to show the existence of $\tau \geq 0$ such that for every $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}}) \in Z$,

$$\min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}^{\text{cv}}(Z)} f_Z^{\text{cv}}(\mathbf{x},\mathbf{y}) \leq \tau w(Z)^\beta.$$

Consider $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}}) \in Z$. By virtue of the assumption that the dual lower bounding scheme, with the dual variables fixed to $((\boldsymbol{\mu}_Z, \boldsymbol{\lambda}_Z))_{Z \in \mathbb{I}(X \times Y)}$, has convergence of order β_d at $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$, we have

$$\min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_Z^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] \leq \tau_d w(Z)^{\beta_d}. \tag{5}$$

Choose $\boldsymbol{\lambda}_{Z,+}, \boldsymbol{\lambda}_{Z,-} \in \mathbb{R}_+^{m_E}$ such that $\boldsymbol{\lambda}_Z = \boldsymbol{\lambda}_{Z,+} - \boldsymbol{\lambda}_{Z,-}$, $\|\boldsymbol{\lambda}_{Z,+}\|_\infty \leq \bar{\lambda}$, and $\|\boldsymbol{\lambda}_{Z,-}\|_\infty \leq \bar{\lambda}$.

Let $\bar{\gamma} := \min \left\{ \gamma_f^{\mathrm{cv}}, \min_{j \in \{1,\cdots,m_I\}} \gamma_{g,j}^{\mathrm{cv}}, \min_{k \in \{1,\cdots,m_E\}} \gamma_{h,k} \right\}$. We have

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_Z^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \min_{\mathscr{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \\
\leq &\min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_Z^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \\
&\qquad \sup_{\boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\lambda}_1 \geq \mathbf{0}, \boldsymbol{\lambda}_2 \leq \mathbf{0}} \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}} \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}} \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) \right] \\
\leq &\min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) - \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \\
&\qquad \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) - \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) \right] \\
\leq &\max_{(\mathbf{x},\mathbf{y}) \in Z} \Big[(f(\mathbf{x},\mathbf{y}) - f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})) + \boldsymbol{\mu}_Z^{\mathrm{T}} (\mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})) + \\
&\qquad \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} (\mathbf{h}(\mathbf{x},\mathbf{y}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})) + \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} (\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) - \mathbf{h}(\mathbf{x},\mathbf{y})) \Big] \\
\leq &\max_{(\mathbf{x},\mathbf{y}) \in Z} (f(\mathbf{x},\mathbf{y}) - f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})) + \max_{(\mathbf{x},\mathbf{y}) \in Z} \boldsymbol{\mu}_Z^{\mathrm{T}} (\mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})) + \\
&\qquad \max_{(\mathbf{x},\mathbf{y}) \in Z} \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} (\mathbf{h}(\mathbf{x},\mathbf{y}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})) + \max_{(\mathbf{x},\mathbf{y}) \in Z} \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} (\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) - \mathbf{h}(\mathbf{x},\mathbf{y})) \\
\leq &\tau_f^{\mathrm{cv}} w(Z)^{\gamma_f^{\mathrm{cv}}} + \sum_{j=1}^{m_I} \bar{\mu} \tau_{g,j}^{\mathrm{cv}} w(Z)^{\gamma_{g,j}^{\mathrm{cv}}} + 2 \sum_{k=1}^{m_E} \bar{\lambda} \tau_{h,k} w(Z)^{\gamma_{h,k}} \\
\leq &\left(\tau_f^{\mathrm{cv}} w(X \times Y)^{\gamma_f^{\mathrm{cv}} - \bar{\gamma}} + \sum_{j=1}^{m_I} \bar{\mu} \tau_{g,j}^{\mathrm{cv}} w(X \times Y)^{\gamma_{g,j}^{\mathrm{cv}} - \bar{\gamma}} + 2 \sum_{k=1}^{m_E} \bar{\lambda} \tau_{h,k} w(X \times Y)^{\gamma_{h,k} - \bar{\gamma}} \right) w(Z)^{\bar{\gamma}},
\end{aligned} \tag{6}$$

where Step 1 follows from weak duality and Step 3 follows from Lemma 3.

Therefore, from Equations (5) and (6), we have

$$\begin{aligned}
\min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) = &\min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_Z^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] + \\
&\min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_Z^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \\
\leq &\tau w(Z)^{\beta},
\end{aligned}$$

where the prefactor τ is defined as

$$\tau := \left(\tau_f^{\mathrm{cv}} w(X \times Y)^{\gamma_f^{\mathrm{cv}} - \beta} + \sum_{j=1}^{m_I} \bar{\mu} \tau_{g,j}^{\mathrm{cv}} w(X \times Y)^{\gamma_{g,j}^{\mathrm{cv}} - \beta} + 2 \sum_{k=1}^{m_E} \bar{\lambda} \tau_{h,k} w(X \times Y)^{\gamma_{h,k} - \beta} + \tau_d w(X \times Y)^{\beta_d - \beta} \right).$$

□

4.2 Duality-based branch-and-bound

In this section, we investigate the convergence order of a Lagrangian dual-based lower bounding scheme. Before we define the convergence order of the scheme, the Lagrangian dual problem is introduced and some of its properties are outlined.

The dual problem of Problem (P) that is obtained by dualizing all of the constraints $\mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}$ is given by

$$\begin{aligned} &\sup_{\boldsymbol{\mu},\boldsymbol{\lambda}} q(\boldsymbol{\mu},\boldsymbol{\lambda}) \\ &\text{s.t. } \boldsymbol{\mu} \in \mathbb{R}^{m_I}_+, \boldsymbol{\lambda} \in \mathbb{R}^{m_E}, \end{aligned} \tag{D}$$

where $q : \mathbb{R}^{m_I}_+ \times \mathbb{R}^{m_E} \to \mathbb{R}$, defined by

$$q(\boldsymbol{\mu},\boldsymbol{\lambda}) := \min_{(\mathbf{x},\mathbf{y}) \in X \times Y} f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}), \quad \forall (\boldsymbol{\mu},\boldsymbol{\lambda}) \in \mathbb{R}^{m_I}_+ \times \mathbb{R}^{m_E},$$

is the Lagrangian dual function. Let $\min(\text{P})$ and $\sup(\text{D})$ respectively denote the optimal objective values of Problem (P) and Problem (D). From weak duality, we know that $\min(\text{P}) \geq \sup(\text{D})$, which validates the use of Problem (D) as a lower bounding problem.

The following result shows that the lower bounds obtained by solving the Lagrangian dual Problem (D) are stronger than those obtained by solving any convex relaxation-based lower bounding problem.

Lemma 9 *Consider Problem (P), and suppose $Z \in \mathbb{I}(X \times Y)$. Let f^{cv}_Z and $\mathbf{g}^{\mathrm{cv}}_Z$ denote convex relaxations of f and $\mathbf{g}$, respectively, on Z, and let $\mathbf{h}^{\mathrm{cv}}_Z$ and $\mathbf{h}^{\mathrm{cc}}_Z$ denote convex and concave relaxations, respectively, of $\mathbf{h}$ on Z. Furthermore, assume that strong duality holds for the convex relaxation-based lower bounding problem $\min_{\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})$. Then the lower bound obtained by solving the Lagrangian dual problem is at least as strong as that obtained by solving the above convex relaxation-based lower bounding problem, i.e.,*

$$\sup_{\boldsymbol{\mu} \geq \mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \min_{\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) \geq 0.$$

Proof Since strong duality holds for the convex relaxation-based lower bounding problem, the difference between the lower bounds can be rewritten as

$$\begin{aligned} &\sup_{\boldsymbol{\mu} \geq \mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \\ &\qquad \sup_{\boldsymbol{\mu} \geq \mathbf{0},\boldsymbol{\lambda}_1 \geq \mathbf{0},\boldsymbol{\lambda}_2 \leq \mathbf{0}} \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y}) \right] \\ = &\sup_{\boldsymbol{\mu} \geq \mathbf{0},\boldsymbol{\lambda}_1 \geq \mathbf{0},\boldsymbol{\lambda}_2 \leq \mathbf{0}} \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \\ &\qquad \sup_{\boldsymbol{\mu} \geq \mathbf{0},\boldsymbol{\lambda}_1 \geq \mathbf{0},\boldsymbol{\lambda}_2 \leq \mathbf{0}} \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y}) \right] \\ \geq &\ 0, \end{aligned}$$

where the last step follows from the fact that $\forall (\mathbf{x},\mathbf{y}) \in Z, \boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\lambda}_1 \geq \mathbf{0}, \boldsymbol{\lambda}_2 \leq \mathbf{0}$,

$$f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) \geq f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y}).$$

□

The following result due to Dür [8] establishes the condition under which the dual lower bounding problem detects infeasibility.

Lemma 10 *Consider Problem* (P) *(satisfying Assumption 1). We have*

$$\sup(\mathrm{D}) = +\infty \iff \operatorname{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X\times Y)\right)\cap\left(\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)=\emptyset.$$

Proof The result follows, in part, by replacing $\mathbf{h}(\mathbf{x},\mathbf{y})=\mathbf{0}$ with $\mathbf{h}(\mathbf{x},\mathbf{y})\leq\mathbf{0}$ and $-\mathbf{h}(\mathbf{x},\mathbf{y})\leq\mathbf{0}$ and using Theorem 2 in [8]. □

Definition 14 can be applied to analyze the convergence order of the above duality-based lower bounding scheme as follows.

The scheme of dual lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$(\mathscr{O}(Z))_{Z\in\mathbb{I}(X\times Y)} := \left(\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}}\min_{(\mathbf{x},\mathbf{y})\in Z}\left[f(\mathbf{x},\mathbf{y})+\boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y})+\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})\right]\right)_{Z\in\mathbb{I}(X\times Y)},$$

$$(\mathscr{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)} := \left(\operatorname{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z)\right)\right)_{Z\in\mathbb{I}(X\times Y)}$$

is thus said to have convergence of order $\beta>0$ at

1. a feasible point $(\mathbf{x},\mathbf{y})\in X\times Y$ if there exists $\tau\geq 0$ such that for every $Z\in\mathbb{I}(X\times Y)$ with $(\mathbf{x},\mathbf{y})\in Z$,
$$\min_{(\mathbf{v},\mathbf{w})\in\mathscr{F}(Z)} f(\mathbf{v},\mathbf{w}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}}\min_{(\mathbf{v},\mathbf{w})\in Z}\left[f(\mathbf{v},\mathbf{w})+\boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{v},\mathbf{w})+\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{v},\mathbf{w})\right]\leq\tau w(Z)^{\beta}.$$
2. an infeasible point $(\mathbf{x},\mathbf{y})\in X\times Y$ if there exists $\bar{\tau}\geq 0$ such that for every $Z\in\mathbb{I}(X\times Y)$ with $(\mathbf{x},\mathbf{y})\in Z$,
$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right) - d\left(\operatorname{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z)\right),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)\leq\bar{\tau}w(Z)^{\beta}.$$

We associate with the dual lower bounding scheme, $(\mathscr{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$, at a feasible point $(\mathbf{x},\mathbf{y})$, a scheme of dual variables $((\boldsymbol{\mu}_Z^{(\mathbf{x},\mathbf{y})},\boldsymbol{\lambda}_Z^{(\mathbf{x},\mathbf{y})}))_{Z\in\mathbb{I}(X\times Y)}$ corresponding to the solution of the scheme of dual lower bounding problems $(\mathscr{O}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with $(\mathbf{x},\mathbf{y})\in Z$ (note that $\sup(\mathrm{D})$ may not be attained, in which case we assume that dual variables that yield a dual function value arbitrarily close to the supremum are available). Using Lemma 10, we next show that if the convex relaxation-based lower bounding problem corresponding to Problem (P) that is obtained by replacing the functions in Problem (P) with their envelopes is infeasible, then $\sup(\mathrm{D})=+\infty$.

Lemma 11 *Let $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $j=1,\cdots,m_I$, denote (any) schemes of convex relaxations of $g_1,\cdots,g_{m_I}$ in $X\times Y$ and $(h_{k,Z}^{\mathrm{cv}},h_{k,Z}^{\mathrm{cc}})_{Z\in\mathbb{I}(X\times Y)}$, $k=1,\cdots,m_E$, denote (any) schemes of relaxations of $h_1,\cdots,h_{m_E}$ in $X\times Y$. Then for each $Z\in\mathbb{I}(X\times Y)$, we have*

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)\geq d\left(\operatorname{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z)\right),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)\geq d\left(\mathscr{I}_C(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right),$$

where $\mathscr{I}_C(Z)$ is defined as

$$\mathscr{I}_C(Z) := \{(\mathbf{v},\mathbf{w})\in\mathbb{R}^{m_I}\times\mathbb{R}^{m_E} : \mathbf{v}=\mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}),\mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\leq\mathbf{w}\leq\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) \text{ for some } (\mathbf{x},\mathbf{y})\in Z\}.$$

Proof The first inequality trivially holds. To prove the second inequality, we first notice that

$$d\left(\mathscr{I}_C(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right)=d\left(\bar{\mathscr{I}}_C(Z),\mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right),$$

where $\bar{\mathscr{I}}_C(Z)$ is defined as

$$\bar{\mathscr{I}}_C(Z) := \{(\mathbf{v},\mathbf{w})\in\mathbb{R}^{m_I}\times\mathbb{R}^{m_E} : \mathbf{v}\geq\mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}),\mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\leq\mathbf{w}\leq\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) \text{ for some } (\mathbf{x},\mathbf{y})\in Z\}.$$

Note that $\bar{\mathscr{I}}_C(Z)$ is a convex set since it can be represented as the direct sum of two convex sets.

Since $\operatorname{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z)\right)$ is the smallest convex set that encloses $\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z)$, the desired result follows. □

Theorem 4 *Consider Problem* (P). *Suppose strong duality holds for the scheme of convex relaxation-based lower bounding problems for Problem* (P) *obtained by using the schemes of (convex) envelopes of* f, $\mathbf{g}$, *and* $\mathbf{h}$. *If the assumptions of Theorem 1 hold for the functions* f, $\mathbf{g}$, $\mathbf{h}$, *and the schemes of (convex) envelopes of* f, $\mathbf{g}$, *and* $\mathbf{h}$, *the dual lower bounding scheme is (at least) first-order convergent on* $X \times Y$. *Furthermore, if the assumptions of Theorem 2 hold for the schemes of (convex) envelopes of* f, $\mathbf{g}$, *and* $\mathbf{h}$ *and* $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}}) \in X \times Y$, *the dual lower bounding scheme is (at least) second-order convergent at* $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$.

Proof From Lemma 11, we have that the convergence order of the dual lower bounding scheme at an infeasible point $(\mathbf{x}, \mathbf{y}) \in X \times Y$ is at least as high as the convergence order at $(\mathbf{x}, \mathbf{y})$ of the convex relaxation-based lower bounding scheme obtained by using the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$. Lemma 9 implies that the lower bounds obtained using the dual lower bounding scheme are at least as tight as the lower bounds obtained using the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$. The desired result follows from Definition 14. □

Note that the conclusions of Theorem 4 hold even if the schemes of relaxations of f, $\mathbf{g}$, and $\mathbf{h}$ do not correspond to their envelopes, so long as the (remaining) assumptions of Theorem 4 are satisfied.

Remark 10 The assumption of strong duality is in fact not required to show first-order convergence of the dual lower bounding scheme when all functions in Problem (P) are Lipschitz continuous. For this case, the proof of first-order convergence at infeasible points follows from Lemmata 7, 8, and 11, and the proof of first-order convergence at feasible points follows from Proposition 1 in [8].

Theorem 4 makes no assumptions on the boundedness of schemes of dual variables. This is reflected in the application of the dual lower bounding scheme to Example 10 where the optimal scheme of dual variables can be unbounded (note, however, that first-order convergence of the dual lower bounding scheme at the global minimizer of Example 10 can be achieved using bounded schemes of dual variables when the dual problem is not solved to optimality). Furthermore, Example 6 shows that the convergence order of the dual lower bounding scheme can be as low as two at $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$ when the assumptions of Theorem 2 are satisfied for the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$ (see Lemma 14). The following result shows that in the absence of equality constraints, the dual lower bounding scheme has arbitrarily high convergence order at unconstrained points.

Proposition 1 *Consider Problem* (P) *with* $m_E = 0$. *Suppose* f *and* g_j, $\forall j \in \{1, \cdots, m_I\}$, *are Lipschitz continuous on* $X \times Y$. *Furthermore, suppose* $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) \in X \times Y$ *such that* $\mathbf{g}(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) < \mathbf{0}$ *(i.e.* $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$ *is a Slater point). The dual lower bounding scheme has arbitrarily high convergence order at* $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$.

Proof The proof is relegated to Appendix A.1 since it is similar to the proof of Corollary 2. □

Remark 11 Proposition 1 as stated does not apply to Problem (P) with active constraints; however, it can be modified to demonstrate second-order convergence when Problem (P) contains active convex constraints (note that this includes affine equality constraints) if f is twice continuously differentiable, and strong duality holds for the scheme of relaxations of Problem (P) in which only the active (convex) constraints are included and f is replaced by its scheme of convex envelopes (see Remark 9). Proposition 1 can also be extended to demonstrate arbitrarily high convergence order of the dual lower bounding scheme on a neighborhood of $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$ in a manner similar to Corollary 3.

The next result shows that the dual lower bounding scheme is second-order convergent at KKT points when the functions f, $\mathbf{g}$, and $\mathbf{h}$ in Problem (P) are twice continuously differentiable.

Theorem 5 *Consider Problem* (P). *Suppose* $\mathrm{int}(X \times Y)$ *is nonempty, and* f, $\mathbf{g}$, *and* $\mathbf{h}$ *are twice continuously differentiable on* $\mathrm{int}(X \times Y)$. *Furthermore, suppose there exists* $(\mathbf{x}^*, \mathbf{y}^*) \in \mathrm{int}(X \times Y)$, $\boldsymbol{\mu}^* \in \mathbb{R}^{m_I}_+$, *and* $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ *such that* $(\mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ *is a KKT point for Problem* (P). *The dual lower bounding scheme is at least second-order convergent at* $(\mathbf{x}^*, \mathbf{y}^*)$.

Proof Let $L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) := f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})$ denote the Lagrangian of Problem (P). Since we are concerned about the convergence order at the feasible point $(\mathbf{x}^*,\mathbf{y}^*)$, it suffices to show the existence of $\tau \geq 0$ such that for every $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^*,\mathbf{y}^*) \in Z$,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \leq \tau w(Z)^2.$$

We have

$$\begin{aligned}
\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) &\geq \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) \\
&= \min_{(\mathbf{x},\mathbf{y})\in Z} \Big[L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) + \nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \\
&\qquad + \nabla_{\mathbf{y}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{y}-\mathbf{y}^*) - O(w(Z)^2) \Big] \\
&= \min_{(\mathbf{x},\mathbf{y})\in Z} \big[f(\mathbf{x}^*,\mathbf{y}^*) - O(w(Z)^2) \big] \\
&\geq f(\mathbf{x}^*,\mathbf{y}^*) - O(w(Z)^2).
\end{aligned}$$

Note that Step 3 above uses $L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = f(\mathbf{x}^*,\mathbf{y}^*)$, $\nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$, and $\nabla_{\mathbf{y}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$ by virtue of the assumption that $(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is a KKT point for Problem (P). Therefore,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \leq O(w(Z)^2),$$

which establishes the existence of τ for all $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^*,\mathbf{y}^*) \in Z$ by analogy to Lemma 5 since the dual lower bounding scheme is at least first-order convergent at $(\mathbf{x}^*,\mathbf{y}^*)$. □

A corollary of Theorems 3 and 5 is that second-order convergence at KKT points is guaranteed for convex relaxation-based lower bounding schemes in which second-order pointwise convergent schemes of relaxations are used.

Corollary 4 *Consider Problem* (P). *Suppose* $\mathrm{int}(X \times Y)$ *is nonempty and* f, $\mathbf{g}$, *and* $\mathbf{h}$ *are twice continuously differentiable on* $\mathrm{int}(X \times Y)$. *Furthermore, suppose there exists* $(\mathbf{x}^*,\mathbf{y}^*) \in \mathrm{int}(X \times Y)$, $\boldsymbol{\mu}^* \in \mathbb{R}^{m_I}_+$, *and* $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ *such that* $(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ *is a KKT point for Problem* (P). *Let* $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$, $(g^{\mathrm{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1,\cdots,m_I$, *denote continuous schemes of convex relaxations of* $f, g_1,\cdots,g_{m_I}$, *respectively, in* $X \times Y$ *with pointwise convergence orders* $\gamma^{\mathrm{cv}}_f \geq 2$, $\gamma^{\mathrm{cv}}_{g,1} \geq 2,\cdots,\gamma^{\mathrm{cv}}_{g,m_I} \geq 2$ *and corresponding constants* τ^{cv}_f, $\tau^{\mathrm{cv}}_{g,1},\cdots,\tau^{\mathrm{cv}}_{g,m_I}$. *Let* $(h^{\mathrm{cv}}_{k,Z}, h^{\mathrm{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1,\cdots,m_E$, *denote continuous schemes of relaxations of* $h_1,\cdots,h_{m_E}$, *respectively, in* $X \times Y$ *with pointwise convergence orders* $\gamma_{h,1} \geq 2,\cdots,\gamma_{h,m_E} \geq 2$ *and corresponding constants* $\tau_{h,1},\cdots,\tau_{h,m_E}$. *Then, the resulting scheme of convex relaxation-based lower bounding problems for Problem* (P) *is at least second-order convergent at* $(\mathbf{x}^*,\mathbf{y}^*)$.

Proof The result holds as a consequence of Theorems 3 and 5, by using $\boldsymbol{\mu}_Z = \boldsymbol{\mu}^*$, $\boldsymbol{\lambda}_Z = \boldsymbol{\lambda}^*$, $\bar{\mu} = \|\boldsymbol{\mu}^*\|_\infty$, $\bar{\lambda} = \|\boldsymbol{\lambda}^*\|_\infty$ in Theorem 3. □

The following example shows that the convergence order may be as low as two under the assumptions of Theorem 5.

Example 11 Let $X = [-2,2], Y = [0,3]$, $m_I = 1$, and $m_E = 1$ with $f(x,y) = x+y$, $g(x,y) = -y^2+y+2$, and $h(x,y) = x$. Consider intervals $[0,0] \times [2-\varepsilon, 2+\varepsilon] =: Z \in \mathbb{I}(X \times Y)$ with $0 < \varepsilon \leq 1$. Note that $w(Z) = 2\varepsilon$, and

that $\left(0, 2, \frac{1}{3}, -1\right)$ is a KKT point for Problem (P). The optimal objective value of Problem (P) on Z is 2, while the optimal objective value of the Lagrangian dual-based lower bounding problem on Z can be derived as

$$\begin{aligned}
\mathscr{O}(Z) &= \sup_{\mu \geq 0, \lambda} \min_{(x,y) \in Z} x + y + \mu\left(-y^2 + y + 2\right) + \lambda x \\
&= \sup_{\mu \geq 0} \min\left\{(2-\varepsilon) + \mu\left(-(2-\varepsilon)^2 + (2-\varepsilon) + 2\right), (2+\varepsilon) + \mu\left(-(2+\varepsilon)^2 + (2+\varepsilon) + 2\right)\right\} \\
&= \sup_{\mu \geq 0} \min\left\{(2-\varepsilon) + \mu\left(3\varepsilon - \varepsilon^2\right), (2+\varepsilon) + \mu\left(-3\varepsilon - \varepsilon^2\right)\right\} \\
&= (2-\varepsilon) + \frac{1}{3}\left(3\varepsilon - \varepsilon^2\right) \\
&= 2 - \frac{\varepsilon^2}{3},
\end{aligned}$$

where Step 2 follows from the fact that the minimum of a concave function on an interval is attained at one of its endpoints, and the value of μ in Step 4 is obtained by equating the two arguments of the inner min function in Step 3. Convergence of the dual lower bounding scheme at the point $(0,2)$ is, therefore, at most second-order.

Finally, we show that the dual lower bounding scheme is (at least) first-order convergent even when the dual problem is not solved to optimality.

Theorem 6 *Consider Problem (P). Suppose f, g_j, $j = 1, \cdots, m_I$, and h_k, $k = 1, \cdots, m_E$, are Lipschitz continuous on $X \times Y$ with Lipschitz constants M_f, $M_{g,1}, \cdots, M_{g,m_I}$, $M_{h,1}, \cdots, M_{h,m_E}$, respectively. Furthermore, suppose the dual lower bounding scheme involves at most $n_d \geq 1$ iterations of an algorithm applied to the dual at each node of the branch-and-bound tree. In addition, suppose the branch-and-bound algorithm uses first-order (Hausdorff) convergent schemes of constant relaxations $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, on $X \times Y$ to overestimate $\left(\overline{g}_j(Z)\right)_{Z \in \mathbb{I}(X \times Y)}$ and first-order (Hausdorff) convergent schemes of constant relaxations $\left(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, on $X \times Y$ to overestimate $\left(\overline{h}_k(Z)\right)_{Z \in \mathbb{I}(X \times Y)}$ (such schemes of constant relaxations can be obtained, for example, using interval arithmetic [23]), sets $\mu_j = 0$ at each iteration of the algorithm applied to the dual on Z if $g^{\mathrm{U}}_{j,Z} < 0$ (i.e., when inequality constraint j is determined to be inactive on Z by $g^{\mathrm{U}}_{j,Z}$), and determines the dual lower bounding problem on Z to be infeasible either when $g^{\mathrm{L}}_{j,Z} > 0$ for any $j \in \{1, \cdots, m_I\}$ (i.e., when inequality constraint j is determined to be unsatisfiable on Z by $g^{\mathrm{L}}_{j,Z}$), or when $0 \notin \left[h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right]$ for any $k \in \{1, \cdots, m_E\}$ (i.e., when equality constraint k is determined to be unsatisfiable on Z by $(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z})$). Assume that the absolute values of the schemes of dual variables generated by the dual lower bounding scheme are bounded from above by M_∞. Then the dual lower bounding scheme is at least first-order convergent on $X \times Y$.*

Proof From the assumption that $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, and $\left(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, are first-order convergent on $X \times Y$, the determination of infeasibility of the dual lower bounding problem on Z if $g^{\mathrm{L}}_{j,Z} > 0$ for any $j \in \{1, \cdots, m_I\}$, or if $0 \notin \left[h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right]$ for any $k \in \{1, \cdots, m_E\}$, Proposition 1 in [6], and Lemma 8, we conclude that the dual lower bounding scheme has at least first-order convergence at infeasible points (although the dual lower bounding scheme detects infeasibility of infeasible points in $X \times Y$ at least as quickly as any convex relaxation-based lower bounding scheme (see Lemma 11), we assume that the schemes $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, and $\left(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, are available to detect infeasibility since we are only allowed to use at most n_d iterations of an algorithm applied to the dual).

Next, suppose $\mathscr{F}(X \times Y) \neq \emptyset$ and $Z \in \mathbb{I}(X \times Y)$ with $Z \cap \mathscr{F}(X \times Y) \neq \emptyset$. Let J_Z denote the set of inequality constraints that are potentially active at some point in Z as determined by $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)$, i.e. $J_Z := \left\{j \in \{1, \cdots, m_I\} : g^{\mathrm{U}}_{j,Z} \geq 0\right\}$. Let $(\overline{\boldsymbol{\mu}}_Z, \overline{\boldsymbol{\lambda}}_Z) \in \mathbb{R}^{m_I}_+ \times \mathbb{R}^{m_E}$ denote the pair of dual variables corresponding to

the dual lower bound on Z after at most n_d iterations of an algorithm applied to the dual with $\bar{\mu}_{j,Z} = 0$, $\forall j \in \{1,\cdots,m_I\}\backslash J_Z$, and let $(\bar{\mathbf{x}}_Z, \bar{\mathbf{y}}_Z) \in \underset{(\mathbf{x},\mathbf{y})\in Z}{\arg\min}\, f(\mathbf{x},\mathbf{y}) + \bar{\boldsymbol{\mu}}_Z^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \bar{\boldsymbol{\lambda}}_Z^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})$. Note that the condition $\bar{\mu}_{j,Z} = 0$, $\forall j \in \{1,\cdots,m_I\}\backslash J_Z$ can be guaranteed by a suitable initialization of the dual variables and by suitably modifying the dual variables generated by the algorithm applied to the dual (this modification of the dual lower bounding scheme is once again necessitated by the assumption that at most n_d iterations of an algorithm applied to the dual are used). For each $j \in J_Z$, we have

$$g_{j,Z}^{\mathrm{U}} - \max_{(\mathbf{x},\mathbf{y})\in Z} g_j(\mathbf{x},\mathbf{y}) \leq \tau_j w(Z)$$

for some constant $\tau_j \geq 0$ by virtue of the fact that the scheme of constant concave relaxations $\left(g_{j,Z}^{\mathrm{U}}\right)_{Z\in\mathbb{I}(X\times Y)}$ has first-order convergence on $X \times Y$. Since $g_{j,Z}^{\mathrm{U}} \geq 0$, $\forall j \in J_Z$, and g_j is Lipschitz continuous on $X \times Y$, this implies

$$g_j(\mathbf{x},\mathbf{y}) \geq -\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right) w(Z), \quad \forall (\mathbf{x},\mathbf{y}) \in Z,\ \forall j \in J_Z.$$

Let $(\mathbf{x}_Z^*, \mathbf{y}_Z^*) \in \underset{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)}{\arg\min}\ f(\mathbf{x},\mathbf{y})$. We have

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in Z}\left[f(\mathbf{x},\mathbf{y}) + \bar{\boldsymbol{\mu}}_Z^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \bar{\boldsymbol{\lambda}}_Z^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})\right] \\
&= f(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - \left[f(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z) + \bar{\boldsymbol{\mu}}_Z^{\mathrm{T}}\mathbf{g}(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z) + \bar{\boldsymbol{\lambda}}_Z^{\mathrm{T}}\mathbf{h}(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\right] \\
&= \left(f(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - f(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\right) - \sum_{j\in J_Z}\bar{\mu}_{j,Z}\, g_j(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z) + \bar{\boldsymbol{\lambda}}_Z^{\mathrm{T}}\left(\mathbf{h}(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - \mathbf{h}(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\right) \\
&\leq M_f\|(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - (\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\| + \sum_{j\in J_Z}\bar{\mu}_{j,Z}\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right) w(Z) + \sum_{k=1}^{m_E}\left|\bar{\lambda}_{k,Z}\right| M_{h,k}\|(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - (\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\| \\
&\leq \left(M_f\sqrt{n_x+n_y} + \sum_{j\in J_Z} M_\infty\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right) + \sum_{k=1}^{m_E} M_\infty M_{h,k}\sqrt{n_x+n_y}\right) w(Z) \\
&\leq \left(M_f\sqrt{n_x+n_y} + \sum_{j=1}^{m_I} M_\infty\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right) + \sum_{k=1}^{m_E} M_\infty M_{h,k}\sqrt{n_x+n_y}\right) w(Z),
\end{aligned}$$

which establishes the desired result. □

5 Reduced-Space Branch-and-Bound Algorithms

In this section, we present some results on the convergence orders of some widely-applicable reduced-space lower bounding schemes in the literature [9, 10] for Problem (P) when only the set Y may be partitioned during the course of the algorithm. This section is divided into two parts. First, we consider a convex relaxation-based reduced-space lower bounding scheme for a subclass of Problem (P) [10] and investigate its convergence order. Next, we look at the convergence order of a duality-based reduced-space lower bounding scheme [9, Section 3.3] for Problem (P). Algorithm 1 outlines a generic reduced-space branch-and-bound algorithm for Problem (P).

Algorithm 1 A generic reduced-space branch-and-bound algorithm

Initialize:

1. Iteration counter $k = 0$, bounds X^0 and Y^0 on $\mathbf{x}$ and $\mathbf{y}$, respectively, after the optional application of preprocessing techniques to the input data, and tolerances $\varepsilon > 0, \varepsilon^l > 0$, and $\varepsilon^u > 0$ such that $\varepsilon^l + \varepsilon^u \leq \varepsilon$.
2. Domain of the root node $M^0 := X^0 \times Y^0$, and the initial partition $\mathscr{P}^0 = \{M^0\}$.
3. Current best feasible point $\{(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})\} = \emptyset$, upper bound $UBD = +\infty$, lower bound for the root node $LBD^0 = -\infty$, minimum $\{(\mathbf{x}^*, \mathbf{y}^*)\} = \emptyset$, and the optimal objective function value $UBD^* = +\infty$.

repeat

1. If $\mathscr{P}^k = \emptyset$, **terminate**. Otherwise, pick $n \in \left\{n \in \mathbb{N} \cup \{0\} : M^n \in \mathscr{P}^k\right\}$ using some node selection heuristic and set $\mathscr{P}^{k+1} \leftarrow \mathscr{P}^k \backslash \{M^n\}$.
2. (Optional) Solve the (reduced-space) upper bounding problem on M^n with a tolerance of ε^u to try and determine a feasible point. Update UBD, $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$ if a feasible solution better than the current best solution is obtained.
3. (Optional) Apply finite reduced-space bounds-tightening techniques to obtain $\bar{X}^n \subset X^n$, $\bar{Y}^n \subset Y^n$. Set $X^n \leftarrow \bar{X}^n$, $Y^n \leftarrow \bar{Y}^n$. If either X^n or Y^n is empty, **goto** Step 6.
4. Solve the reduced-space lower bounding problem on M^n to ε^l-optimality to obtain the lower bound LBD^n (if the lower bounding problem on M^n is infeasible, set $LBD^n = +\infty$). If node n can be fathomed, **goto** Step 6.
5. Partition M^n into M^{n_1} and M^{n_2} by branching only on the Y-space. Set $\mathscr{P}^{k+1} \leftarrow \mathscr{P}^{k+1} \cup \{M^{n_1}\} \cup \{M^{n_2}\}$, $LBD^{n_1} = LBD^{n_2} = LBD^n$.
6. Set $\mathscr{P}^{k+1} \leftarrow \mathscr{P}^{k+1} \backslash \left\{M^p \in \mathscr{P}^{k+1} : LBD^p \geq UBD - \varepsilon\right\}$, $k \leftarrow k+1$.

until $\mathscr{P}^k = \emptyset$

Set $UBD^* = UBD$ and $(\mathbf{x}^*, \mathbf{y}^*) = (\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$ if $UBD < +\infty$.

It should be noted that Algorithm 1 merely provides the backbone of a generic reduced-space branch-and-bound algorithm. In practice, the order in which the subproblems are solved may vary and additional subproblems may be solved to speed up the convergence of the algorithm. The reader is directed to references [10] and [9] for two widely-applicable instances of Algorithm 1, and for examples of their application. In the remainder of this section, we investigate the convergence orders of the reduced-space lower bounding schemes described in [10] and [9].

5.1 Convex relaxation-based branch-and-bound for problems with special structure

Epperly and Pistikopoulos [10] proposed a reduced-space branch-and-bound algorithm for Problem (P) when $m_E = 0$ (note that this condition can be relaxed as detailed below), and the functions f and $g_j, \forall j \in \{1, \cdots, m_I\}$, in Problem (P) are each of the form

$$w(\mathbf{x}, \mathbf{y}) = w^A(\mathbf{x}) + \sum_{i \in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y}), \tag{W}$$

where Q is a finite set of indices, and the functions $w^A, \mathbf{w}^B, \mathbf{w}^C$, and w^D satisfy:

1. w^A and $\mathbf{w}^B$ are convex on X.
2. $\mathbf{w}^C$ and w^D are continuous on Y.
3. Strongly consistent convex and concave relaxations are available for $\mathbf{w}^C$ and w^D on Y.
4. $\mathbf{w}^B$ and $\mathbf{w}^C$ have continuous tight bounds.
5. For each $i \in Q$, at least one of the following two conditions must hold:
   a. $w_i^B(\mathbf{x}) = \mathbf{c}_i^{\mathrm{T}} \mathbf{x}$ for some constant $\mathbf{c}_i \in \mathbb{R}^{n_x}$,
   b. $w_i^C(\mathbf{y}) \geq 0$ for all $\mathbf{y} \in Y$.

Epperly and Pistikopoulos [10] state that equality constraints can be equivalently reformulated using pairs of inequalities; however, the above assumptions restrict the functional forms of the equality constraints h_k, $k = 1, \cdots, m_E$, to

$$h_k(\mathbf{x}, \mathbf{y}) = \sum_{i \in Q} \left(\mathbf{c}_i^{\mathrm{T}} \mathbf{x}\right) w_i^C(\mathbf{y}) + w^D(\mathbf{y}). \tag{W$_{\mathrm{eq}}$}$$

Suppose for each $Z \in \mathbb{I}Y$, we associate an interval $X(Z)$ such that $\square X \supset X(Z) \supset \mathscr{F}_X(Z)$, where $\square X$ denotes the interval hull of X (note that we make the implicit assumption (see Remark 2) that X is an interval in this section). Assumption 3 can be restated as follows: there exists a continuous scheme, $(w_{i,Z}^{C,\mathrm{cv}}, w_{i,Z}^{C,\mathrm{cc}})_{Z \in \mathbb{I}Y}$, of relaxations of w_i^C, $i \in Q$, in Y with pointwise convergence order $\gamma_i^C > 0$, and there exists a continuous scheme of convex relaxations, $(w_Z^{D,\mathrm{cv}})_{Z \in \mathbb{I}Y}$, of w^D in Y with pointwise convergence order $\gamma^{D,\mathrm{cv}} > 0$. Assumption 4 can be replaced by the following: there exist schemes of constant relaxations $(w_{i,Z}^{B,\mathrm{L}}, w_{i,Z}^{B,\mathrm{U}})_{Z \in \mathbb{I}\square X}$ and $(w_{i,Z}^{C,\mathrm{L}}, w_{i,Z}^{C,\mathrm{U}})_{Z \in \mathbb{I}Y}$, $i \in Q$, of w_i^B and w_i^C in X and Y, respectively, with (Hausdorff) convergence orders $\beta_i^{B,\mathrm{c}} > 0$ and $\beta_i^{C,\mathrm{c}} > 0$. In addition, we assume that the range order of w_i^C, $\forall i \in Q$, is greater than zero on Y (cf. Lemma 12).

Under the above assumptions, Epperly and Pistikopoulos [10] show that underestimating each function $w(\mathbf{x}, \mathbf{y})$ of the form (W) using the scheme $(w_{X(Z)\times Z}^{\mathrm{cv}})_{Z \in \mathbb{I}Y}$ of convex relaxations defined by

$$w_{X(Z)\times Z}^{\mathrm{cv}}(\mathbf{x}, \mathbf{y}) = w^A(\mathbf{x}) + \sum_{i \in Q} w_{i,X(Z)\times Z}^{BC,\mathrm{cv}}(\mathbf{x}, \mathbf{y}) + w_Z^{D,\mathrm{cv}}(\mathbf{y}), \tag{W$^{\mathrm{cv}}$}$$

where, for each $i \in Q$, the scheme of convex relaxations $(w_{i,X(Z)\times Z}^{BC,\mathrm{cv}})_{Z \in \mathbb{I}Y}$ is obtained using McCormick's product rule [20] as

$$w_{i,X(Z)\times Z}^{BC,\mathrm{cv}}(\mathbf{x}, \mathbf{y}) = \begin{cases} \max \left\{ \begin{array}{l} w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}, \\ w_{i,X(Z)}^{B,\mathrm{L}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{L}} - w_{i,X(Z)}^{B,\mathrm{L}} w_{i,Z}^{C,\mathrm{L}} \end{array} \right\}, & \text{if } w_{i,X(Z)}^{B,\mathrm{L}} \geq 0 \\ \max \left\{ \begin{array}{l} w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}, \\ w_{i,X(Z)}^{B,\mathrm{L}} w_{i,Z}^{C,\mathrm{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{L}} - w_{i,X(Z)}^{B,\mathrm{L}} w_{i,Z}^{C,\mathrm{L}} \end{array} \right\}, & \text{if } w_{i,X(Z)}^{B,\mathrm{U}} < 0 \\ \max \left\{ \begin{array}{l} w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}, \\ w_{i,X(Z)}^{B,\mathrm{L}} w_{i,Z}^{C,\mathrm{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{L}} - w_{i,X(Z)}^{B,\mathrm{L}} w_{i,Z}^{C,\mathrm{L}} \end{array} \right\}, & \text{otherwise} \end{cases},$$

yields a convergent reduced-space lower bounding scheme with any accumulation point of the sequence of lower bounding solutions solving Problem (P) when the subdivision process is exhaustive on Y and the selection procedure is bound improving.

Before we investigate the convergence order of the reduced-space lower bounding scheme in [10], we look at the propagation of the convergence orders of the relaxation schemes $(w_{i,Z}^{B,\mathrm{L}}, w_{i,Z}^{B,\mathrm{U}})_{Z \in \mathbb{I}\square X}$, $(w_{i,Z}^{C,\mathrm{cv}}, w_{i,Z}^{C,\mathrm{cc}})_{Z \in \mathbb{I}Y}$, $(w_{i,Z}^{C,\mathrm{L}}, w_{i,Z}^{C,\mathrm{U}})_{Z \in \mathbb{I}Y}$, $\forall i \in Q$, and $(w_Z^{D,\mathrm{cv}})_{Z \in \mathbb{I}Y}$ to the convergence order of the reduced-space scheme of convex relaxations $(w_{X(Z)\times Z}^{\mathrm{cv}})_{Z \in \mathbb{I}Y}$. Note that unless otherwise specified, we simply use $X(Z) = \square X (= X)$, $\forall Z \in \mathbb{I}Y$. The following result provides sufficient conditions for the scheme of convex relaxations defined by (W$^{\mathrm{cv}}$) to have pointwise convergence of a given order on Y.

Lemma 12 *Let $X \subset \mathbb{R}^{n_x}$, $Y \subset \mathbb{R}^{n_y}$ be nonempty compact convex sets and $f : X \times Y \to \mathbb{R}$ be a function of the form* (W) *such that*

$$f : X \times Y \ni (\mathbf{x}, \mathbf{y}) \longmapsto w^A(\mathbf{x}) + \sum_{i \in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y}).$$

Assume that w^A, w_i^B, $\forall i \in Q$, and w^D are continuous, and for each $i \in Q$, w_i^C has range of order $\alpha_i^C \geq 1$ on Y with corresponding constant $\tau_i^{C,\mathrm{r}}$. Let $(w_{i,Z}^{C,\mathrm{cv}}, w_{i,Z}^{C,\mathrm{cc}})_{Z \in \mathbb{I}Y}$ and $(w_Z^{D,\mathrm{cv}})_{Z \in \mathbb{I}Y}$ respectively denote continuous

schemes of relaxations of w_i^C, $i \in Q$, and w^D in Y with pointwise convergence orders $\gamma_i^C \geq 1$ and $\gamma^{D,\mathrm{cv}} \geq 1$ and corresponding constants τ_i^C and $\tau^{D,\mathrm{cv}}$. Let $(w_{i,Z}^{B,\mathrm{L}}, w_{i,Z}^{B,\mathrm{U}})_{Z \in \mathbb{I}\square X}$ and $(w_{i,Z}^{C,\mathrm{L}}, w_{i,Z}^{C,\mathrm{U}})_{Z \in \mathbb{I}Y}$ respectively denote schemes of constant relaxations of w_i^B in $\square X$ and w_i^C in Y, $\forall i \in Q$, with (Hausdorff) convergence orders $\beta_i^{B,\mathrm{c}} > 0$ and $\beta_i^{C,\mathrm{c}} \geq 1$ and corresponding constants $\tau_i^{B,\mathrm{c}}$ and $\tau_i^{C,\mathrm{c}}$. Then the continuous scheme of convex relaxations $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z \in \mathbb{I}Y}$ of the form (W^{cv}) defined by

$$f_{X(Z)\times Z}^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) := w^A(\mathbf{x}) + \sum_{i\in Q} w_{i,X(Z)\times Z}^{BC,\mathrm{cv}}(\mathbf{x},\mathbf{y}) + w_Z^{D,\mathrm{cv}}(\mathbf{y}), \quad \forall (\mathbf{x},\mathbf{y}) \in X(Z)\times Z,$$

has pointwise convergence of order at least $\min\left\{\min_{i\in Q}\left\{\min\left\{\alpha_i^C, \beta_i^{C,\mathrm{c}}, \gamma_i^C\right\}\right\}, \gamma^{D,\mathrm{cv}}\right\}$ *on* Y.

Proof From Equation (W^{cv}), we have for each $(\mathbf{x},\mathbf{y}) \in X(Z)\times Z$:

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) - f_{X(Z)\times Z}^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) &= \left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y})\right) - \left(w^A(\mathbf{x}) + \sum_{i\in Q} w_{i,X(Z)\times Z}^{BC,\mathrm{cv}}(\mathbf{x},\mathbf{y}) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\right) \\ &= \sum_{i\in Q}\left(w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - w_{i,X(Z)\times Z}^{BC,\mathrm{cv}}(\mathbf{x},\mathbf{y})\right) + \left(w^D(\mathbf{y}) - w_Z^{D,\mathrm{cv}}(\mathbf{y})\right). \end{aligned}$$

Depending on whether $w_{i,X(Z)}^{B,\mathrm{L}} \geq 0$, $w_{i,X(Z)}^{B,\mathrm{U}} < 0$, or $0 \in \left(w_{i,X(Z)}^{B,\mathrm{L}}, w_{i,X(Z)}^{B,\mathrm{U}}\right]$ for each $i \in Q$, we have that $\left(w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - w_{i,X(Z)\times Z}^{BC,\mathrm{cv}}(\mathbf{x},\mathbf{y})\right)$ is bounded from above either by

$$\left[w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - \left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\right)\right],$$

or by

$$\left[w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - \left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\right)\right]$$

for each $(\mathbf{x},\mathbf{y}) \in X(Z)\times Z$. Consequently, it is sufficient to show the existence of constants $\tau_1, \tau_2 \geq 0$ such that

$$\begin{aligned} \max_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} \Bigg| &\left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y})\right) - \\ &\left(w^A(\mathbf{x}) + \sum_{i\in Q}\left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\right) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\right)\Bigg| \leq \tau_1 w(Z)^\gamma \end{aligned}$$

and

$$\begin{aligned} \max_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} \Bigg| &\left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y})\right) - \\ &\left(w^A(\mathbf{x}) + \sum_{i\in Q}\left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\right) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\right)\Bigg| \leq \tau_2 w(Z)^\gamma, \end{aligned}$$

where $\gamma := \min\left\{\min_{i\in Q}\left\{\min\left\{\alpha_i^C, \beta_i^{C,\mathrm{c}}, \gamma_i^C\right\}\right\}, \gamma^{D,\mathrm{cv}}\right\}$, to prove that $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$ converges pointwise to f with order γ on Y. The ensuing arguments prove the existence of τ_1; the existence of τ_2 can be proven analogously.

We have $\forall(\mathbf{x},\mathbf{y}) \in X(Z) \times Z$:

$$\begin{aligned}
&\left(\left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y})\right) - \right. \\
&\qquad \left.\left(w^A(\mathbf{x}) + \sum_{i\in Q}\left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\right) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\right)\right) \\
&= \left(\sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - \left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\right)\right) + \left(w^D(\mathbf{y}) - w_Z^{D,\mathrm{cv}}(\mathbf{y})\right). \qquad (7)
\end{aligned}$$

Note that $\left|w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right|$ can be bounded from above as

$$\begin{aligned}
\left|w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right| &= \left|\left(w_i^C(\mathbf{y}) - \max_{\mathbf{y}\in Z} w_i^C(\mathbf{y})\right) + \left(\max_{\mathbf{y}\in Z} w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right)\right| \\
&\le \left|w_i^C(\mathbf{y}) - \max_{\mathbf{y}\in Z} w_i^C(\mathbf{y})\right| + \left|\max_{\mathbf{y}\in Z} w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right| \\
&\le \left(\tau_i^{C,\mathrm{r}} w(Z)^{\alpha_i^C - \min\left\{\alpha_i^C,\beta_i^{C,\mathrm{c}}\right\}} + \tau_i^{C,\mathrm{c}} w(Z)^{\beta_i^{C,\mathrm{c}} - \min\left\{\alpha_i^C,\beta_i^{C,\mathrm{c}}\right\}}\right) w(Z)^{\min\left\{\alpha_i^C,\beta_i^{C,\mathrm{c}}\right\}} \\
&\le M_i^C w(Z)^{\beta_i^{C,\mathrm{r}}}, \quad \forall \mathbf{y} \in Z,
\end{aligned}$$

with $M_i^C := \tau_i^{C,\mathrm{r}} w(Y)^{\alpha_i^C - \beta_i^{C,\mathrm{r}}} + \tau_i^{C,\mathrm{c}} w(Y)^{\beta_i^{C,\mathrm{c}} - \beta_i^{C,\mathrm{r}}}$ and $\beta_i^{C,\mathrm{r}} := \min\left\{\alpha_i^C, \beta_i^{C,\mathrm{c}}\right\}$.

The first term in Equation (7) can be bounded as

$$\begin{aligned}
&\sum_{i\in Q}\left(w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - \left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\right)\right) \\
&= \sum_{i\in Q}\left[\left(w_i^B(\mathbf{x}) - w_{i,X(Z)}^{B,\mathrm{U}}\right)\left(w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right) + w_{i,X(Z)}^{B,\mathrm{U}}\left(w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y})\right)\right] \\
&\le \sum_{i\in Q}\left|w_i^B(\mathbf{x}) - w_{i,X(Z)}^{B,\mathrm{U}}\right|\left|w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right| + \left|w_{i,X(Z)}^{B,\mathrm{U}}\left(w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y})\right)\right| \\
&\le \sum_{i\in Q} M_i^{BC} w(Z)^{\gamma_i^{BC}} \\
&\le M^{BC} w(Z)^{\gamma^{BC}}, \quad \forall(\mathbf{x},\mathbf{y}) \in X(Z) \times Z, \qquad (8)
\end{aligned}$$

where the constants M^{BC}, γ^{BC}, and M_i^{BC}, γ_i^{BC}, $\forall i \in Q$, can be computed as

$$\begin{aligned}
M^{BC} &:= \sum_{i\in Q} M_i^{BC} w(Y)^{\gamma_i^{BC} - \gamma^{BC}}, \quad \gamma^{BC} := \min_{i\in Q} \gamma_i^{BC}, \quad \gamma_i^{BC} := \min\left\{\beta_i^{C,\mathrm{r}}, \gamma_i^C\right\}, \\
M_i^{BC} &:= \left[M_i^{B,1} M_i^C w(Y)^{\beta_i^{C,\mathrm{r}} - \gamma_i^{BC}} + M_i^{B,2} \tau_i^C w(Y)^{\gamma_i^C - \gamma_i^{BC}}\right], \\
M_i^{B,1} &:= \max_{\mathbf{x}\in X} w_i^B(\mathbf{x}) - \min_{\mathbf{x}\in X} w_i^B(\mathbf{x}) + \tau_i^{B,\mathrm{c}} w(X)^{\beta_i^{B,\mathrm{c}}}, \quad M_i^{B,2} := \max_{\mathbf{x}\in X} w_i^B(\mathbf{x}) + \tau_i^{B,\mathrm{c}} w(X)^{\beta_i^{B,\mathrm{c}}}.
\end{aligned}$$

The second term in Equation (7) is simply bounded as

$$w^D(\mathbf{y}) - w_Z^{D,\mathrm{cv}}(\mathbf{y}) \leq \tau^{D,\mathrm{cv}} w(Z)^{\gamma^{D,\mathrm{cv}}}, \quad \forall \mathbf{y} \in Z. \tag{9}$$

From Equations (8) and (9), we have

$$\begin{aligned}
&\max_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} \left| \left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y}) \right) - \right. \\
&\qquad \left. \left(w^A(\mathbf{x}) + \sum_{i\in Q} \left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}} \right) + w_Z^{D,\mathrm{cv}}(\mathbf{y}) \right) \right| \\
&\leq \left(M^{BC} w(Y)^{\gamma^{BC}-\gamma} + \tau^{D,\mathrm{cv}} w(Y)^{\gamma^{D,\mathrm{cv}}-\gamma} \right) w(Z)^{\gamma},
\end{aligned}$$

which proves the existence of τ_1. □

The following remark is in order.

Remark 12

1. Suppose w_i^C is Lipschitz continuous on Y for each $i \in Q$. We then have $\alpha_i^C \geq 1, \forall i \in Q$. If $\gamma_i^C \geq 1$ and $\beta_i^{C,\mathrm{c}} \geq 1, \forall i \in Q$, and $\gamma^{D,\mathrm{cv}} \geq 1$, we have from Lemma 12 that $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$ has at least first-order convergence on Y.
2. Let $X = [1,2]$, $Y = [-1,1]$, and $f(x,y) = xy$. For any $[-\varepsilon,\varepsilon] =: Z \in \mathbb{I}Y$ with $\varepsilon > 0$, consider the scheme of convex relaxations $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$ of f in Y with

$$f_{X(Z)\times Z}^{\mathrm{cv}}(x,y) = \max\{y - \varepsilon x + \varepsilon, 2y + \varepsilon x - 2\varepsilon\}.$$

The above scheme corresponds to the tightest possible scheme of convex relaxations in the reduced-space, but has at most first-order pointwise convergence on Y. This is in contrast to Theorem 10 in [5] where the scheme of convex envelopes of any twice continuously differentiable function was shown to have pointwise convergence order of at least two on $X \times Y$. Note that if $Q = \emptyset$, the pointwise convergence order of the scheme of convex relaxations $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$ is dictated by the pointwise convergence order of the scheme $(w_Z^{D,\mathrm{cv}})_{Z\in\mathbb{I}Y}$, and second-order pointwise convergence of $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$ can be achieved by using the scheme of convex envelopes of w^D if it is twice continuously differentiable. Also note that Theorem 2 in [5], which states that the pointwise convergence order of a scheme of relaxations of a nonlinear twice continuously differentiable function can be at most two on $X \times Y$, naturally holds over Y as well.

The following result establishes a lower bound on the convergence order of the reduced-space lower bounding scheme proposed in [10] at infeasible points.

Lemma 13 *Consider Problem* (P)*, and suppose functions* g_j, $j = 1,\cdots,m_I$, *are each of the form* (W) *and functions* h_k, $k = 1,\cdots,m_E$, *are each of the form* ($\mathrm{W_{eq}}$)*. Let* $(g_{j,X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$, $j = 1,\cdots,m_I$, *denote continuous schemes of convex relaxations of* $g_1,\cdots,g_{m_I}$, *respectively, in* Y *with pointwise convergence orders* $\gamma_{g,1}^{\mathrm{cv}} > 0,\cdots,\gamma_{g,m_I}^{\mathrm{cv}} > 0$ *and corresponding constants* $\tau_{g,1}^{\mathrm{cv}},\cdots,\tau_{g,m_I}^{\mathrm{cv}}$*, and let* $(h_{k,X(Z)\times Z}^{\mathrm{cv}}, h_{k,X(Z)\times Z}^{\mathrm{cc}})_{Z\in\mathbb{I}Y}$, $k = 1,\cdots,m_E$, *denote continuous schemes of relaxations of* $h_1,\cdots,h_{m_E}$, *respectively, in* Y *with pointwise convergence orders* $\gamma_{h,1} > 0,\cdots,\gamma_{h,m_E} > 0$ *and corresponding constants* $\tau_{h,1},\cdots,\tau_{h,m_E}$*. Then, there exists* $\bar{\tau} \geq 0$ *such that for every* $Z \in \mathbb{I}Y$

$$d\left(\overline{\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}}(X(Z)\times Z), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) - d\left(\mathscr{S}_C(Z), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) \leq \bar{\tau} w(Z)^{\beta},$$

where $\mathscr{I}_C(Z)$ *is defined as*

$$\mathscr{I}_C(Z) := \Big\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}), \mathbf{h}^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\mathrm{cc}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \\ \text{for some } (\mathbf{x},\mathbf{y}) \in X(Z)\times Z\Big\}$$

and β *is defined as*

$$\beta := \min\left\{\min_{j\in\{1,\cdots,m_I\}} \gamma^{\mathrm{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}} \gamma_{h,k}\right\}.$$

Proof The proof is similar to that of Lemma 8 and is therefore omitted. □

Definition 15 (Feasible Point in the Reduced-Space) Consider Problem (P). A point $\mathbf{y} \in Y$ is said to be feasible (in the reduced-space) if there exists $\mathbf{x} \in X$ such that $(\mathbf{x},\mathbf{y})$ is feasible for Problem (P).

The following result establishes first-order convergence of the reduced-space lower bounding scheme proposed in [10] at a feasible point in the reduced-space when first-order pointwise convergent schemes of relaxations are used and the reduced-space dual lower bounding scheme (see Section 5.2) is first-order convergent.

Theorem 7 *Consider Problem* (P). *Suppose the functions* f *and* g_j, $j = 1,\cdots,m_I$, *are each of the form* (W) *and functions* h_k, $k = 1,\cdots,m_E$, *are each of the form* ($\mathrm{W_{eq}}$). *Let* $\mathbf{y}^{\mathrm{f}} \in Y$ *be a feasible point in the reduced-space for Problem* (P). *Suppose the reduced-space dual lower bounding scheme (see Section 5.2) has convergence of order* β_d *at* $\mathbf{y}^{\mathrm{f}}$ *and a corresponding scheme of dual variables* $\left(\left(\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}_Z, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}_Z\right)\right)_{Z\in\mathbb{I}Y}$ *(not necessarily optimal, but which yield* β_d*-order convergence at* $\mathbf{y}^{\mathrm{f}}$*) with* $\left(\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}_Z, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}_Z\right) \in \mathbb{R}^{m_I}_+ \times \mathbb{R}^{m_E}$, $\left\|\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}_Z\right\|_\infty \leq \bar{\mu}$ *and* $\left\|\boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}_Z\right\|_\infty \leq \bar{\lambda}$, $\forall Z \in \mathbb{I}Y$, *for some constants* $\bar{\mu},\bar{\lambda} \geq 0$. *Let* $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $(g^{\mathrm{cv}}_{j,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $j = 1,\cdots,m_I$, *denote continuous schemes of convex relaxations of* f, $g_1,\cdots,g_{m_I}$, *respectively, in* Y *with pointwise convergence orders* $\gamma^{\mathrm{cv}}_f \geq 1$, $\gamma^{\mathrm{cv}}_{g,1} \geq 1,\cdots,\gamma^{\mathrm{cv}}_{g,m_I} \geq 1$ *and corresponding constants* τ^{cv}_f, $\tau^{\mathrm{cv}}_{g,1},\cdots,\tau^{\mathrm{cv}}_{g,m_I}$. *Let* $(h^{\mathrm{cv}}_{k,X(Z)\times Z}, h^{\mathrm{cc}}_{k,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $k = 1,\cdots,m_E$, *denote continuous schemes of relaxations of* $h_1,\cdots,h_{m_E}$, *respectively, in* Y *with pointwise convergence orders* $\gamma_{h,1} \geq 1,\cdots,\gamma_{h,m_E} \geq 1$ *and corresponding constants* $\tau_{h,1},\cdots,\tau_{h,m_E}$. *Then, the scheme of lower bounding problems* $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ *with*

$$(\mathscr{O}(Z))_{Z\in\mathbb{I}Y} := \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}Y},$$

$$(\mathscr{I}_C(Z))_{Z\in\mathbb{I}Y} := \Big(\Big\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}), \mathbf{h}^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\mathrm{cc}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \\ \text{for some } (\mathbf{x},\mathbf{z}) \in X(Z)\times Z\Big\}\Big)_{Z\in\mathbb{I}Y}$$

is at least $\min\left\{\min\left\{\gamma^{\mathrm{cv}}_f, \min_{j\in\{1,\cdots,m_I\}} \gamma^{\mathrm{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}} \gamma_{h,k}\right\}, \beta_d\right\}$*-order convergent at* $\mathbf{y}^{\mathrm{f}}$.

Proof The proof is similar to that of Theorem 3 and is therefore omitted. □

Definition 16 (Unconstrained Point in the Reduced-Space) Consider Problem (P) with $m_E = 0$. A point $\mathbf{y} \in Y$ is said to be unconstrained (in the reduced-space) if there exists $\delta > 0$ such that $\forall \mathbf{z} \in Y$ with $\|\mathbf{z}-\mathbf{y}\| < \delta$, we have $\mathbf{g}(\mathbf{x},\mathbf{z}) < \mathbf{0}$, $\forall \mathbf{x} \in X$.

The next result establishes first-order convergence of the reduced-space lower bounding scheme proposed in [10] at unconstrained points in the reduced-space when a first-order convergent scheme of relaxations of the objective is used by the (convergent) lower bounding scheme.

Proposition 2 *Consider Problem (P) with $m_E = 0$. Suppose the functions f and g_j, $j = 1,\cdots,m_I$, are each of the form (W). Let $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ denote a continuous scheme of convex relaxations of f in Y with convergence order $\beta^{\mathrm{cv}}_f > 0$ and corresponding constant τ^{cv}_f, $(g^{\mathrm{cv}}_{j,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $j = 1,\cdots,m_I$, denote continuous schemes of convex relaxations of $g_1,\cdots,g_{m_I}$, respectively, in Y with pointwise convergence orders $\gamma^{\mathrm{cv}}_{g,1} > 0,\cdots,\gamma^{\mathrm{cv}}_{g,m_I} > 0$ and corresponding constants $\tau^{\mathrm{cv}}_{g,1},\cdots,\tau^{\mathrm{cv}}_{g,m_I}$.*

Suppose $\mathbf{y}^{\mathrm{S}} \in Y$ is an unconstrained point in the reduced-space, and the scheme of lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ with

$$(\mathscr{O}(Z))_{Z\in\mathbb{I}Y} := \left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}Y},$$
$$(\mathscr{I}_C(Z))_{Z\in\mathbb{I}Y} := \left(\overline{\mathbf{g}}^{\mathrm{cv}}_{X(Z)\times Z}(X(Z)\times Z)\right)_{Z\in\mathbb{I}Y}$$

has convergence of order $\beta \in (0,\beta^{\mathrm{cv}}_f]$ at $\mathbf{y}^{\mathrm{S}}$. Then the scheme of lower bounding problems $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ is at least β^{cv}_f-order convergent at $\mathbf{y}^{\mathrm{S}}$.

Proof The proof is relegated to Appendix A.2 since it is similar to the proof of Corollary 2. □

Note that Proposition 2 can be generalized in a manner similar to Corollary 3 to show that the above lower bounding scheme has β^{cv}_f-order convergence on a neighborhood of $\mathbf{y}^{\mathrm{S}}$.

The following example shows that the convergence order of the reduced-space lower bounding scheme is dictated by the convergence order, β^{cv}_f, of the scheme $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ under the assumptions of Proposition 2.

Example 12 Let $X = [-1,0.1], Y = [-1,1]$, $m_I = 1$, and $m_E = 0$ with $f(x,y) = x^2+y^2$ and $g(x,y) = x+y-0.5$. For any $[y^{\mathrm{L}},y^{\mathrm{U}}] =: Z \in \mathbb{I}Y$, let

$$f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) = \begin{cases} x^2-(y^{\mathrm{U}}-y^{\mathrm{L}})^3, & \text{if } 0\in[y^{\mathrm{L}},y^{\mathrm{U}}] \\ x^2+\min\left\{(y^{\mathrm{L}})^2,(y^{\mathrm{U}})^2\right\}-(y^{\mathrm{U}}-y^{\mathrm{L}})^3, & \text{otherwise} \end{cases},$$
$$g^{\mathrm{cv}}_{X(Z)\times Z}(x,y) = x+y-0.5.$$

The scheme $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ has first-order pointwise convergence on Y and third-order convergence on Y, while the scheme $(g^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ has arbitrarily high pointwise convergence order on Y.

Let $y^{\mathrm{L}} = -\varepsilon$, $y^{\mathrm{U}} = \varepsilon$ with $0 < \varepsilon \leq 0.1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-8\varepsilon^3$. Convergence at the point $y = 0$ is, therefore, at most third-order.

It is natural to wonder at this stage whether the reduced-space lower bounding scheme in [10] has 'similar convergence properties' to the full-space lower bounding scheme that was analyzed in Section 4.1. Example 16 presents a case where the reduced-space lower bounding scheme in [10] only has first-order convergence at a constrained minimizer that is a KKT point (cf. Example 6, Theorem 2 and Corollary 4). The following example shows that the reduced-space lower bounding scheme in [10] may have a convergence order as low as one even for unconstrained problems with smooth objective functions.

Example 13 Consider the following instance of Problem (P):

$$\min_{x,y}\ 2x^2+x^2y-xy^2+(y-0.5)^2$$
$$\text{s.t. } x\in[-1,1], y\in[0,1].$$

The global minimum, (x^*,y^*), of the above 'unconstrained problem' is $x^* = \frac{2\sqrt{21}}{3}-3$, $y^* = \frac{\sqrt{21}}{3}-1$ with optimal objective value $v^* = 2(x^*)^2+(x^*)^2y^*-x^*(y^*)^2+(y^*-0.5)^2$.

Consider $[y^* - \varepsilon, y^* + \varepsilon] =: Z \in \mathbb{I}Y$ with $\varepsilon \in (0, 0.25]$. The reduced-space lower bounding scheme in [10] yields

$$\begin{aligned}
\mathscr{O}(Z) = \min_{x,y,w_1,w_2} \; & 2x^2 + w_1 + w_2 + (y - 0.5)^2 \\
\text{s.t.} \; & w_1 \geq x^2(y^* - \varepsilon), \\
& w_1 \geq y + x^2(y^* + \varepsilon) - (y^* + \varepsilon), \\
& w_2 \geq y^2 - x(y^* + \varepsilon)^2 - (y^* + \varepsilon)^2, \\
& w_2 \geq (y^*)^2 - 2y^* y - \varepsilon^2 - x(y^* - \varepsilon)^2 + (y^* - \varepsilon)^2, \\
& x \in [-1, 1], y \in [y^* - \varepsilon, y^* + \varepsilon].
\end{aligned}$$

Note that the point $(x^{\mathrm{f}}_Z, y^{\mathrm{f}}_Z, w^{\mathrm{f}}_{1,Z}, w^{\mathrm{f}}_{2,Z}) = (x^*, y^*, (x^*)^2(y^* - \varepsilon), -(y^*)^2 - \varepsilon^2 - x^*(y^* - \varepsilon)^2 + (y^* - \varepsilon)^2)$ is feasible for the lower bounding scheme with objective value $2(x^*)^2 + w^{\mathrm{f}}_{1,Z} + w^{\mathrm{f}}_{2,Z} + (y^* - 0.5)^2 = v^* + 2x^* y^* \varepsilon - (x^*)^2 \varepsilon - x^* \varepsilon^2 - 2y^* \varepsilon$. Therefore, we have

$$\begin{aligned}
\min_{(x,y) \in \mathscr{F}(Z)} f(x,y) \; - \min_{(x,y) \in \mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z) \times Z}(x,y) &\geq (x^*)^2 \varepsilon + x^* \varepsilon^2 + 2y^* \varepsilon - 2x^* y^* \varepsilon \\
&= \left(0.5(x^*)^2 + 0.5x^* \varepsilon + y^* - x^* y^*\right) w(Z) \\
&= 0.5\,(1 + \varepsilon x^*)\, w(Z) \\
&\geq 0.5 w(Z),
\end{aligned}$$

which establishes that the reduced-space lower bounding scheme in [10] has at most first-order convergence at (the reduced-space minimizer) y^*.

Remark 13 Example 13 provides an instance of Problem (P) for which the minimum is unconstrained but the reduced-space lower bounding scheme in [10] is only first-order convergent at the reduced-space minimizer. Therefore, the lower bounding scheme in [10] could face severe clustering for this example [7, 39]. Note that this is in contrast to the full-space lower bounding schemes in Section 4 which can achieve at least second-order convergence at the above minimizer and thereby mitigate clustering.

The presence of the terms $x^2 y$ and $-xy^2$ in the objective function in Example 13 plays a crucial role in limiting the convergence order of the reduced-space lower bounding scheme in [10] (see Remark 12). Additionally, the analysis in Example 13 implies that the scheme of relaxations of its objective function has at most first-order Hausdorff convergence on Y. Theorem 10 in Section 5.2 implies that the reduced-space lower bounding scheme in [10] has second-order convergence at KKT points when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$.

5.2 Duality-based branch-and-bound

Dür and Horst [9, Section 3.3] outlined a reduced-space branch-and-bound algorithm in which they used Lagrangian duality to obtain lower bounds (also see [3, 8]). Dür and Horst [9] prove that when a constraint qualification holds for the reduced-space convex relaxation-based lower bounding scheme with each function in Problem (P) replaced by its (convex) envelope on $X \times Z$ (for each $Z \in \mathbb{I}Y$), the subdivision process is exhaustive on Y, and the selection procedure is bound improving, then any accumulation point of the sequence of reduced-space dual lower bounding solutions solves Problem (P).

The reduced-space Lagrangian dual lower bounding problem is in essence the same as its full-space counterpart Problem (D), with the major difference being that we only branch on the Y-space in the reduced-space dual lower bounding scheme to converge. We associate with the reduced-space dual lower bounding scheme, $(\mathscr{L}(Z))_{Z \in \mathbb{I}Y}$, at a feasible point in the reduced-space $\mathbf{y}$, a scheme of dual variables $((\boldsymbol{\mu}^{\mathbf{y}}_Z, \boldsymbol{\lambda}^{\mathbf{y}}_Z))_{Z \in \mathbb{I}Y}$ corresponding to the solution of the scheme of dual lower bounding problems $(\mathscr{O}(Z))_{Z \in \mathbb{I}Y}$ with $\mathbf{y} \in Z$. Dür and Horst [9, Section 4] also outlined classes of problems for which the reduced-space dual lower bounding problem can be solved to optimality. The following result, analogous to Theorem 4, holds.

Theorem 8 *Consider Problem (P). Suppose strong duality holds for the reduced-space convex relaxation-based lower bounding scheme for Problem (P) obtained by using the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$. Then, the reduced-space dual lower bounding scheme has at least as high a convergence order as the reduced-space convex relaxation-based lower bounding scheme obtained by using schemes of (convex) envelopes.*

Proof The proof is similar to that of Theorem 4 and is therefore omitted. □

The following result from [9] states that when the constraints in Problem (P) are affine on $X \times Y$, the lower bounding scheme corresponding to schemes of (convex) envelopes provides the same scheme of lower bounds as that obtained using the dual lower bounding scheme.

Lemma 14 *Consider Problem (P), and suppose the constraints in Problem (P) are affine in $\mathbf{x}$ and $\mathbf{y}$, i.e. $\mathbf{g} : (\mathbf{x},\mathbf{y}) \mapsto \mathbf{A}_g\mathbf{x} + \mathbf{B}_g\mathbf{y} - \mathbf{c}_g$ and $\mathbf{h} : (\mathbf{x},\mathbf{y}) \mapsto \mathbf{A}_h\mathbf{x} + \mathbf{B}_h\mathbf{y} - \mathbf{c}_h$. In addition, suppose Problem (P) is feasible and strong duality holds for Problem (P) for $\mathbf{y}$ restricted to any feasible point in Y. Then the lower bound obtained by solving the dual problem on $Z \in \mathbb{I}Y$ is the same as the lower bound obtained by solving the relaxation of the original problem on Z with the objective function f replaced by its convex envelope on $X \times Z$.*

Proof See Proposition 2.1 in [9]. □

Lemma 13 (in conjunction with Lemma 11) guarantees that the reduced-space dual lower bounding scheme has at least first-order convergence at infeasible points for the subclass of Problem (P) for which the algorithm of Epperly and Pistikopoulos is applicable when the functions w_i^C, $\forall i \in Q$, and w^D corresponding to each of the constraints are Lipschitz continuous. The following result shows that first-order convergence at infeasible points is guaranteed for a more general class of problems in the form of Problem (P) even when constraint propagation techniques are not used.

Lemma 15 *Let $X \subset \mathbb{R}^{n_x}$, $Y \subset \mathbb{R}^{n_y}$ be nonempty compact convex sets, $f : X \times Y \to \mathbb{R}$ be Lipschitz continuous on $X \times Y$ with Lipschitz constant M_f. Suppose f is partially convex with respect to $\mathbf{x}$, i.e. $f(\cdot,\mathbf{y})$ is convex on X for each $\mathbf{y} \in Y$. For any $Z \in \mathbb{I}Y$, let $f_{X\times Z}^{\text{cv,env}} : X \times Z \to \mathbb{R}$ denote the convex envelope of f on $X \times Z$. Assume that for each $\bar{\mathbf{x}} \in X$, there exists a subgradient $\mathbf{s}(\mathbf{y};\bar{\mathbf{x}}) \in \partial_{\mathbf{x}} f(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\bar{\mathbf{x}}}$ such that each $s_i(\mathbf{y};\bar{\mathbf{x}})$, $i = 1,\cdots,n_x$, is Lipschitz continuous on Y with Lipschitz constant M_s. Then, the reduced-space scheme of convex envelopes $\left(f_{X\times Z}^{\text{cv,env}}\right)_{Z\in\mathbb{I}Y}$ has pointwise convergence of order at least one on Y.*

Proof We wish to prove the existence of a constant $\tau \geq 0$ such that

$$\max_{(\mathbf{x},\mathbf{y})\in X\times Z} |f(\mathbf{x},\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x},\mathbf{y})| \leq \tau w(Z), \quad \forall Z \in \mathbb{I}Y.$$

Note that the existence of the maximum in the above expression follows from the (Lipschitz) continuity of f, Lemma 4, and the compactness of $X \times Y$. Consider $Z \in \mathbb{I}Y$, and let $(\mathbf{x}_Z^*,\mathbf{y}_Z^*) \in \underset{(\mathbf{x},\mathbf{y})\in X\times Z}{\arg\max} |f(\mathbf{x},\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x},\mathbf{y})|$. We have

$$\begin{aligned}\max_{(\mathbf{x},\mathbf{y})\in X\times Z} |f(\mathbf{x},\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x},\mathbf{y})| &= f(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y}_Z^*) \\ &= \max_{\mathbf{y}\in Z} |f(\mathbf{x}_Z^*,\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y})|. \end{aligned} \tag{10}$$

Since $f(\cdot,\mathbf{y})$ is convex on X for each $\mathbf{y} \in Y$, we have

$$\begin{aligned} f(\mathbf{x},\mathbf{y}) &\geq f(\mathbf{x}_Z^*,\mathbf{y}) + \mathbf{s}(\mathbf{y};\mathbf{x}_Z^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_Z^*) \\ &= f(\mathbf{x}_Z^*,\mathbf{y}) + w_Z(\mathbf{x},\mathbf{y}) \\ &\geq f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y}) + w_{X\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y}), \quad \forall (\mathbf{x},\mathbf{y}) \in X \times Z, \end{aligned}$$

where $\mathbf{s}(\mathbf{y};\mathbf{x}_Z^*) \in \partial_{\mathbf{x}} f(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{x}_Z^*}$ is a subgradient of $f(\cdot,\mathbf{y})$ at $\mathbf{x}_Z^*$ such that $s_i(\mathbf{y};\mathbf{x}_Z^*)$, $\forall i \in \{1,\cdots,n_x\}$, is Lipschitz continuous on Z with Lipschitz constant M_s, $f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\cdot)$ denotes the convex envelope of $f(\mathbf{x}_Z^*,\cdot)$ on Z, $w_Z(\mathbf{x},\mathbf{y}) := \mathbf{s}(\mathbf{y};\mathbf{x}_Z^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_Z^*)$ is a function of the form (W), and $w_{X\times Z}^{\text{cv}}$ is a convex relaxation of w_Z on $X\times Z$ of the form (W^{cv}) with first-order (pointwise) convergent schemes of estimators of $\mathbf{s}(\mathbf{y};\mathbf{x}_Z^*)$ used in its construction.

Since f is Lipschitz continuous on $X\times Y$ and $f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\cdot)$ is the convex envelope of $f(\mathbf{x}_Z^*,\cdot)$ on Z, we have from Lemma 7 that

$$\max_{\mathbf{y}\in Z} |f(\mathbf{x}_Z^*,\mathbf{y}) - f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y})| \leq M_f w(Z).$$

Using Lemma 12 with $w_i^B(\mathbf{x}) = (x_i - x_{i,Z}^*)$, $w_i^C(\mathbf{y}) = s_i(\mathbf{y};\mathbf{x}_Z^*)$, $w_{i,X}^{B,\text{L}} = \min_{\mathbf{x}\in X}(x_i - x_{i,Z}^*)$, $w_{i,X}^{B,\text{U}} = \max_{\mathbf{x}\in X}(x_i - x_{i,Z}^*)$, $w_{i,Z}^{C,\text{cv}}(\mathbf{y}) = w_{i,Z}^{C,\text{L}} = \min_{\mathbf{y}\in Z} s_i(\mathbf{y};\mathbf{x}_Z^*)$, and $w_{i,Z}^{C,\text{cc}}(\mathbf{y}) = w_{i,Z}^{C,\text{U}} = \max_{\mathbf{y}\in Z} s_i(\mathbf{y};\mathbf{x}_Z^*)$, we can show the existence of a constant $\bar{\tau} \geq 0$ such that

$$\max_{(\mathbf{x},\mathbf{y})\in X\times Z} \left| w_Z(\mathbf{x},\mathbf{y}) - w_{X\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y}) \right| \leq \bar{\tau} w(Z).$$

From the above two inequalities, we have

$$\max_{(\mathbf{x},\mathbf{y})\in X\times Z} \left| \left(f(\mathbf{x}_Z^*,\mathbf{y}) + w_Z(\mathbf{x},\mathbf{y})\right) - \left(f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y}) + w_{X\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y})\right) \right| \leq \left(M_f + \bar{\tau}\right) w(Z).$$

Using $w_Z(\mathbf{x}_Z^*,\mathbf{y}) = 0$, we obtain

$$\max_{\mathbf{y}\in Z} \left| f(\mathbf{x}_Z^*,\mathbf{y}) - \left(f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y}) + w_{X\times Z}^{\text{cv}}(\mathbf{x}_Z^*,\mathbf{y})\right) \right| \leq \left(M_f + \bar{\tau}\right) w(Z).$$

Since the convex envelope of f on $X\times Z$, $f_{X\times Z}^{\text{cv,env}}$, is, by definition, tighter than the convex relaxation $f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\cdot) + w_{X\times Z}^{\text{cv}}$ at $\mathbf{x}_Z^*$, we have from Equation (10) that

$$\max_{\mathbf{y}\in Z} |f(\mathbf{x}_Z^*,\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y})| \leq \left(M_f + \bar{\tau}\right) w(Z),$$

which proves the existence of τ. □

Note that the assumptions of Lemma 15 are readily satisfied if f is a Lipschitz continuous function of the form (W) that is composed of continuous functions w^A, w_i^B, $\forall i \in Q$, and w^D and Lipschitz continuous functions w_i^C, $\forall i \in Q$. An instance for which the assumptions of Lemma 15 are not satisfied is $f(x,y) = |y||x+y+1|$ with $X = [-1,1]$ and $Y = [-1,1]$. The following examples provide instances for which the assumptions of Lemma 15 are satisfied, but where the functions involved are not in the form (W).

Example 14 Let $X = [-1,1]$, $Y = [-1,1]$, and $f(x,y) = \exp(xy)$. We have $M_f = \sqrt{2}\exp(1)$, $s(y;x) = y\exp(xy)$, and $M_s = 2\exp(1)$ satisfying the assumptions of Lemma 15.

Example 15 Let $X = [-1,1]$, $Y = [-1,1]$, and $f(x,y) = -|y|\sqrt{x+y+3}$. We have $M_f = 4$, $s(y;x) = -\frac{|y|}{2\sqrt{x+y+3}}$, and $M_s = 1$ satisfying the assumptions of Lemma 15.

The next result shows that the reduced-space dual lower bounding scheme has arbitrarily high convergence order at unconstrained points in the reduced-space.

Proposition 3 *Consider Problem* (P) *with* $m_E = 0$. *Suppose* $\mathbf{y}^{\text{S}} \in Y$ *is an unconstrained point in the reduced-space. Furthermore, suppose the reduced-space dual lower bounding scheme has convergence of order* $\beta > 0$ *at* $\mathbf{y}^{\text{S}}$. *Then the reduced-space dual lower bounding scheme has arbitrarily high convergence order at* $\mathbf{y}^{\text{S}}$.

Proof The proof is relegated to Appendix A.3 since it is similar to the proof of Proposition 1. □

The following result establishes first-order convergence of the reduced-space dual lower bounding scheme even in the absence of constraint propagation.

Theorem 9 *Consider Problem (P). Suppose f, g_j, $j=1,\cdots,m_I$, and h_k, $k=1,\cdots,m_E$, are Lipschitz continuous on $X\times Y$ with Lipschitz constants $M_f, M_{g,1},\cdots,M_{g,m_I}, M_{h,1},\cdots,M_{h,m_E}$, respectively, and assume that the assumptions of Lemma 15 hold for* $\mathbf{g}$ *and* $\mathbf{h}$*. Assume, in addition, that Problem (P) is feasible, and that strong duality holds for Problem (P) for* $\mathbf{y}$ *restricted to any feasible point in Y. Furthermore, suppose the set of optimal dual variables for Problem (P) restricted to any feasible $\mathbf{y}\in Y$ is bounded (with the bound independent of* $\mathbf{y}$*). Then the reduced-space dual lower bounding scheme is at least first-order convergent on Y.*

Proof Lemmata 11, 13, and 15 imply that the dual lower bounding scheme is at least first-order convergent at any infeasible point $\mathbf{y}\in Y$ with the prefactor independent of $\mathbf{y}$ (note that the conclusion of Lemma 13 does not depend on the schemes of relaxations of the constraints being in the form (W^{cv})).

Define $F(\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) := \min_{\mathbf{x}\in X} f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})$. We first show that $F(\cdot,\boldsymbol{\mu},\boldsymbol{\lambda})$ is Lipschitz continuous on Y for any $(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{m_I}_+\times\mathbb{R}^{m_E}$. Consider $(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{m_I}_+\times\mathbb{R}^{m_E}$ and $\mathbf{y}_1,\mathbf{y}_2\in Y$. We have

$$
\begin{aligned}
&|F(\mathbf{y}_1,\boldsymbol{\mu},\boldsymbol{\lambda}) - F(\mathbf{y}_2,\boldsymbol{\mu},\boldsymbol{\lambda})| \\
&= \left| \left(\min_{\mathbf{x}\in X} f(\mathbf{x},\mathbf{y}_1) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}_1) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}_1) \right) - \left(\min_{\mathbf{x}\in X} f(\mathbf{x},\mathbf{y}_2) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}_2) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}_2) \right) \right| \\
&\le \max_{\mathbf{x}\in X} \left| (f(\mathbf{x},\mathbf{y}_1) - f(\mathbf{x},\mathbf{y}_2)) + \boldsymbol{\mu}^{\mathrm{T}}(\mathbf{g}(\mathbf{x},\mathbf{y}_1) - \mathbf{g}(\mathbf{x},\mathbf{y}_2)) + \boldsymbol{\lambda}^{\mathrm{T}}(\mathbf{h}(\mathbf{x},\mathbf{y}_1) - \mathbf{h}(\mathbf{x},\mathbf{y}_2)) \right| \\
&\le \max_{\mathbf{x}\in X} |f(\mathbf{x},\mathbf{y}_1) - f(\mathbf{x},\mathbf{y}_2)| + \max_{\mathbf{x}\in X} \left| \boldsymbol{\mu}^{\mathrm{T}}(\mathbf{g}(\mathbf{x},\mathbf{y}_1) - \mathbf{g}(\mathbf{x},\mathbf{y}_2)) \right| + \max_{\mathbf{x}\in X} \left| \boldsymbol{\lambda}^{\mathrm{T}}(\mathbf{h}(\mathbf{x},\mathbf{y}_1) - \mathbf{h}(\mathbf{x},\mathbf{y}_2)) \right| \\
&\le \left(M_f + \sum_{j=1}^{m_I} |\mu_j| M_{g,j} + \sum_{k=1}^{m_E} |\lambda_k| M_{h,k} \right) \|\mathbf{y}_1 - \mathbf{y}_2\|,
\end{aligned}
$$

where Step 2 follows from Lemma 3, and Step 4 follows from the Lipschitz continuity of the functions involved.

Suppose $\mathscr{F}(Y)\neq\emptyset$ and $Z\in\mathbb{I}Y$ such that $Z\cap\mathscr{F}(Y)\neq\emptyset$. Since strong duality holds for Problem (P) with $\mathbf{y}$ restricted to any feasible point in Y, Problem (P) can be equivalently expressed on Z as

$$
\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) = \min_{\mathbf{y}\in Z} \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{m_I}_+\times\mathbb{R}^{m_E}} F(\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}).
$$

By strong duality and $Z\cap\mathscr{F}(Y)\neq\emptyset$, there exists a minimizer $(\mathbf{y}^*_Z,\boldsymbol{\mu}^*_Z,\boldsymbol{\lambda}^*_Z)$ of the above 'dual form' of Problem (P) when $\mathbf{y}$ is restricted to Z. We have

$$
\begin{aligned}
&\left| \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{m_I}_+\times\mathbb{R}^{m_E}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) \right] \right| \\
&= \left| F(\mathbf{y}^*_Z,\boldsymbol{\mu}^*_Z,\boldsymbol{\lambda}^*_Z) - \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{m_I}_+\times\mathbb{R}^{m_E}} \min_{\mathbf{y}\in Z} F(\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \right| \\
&\le \left| F(\mathbf{y}^*_Z,\boldsymbol{\mu}^*_Z,\boldsymbol{\lambda}^*_Z) - \min_{\mathbf{y}\in Z} F(\mathbf{y},\boldsymbol{\mu}^*_Z,\boldsymbol{\lambda}^*_Z) \right| \\
&= |F(\mathbf{y}^*_Z,\boldsymbol{\mu}^*_Z,\boldsymbol{\lambda}^*_Z) - F(\bar{\mathbf{y}}_Z,\boldsymbol{\mu}^*_Z,\boldsymbol{\lambda}^*_Z)| \\
&\le \left(M_f + \sum_{j=1}^{m_I} |\mu^*_{j,Z}| M_{g,j} + \sum_{k=1}^{m_E} |\lambda^*_{k,Z}| M_{h,k} \right) \|\mathbf{y}^*_Z - \bar{\mathbf{y}}_Z\| \\
&\le \left(M_f + \sum_{j=1}^{m_I} M_\infty M_{g,j} + \sum_{k=1}^{m_E} M_\infty M_{h,k} \right) \sqrt{n_y}\, w(Z),
\end{aligned}
$$

where $\bar{\mathbf{y}}_Z \in \arg\min_{\mathbf{y}\in Z} F(\mathbf{y},\boldsymbol{\mu}^*_Z,\boldsymbol{\lambda}^*_Z)$, $M_\infty := \sup_{\mathbf{y}\in Y} \max\left\{ \|\boldsymbol{\mu}^*(\mathbf{y})\|_\infty, \|\boldsymbol{\lambda}^*(\mathbf{y})\|_\infty \right\}$ is an upper bound on the norm of pairs of optimal dual variables $(\boldsymbol{\mu}^*(\mathbf{y}),\boldsymbol{\lambda}^*(\mathbf{y}))\in \arg\max_{\boldsymbol{\mu}\ge\mathbf{0},\boldsymbol{\lambda}} F(\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda})$, and the penultimate step follows from the Lipschitz continuity of $F(\cdot,\boldsymbol{\mu},\boldsymbol{\lambda})$ on Y. □

The assumption that the set of optimal dual variables for Problem (P) restricted to any feasible $\mathbf{y} \in Y$ is bounded can be replaced with the less restrictive assumption that there exists an optimal dual variable pair $(\boldsymbol{\mu}^*(\mathbf{y}), \boldsymbol{\lambda}^*(\mathbf{y})) \in \arg\max_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} F(\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda})$ for each $\mathbf{y} \in Y$ such that $\sup_{\mathbf{y}\in Y} \max\left\{\|\boldsymbol{\mu}^*(\mathbf{y})\|_\infty, \|\boldsymbol{\lambda}^*(\mathbf{y})\|_\infty\right\} \leq M_\infty$.

A corollary of Theorems 7 and 9 is that first-order convergence is guaranteed for the convex relaxation-based reduced-space lower bounding scheme in [10] when first-order pointwise convergent schemes of relaxations on Y are used in its construction. Instead of proving first-order convergence of the lower bounding scheme in [10] at feasible points under the assumption that schemes of bounded optimal dual variables exist, we show that the reduced-space lower bounding scheme in [10] enjoys first-order convergence at any feasible point in the reduced-space under the (less restrictive) assumption that strong duality holds for Problem (P) with $\mathbf{y}$ fixed to the feasible point.

Corollary 5 *Consider Problem* (P). *Suppose the functions* f *and* g_j, $\forall j \in \{1,\cdots,m_I\}$, *are Lipschitz continuous on* $X \times Y$ *with Lipschitz constants* $M_f, M_{g,1},\cdots,M_{g,m_I}$, *respectively, and are each of the form* (W). *Furthermore, suppose functions* h_k, $k=1,\cdots,m_E$, *are Lipschitz continuous on* $X\times Y$ *with Lipschitz constants* $M_{h,1},\cdots,M_{h,m_E}$, *respectively, and are each of the form* ($\mathrm{W_{eq}}$). *Suppose* $\mathbf{y}^{\mathrm{f}} \in Y$ *is a feasible point in the reduced-space and strong duality holds for Problem* (P) *when* $\mathbf{y}$ *is fixed to* $\mathbf{y}^{\mathrm{f}}$. *Let* $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $(g^{\mathrm{cv}}_{j,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $j=1,\cdots,m_I$, *denote continuous schemes of convex relaxations of* $f, g_1,\cdots,g_{m_I}$, *respectively, in* Y *with pointwise convergence orders* $\gamma^{\mathrm{cv}}_f \geq 1$, $\gamma^{\mathrm{cv}}_{g,1} \geq 1,\cdots,\gamma^{\mathrm{cv}}_{g,m_I} \geq 1$ *and corresponding constants* τ^{cv}_f, $\tau^{\mathrm{cv}}_{g,1},\cdots,\tau^{\mathrm{cv}}_{g,m_I}$. *Let* $(h^{\mathrm{cv}}_{k,X(Z)\times Z}, h^{\mathrm{cc}}_{k,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $k=1,\cdots,m_E$, *denote continuous schemes of relaxations of* $h_1,\cdots,h_{m_E}$, *respectively, in* Y *with pointwise convergence orders* $\gamma_{h,1} \geq 1,\cdots,\gamma_{h,m_E} \geq 1$ *and corresponding constants* $\tau_{h,1},\cdots,\tau_{h,m_E}$. *Then, the scheme of lower bounding problems* $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ *proposed in [10] is at least first-order convergent at* $\mathbf{y}^{\mathrm{f}}$.

Proof Let $\left(\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) \in \arg\max_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} F(\mathbf{y}^{\mathrm{f}},\boldsymbol{\mu},\boldsymbol{\lambda})$ be an optimal pair of dual variables for $\mathbf{y}$ fixed to $\mathbf{y}^{\mathrm{f}}$ in Problem (P). Suppose $Z \in \mathbb{I}Y$ with $\mathbf{y}^{\mathrm{f}} \in Z$. Similar to the proof of Theorem 9, we have

$$
\begin{aligned}
&\left| \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{m_I}_+\times\mathbb{R}^{m_E}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) \right] \right| \\
&\leq \left| F\left(\mathbf{y}^{\mathrm{f}}, \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) - \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}^{m_I}_+\times\mathbb{R}^{m_E}} \min_{\mathbf{y}\in Z} F(\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \right| \\
&\leq \left| F\left(\mathbf{y}^{\mathrm{f}}, \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) - \min_{\mathbf{y}\in Z} F\left(\mathbf{y}, \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) \right| \\
&\leq \tau^{\mathrm{f}} w(Z),
\end{aligned}
$$

for some constant $\tau^{\mathrm{f}} \geq 0$. The result then holds as a consequence of Theorem 7 by using $\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}_Z = \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}$, $\boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}_Z = \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}$, $\bar{\mu} = \left\|\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}\right\|_\infty$, and $\bar{\lambda} = \left\|\boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right\|_\infty$ in Theorem 7. □

The following example shows that the convergence order of the reduced-space dual lower bounding scheme may be as low as one at constrained minima.

Example 16 Consider the following instance of Problem (P):

$$
\begin{aligned}
\min_{x,y} \; & -xy \\
\text{s.t. } & x+y \leq 1, \\
& x \in [-1,1], y \in [0,1].
\end{aligned}
$$

The optimal solution is $(x^*, y^*) = (0.5, 0.5)$ with optimal objective value -0.25. When the inequality constraint is dualized, the following dual function is obtained:

$$\begin{aligned} q(\mu) = \min_{x,y} \; & -xy + \mu\,(x+y-1) \\ \text{s.t. } & x \in [-1,1], y \in [0,1]. \end{aligned}$$

Consider $[y^{\mathrm{L}}, y^{\mathrm{U}}] = [0.5-\varepsilon, 0.5+\varepsilon] =: Z \in \mathbb{I}Y$ with $\varepsilon \in (0, 0.5]$. In order to derive the dual function

$$q(\mu) = \min_{\substack{x\in[-1,1] \\ y\in[y^{\mathrm{L}}, y^{\mathrm{U}}]}} -xy + \mu(x+y-1)$$

as an explicit function of μ, we partition the domain of μ to obtain

$$q(\mu) = \begin{cases} (\mu-1)y^{\mathrm{U}}, & \text{if } \mu \leq y^{\mathrm{L}} \\ \min\{(\mu-1)y^{\mathrm{U}}, (\mu+1)y^{\mathrm{L}} - 2\mu\}, & \text{if } y^{\mathrm{L}} \leq \mu \leq y^{\mathrm{U}} \\ (\mu+1)y^{\mathrm{L}} - 2\mu, & \text{if } \mu \geq y^{\mathrm{U}} \end{cases}$$

when the bounds on x are taken to be $[-1,1]$ irrespective of the bounds on y. The dual lower bound can therefore be derived as:

$$\sup_{\mu \geq 0} q(\mu) = \frac{(y^{\mathrm{L}}-1)y^{\mathrm{U}}}{1+0.5(y^{\mathrm{U}}-y^{\mathrm{L}})}.$$

Therefore, for $[y^{\mathrm{L}}, y^{\mathrm{U}}] = [0.5-\varepsilon, 0.5+\varepsilon]$, the dual lower bound can be derived as

$$\sup_{\mu \geq 0} q(\mu) = \frac{(-0.5-\varepsilon)(0.5+\varepsilon)}{1+\varepsilon} = -\frac{(0.5+\varepsilon)^2}{1+\varepsilon}.$$

Consequently,

$$\min_{(x,y)\in\mathscr{F}(Z)} -xy \; - \; \sup_{\mu\geq 0} q(\mu) = -0.25 + \frac{(0.5+\varepsilon)^2}{1+\varepsilon} = \frac{0.75\varepsilon+\varepsilon^2}{1+\varepsilon} \geq 0.75\varepsilon,$$

which implies that the dual lower bounding scheme is at most first-order convergent at y^*.

Remark 14 Example 16 provides an instance of Problem (P) for which both the reduced-space dual lower bounding scheme [9] and the reduced-space lower bounding scheme in [10] (this follows from Lemma 14) are only first-order convergent at the minimizer. Furthermore, for each $y \in [0,1]$, the reduced-space objective function $v : [0,1] \to \mathbb{R}$ can be derived as

$$\begin{aligned} v(y) = \min_{x} \; & -xy \\ \text{s.t. } & x+y \leq 1, \\ & x \in [-1,1], \end{aligned}$$

which reduces to $v(y) = y^2 - y$. It can be seen that $y^* = 0.5$ is an unconstrained minimum of the reduced-space objective $v(y)$, which implies that at least second-order convergence of the reduced-space lower bounding scheme is typically required at y^* to mitigate clustering [7, 39].

Therefore, neither reduced-space lower bounding scheme can be expected to avoid clustering for this example. Note that this is in contrast to the full-space lower bounding schemes in Section 4 which can achieve at least second-order convergence at (x^*, y^*) and thereby mitigate clustering [15].

Note, however, that the use of constraint propagation techniques by reduced-space lower bounding schemes can potentially increase their convergence order as demonstrated by Examples 17 and 18. This demonstrates the importance of constraint propagation techniques in reduced-space lower bounding schemes, which has not been emphasized in [9, 10].

Example 17 Consider the instance of Problem (P) in Example 16 with $Z = [y^{\mathrm{L}}, y^{\mathrm{U}}] \subset [0,1]$, $y^{\mathrm{L}} \leq 0.5$, $y^{\mathrm{U}} \geq 0.5$. Suppose we use constraint propagation to derive $X(Z) = [-1, 1-y^{\mathrm{L}}]$. The dual function can be derived as

$$q(\mu) = \begin{cases} \mu(y^{\mathrm{U}} - y^{\mathrm{L}}) + y^{\mathrm{U}}(y^{\mathrm{L}} - 1), & \text{if } \mu \leq y^{\mathrm{L}} \\ \min\{\mu(y^{\mathrm{U}} - y^{\mathrm{L}}) + y^{\mathrm{U}}(y^{\mathrm{L}} - 1), (\mu+1)y^{\mathrm{L}} - 2\mu\}, & \text{if } y^{\mathrm{L}} \leq \mu \leq y^{\mathrm{U}} \\ (\mu+1)y^{\mathrm{L}} - 2\mu, & \text{if } \mu \geq y^{\mathrm{U}} \end{cases},$$

which yields the dual lower bound

$$\sup_{\mu \geq 0} q(\mu) = \frac{(y^{\mathrm{L}} + y^{\mathrm{U}} - y^{\mathrm{L}} y^{\mathrm{U}})(y^{\mathrm{L}} - 2)}{2 + y^{\mathrm{U}} - 2y^{\mathrm{L}}} + y^{\mathrm{L}}.$$

Consider $y^{\mathrm{L}} = 0.5 - \varepsilon$, $y^{\mathrm{U}} = 0.5 + \varepsilon$ for some $\varepsilon \in (0, 0.5)$. The dual lower bound reduces to

$$\sup_{\mu \geq 0} q(\mu) = \frac{-\varepsilon^3 - 4.5\varepsilon^2 - 0.75\varepsilon - 0.375}{1.5 + 3\varepsilon}.$$

Consequently,

$$\begin{aligned} \min_{(x,y) \in \mathscr{F}(Z)} -xy - \sup_{\mu \geq 0} q(\mu) &= -0.25 + \frac{\varepsilon^3 + 4.5\varepsilon^2 + 0.75\varepsilon + 0.375}{1.5 + 3\varepsilon} \\ &= -0.25 + \frac{1}{1.5}\left(\varepsilon^3 + 4.5\varepsilon^2 + 0.75\varepsilon + 0.375\right)(1 + 2\varepsilon)^{-1} \\ &= -0.25 + \frac{1}{1.5}\left(\varepsilon^3 + 4.5\varepsilon^2 + 0.75\varepsilon + 0.375\right)\left(1 - 2\varepsilon + 4\varepsilon^2 + O(\varepsilon^3)\right) \\ &= 3\varepsilon^2 + O(\varepsilon^3) \\ &\leq \tau\varepsilon^2, \end{aligned}$$

for some constant $\tau > 0$ (where we may assume that the above inequality holds for $\varepsilon = 0.5$ as well).

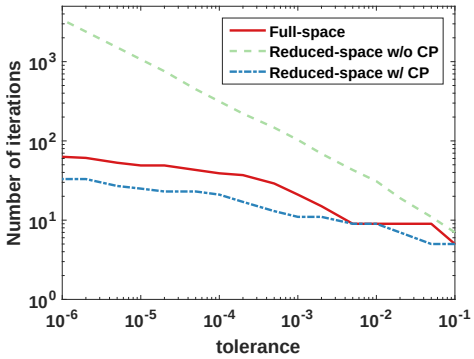
Consider any nondegenerate interval $Z = [y^{\mathrm{L}}, y^{\mathrm{U}}] \subset [0,1]$ with $0.5 \in Z$ and construct $\bar{Z} \supset Z$ such that $\bar{Z} = [y^* - \varepsilon, y^* + \varepsilon]$ with $\varepsilon = \max\{y^{\mathrm{U}} - y^*, y^* - y^{\mathrm{L}}\}$. We have

$$\begin{aligned} \min_{(x,y) \in \mathscr{F}(Z)} -xy - \sup_{\mu \geq 0} \min_{(x,y) \in X(Z) \times Z} [-xy + \mu g(x,y)] &\leq \min_{(x,y) \in \mathscr{F}(\bar{Z})} -xy - \sup_{\mu \geq 0} \min_{(x,y) \in X(Z) \times \bar{Z}} [-xy + \mu g(x,y)] \\ &\leq \tau\varepsilon^2 \\ &\leq \tau w(Z)^2, \end{aligned}$$

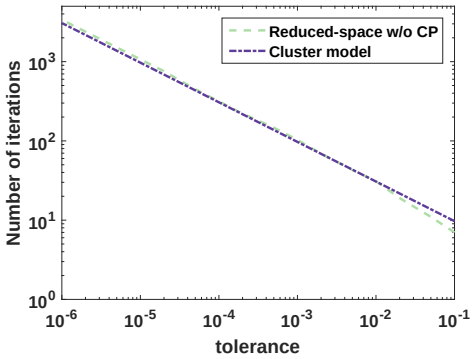
which implies that the reduced-space dual lower bounding scheme with constraint propagation is second-order convergent at y^*.

Figure 1 illustrates the performance of the lower bounding schemes considered in this work in a bare-bones branch-and-bound framework for Examples 16 and 17. The branch-and-bound framework was implemented in MATLAB®, and the (convex) lower bounding problems were solved using the CVX [12] package. The lowest lower bound node selection rule and the interval bisection branching rule (which bisects the domain of the variable whose interval has the largest width) were used by the branch-and-bound algorithm. Since Example 16 is not particularly challenging, it is assumed that a local solver finds its global solution at the root node of the branch-and-bound tree (i.e., the upper bound is set to the optimal objective value at the root node). In addition, the bounds on x and y were modified to $\left[-1, 1 - \frac{\sqrt{3}}{100}\right]$ and $\left[\frac{\sqrt{2}}{100}, 1\right]$, respectively, to prevent the full-space lower bounding schemes from branching at the optimal solution and (fortuitously) converging early (this modification enables a truer characterization of the convergence rates of the lower bounding schemes).

Figure 1a plots the number of iterations of the branch-and-bound algorithm versus the (absolute) termination tolerance for the full-space lower bounding schemes, the reduced-space lower bounding schemes without



(a) Comparison of the number of branch-and-bound iterations versus the termination tolerance between the different lower bounding schemes



(b) Comparison of the number of branch-and-bound iterations of the reduced-space lower bounding schemes without constraint propagation with the predictions from the cluster problem model for different termination tolerances

Fig. 1: (Left Plot) Plots of the number of iterations of the branch-and-bound algorithm versus the absolute termination tolerance for the lower bounding schemes considered in this work for Example 16. The solid line indicates the number of iterations of the full-space lower bounding schemes, the dashed line indicates the number of iterations of the reduced-space lower bounding schemes without constraint propagation, and the dash-dotted line indicates the number of iterations of the reduced-space lower bounding schemes with constraint propagation. (Right Plot) Comparison of the number of iterations of the reduced-space branch-and-bound algorithms without constraint propagation for Example 16 with the corresponding cluster problem model. The dashed line indicates the number of iterations of the reduced-space lower bounding schemes without constraint propagation, and the dash-dotted line indicates the predicted number of iterations from the cluster problem model.

constraint propagation (see Example 16), and the reduced-space lower bounding schemes with constraint propagation (see Example 17). Note that both full-space (reduced-space) lower bounding schemes considered in this work result in the same lower bound for this problem (see Lemma 14). It can be seen that the full-space lower bounding schemes and the reduced-space lower bounding schemes with constraint propagation perform significantly better than the reduced-space lower bounding schemes without constraint propagation for small tolerances, and that they exhibit a much more favorable scaling with a decrease in the termination tolerance as well. Furthermore, the advantage of using constraint propagation techniques in the reduced-space lower bounding schemes is evident, and its use puts the reduced-space lower bounding schemes at an advantage compared to the full-space lower bounding schemes. Figure 1b illustrates that the dependence of the number of iterations on the termination tolerance for the reduced-space lower bounding schemes without constraint propagation is in good agreement with their associated cluster problem models (see [15, Corollary 4] for the details of the cluster problem model). Note that the prediction of the number of iterations from the cluster problem model in Figure 1b is obtained by fitting the prefactor in the cluster model (i.e., intercept of the line in the plot; the slope of the line is determined by the cluster model using the estimate of the convergence order of the lower bounding scheme obtained from this work) against the number of iterations obtained from the computational experiments. It is worth mentioning at this stage that only basic versions of the lower bounding schemes considered in this work have been used to generate Figure 1; the performance of the lower bounding schemes may be significantly different if they are implemented within a state-of-the-art branch-and-bound framework that solves additional subproblems to speed up their convergence.

The following example illustrates another instance of Problem (P) for which constraint propagation plays a crucial rule in boosting the convergence order of the convex relaxation-based reduced-space lower bounding scheme in [10].

Example 18 Consider the following instance of Problem (P):

$$\begin{aligned}\min_{x,y}\ & \exp(x)-4x+y\\ \text{s.t. } & x^2+x\exp(3-y)\le 10,\\ & x\in[0.5,2], y\in[-1,1].\end{aligned}$$

The optimal solution of the above problem, which is a constrained minimum, is $(x^*,y^*)\approx(1.029,0.838)$ $\left(\text{the 'exact' optimal solution can be determined as follows: } x^* \text{ is the (unique real) root of the function } (4-\exp(x))(10x-x^3)-x^2-10 \text{ in } [0.5,2], \text{ and } y^*:=3-\ln\left(\frac{10-(x^*)^2}{x^*}\right)\right)$ with optimal objective value approximately equal to -0.480. The reader can verify that (x^*,y^*,μ^*) is a KKT point for Problem (P), where $\mu^*:=\frac{1}{x^*\exp(3-y^*)}$. This implies, in particular, that the full-space lower bounding schemes in Section 4 can be designed to be at least second-order convergent at (x^*,y^*) (see Theorem 5 and Corollary 4). The reader can also verify that second-order convergence of the lower bounding scheme may be sufficient to mitigate the cluster problem around (x^*,y^*) [15].

Since all of the functions in the above instance of Problem (P) are in the form (W), both the reduced-space lower bounding schemes considered in this section can be employed to solve it. The ensuing arguments show that the convex relaxation-based reduced-space lower bounding scheme in [10] is only first-order convergent at y^* when constraint propagation techniques are not used.

Consider $[y^L,y^U]:=[y^*-\varepsilon,y^*+\varepsilon]=:Z\in\mathbb{I}Y$ with $0<\varepsilon\le 0.1$. The reduced-space lower bounding scheme in [10] yields

$$\begin{aligned}\mathscr{O}(Z)=\min_{x,y}\ & \exp(x)-4x+y\\ \text{s.t. } & x^2+2\exp(3-y)+x\exp(3-y^L)-2\exp(3-y^L)\le 10,\\ & x^2+0.5\exp(3-y)+x\exp(3-y^U)-0.5\exp(3-y^U)\le 10,\\ & x\in[0.5,2], y\in[y^L,y^U].\end{aligned}$$

Note that the point

$$(x_Z^f,y_Z^f):=\left(\frac{\sqrt{(\exp(3-y^U))^2+40+2(\exp(3-y^U)-\exp(3-y^*))}-\exp(3-y^U)}{2},y^*\right)$$

is feasible for the above lower bounding scheme with objective value $\exp(x_Z^f)-4x_Z^f+y_Z^f$. Furthermore,

$$\begin{aligned}x_Z^f-x^*=&\left(\frac{\sqrt{(\exp(3-y^U))^2+40+2(\exp(3-y^U)-\exp(3-y^*))}-\exp(3-y^U)}{2}-\right.\\ &\left.\frac{\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)}{2}\right)\\ \ge\ & 0.2\varepsilon+o(\varepsilon),\end{aligned}$$

where the details pertaining to the derivation of the above inequality are presented in Appendix A.4. Therefore, we have

$$\begin{aligned}\min_{(x,y)\in\mathscr{F}(Z)} f(x,y) \; - \min_{(x,y)\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) &\geq \left(\exp(x^*) - 4x^* + y^*\right) - \left(\exp(x^{\mathrm{f}}_Z) - 4x^{\mathrm{f}}_Z + y^{\mathrm{f}}_Z\right)\\ &= \left(\exp(x^*) - \exp(x^{\mathrm{f}}_Z)\right) + 4\left(x^{\mathrm{f}}_Z - x^*\right)\\ &= \left(4 - \exp(x^*)\right)\left(x^{\mathrm{f}}_Z - x^*\right) + o\left(\left|x^{\mathrm{f}}_Z - x^*\right|\right)\\ &\geq \left(x^{\mathrm{f}}_Z - x^*\right) + o\left(\left|x^{\mathrm{f}}_Z - x^*\right|\right)\\ &\geq 0.2\varepsilon + o(\varepsilon)\\ &\geq 0.05 w(Z)\end{aligned}$$

for $\varepsilon \ll 1$, which establishes that the reduced-space lower bounding scheme in [10] has at most first-order convergence at y^* (note that first-order convergence of the scheme follows from Corollary 5). This is rather unfortunate because y^* can be seen to be an unconstrained minimizer of the reduced-space objective function $v : [-1,1] \to \mathbb{R}$, which can be derived (around $y = y^*$) to be

$$v(y) = \exp\left(x^*(y)\right) - 4x^*(y) + y, \quad \forall y \in [0.5, 1],$$

where $x^* : [0.5,1] \ni y \longmapsto [0.5,2]$ is given by

$$x^*(y) := \frac{\sqrt{\left(\exp(3-y)\right)^2 + 40} - \exp(3-y)}{2},$$

which implies that at least second-order convergence of the reduced-space lower bounding scheme at y^* is typically required to mitigate clustering [7, 39].

We next show that when constraint propagation is used to infer (exact) bounds for x on Z, second-order convergence of the reduced-space lower bounding scheme in [10] can be achieved. Note that for $[y^{\mathrm{L}}, y^{\mathrm{U}}] := [y^* - \varepsilon, y^* + \varepsilon] =: Z \in \mathbb{I}Y$ with $0 < \varepsilon \leq 0.1$, the best possible (interval) bounds that can be obtained for x are $x \in X(Z) := [x^{\mathrm{L}}_Z, x^{\mathrm{U}}_Z]$ with

$$x^{\mathrm{L}}_Z = 0.5, \quad x^{\mathrm{U}}_Z = \frac{\sqrt{\left(\exp(3-y^{\mathrm{U}})\right)^2 + 40} - \exp(3-y^{\mathrm{U}})}{2}.$$

The reduced-space lower bounding scheme in [10] with constraint propagation yields

$$\begin{aligned}\mathscr{O}(Z) = \min_{x,y}\; & \exp(x) - 4x + y\\ \text{s.t. } & x^2 + x^{\mathrm{U}}_Z \exp(3-y) + x\exp(3-y^{\mathrm{L}}) - x^{\mathrm{U}}_Z \exp(3-y^{\mathrm{L}}) \leq 10,\\ & x^2 + 0.5\exp(3-y) + x\exp(3-y^{\mathrm{U}}) - 0.5\exp(3-y^{\mathrm{U}}) \leq 10,\\ & x \in [0.5, x^{\mathrm{U}}_Z], y \in [y^{\mathrm{L}}, y^{\mathrm{U}}].\end{aligned}$$

By noticing that the first constraint in the above relaxation of Problem (P) is always active at the solution of the relaxed problem, we can reformulate the reduced-space lower bounding problem as

$$\mathscr{O}(Z) = \min_{y\in[y^{\mathrm{L}}, y^{\mathrm{U}}]} \exp\left(\bar{x}_Z(y)\right) - 4\left(\bar{x}_Z(y)\right) + y,$$

where $\bar{x}_Z : Z \ni y \longmapsto [0.5, x^{\mathrm{U}}_Z]$ is given by

$$\bar{x}_Z(y) := \frac{\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2 + 40 + 4x^{\mathrm{U}}_Z\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} - \exp(3-y^{\mathrm{L}})}{2}.$$

We have (see Appendix A.4 for details)

$$\begin{aligned}\bar{x}_Z(y)-x^*(y) &= \frac{\left(\exp(3-y^{\mathrm{L}})+\exp(3-y)+4x_Z^{\mathrm{U}}\right)\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}{2\left(\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{\left(\exp(3-y)\right)^2+40}\right)}- \\ &\qquad \frac{\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}{2} \\ &= \frac{\alpha\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}{2\left(\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{\left(\exp(3-y)\right)^2+40}\right)},\end{aligned}$$

with $0 \le \alpha \le \hat{\tau}\varepsilon + O(\varepsilon^2)$ for some $\hat{\tau} \ge 0$. Consequently, we have $\forall y \in Z$ that

$$\begin{aligned}\bar{x}_Z(y)-x^*(y) &\le \frac{\left(\hat{\tau}\varepsilon+O(\varepsilon^2)\right)\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}{2\left(\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{\left(\exp(3-y)\right)^2+40}\right)} \\ &\le \bar{\tau}\varepsilon^2+O(\varepsilon^3)\end{aligned}$$

for some $\bar{\tau} \ge 0$, since $\exp(3-y^{\mathrm{L}})-\exp(3-y)$ is $O(\varepsilon)$. Note that $\bar{x}_Z(y) \ge x^*(y)$, $\forall Z$. Therefore, on intervals $[y^*-\varepsilon, y^*+\varepsilon] =: Z \in \mathbb{I}Y$ with $0 < \varepsilon \le 0.1$, we have

$$\begin{aligned}\min_{(x,y)\in\mathscr{F}(Z)} f(x,y) \;-\; \min_{(x,y)\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) &= \min_{y\in Z} f(x^*(y),y) - \min_{y\in Z} f(\bar{x}_Z(y),y) \\ &\le \max_{y\in Z} \left| f(x^*(y),y) - f(\bar{x}_Z(y),y)\right| \\ &= \max_{y\in Z} \left|\exp(x^*(y)) - \exp(\bar{x}_Z(y)) + 4\bar{x}_Z(y) - 4x^*(y)\right| \\ &= \max_{y\in Z} \left|\left(4-\exp(x^*(y))\right)\left(\bar{x}_Z(y)-x^*(y)\right) + o\left(\bar{x}_Z(y)-x^*(y)\right)\right| \\ &\le \max_{y\in Z} \left|2\left(\bar{x}_Z(y)-x^*(y)\right) + o\left(\bar{x}_Z(y)-x^*(y)\right)\right| \\ &\le 2\bar{\tau}\varepsilon^2 + o\left(\varepsilon^2\right) \\ &\le \bar{\tau} w(Z)^2\end{aligned}$$

for $\varepsilon \ll 1$, which establishes second-order convergence of the scheme at y^* when restricted to symmetric intervals around y^*.

Consider any nondegenerate interval $Z = [y^{\mathrm{L}}, y^{\mathrm{U}}] \in \mathbb{I}Y$ with $y^* \in Z$ and $w(Z) \le 0.1$, and construct $\bar{Z} \supset Z$ such that $\bar{Z} = [y^*-\varepsilon, y^*+\varepsilon]$ with $\varepsilon = \max\{y^{\mathrm{U}}-y^*, y^*-y^{\mathrm{L}}\}$. We have

$$\begin{aligned}\min_{(x,y)\in\mathscr{F}(Z)} f(x,y) \;-\; \min_{(x,y)\in\mathscr{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) &\le \min_{(x,y)\in\mathscr{F}(\bar{Z})} f(x,y) \;-\; \min_{(x,y)\in\mathscr{F}^{\mathrm{cv}}(\bar{Z})} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) \\ &\le \bar{\tau} w(\bar{Z})^2 \\ &\le 4\bar{\tau} w(Z)^2,\end{aligned}$$

which implies that the convex relaxation-based reduced-space dual lower bounding scheme with constraint propagation is second-order convergent at y^*.

Finally, we show that the reduced-space dual lower bounding scheme in [9] has at least second-order convergence at y^* even when constraint propagation is not used to infer bounds on x. Consider $[y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}Y$ with $w(Z) \le 0.1$. The feasible region of the original problem on Z is given by

$$\mathscr{F}(Z) = \left\{(x,y) \in [0.5,2] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] : x \le x^*(y)\right\}.$$

The convex hull of the feasible region on Z is given by

$$\mathrm{conv}(\mathscr{F}(Z)) = \left\{(x,y) \in [0.5,2] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] : x \leq x_Z^{*,\mathrm{cc}}(y)\right\},$$

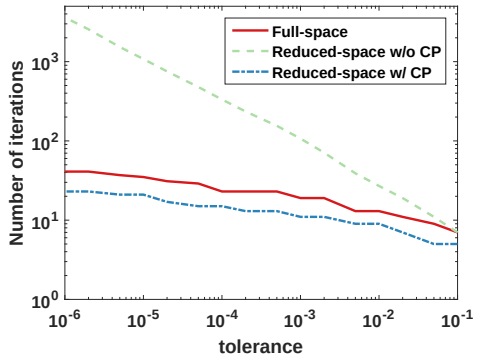
where $x_Z^{*,\mathrm{cc}}$ denotes the concave envelope of x^* on Z. It is not hard to see that $d_H(\mathscr{F}(Z), \mathrm{conv}(\mathscr{F}(Z))) \leq \tilde{\tau} w(Z)^2$ for some $\tilde{\tau} \geq 0$ (this partly follows from the fact that x^* is twice continuously differentiable on $[y^* - 0.1, y^* + 0.1]$ and the fact that $\left(x_Z^{*,\mathrm{cc}}\right)_{Z \in \mathbb{I}Y}$ converges pointwise to x^* with order at least two on $[y^* - 0.1, y^* + 0.1]$). Since the dual lower bounding scheme produces a lower bound that is at least as tight as any convex relaxation-based scheme, we have

$$\begin{aligned}
\min_{(x,y)\in\mathscr{F}(Z)} f(x,y) \; - \; \sup_{\mu\geq 0} \min_{(x,y)\in X(Z)\times Z} \left[f(x,y) + \mu g(x,y)\right] &\leq \min_{(x,y)\in\mathscr{F}(Z)} f(x,y) - \min_{(x,y)\in\mathrm{conv}(\mathscr{F}(Z))} f(x,y) \\
&= f(x^*,y^*) - f(\tilde{x}_Z,\tilde{y}_Z) \\
&\leq f(\hat{x}_Z,\hat{y}_Z) - f(\tilde{x}_Z,\tilde{y}_Z) \\
&\leq M_f \|(\hat{x}_Z,\hat{y}_Z) - (\tilde{x}_Z,\tilde{y}_Z)\| \\
&\leq M_f \tilde{\tau} w(Z)^2,
\end{aligned}$$

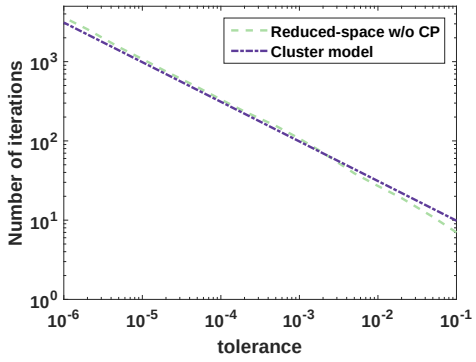
where $(\tilde{x}_Z,\tilde{y}_Z) \in \underset{(x,y)\in\mathrm{conv}(\mathscr{F}(Z))}{\arg\min} f(x,y)$, $(\hat{x}_Z,\hat{y}_Z) \in \mathscr{F}(Z)$ is chosen such that $\|(\hat{x}_Z,\hat{y}_Z) - (\tilde{x}_Z,\tilde{y}_Z)\| \leq \tilde{\tau} w(Z)^2$, and M_f denotes the Lipschitz constant of f on $[0.5,2] \times [-1,1]$. Since the Lagrangian dual-based reduced-space lower bounding scheme is at least first-order convergent at y^* from Theorem 9, it is at least second-order convergent at y^* by analogy to Lemma 5.

Figure 2 illustrates the performance of the convex relaxation-based full-space and reduced-space lower bounding schemes in the bare-bones branch-and-bound implementation for Example 18 (note that we do not consider the Lagrangian dual-based full-space and reduced-space lower bounding schemes for the numerical experiments for this example because we do not have closed-form expressions for the lower bounds obtained using those schemes). Once again, the convex lower bounding problems were solved using the CVX [12] package, and the lowest lower bound node selection rule and the interval bisection branching rule were used by the branch-and-bound algorithm. Since Example 18 is not particularly challenging, we assume that a local solver finds its global solution at the root node of the branch-and-bound tree (i.e., the upper bound is set to the optimal objective value of the problem at the root node).

Figure 2a plots the number of iterations of the branch-and-bound algorithm versus the (absolute) termination tolerance for the full-space convex relaxation-based lower bounding scheme, the reduced-space convex relaxation-based lower bounding scheme without constraint propagation, and the reduced-space convex relaxation-based lower bounding scheme with constraint propagation. It can be seen that the full-space lower bounding scheme and the reduced-space lower bounding scheme with constraint propagation perform significantly better (for small tolerances) and exhibit a much more favorable scaling with a decrease in the termination tolerance compared to the reduced-space lower bounding scheme without constraint propagation. Furthermore, there is a clear advantage in using constraint propagation techniques in the reduced-space lower bounding scheme, and its use makes the performance of the reduced-space lower bounding scheme superior to that of the full-space lower bounding scheme for this example. Figure 2b shows that the number of iterations versus the termination tolerance for the reduced-space lower bounding scheme without constraint propagation closely follows the prediction from its associated cluster problem model (see [15] for the details of the cluster problem model). Note, once again, that the prediction of the number of iterations from the cluster problem model in Figure 2b is obtained by fitting the prefactor in the cluster model against the number of iterations obtained from the computational experiments. We wish to reiterate that only basic versions of the convex relaxation-based lower bounding schemes have been used to generate Figure 2; the performance of the lower bounding schemes may be significantly different if they are implemented within a state-of-the-art branch-and-bound framework that solves additional subproblems to speed up their convergence.



(a) Comparison of the number of branch-and-bound iterations versus the termination tolerance between the convex relaxation-based lower bounding schemes



(b) Comparison of the number of branch-and-bound iterations of the convex relaxation-based reduced-space lower bounding scheme without constraint propagation with the predictions from the cluster problem model for different termination tolerances

Fig. 2: (Left Plot) Plots of the number of iterations of the branch-and-bound algorithm versus the absolute termination tolerance for the full-space and reduced-space convex relaxation-based lower bounding schemes considered in this work for Example 18. The solid line indicates the number of iterations of the convex relaxation-based full-space lower bounding scheme, the dashed line indicates the number of iterations of the convex relaxation-based reduced-space lower bounding scheme without constraint propagation, and the dash-dotted line indicates the number of iterations of the convex relaxation-based reduced-space lower bounding scheme with constraint propagation. (Right Plot) Comparison of the number of iterations of the convex relaxation-based reduced-space branch-and-bound algorithm without constraint propagation with the corresponding cluster problem model. The dashed line indicates the number of iterations of the convex relaxation-based reduced-space lower bounding scheme without constraint propagation, and the dash-dotted line indicates the predicted number of iterations from the cluster problem model.

The following result shows that the reduced-space dual lower bounding scheme is second-order convergent at KKT points even in the absence of constraint propagation when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$.

Theorem 10 *Consider Problem* (P)*, and suppose* $f, g_j,\ j = 1,\cdots,m_I$*, and* $h_k, k = 1,\cdots,m_E$*, are separable in* $\mathbf{x}$ *and* $\mathbf{y}$*. Suppose* $\text{int}(X \times Y)$ *is nonempty, and* f*,* $\mathbf{g}$*, and* $\mathbf{h}$ *are twice continuously differentiable on* $\text{int}(X \times Y)$*. Furthermore, suppose there exists* $(\mathbf{x}^*, \mathbf{y}^*) \in \text{int}(X \times Y)$*,* $\boldsymbol{\mu}^* \in \mathbb{R}^{m_I}_+$*,* $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ *such that* $(\mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ *is a KKT point for Problem* (P)*. The reduced-space dual lower bounding scheme is at least second-order convergent at* $\mathbf{y}^*$*.*

Proof Let $L(\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda}) := f(\mathbf{x}, \mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x}, \mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x}, \mathbf{y})$ denote the Lagrangian of Problem (P). Since we are concerned about the convergence order at the reduced-space feasible point $\mathbf{y}^*$, it suffices to show the existence of $\tau \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y}^* \in Z$,

$$\min_{(\mathbf{x},\mathbf{y}) \in \mathscr{F}(Z)} f(\mathbf{x}, \mathbf{y}) - \sup_{\boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y}) \in X \times Z} L(\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda}) \leq \tau w(Z)^2.$$

We have

$$\begin{aligned}
\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) &\geq \min_{(\mathbf{x},\mathbf{y})\in X\times Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) \\
&\geq \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \left[L(\mathbf{x}^*,\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)+\nabla_{\mathbf{x}}L(\mathbf{x}^*,\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*)\right] \\
&= \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \Big[L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)+\nabla_{\mathbf{x}}L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \\
&\qquad +\left(\nabla_{\mathbf{y}}\left(\nabla_{\mathbf{x}}L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*)\right)\right)^{\mathrm{T}}(\mathbf{y}-\mathbf{y}^*) \\
&\qquad +\nabla_{\mathbf{y}}L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{y}-\mathbf{y}^*)-O(w(Z)^2)\Big] \\
&= \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \left[f(\mathbf{x}^*,\mathbf{y}^*)-O(w(Z)^2)\right] \\
&\geq f(\mathbf{x}^*,\mathbf{y}^*)-O(w(Z)^2).
\end{aligned}$$

Note that we have used the fact that L is partly convex with respect to $\mathbf{x}$ in Step 2, that $L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = f(\mathbf{x}^*,\mathbf{y}^*)$, $\nabla_{\mathbf{x}}L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)=\mathbf{0}$, $\nabla_{\mathbf{y}}L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)=\mathbf{0}$ in Step 4 since it is assumed that $(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is a KKT point for Problem (P), and that $\nabla_{\mathbf{y}}\left(\nabla_{\mathbf{x}}L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*)\right)=\mathbf{0}$ in Step 4 by virtue of the assumption that the Lagrangian is separable in $\mathbf{x}$ and $\mathbf{y}$. Therefore,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \leq O(w(Z)^2),$$

which establishes the existence of τ for all $Z\in\mathbb{I}Y$ with $\mathbf{y}^*\in Z$ by analogy to Lemma 5. □

Note that the assumption of separability in Theorem 10 can be replaced with the weaker assumption that $\nabla^2_{\mathbf{xy}}L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is the zero matrix.

Remark 15 Similar to Corollary 5, it can be shown that the reduced-space lower bounding scheme in [10] has second-order convergence at KKT points even in the absence of constraint propagation when all of the functions in Problem (P) are separable in $\mathbf{x}$ and $\mathbf{y}$ and second-order pointwise convergent schemes of relaxations are used. Furthermore, under the above assumption of separability, the reduced-space lower bounding schemes in [10] and [9] can be shown to possess second-order convergence at infeasible points and unconstrained points in the reduced-space under suitable assumptions on the lower bounding schemes (see Remark 12). Consequently, the convergence properties of the reduced-space lower bounding schemes considered in this section are similar to their counterpart full-space lower bounding schemes in Section 4 when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$. Example 11 provides an instance wherein the convergence order is exactly two at $\mathbf{y}^*$ under the assumptions of Theorem 10.

6 Conclusion

A definition of convergence order for constrained problems has been introduced. The definition reduces to previously developed notions of convergence order for the case of unconstrained problems. An analysis of the convergence order of some full-space and reduced-space branch-and-bound algorithms has been performed.

It has been shown that convex relaxation-based full-space lower bounding schemes enjoy first-order convergence under mild assumptions and second-order convergence at KKT points when second-order pointwise convergent schemes of relaxations of the objective and the constraints are used. Furthermore, the importance of a sufficiently high convergence order at nearly-feasible points has been demonstrated. Lagrangian dual-based full-space lower bounding schemes have been shown to have at least as large a convergence order as convex relaxation-based lower bounding schemes. In addition, it has been shown that Lagrangian dual-based lower bounding schemes where the dual function is not exactly optimized still enjoy first-order convergence.

The convergence order of the reduced-space convex relaxation-based lower bounding scheme of Epperly and Pistikopoulos has been investigated, and it has been shown that the scheme enjoys first-order convergence under certain assumptions. However, their scheme can have as low as first-order convergence even at unconstrained points which can lead to clustering. It has also been shown that the reduced-space dual lower bounding scheme enjoys first-order convergence and that its convergence order may be as low as one for constrained problems. In that regard, the importance of constraint propagation in boosting the convergence order of reduced-space lower bounding schemes has been demonstrated. Furthermore, it has been shown that when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$, the above reduced-space lower bounding schemes can achieve second-order convergence at KKT points, at unconstrained points in the reduced-space, and at infeasible points.

Future work involves determining whether full-space lower bounding schemes can achieve second-order convergence on a neighborhood of constrained minima that are KKT points (such a result may be required to mitigate the cluster problem at such constrained minima - see [15, Proposition 2], for instance), analyzing the convergence orders of some other widely-applicable reduced-space lower bounding schemes in the literature (see, for example, [33]), and determining sufficient conditions on the constraint propagation scheme to ensure second-order convergence of reduced-space lower bounding schemes at constrained minima that satisfy certain regularity conditions.

Acknowledgements The authors would like to thank Garrett Dowdy and Peter Stechlinski for helpful discussions.

A Proofs

A.1 Proof of Proposition 1

Proposition 1 *Consider Problem (P) with $m_E = 0$. Suppose f and g_j, $\forall j \in \{1,\cdots,m_I\}$, are Lipschitz continuous on $X \times Y$. Furthermore, suppose $(\mathbf{x}^S,\mathbf{y}^S) \in X \times Y$ such that $\mathbf{g}(\mathbf{x}^S,\mathbf{y}^S) < \mathbf{0}$ (i.e. $(\mathbf{x}^S,\mathbf{y}^S)$ is a Slater point). The dual lower bounding scheme has arbitrarily high convergence order at $(\mathbf{x}^S,\mathbf{y}^S)$.*

Proof The arguments below are closely related to the proof of Corollary 2.

Since we wish to prove that the dual lower bounding scheme has arbitrarily high convergence order at the feasible point $(\mathbf{x}^S,\mathbf{y}^S)$, it suffices to show that for each $\beta > 0$, there exists $\tau \geq 0$, $\delta > 0$ such that for every $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^S,\mathbf{y}^S) \in Z$ and $w(Z) \leq \delta$,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^T\mathbf{g}(\mathbf{x},\mathbf{y})\right] \leq \tau w(Z)^\beta,$$

and the desired result follows by analogy to Lemma 5 by observing that the dual lower bounding scheme is at least first-order convergent at $(\mathbf{x}^S,\mathbf{y}^S)$.

Let $g_j(\mathbf{x}^S,\mathbf{y}^S) = -\varepsilon_j < 0$, $\forall j \in \{1,\cdots,m_I\}$. Since g_j is continuous for each $j \in \{1,\cdots,m_I\}$, there exists $\delta_j > 0$, $\forall j \in \{1,\cdots,m_I\}$, such that $\|(\mathbf{x},\mathbf{y}) - (\mathbf{x}^S,\mathbf{y}^S)\|_\infty < \delta_j$ implies $|g_j(\mathbf{x},\mathbf{y}) - g_j(\mathbf{x}^S,\mathbf{y}^S)| < \frac{\varepsilon_j}{2}$ (see Lemma 2).

Define $\delta := \min_{j\in\{1,\cdots,m_I\}} \delta_j$, and note that $\delta > 0$. Consider $Z \in \mathbb{I}(X \times Y)$ such that $(\mathbf{x}^S,\mathbf{y}^S) \in Z$ and $w(Z) \leq \delta$. For each $(\mathbf{x},\mathbf{y}) \in Z$, $j \in \{1,\cdots,m_I\}$ we have $|g_j(\mathbf{x},\mathbf{y}) - g_j(\mathbf{x}^S,\mathbf{y}^S)| < \frac{\varepsilon_j}{2}$. Therefore, for each $j \in \{1,\cdots,m_I\}$, $g_j(\mathbf{x},\mathbf{y}) < -\frac{\varepsilon_j}{2} < 0$, $\forall (\mathbf{x},\mathbf{y}) \in Z$. Consequently,

$$\begin{aligned}\sup_{\boldsymbol{\mu}\geq\mathbf{0}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^T\mathbf{g}(\mathbf{x},\mathbf{y})\right] &\geq \min_{(\mathbf{x},\mathbf{y})\in Z} f(\mathbf{x},\mathbf{y}) \\ &= \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y})\end{aligned}$$

since Problem (P) is effectively unconstrained over the small intervals Z around $(\mathbf{x}^S,\mathbf{y}^S)$, which implies $\tau = 0$ and $\delta = \min_{j\in\{1,\cdots,m_I\}} \delta_j$ satisfy the requirements. □

A.2 Proof of Proposition 2

Proposition 2 *Consider Problem (P) with $m_E = 0$. Suppose the functions f and g_j, $j = 1,\cdots,m_I$, are each of the form (W). Let $(f^{cv}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ denote a continuous scheme of convex relaxations of f in Y with convergence order $\beta^{cv}_f > 0$ and corresponding*

constant τ_f^{cv}, $(g_{j,X(Z)\times Z}^{\text{cv}})_{Z\in\mathbb{I}Y}$, $j=1,\cdots,m_I$, *denote continuous schemes of convex relaxations of* $g_1,\cdots,g_{m_I}$, *respectively, in* Y *with pointwise convergence orders* $\gamma_{g,1}^{\text{cv}}>0,\cdots,\gamma_{g,m_I}^{\text{cv}}>0$ *and corresponding constants* $\tau_{g,1}^{\text{cv}},\cdots,\tau_{g,m_I}^{\text{cv}}$.

Suppose $\mathbf{y}^{\text{S}}\in Y$ *is an unconstrained point in the reduced-space, and the scheme of lower bounding problems* $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ *with*

$$(\mathscr{O}(Z))_{Z\in\mathbb{I}Y}:=\left(\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\text{cv}}(Z)} f_{X(Z)\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}Y},$$
$$(\mathscr{I}_C(Z))_{Z\in\mathbb{I}Y}:=\left(\overline{\mathbf{g}}_{X(Z)\times Z}^{\text{cv}}(X(Z)\times Z)\right)_{Z\in\mathbb{I}Y}$$

has convergence of order $\beta\in(0,\beta_f^{\text{cv}}]$ *at* $\mathbf{y}^{\text{S}}$. *Then the scheme of lower bounding problems* $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ *is at least* β_f^{cv}*-order convergent at* $\mathbf{y}^{\text{S}}$.

Proof The proof is similar to the proof of Corollary 2.

Since $\mathbf{y}^{\text{S}}$ is an unconstrained point in the reduced-space and g_j is continuous for each $j\in\{1,\cdots,m_I\}$ by virtue of Assumption 1, $\exists\delta>0$ such that $\forall\mathbf{z}\in Y$ with $\|\mathbf{z}-\mathbf{y}^{\text{S}}\|_\infty\leq\delta$ (see Lemma 2), we have $\mathbf{g}(\mathbf{x},\mathbf{z})<\mathbf{0}$, $\forall\mathbf{x}\in X$.

Consider $Z\in\mathbb{I}Y$ with $\mathbf{y}^{\text{S}}\in Z$ and $w(Z)\leq\delta$. We have $\overline{\mathbf{g}}(X(Z)\times Z)\subset\mathbb{R}^{m_I}_-$ and $\overline{\mathbf{g}}_{X(Z)\times Z}^{\text{cv}}(X(Z)\times Z)\subset\mathbb{R}^{m_I}_-$. Consequently,

$$\begin{aligned}\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y})\ -\ \min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}^{\text{cv}}(Z)} f_{X(Z)\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y}) &= \min_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} f(\mathbf{x},\mathbf{y})\ -\ \min_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} f_{X(Z)\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y})\\ &\leq \tau_f^{\text{cv}} w(Z)^{\beta_f^{\text{cv}}}.\end{aligned}$$

The desired result follows by analogy to Lemma 5 based on the assumption that $(\mathscr{L}(Z))_{Z\in\mathbb{I}Y}$ is at least β-order convergent at $\mathbf{y}^{\text{S}}$. □

A.3 Proof of Proposition 3

Proposition 3 *Consider Problem* (P) *with* $m_E=0$. *Suppose* $\mathbf{y}^{\text{S}}\in Y$ *is an unconstrained point in the reduced-space. Furthermore, suppose the reduced-space dual lower bounding scheme has convergence of order* $\beta>0$ *at* $\mathbf{y}^{\text{S}}$. *Then the reduced-space dual lower bounding scheme has arbitrarily high convergence order at* $\mathbf{y}^{\text{S}}$.

Proof The proof is closely related to the proof of Proposition 1.

Since $\mathbf{y}^{\text{S}}$ is an unconstrained point in the reduced-space and g_j is continuous for each $j\in\{1,\cdots,m_I\}$ by virtue of Assumption 1, there exists $\delta>0$ such that $\forall\mathbf{z}\in Y$ satisfying $\|\mathbf{z}-\mathbf{y}^{\text{S}}\|_\infty\leq\delta$ (see Lemma 2), we have $\mathbf{g}(\mathbf{x},\mathbf{z})<0$, $\forall\mathbf{x}\in X$.

Consider $Z\in\mathbb{I}Y$ with $w(Z)\leq\delta$. Since $\overline{\mathbf{g}}(X(Z)\times Z)\subset\mathbb{R}^{m_I}_-$, Problem (P) can be reformulated as

$$\min_{(\mathbf{x},\mathbf{y})\in\mathscr{F}(Z)} f(\mathbf{x},\mathbf{y})=\min_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} f(\mathbf{x},\mathbf{y}).$$

The dual lower bound can be bounded from below as

$$\sup_{\boldsymbol{\mu}\geq\mathbf{0}}\min_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z}\left[f(\mathbf{x},\mathbf{y})+\boldsymbol{\mu}^{\text{T}}\mathbf{g}(\mathbf{x},\mathbf{y})\right]\geq\min_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} f(\mathbf{x},\mathbf{y}).$$

The desired result follows by analogy to Lemma 5 and the assumption that the dual lower bounding scheme is at least β-order convergent at $\mathbf{y}^{\text{S}}$. □

A.4 Proof of Arguments in Example 18

Proof We first show that $x_Z^{\mathrm{f}} - x^* \geq 0.2\varepsilon + o(\varepsilon)$.

$$
\begin{aligned}
x_Z^{\mathrm{f}} - x^* &= \left(\frac{\sqrt{(\exp(3-y^{\mathrm{U}}))^2 + 40 + 2\left(\exp(3-y^{\mathrm{U}}) - \exp(3-y^*)\right)} - \exp(3-y^{\mathrm{U}})}{2} - \frac{\sqrt{(\exp(3-y^*))^2 + 40} - \exp(3-y^*)}{2} \right) \\
&= \frac{\left(\left(\exp(3-y^{\mathrm{U}})\right)^2 - \left(\exp(3-y^*)\right)^2 \right) + 2\left(\exp(3-y^{\mathrm{U}}) - \exp(3-y^*)\right)}{2\left(\sqrt{(\exp(3-y^{\mathrm{U}}))^2 + 40 + 2\left(\exp(3-y^{\mathrm{U}}) - \exp(3-y^*)\right)} + \sqrt{(\exp(3-y^*))^2 + 40} \right)} + \frac{\left(\exp(3-y^*) - \exp(3-y^{\mathrm{U}})\right)}{2} \\
&= \frac{\left(\exp(3-y^{\mathrm{U}}) + \exp(3-y^*) + 2\right)\left(\exp(3-y^{\mathrm{U}}) - \exp(3-y^*)\right)}{2\left(\sqrt{(\exp(3-y^{\mathrm{U}}))^2 + 40 + 2\left(\exp(3-y^{\mathrm{U}}) - \exp(3-y^*)\right)} + \sqrt{(\exp(3-y^*))^2 + 40} \right)} + \frac{\left(\exp(3-y^*) - \exp(3-y^{\mathrm{U}})\right)}{2} \\
&\geq \frac{\left(\exp(3-y^*) + \exp(3-y^*) + 2\right)\left(\exp(3-y^{\mathrm{U}}) - \exp(3-y^*)\right)}{2\left(\sqrt{(\exp(3-y^*-0.1))^2 + 40 + 2\left(\exp(3-y^*-0.1) - \exp(3-y^*)\right)} + \sqrt{(\exp(3-y^*))^2 + 40} \right)} + \\
&\qquad \frac{\left(\exp(3-y^*) - \exp(3-y^{\mathrm{U}})\right)}{2} \\
&\geq 0.025\left(\exp(3-y^*) - \exp(3-y^{\mathrm{U}})\right) \\
&= 0.025\exp(3-y^*)\varepsilon + o(\varepsilon) \\
&\geq 0.2\varepsilon + o(\varepsilon).
\end{aligned}
$$

Next, we derive an expression for $\bar{x}_Z(y) - x^*(y)$.

$$
\begin{aligned}
\bar{x}_Z(y) - x^*(y) &= \left(\frac{\sqrt{(\exp(3-y^{\mathrm{L}}))^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} - \exp(3-y^{\mathrm{L}})}{2} - \frac{\sqrt{(\exp(3-y))^2 + 40} - \exp(3-y)}{2} \right) \\
&= \frac{\left(\left(\exp(3-y^{\mathrm{L}})\right)^2 - \left(\exp(3-y)\right)^2 \right) + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)}{2\left(\sqrt{(\exp(3-y^{\mathrm{L}}))^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{(\exp(3-y))^2 + 40} \right)} + \\
&\qquad \frac{\left(\exp(3-y) - \exp(3-y^{\mathrm{L}})\right)}{2} \\
&= \frac{\left(\exp(3-y^{\mathrm{L}}) + \exp(3-y) + 4x_Z^{\mathrm{U}}\right)\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)}{2\left(\sqrt{(\exp(3-y^{\mathrm{L}}))^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{(\exp(3-y))^2 + 40} \right)} - \\
&\qquad \frac{\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)}{2}.
\end{aligned} \tag{11}
$$

We next establish the dependence of the different terms in Equation (11) on ε. We first derive an expression for $\exp(3-y^{\mathrm{L}}) + \exp(3-y) + 4x_Z^{\mathrm{U}}$.

$$
\begin{aligned}
&\exp(3-y^{\mathrm{L}}) + \exp(3-y) + 4x_Z^{\mathrm{U}} \\
&= \exp(3-y^*+\varepsilon) + \exp(3-y) + 2\sqrt{(\exp(3-y^*-\varepsilon)^2 + 40} - 2\exp(3-y^*-\varepsilon) \\
&= \exp(3-y^*) + \exp(3-y) + \varepsilon\exp(3-y^*) + O(\varepsilon^2) + 2\sqrt{(\exp(3-y^*))^2\left[1 - 2\varepsilon + O(\varepsilon^2)\right] + 40} \\
&\qquad - 2\exp(3-y^*)\left[1 - \varepsilon + O(\varepsilon^2)\right] \\
&= 2\sqrt{(\exp(3-y^*))^2 + 40} + \exp(3-y) - \exp(3-y^*) + 3\exp(3-y^*)\varepsilon - \frac{2\left(\exp(3-y^*)\right)^2\varepsilon}{\sqrt{(\exp(3-y^*))^2 + 40}} + O(\varepsilon^2).
\end{aligned}
$$

Next, we derive an expression for $4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)$.

$$\begin{aligned}
4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right) &= \left(2\sqrt{(\exp(3-y^*-\varepsilon)^2+40}-2\exp(3-y^*-\varepsilon)\right)\left(\exp(3-y^*+\varepsilon)-\exp(3-y)\right)\\
&= \left(2\sqrt{(\exp(3-y^*))^2+40}-2\exp(3-y^*)\right)\left(\exp(3-y^*+\varepsilon)-\exp(3-y)\right)+O(\varepsilon^2)\\
&= \left(2\sqrt{(\exp(3-y^*))^2+40}-2\exp(3-y^*)\right)\left(\exp(3-y^*)-\exp(3-y)+\exp(3-y^*)\varepsilon\right)+O(\varepsilon^2).
\end{aligned}$$

Finally, we consider $\sqrt{(\exp(3-y^{\mathrm{L}}))^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{(\exp(3-y))^2+40}$.

$$\begin{aligned}
&\sqrt{(\exp(3-y^{\mathrm{L}}))^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{(\exp(3-y))^2+40}\\
&=\sqrt{(\exp(3-y^*+\varepsilon))^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{(\exp(3-y))^2+40}\\
&=\sqrt{(\exp(3-y^*))^2+40}\sqrt{1+\frac{4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)+2\left(\exp(3-y^*)\right)^2\varepsilon+O(\varepsilon^2)}{(\exp(3-y^*))^2+40}}+\sqrt{(\exp(3-y))^2+40}\\
&=\sqrt{(\exp(3-y^*))^2+40}+\sqrt{(\exp(3-y))^2+40}+\frac{(\exp(3-y^*))^2\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}+\\
&\qquad\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)\left(\exp(3-y^*)-\exp(3-y)+\exp(3-y^*)\varepsilon\right)}{\sqrt{(\exp(3-y^*))^2+40}}+O(\varepsilon^2).
\end{aligned}$$

Substituting the above expressions in Equation (11), we get

$$\bar{x_Z}(y)-x^*(y)=\frac{\alpha\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}{2\left(\sqrt{(\exp(3-y^{\mathrm{L}}))^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{(\exp(3-y))^2+40}\right)},$$

with

$$\begin{aligned}
\alpha :=& \sqrt{(\exp(3-y^*))^2+40}-\sqrt{(\exp(3-y))^2+40}+\exp(3-y)-\exp(3-y^*)-\\
&\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)\left(\exp(3-y^*)-\exp(3-y)\right)}{\sqrt{(\exp(3-y^*))^2+40}}+3\exp(3-y^*)\varepsilon-\\
&\frac{3(\exp(3-y^*))^2\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}-\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)\exp(3-y^*)\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}+O(\varepsilon^2)\\
=&\left(\frac{\exp(3-y^*)+\exp(3-y)}{\sqrt{(\exp(3-y^*))^2+40}+\sqrt{(\exp(3-y))^2+40}}-1\right)\left(\exp(3-y^*)-\exp(3-y)\right)-\\
&\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)\left(\exp(3-y^*)-\exp(3-y)\right)}{\sqrt{(\exp(3-y^*))^2+40}}+3\exp(3-y^*)\varepsilon-\\
&\frac{3(\exp(3-y^*))^2\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}-\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)\exp(3-y^*)\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}+O(\varepsilon^2)\\
\leq& \hat{\tau}\varepsilon+O(\varepsilon^2)
\end{aligned}$$

for some $\hat{\tau}\geq 0$ since $y\in Z=[y^{\mathrm{L}},y^{\mathrm{U}}]$ with $w(Z)=O(\varepsilon)$ and each term in the expression for α is $O(\varepsilon)$. Note that $\alpha\geq 0$ (since $\bar{x_Z}(y)\geq x^*(y)$). □

References

1. Adjiman, C.S., Floudas, C.A.: Rigorous convex underestimators for general twice-differentiable problems. Journal of Global Optimization **9**(1), 23–40 (1996)
2. Belotti, P., Lee, J., Liberti, L., Margot, F., Wächter, A.: Branching and bounds tightening techniques for non-convex MINLP. Optimization Methods & Software **24**(4-5), 597–634 (2009)
3. Ben-Tal, A., Eiger, G., Gershovitz, V.: Global minimization by reducing the duality gap. Mathematical Programming **63**(1-3), 193–212 (1994)
4. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. Society for Industrial and Applied Mathematics (2001). DOI 10.1137/1.9780898718829
5. Bompadre, A., Mitsos, A.: Convergence rate of McCormick relaxations. Journal of Global Optimization **52**(1), 1–28 (2012)
6. Bompadre, A., Mitsos, A., Chachuat, B.: Convergence analysis of Taylor models and McCormick-Taylor models. Journal of Global Optimization **57**(1), 75–114 (2013)
7. Du, K., Kearfott, R.B.: The cluster problem in multivariate global optimization. Journal of Global Optimization **5**(3), 253–265 (1994)
8. Dür, M.: Dual bounding procedures lead to convergent Branch–and–Bound algorithms. Mathematical Programming **91**(1), 117–125 (2001)
9. Dür, M., Horst, R.: Lagrange duality and partitioning techniques in nonconvex global optimization. Journal of Optimization Theory and Applications **95**(2), 347–369 (1997)
10. Epperly, T.G.W., Pistikopoulos, E.N.: A reduced space branch and bound algorithm for global optimization. Journal of Global Optimization **11**(3), 287–311 (1997)
11. Floudas, C.A., Gounaris, C.E.: A review of recent advances in global optimization. Journal of Global Optimization **45**(1), 3–38 (2009)
12. Grant, M., Boyd, S.: CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx (2014)
13. Horst, R., Tuy, H.: Global Optimization: Deterministic Approaches, 3rd edn. Springer-Verlag Berlin Heidelberg (1996)
14. Hunter, J.K.: An introduction to real analysis. Department of Mathematics, University of California at Davis (2014)
15. Kannan, R., Barton, P.I.: The cluster problem in constrained global optimization. Submitted (2016)
16. Khan, K.A.: Sensitivity analysis for nonsmooth dynamic systems. Ph.D. thesis, Massachusetts Institute of Technology (2015)
17. Khan, K.A., Watson, H.A.J., Barton, P.I.: Differentiable McCormick relaxations. Journal of Global Optimization **67**(4), 687–729 (2017)
18. Krawczyk, R., Nickel, K.: Die zentrische form in der Intervallarithmetik, ihre quadratische Konvergenz und ihre Inklusionsisotonie. Computing **28**(2), 117–137 (1982)
19. Liberti, L., Pantelides, C.C.: Convex envelopes of monomials of odd degree. Journal of Global Optimization **25**(2), 157–168 (2003)
20. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: Part I: Convex underestimating problems. Mathematical Programming **10**(1), 147–175 (1976)
21. Misener, R., Floudas, C.A.: ANTIGONE: Algorithms for coNTinuous/Integer Global Optimization of Nonlinear Equations. Journal of Global Optimization **59**(2-3), 503–526 (2014)
22. Mitsos, A., Chachuat, B., Barton, P.I.: McCormick-based relaxations of algorithms. SIAM Journal on Optimization **20**(2), 573–601 (2009)
23. Moore, R.E., Kearfott, R.B., Cloud, M.J.: Introduction to Interval Analysis. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2009)
24. Najman, J., Mitsos, A.: Convergence analysis of multivariate McCormick relaxations. Journal of Global Optimization **66**(4), 597–628 (2016)
25. Ratschek, H., Rokne, J.: Computer methods for the range of functions. Mathematics and its Applications. Ellis Horwood Ltd. (1984)
26. Rockafellar, R.T.: Convex Analysis. Princeton University Press (1970)
27. Rote, G.: The convergence rate of the sandwich algorithm for approximating convex functions. Computing **48**(3-4), 337–361 (1992)
28. Sahlodin, A.M., Chachuat, B.: Convex/concave relaxations of parametric ODEs using Taylor models. Computers & Chemical Engineering **35**(5), 844–857 (2011)
29. Schechter, M.: Principles of functional analysis, vol. 36, 2nd edn. American Mathematical Society (2001)
30. Schöbel, A., Scholz, D.: The theoretical and empirical rate of convergence for geometric branch-and-bound methods. Journal of Global Optimization **48**(3), 473–495 (2010)
31. Scholz, D.: Theoretical rate of convergence for interval inclusion functions. Journal of Global Optimization **53**(4), 749–767 (2012)
32. Sopasakis, P., Giraudo, D.: Basic properties of the point-to-set distance function. Mathematics Stack Exchange. http://math.stackexchange.com/questions/107478/basic-properties-of-the-point-to-set-distance-function (Version: 2012-02-10. Accessed: 2017-04-11)
33. Stuber, M.D., Scott, J.K., Barton, P.I.: Convex and concave relaxations of implicit functions. Optimization Methods and Software **30**(3), 1–37 (2015)
34. Tawarmalani, M., Sahinidis, N.V.: Convex extensions and envelopes of lower semi-continuous functions. Mathematical Programming **93**(2), 247–263 (2002)

35. Tawarmalani, M., Sahinidis, N.V.: Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. Mathematical Programming **99**(3), 563–591 (2004)
36. Tawarmalani, M., Sahinidis, N.V.: A polyhedral branch-and-cut approach to global optimization. Mathematical Programming **103**(2), 225–249 (2005)
37. Tsoukalas, A., Mitsos, A.: Multivariate McCormick relaxations. Journal of Global Optimization **59**(2-3), 633–662 (2014)
38. Wechsung, A.: Global optimization in reduced space. Ph.D. thesis, Massachusetts Institute of Technology (2014)
39. Wechsung, A., Schaber, S.D., Barton, P.I.: The cluster problem revisited. Journal of Global Optimization **58**(3), 429–438 (2014)