J Glob Optim manuscript No.
(will be inserted by the editor)

Extended McCormick Relaxation Rules for Handling Empty Arguments Representing Infeasibility

Jason Ye · Joseph K. Scott

Received: date / Accepted: date

Abstract McCormick's relaxation technique is one of the most versatile and commonly used methods for computing the convex relaxations necessary for deterministic global optimization. The core of the method is a set of rules for propagating relaxations through basic arithmetic operations. Computationally, each rule operates on four-tuples describing each input argument in terms of a lower bound value, an upper bound value, a convex relaxation value, and a concave relaxation value. We call such tuples *McCormick objects*. This paper extends McCormick's rules to accommodate input objects that are *empty* (i.e., the convex relaxation value lies above the concave, or both relaxation values lie outside the bounds). Empty McCormick objects provide a natural way to represent infeasibility and are readily generated by McCormick-based domain reduction techniques. The standard McCormick rules are strictly undefined for empty inputs and applying them anyway can yield relaxations that are non-convex/concave on infeasible parts of their domains. In contrast, our extended rules always produce relaxations that are well-defined and convex/concave on their entire domain. This capability has important applications in reduced-space global optimization, global dynamic optimization, and domain reduction.

Keywords McCormick relaxations · Reduced space optimization · Global dynamic optimization · Constraint-based refinement

This material is based upon work supported by the National Science Foundation under Grant No. 1949747.

Jason Ye
Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332, USA
E-mail: jye76@gatech.edu

Joseph K. Scott
Department of Chemical and Biomolecular Engineering, Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332, USA
E-mail: joseph.scott@chbe.gatech.edu

1 Introduction

In 1976, McCormick proposed methods for propagating pairs of convex and concave relaxations through elementary operations including binary addition, binary multiplication, and common univariate functions [4]. By applying these rules recursively, convex and concave relaxations can be constructed for any function that can be built by composition of these elementary operations (i.e., *factorable functions*). As a result, McCormick's technique is one of the most versatile and commonly used methods for computing the convex relaxations necessary for deterministic global optimization. It has proven particularly useful for constructing relaxations of problems where the objective and/or constraints are defined either implicitly (e.g., by the solutions of systems of differential or algebraic equations [11, 12, 18, 10], or as the expected value of a random function [14]) or explicitly but in terms of complex or deeply nested expression trees (e.g., by the output of certain classes of algorithms [5, 21] or neural networks [8]). Consequently, McCormick's method has been heavily studied in recent years. Important contributions include the convergence analysis in [1] and the generalizations in [2, 7, 13, 19].

To state the objective of this paper precisely, it is helpful to outline McCormick's procedure in more detail. Let $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{x} : P \subset \mathbb{R}^{n_p} \to D$ and define the composite function $\mathbf{g} : P \to \mathbb{R}^m$ by $\mathbf{g}(\mathbf{p}) = \mathbf{f}(\mathbf{x}(\mathbf{p}))$, $\forall \mathbf{p} \in P$. Let $X = [\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] \subset D$ be an interval and let $\mathbf{x}^{\mathrm{cv}}, \mathbf{x}^{\mathrm{cc}} : P \to \mathbb{R}^n$. Suppose that X satisfies $\mathbf{x}(\mathbf{p}) \in X$, $\forall \mathbf{p} \in P$, and $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$ are, respectively, convex and concave relaxations of $\mathbf{x}$ on P (i.e., $\mathbf{x}^{\mathrm{cv}}$ is convex, $\mathbf{x}^{\mathrm{cc}}$ is concave, and $\mathbf{x}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{\mathrm{cc}}(\mathbf{p})$, $\forall \mathbf{p} \in P$). Under these assumptions, McCormick's relaxation technique provides a procedure for propagating this data through the operations defining $\mathbf{f}$ to compute analogous bounds and relaxations for $\mathbf{g}$. Computationally, this is done pointwise for each fixed $\mathbf{p} \in P$, with the data for $\mathbf{x}$ and $\mathbf{g}$ represented by tuples $\mathscr{X}(\mathbf{p}) = (\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cv}}(\mathbf{p}), \mathbf{x}^{\mathrm{cc}}(\mathbf{p}))$ and $\mathscr{G}(\mathbf{p}) = (\mathbf{g}^{\mathrm{L}}, \mathbf{g}^{\mathrm{U}}, \mathbf{g}^{\mathrm{cv}}(\mathbf{p}), \mathbf{g}^{\mathrm{cc}}(\mathbf{p}))$, referred to here as *McCormick objects*.

Naturally, it is always assumed that McCormick objects satisfy $\mathbf{x}^{\mathrm{L}} \leq \mathbf{x}^{\mathrm{U}}$, $\mathbf{x}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x}^{\mathrm{cc}}(\mathbf{p})$, and $[\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] \cap [\mathbf{x}^{\mathrm{cv}}(\mathbf{p}), \mathbf{x}^{\mathrm{cc}}(\mathbf{p})] \neq \emptyset$, for all $\mathbf{p} \in P$. These conditions are not only sensible, they are critical to the theoretical arguments establishing the validity of McCormick's relaxations as well as important properties such as inclusion monotonicity [9, 13]. In contrast, the objective of this paper is to present an extension of McCormick's method to handle the case where $\mathscr{X}(\mathbf{p})$ may violate $\mathbf{x}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x}^{\mathrm{cc}}(\mathbf{p})$ or $[\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] \cap [\mathbf{x}^{\mathrm{cv}}(\mathbf{p}), \mathbf{x}^{\mathrm{cc}}(\mathbf{p})] \neq \emptyset$ for some $\mathbf{p}$. Such McCormick objects are called *empty* and are intended to represent the case where the point $\mathbf{p}$ is infeasible in a related optimization problem to be described. Our goal is to extend McCormick's rules to this setting in such a way that the output $\mathscr{G}(\mathbf{p})$ satisfies: (i) $\mathbf{g}^{\mathrm{cv}}$ is well-defined and convex on all of P, (ii) $\mathbf{g}^{\mathrm{cc}}$ is well-defined and concave on all of P, and (iii) we have

$$\mathbf{g}^{\mathrm{L}} \leq \mathbf{g}(\mathbf{p}) \leq \mathbf{g}^{\mathrm{U}} \text{ and } \mathbf{g}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{g}(\mathbf{p}) \leq \mathbf{g}^{\mathrm{cc}}(\mathbf{p}), \quad \forall \mathbf{p} \in P^*, \tag{1}$$

where P^* is the *feasible set*

$$P^* \equiv \{\mathbf{p} \in P : \mathbf{x}(\mathbf{p}) \in X \text{ and } \mathbf{x}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{\mathrm{cc}}(\mathbf{p})\}. \tag{2}$$

The motivation for this extension is to enable advances in McCormick-based algorithms for global dynamic optimization [11], reduced-space global optimization

[18,20], and domain reduction for nonconvex NLPs [22]. All of these methods use McCormick relaxations to construct convex enclosures of feasible sets in a similar way, and while effective methods for tightening these enclosures exist, they can lead to empty McCormick objects that are problematic for subsequent calculations.

To be more specific, in these methods the decision variables are partitioned into two vectors, $(\mathbf{p}, \mathbf{x})$, and some or all of the constraints are used to define $\mathbf{x}$ as a function of $\mathbf{p}$, $\mathbf{x}(\mathbf{p})$. In dynamic optimization, $\mathbf{x}$ contains the state variables at some time, $\mathbf{p}$ contains the remaining decisions, and $\mathbf{x}(\mathbf{p})$ is defined as the parametric solution of the dynamic system. In reduced-space global optimization, $\mathbf{x}$ and $\mathbf{p}$ are dependent and independent variables, respectively, and $\mathbf{x}(\mathbf{p})$ is the (explicit or implicit) solution of a system of equality constraints. Finally, in the domain-reduction strategy in [22], the partition is arbitrary and $\mathbf{x}(\mathbf{p})$ is defined as the set-valued map taking $\mathbf{p}$ into the set of $\mathbf{x}$ values that are feasible for $\mathbf{p}$. Given $\mathbf{x}(\mathbf{p})$, each method then constructs convex and concave relaxations of $\mathbf{x}$ to obtain an enclosure of the feasible set of the form $\{(\mathbf{p}, \mathbf{x}) \in P \times X : \mathbf{x}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x} \leq \mathbf{x}^{\mathrm{cc}}(\mathbf{p})\}$. Once initial values of $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$ have been computed, there is an opportunity to tighten them based on problem constraints or optimality cuts using iterative algorithms akin to the interval contractor and Newton methods commonly used for domain reduction [18,22]. In addition to providing enhanced domain reduction, these tightened relaxations can also be used as inputs to a McCormick relaxation of the objective function, leading to a tighter relaxation that incorporates feasibility information [22]. This type of relaxation refinement has been shown to lead to significantly faster branch-and-bound convergence for several standard and reduced-space global optimization problems in [17,22]. Moreover, although it has not been applied in global dynamic optimization due to technical difficulties described below, a similar approach has been applied for interval-based bounding methods for dynamic systems and shown to lead to substantially tighter enclosures [15,16].

Unfortunately, refining $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$ can lead to empty McCormick objects for some $\mathbf{p}$, indicating that these values are infeasible. This is problematic because any subsequent computation with the object $(\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cv}}(\mathbf{p}), \mathbf{x}^{\mathrm{cc}}(\mathbf{p}))$, such as relaxing the objective function, is undefined. With the exception of [22], the refinement procedures used in previous studies were guaranteed to yield nonempty McCormick objects due to special properties of the constraints being used [10,18]. In [22], the occurrence of an empty object $(v^{\mathrm{L}}, v^{\mathrm{U}}, v^{\mathrm{cv}}, v^{\mathrm{cc}})$ was handled by immediately setting $v^{\mathrm{cv}} = v^{\mathrm{cc}} = \mathrm{NaN}$, indicating the empty set, with the result that all subsequent computations with this object also returned NaN. While this is workable, it has some critical drawbacks. First, when these relaxations are used to form lower bounding problems for branch-and-bound, the objective and/or constraints may have NaN values at some points in the search space, which is problematic for numerical solvers. Second, obtaining a NaN value at some $\mathbf{p} \in P$ only indicates that $\mathbf{p}$ is infeasible, but provides no information about how to explicitly reduce the domain P to exclude $\mathbf{p}$.

Empty objects create even more serious problems in McCormick-based methods for global dynamic optimization [10,11]. In these methods, $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$ are computed as the solutions of a relaxed dynamic system solved using a standard numerical integration code. At each time step, it is necessary to compute relaxations of the differential equations using the current $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$ as input. However, due to technical details of the relaxation theory, these inputs are first modified by assigning either $x_i^{\text{cv}} \leftarrow x_i^{\text{cc}}$ or $x_i^{\text{cc}} \leftarrow x_i^{\text{cv}}$ for some i. This commonly leads to empty objects, even when the relaxations are being computed for values of $\mathbf{p}$ that are feasible in the original problem. Returning a NaN value as in [22] would cause the integration code to terminate, and interpreting this outcome as infeasible would be erroneous. In the current implementations of these methods, hybrid switching conditions must be used to prevent the occurrence of empty objects, leading to higher complexity, higher computational cost, and decreased numerical reliability. Moreover, there is no known sensitivity theory for the class of hybrid systems used, making it impossible to compute valid subgradients for optimization. Finally, there is no clear way to extend the hybrid systems approach to handle the empty objects that would arise from using constraints to refine $\mathbf{x}^{\text{cv}}$ and $\mathbf{x}^{\text{cc}}$ in each time step, as is done to great effect in the interval-based methods in [15,16]. Thus, better methods for handling empty McCormick objects are critical for advancing global dynamic optimization algorithms.

In this paper, we extend McCormick's relaxation method to potentially empty inputs. When the extended method is used to construct relaxations of the general composition $g(\mathbf{p}) = f(\mathbf{x}(\mathbf{p}))$ from relaxations of $\mathbf{x}(\mathbf{p})$, it yields functions g^{cv} and g^{cc} that are convex and concave on the entire $\mathbf{p}$ domain, respectively, and which bound $g(\mathbf{p})$ for all feasible $\mathbf{p}$. This is in contrast to the standard McCormick rules, which are strictly undefined for infeasible $\mathbf{p}$ and can give nonconvex results if applied despite this. The extended relaxations are also shown to be inclusion monotonic and agree with the standard relaxations for all nonempty inputs. After some foundational definitions and results in §2, the extended rules for elementary operations are presented in §3 and their properties are established. In §4, we show that these rules can be composed to compute relaxations of any factorable function with the desired properties. Section 5 presents a simple relaxation refinement method based on linear constraints to demonstrate how empty objects can be generated and propagated through subsequent computations, leading to tighter relaxations. Finally, concluding remarks are given in §6.

2 Extended McCormick Analysis

By analogy to interval arithmetic, McCormick's relaxation rules can be viewed as arithmetic operations on *McCormick objects*. This view has several advantages and closely resembles the implementation of McCormick relaxations in code. This section develops the basic definitions required by this view. We largely follow the development in [9], but with the necessary modifications to accommodate empty McCormick objects.

Definition 1 For any $\mathbf{x}^{\text{L}}, \mathbf{x}^{\text{U}} \in \mathbb{R}^n$, let $[\mathbf{x}^{\text{L}}, \mathbf{x}^{\text{U}}]$ denote the interval $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^{\text{L}} \leq \mathbf{x} \leq \mathbf{x}^{\text{U}}\}$. Let $\mathbb{IR}^n$ denote the space of all nonempty intervals in $\mathbb{R}^n$. Moreover, for any $D \subset \mathbb{R}^n$, define $\mathbb{I}D \equiv \{X \in \mathbb{IR}^n : X \subset D\}$.

Computing a McCormick relaxation of a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ on some $X \in \mathbb{IR}^n$ results in four objects: upper and lower bounds on $\mathbf{f}(\mathbf{x})$ for all $\mathbf{x} \in X$, denoted by $\mathbf{f}^{\text{L}}$ and $\mathbf{f}^{\text{U}}$, and convex and concave relaxations of $\mathbf{f}$ on X, denoted by $\mathbf{f}^{\text{cv}}(\mathbf{x})$ and $\mathbf{f}^{\text{cc}}(\mathbf{x})$. At any single $\mathbf{x} \in X$, this data can be compactly represented by the tuple $\mathcal{F}(\mathbf{x}) = (\mathbf{f}^{\text{L}}, \mathbf{f}^{\text{U}}, \mathbf{f}^{\text{cv}}(\mathbf{x}), \mathbf{f}^{\text{cc}}(\mathbf{x}))$. This motivates the following definition.

Definition 2 Define the space of McCormick objects by

$$\mathbb{MR}^n \equiv \left\{(\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cv}}, \mathbf{x}^{\mathrm{cc}}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n : \mathbf{x}^{\mathrm{L}} \le \mathbf{x}^{\mathrm{U}}\right\}. \tag{3}$$

Moreover, for $D \subset \mathbb{R}^n$, define $\mathbb{M}D \equiv \left\{(\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cv}}, \mathbf{x}^{\mathrm{cc}}) \in \mathbb{MR}^n : [\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] \subset D\right\}$.

McCormick objects are denoted by script capital letters throughout. Moreover, for any $\mathcal{X} \in \mathbb{MR}^n$, the superscripts L, U, cv, and cc will be used to refer to the components of $\mathcal{X}$, as in $\mathcal{X} = (\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cv}}, \mathbf{x}^{\mathrm{cc}})$.

The relaxation data for $\mathbf{f}$ described above can now be viewed as a function of the form $\mathcal{F} : X \to \mathbb{MR}^m$. More generally, operations on McCormick objects can be viewed as functions of the form $\mathcal{G} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$. We refer to functions mapping into the space $\mathbb{MR}^m$ as *McCormick functions*.

We now define *empty* McCormick objects, which are central to this paper.

Definition 3 Define the *enclosure* of a McCormick object $\mathcal{X} \in \mathbb{MR}^n$ as the set

$$\mathrm{Encl}(\mathcal{X}) \equiv \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^{\mathrm{L}} \le \mathbf{x} \le \mathbf{x}^{\mathrm{U}},\ \mathbf{x}^{\mathrm{cv}} \le \mathbf{x} \le \mathbf{x}^{\mathrm{cc}}\}.$$

A McCormick object $\mathcal{X} \in \mathbb{MR}^n$ is said to be *empty* if $\mathrm{Encl}(\mathcal{X}) = \emptyset$ and nonempty otherwise. Finally, a function $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ is said to *preserve nonemptiness on* $\mathcal{D}$ if $\mathcal{F}(\mathcal{X})$ is nonempty for every nonempty $\mathcal{X} \in \mathcal{D}$.

In [9], the space $\mathbb{MR}^n$ was defined to include only nonempty objects. Therefore, none of the methods or results therein apply to empty objects. The extension of $\mathbb{MR}^n$ to include empty objects in Definition 2 is therefore a major distinction between the present work and [9]. The next definition extends the notion of inclusion (i.e., $\mathcal{X}_1 \subset \mathcal{X}_2$) used in [9] to a more general order relation that is well-defined for empty objects.

Definition 4 Let $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{MR}^n$. We say that $\mathcal{X}_1 \preceq \mathcal{X}_2$ if:

1. $[\mathbf{x}_1^{\mathrm{L}}, \mathbf{x}_1^{\mathrm{U}}] \subset [\mathbf{x}_2^{\mathrm{L}}, \mathbf{x}_2^{\mathrm{U}}]$,
2. $\mathbf{x}_1^{\mathrm{cv}} \ge \mathbf{x}_2^{\mathrm{cv}}$,
3. $\mathbf{x}_1^{\mathrm{cc}} \le \mathbf{x}_2^{\mathrm{cc}}$.

Remark 1 It is straightforward to show that $\preceq$ is a partial order on $\mathbb{MR}^n$; i.e., it is reflexive, antisymmetric, and transitive.

Definitions 5–9 introduce several properties of McCormick functions, culminating in the central notion of a *relaxation function*. We then show in Theorem 1 that the cv and cc components of a relaxation function provide convex and concave relaxations with the desired properties outlined in §1.

Definition 5 For any $\mathcal{X}, \mathcal{Y} \in \mathbb{MR}^n$, we say that $\mathcal{X}$ and $\mathcal{Y}$ are *coherent*, or $\mathcal{X}$ *is coherent to* $\mathcal{Y}$, if $[\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] = [\mathbf{y}^{\mathrm{L}}, \mathbf{y}^{\mathrm{U}}]$. A set $\mathcal{D} \subset \mathbb{MR}^n$ is *closed under coherence* if, for every coherent $\mathcal{X}, \mathcal{Y} \in \mathbb{MR}^n$, $\mathcal{X} \in \mathcal{D}$ implies that $\mathcal{Y} \in \mathcal{D}$. If $\mathcal{D}$ is closed under coherence, then $Q \in \mathbb{IR}^n$ is said to be *represented in* $\mathcal{D}$ if there exists $\mathcal{X} \in \mathcal{D}$ with $[\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] = Q$. A function $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ is coherent if $\mathcal{D}$ is closed under coherence and $\mathcal{F}(\mathcal{X})$ is coherent to $\mathcal{F}(\mathcal{Y})$ for every coherent $\mathcal{X}, \mathcal{Y} \in \mathcal{D}$.

Definition 6 A function $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ is *(fully) inclusion monotonic* on $\mathcal{D}$ if $\mathcal{F}(\mathcal{X}_1) \preceq \mathcal{F}(\mathcal{X}_2)$ for every $\mathcal{X}_1, \mathcal{X}_2 \in \mathcal{D}$ satisfying $\mathcal{X}_1 \preceq \mathcal{X}_2$. $\mathcal{F}$ is *coherently inclusion monotonic* on $\mathcal{D}$ if $\mathcal{F}$ is coherent and $\mathcal{F}(\mathcal{X}_1) \preceq \mathcal{F}(\mathcal{X}_2)$ for every coherent $\mathcal{X}_1, \mathcal{X}_2 \in \mathcal{D}$ satisfying $\mathcal{X}_1 \preceq \mathcal{X}_2$.

Definition 7 For every $\lambda \in [0, 1]$ and every coherent $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{MR}^n$ with common interval part $X = [\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}]$, define the convex combination $\mathrm{Conv}(\lambda, \mathcal{X}_1, \mathcal{X}_2) \equiv (\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \lambda\mathbf{x}_1^{\mathrm{cv}} + (1-\lambda)\mathbf{x}_2^{\mathrm{cv}}, \lambda\mathbf{x}_1^{\mathrm{cc}} + (1-\lambda)\mathbf{x}_2^{\mathrm{cc}}) \in \mathbb{MR}^n$.

Definition 8 $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ is coherently concave on $\mathcal{D}$ if $\mathcal{F}$ is coherent and, for every $\lambda \in [0, 1]$ and every coherent $\mathcal{X}, \mathcal{Y} \in \mathcal{D}$,

$$\mathcal{F}(\mathrm{Conv}(\lambda, \mathcal{X}, \mathcal{Y})) \succeq \mathrm{Conv}(\lambda, \mathcal{F}(\mathcal{X}), \mathcal{F}(\mathcal{Y})). \tag{4}$$

Definition 9 Let $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$. A mapping $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ is a *relaxation function* for $\mathbf{f}$ on $\mathcal{D}$ if it is coherently concave on $\mathcal{D}$, coherently inclusion monotonic on $\mathcal{D}$, and every $\mathcal{X} \in \mathcal{D} \cap \mathbb{M}D$ satisfies $\mathbf{f}(\mathbf{x}) \in \mathrm{Encl}(\mathcal{F}(\mathcal{X}))$, $\forall \mathbf{x} \in \mathrm{Encl}(\mathcal{X})$.

As stated in §1, our aim is to develop a method for computing relaxations of composite functions of the form $\mathbf{g}(\mathbf{p}) = \mathbf{f}(\mathbf{x}(\mathbf{p}))$ that satisfy certain properties even when the input objects $\mathcal{X}(\mathbf{p})$ are empty for some $\mathbf{p}$. Theorem 1 shows that this problem is solved by computing a relaxation function for $\mathbf{f}$. In fact, Definition 9 has been designed to include precisely the conditions needed for this result.

Theorem 1 *Let $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{x} : P \subset \mathbb{R}^{n_p} \to D$ and define the composite function $\mathbf{g} : P \to \mathbb{R}^m$ by $\mathbf{g}(\mathbf{p}) = \mathbf{f}(\mathbf{x}(\mathbf{p}))$, $\forall \mathbf{p} \in P$. Let $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ be a relaxation function for $\mathbf{f}$ on $\mathcal{D}$. Let $X = [\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] \in \mathbb{I}D$ be represented in $\mathcal{D}$, let $\mathbf{x}^{\mathrm{cv}}, \mathbf{x}^{\mathrm{cc}} : P \to \mathbb{R}^n$ be convex and concave on P, respectively, and define $\mathcal{X} : P \to \mathbb{MR}^n$ by $\mathcal{X}(\mathbf{p}) \equiv (\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cv}}(\mathbf{p}), \mathbf{x}^{\mathrm{cc}}(\mathbf{p}))$, $\forall \mathbf{p} \in P$. Finally, define $\mathcal{G} : P \to \mathbb{MR}^m$ by*

$$\mathcal{G}(\mathbf{p}) = (\mathbf{g}^{\mathrm{L}}, \mathbf{g}^{\mathrm{U}}, \mathbf{g}^{\mathrm{cv}}(\mathbf{p}), \mathbf{g}^{\mathrm{cc}}(\mathbf{p})) \equiv \mathcal{F}(\mathcal{X}(\mathbf{p})), \quad \forall \mathbf{p} \in P. \tag{5}$$

Then, $\mathbf{g}^{\mathrm{cv}}$ is convex on P, $\mathbf{g}^{\mathrm{cc}}$ is concave on P, and

$$\mathbf{g}^{\mathrm{L}} \leq \mathbf{g}(\mathbf{p}) \leq \mathbf{g}^{\mathrm{U}} \text{ and } \mathbf{g}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{g}(\mathbf{p}) \leq \mathbf{g}^{\mathrm{cc}}(\mathbf{p}), \quad \forall \mathbf{p} \in P^*, \tag{6}$$

where P^* is the feasible set

$$P^* \equiv \{\mathbf{p} \in P : \mathbf{x}(\mathbf{p}) \in X \text{ and } \mathbf{x}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{\mathrm{cc}}(\mathbf{p})\}. \tag{7}$$

Proof Choose any $\mathbf{p} \in P^*$. Since $\mathcal{F}$ is a relaxation function, $\mathcal{D}$ is closed under coherence. Since $[\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}]$ is represented in $\mathcal{D}$, we have $\mathcal{X}(\mathbf{p}) \in \mathcal{D} \cap \mathbb{M}D$. By (7), $\mathbf{x}(\mathbf{p}) \in \mathrm{Encl}(\mathcal{X}(\mathbf{p}))$. Definition 9 then implies that

$$\mathbf{g}(\mathbf{p}) = \mathbf{f}(\mathbf{x}(\mathbf{p})) \in \mathrm{Encl}(\mathcal{F}(\mathcal{X}(\mathbf{p})) = \mathrm{Encl}(\mathcal{G}(\mathbf{p})). \tag{8}$$

Therefore, (6) holds.

Next, choose any $\mathbf{p}_1, \mathbf{p}_2 \in P$ and $\lambda \in [0, 1]$ and define $\mathbf{p}_\lambda \equiv \lambda\mathbf{p}_1 + (1-\lambda)\mathbf{p}_2$. By the convexity and concavity of $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$,

$$\mathbf{x}^{\mathrm{cv}}(\mathbf{p}_\lambda) \leq \lambda\mathbf{x}^{\mathrm{cv}}(\mathbf{p}_1) + (1-\lambda)\mathbf{x}^{\mathrm{cv}}(\mathbf{p}_2), \tag{9}$$

$$\mathbf{x}^{\mathrm{cc}}(\mathbf{p}_\lambda) \geq \lambda\mathbf{x}^{\mathrm{cc}}(\mathbf{p}_1) + (1-\lambda)\mathbf{x}^{\mathrm{cc}}(\mathbf{p}_2). \tag{10}$$

By Definitions 4 and 7, this implies that

$$\mathrm{Conv}(\lambda, \mathcal{X}(\mathbf{p}_1), \mathcal{X}(\mathbf{p}_2)) \preceq \mathcal{X}(\mathbf{p}_\lambda). \tag{11}$$

The objects $\mathrm{Conv}(\lambda, \mathcal{X}(\mathbf{p}_1), \mathcal{X}(\mathbf{p}_2))$ and $\mathcal{X}(\mathbf{p}_\lambda)$ both have interval part $[\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}]$, so they are coherent and both are in $\mathcal{D}$. Then, since $\mathcal{F}$ is coherently inclusion monotonic on $\mathcal{D}$,

$$\mathcal{F}(\mathrm{Conv}(\lambda, \mathcal{X}(\mathbf{p}_1), \mathcal{X}(\mathbf{p}_2))) \preceq \mathcal{F}(\mathcal{X}(\mathbf{p}_\lambda)). \tag{12}$$

Additionally, since $\mathcal{F}$ is coherently concave,

$$\mathcal{F}(\mathrm{Conv}(\lambda, \mathcal{X}(\mathbf{p}_1), \mathcal{X}(\mathbf{p}_2))) \succeq \mathrm{Conv}(\lambda, \mathcal{F}(\mathcal{X}(\mathbf{p}_1)), \mathcal{F}(\mathcal{X}(\mathbf{p}_2))). \tag{13}$$

Combining, this gives

$$\begin{aligned} \mathcal{G}(\mathbf{p}_\lambda) &= \mathcal{F}(\mathcal{X}(\mathbf{p}_\lambda)), \\ &\succeq \mathrm{Conv}(\lambda, \mathcal{F}(\mathcal{X}(\mathbf{p}_1)), \mathcal{F}(\mathcal{X}(\mathbf{p}_2))), \\ &= \mathrm{Conv}(\lambda, \mathcal{G}(\mathbf{p}_1), \mathcal{G}(\mathbf{p}_2)). \end{aligned} \tag{14}$$

This implies that

$$\mathbf{g}^{\mathrm{cv}}(\mathbf{p}_\lambda) \leq \lambda \mathbf{g}^{\mathrm{cv}}(\mathbf{p}_1) + (1-\lambda)\mathbf{g}^{\mathrm{cv}}(\mathbf{p}_2), \tag{15}$$

$$\mathbf{g}^{\mathrm{cc}}(\mathbf{p}_\lambda) \geq \lambda \mathbf{g}^{\mathrm{cc}}(\mathbf{p}_1) + (1-\lambda)\mathbf{g}^{\mathrm{cc}}(\mathbf{p}_2). \tag{16}$$

Since $\mathbf{p}_1$ and $\mathbf{p}_2$ were chosen arbitrarily, it follows that $\mathbf{g}^{\mathrm{cv}}$ and $\mathbf{g}^{\mathrm{cc}}$ are, respectively, convex and concave on P. □

The following corollary shows that relaxation functions also solve the simpler problem of relaxing $\mathbf{f}$ over an interval X. In this case, there is no place for empty objects to arise and the result is essentially equivalent to Lemma 2.4.11 in [9].

Corollary 1 *Let* $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ *and let* $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ *be a* relaxation function *for* $\mathbf{f}$ *on* $\mathcal{D}$*. For any* $X = [\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] \in \mathbb{I}D$ *that is represented in* $\mathcal{D}$*, define* $\mathcal{X}(\mathbf{x}) \equiv (\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}, \mathbf{x})$*,* $\forall \mathbf{x} \in X$*. Finally, define*

$$(\mathbf{f}^{\mathrm{L}}, \mathbf{f}^{\mathrm{U}}, \mathbf{f}^{\mathrm{cv}}(\mathbf{x}), \mathbf{f}^{\mathrm{cc}}(\mathbf{x})) \equiv \mathcal{F}(\mathcal{X}(\mathbf{x})), \quad \forall \mathbf{x} \in X. \tag{17}$$

Then, $\mathbf{f}^{\mathrm{cv}}$ *is convex on* X*,* $\mathbf{f}^{\mathrm{cc}}$ *is concave on* X*, and*

$$\mathbf{f}^{\mathrm{L}} \leq \mathbf{f}(\mathbf{x}) \leq \mathbf{f}^{\mathrm{U}} \textit{ and } \mathbf{f}^{\mathrm{cv}}(\mathbf{x}) \leq \mathbf{f}(\mathbf{x}) \leq \mathbf{f}^{\mathrm{cc}}(\mathbf{x}), \quad \forall \mathbf{x} \in X. \tag{18}$$

Proof The result follows from applying Theorem 1 with $P = X$ and $\mathbf{x}^{\mathrm{cv}}(\mathbf{p}) = \mathbf{x}^{\mathrm{cc}}(\mathbf{p}) = \mathbf{x}(\mathbf{p}) = \mathbf{p}$ for all $\mathbf{p} \in P$. □

Remark 2 Coherent inclusion monotonicity was not required in the original definition of a relaxation function in [9]. This difference is not directly related to the need to handle empty objects here. Rather, it is related to our focus on relaxing composite functions $\mathbf{g}(\mathbf{p}) = \mathbf{f}(\mathbf{x}(\mathbf{p}))$ as in Theorem 1, which is more general than the relaxation considered in Corollary 1 and is the context in which empty objects arise. Coherent inclusion monotonicity is essential for Theorem 1, which has no analogue in [9], but can be done without in Corollary 1, which is essentially Lemma 2.4.11 in [9].

In light of Theorem 1, the main goal is now to develop a method for computing relaxation functions as per Definition 9. As with the existing McCormick relaxation method, we aim to do this by developing relaxation functions for a library of elementary operations and then composing them recursively to construct relaxation functions for more complex functions. This task is taken up in §3. In the remainder of this section, it remains to establish that the properties of relaxation functions are actually preserved under composition. To facilitate this, we first introduce the much simpler concept of a McCormick extension and show that, in the presence of a few other properties, McCormick extensions are always relaxation functions. This results in Theorem 2, which is analogous to the central result of interval analysis linking interval extensions to inclusion functions [6]. Subsequently, we show in Lemma 2 that all of properties required by Theorem 2 are preserved under compositions.

Definition 10 Let $D \subset \mathbb{R}^n$. A set $\mathcal{D} \subset \mathbb{MR}^n$ is a McCormick extension of D if every $\mathbf{x} \in D$ satisfies $(\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}) \in \mathcal{D}$. Let $\mathbf{f} : D \to \mathbb{R}^m$. A mapping $\mathcal{F} : \mathcal{D} \to \mathbb{MR}^m$ is a McCormick extension of $\mathbf{f}$ if $\mathcal{D}$ is a McCormick extension of D and

$$\mathcal{F}((\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x})) = (\mathbf{f}(\mathbf{x}),\mathbf{f}(\mathbf{x}),\mathbf{f}(\mathbf{x}),\mathbf{f}(\mathbf{x})), \quad \forall \mathbf{x} \in D. \tag{19}$$

Lemma 1 *Let $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and let $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ be a McCormick extension of $\mathbf{f}$. If $\mathcal{F}$ is inclusion monotonic on the set of nonempty inputs $\{\mathcal{X} \in \mathcal{D} \cap \mathbb{M}D : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$, then every $\mathcal{X} \in \mathcal{D} \cap \mathbb{M}D$ satisfies $\mathbf{f}(\mathbf{x}) \in \mathrm{Encl}(\mathcal{F}(\mathcal{X}))$, $\forall \mathbf{x} \in \mathrm{Encl}(\mathcal{X})$. In particular, $\mathcal{F}$ preserves nonemptiness on $\mathcal{D} \cap \mathbb{M}D$.*

Proof Choose any $\mathcal{X} \in \mathcal{D} \cap \mathbb{M}D$. If $\mathrm{Encl}(\mathcal{X}) = \emptyset$, then the result holds trivially. Suppose $\mathrm{Encl}(\mathcal{X}) \neq \emptyset$ and choose any $\mathbf{x} \in \mathrm{Encl}(\mathcal{X})$. Then, $\mathbf{x} \in [\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}] \subset D$ and hence $(\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}) \in \mathcal{D}$ and $\mathcal{F}((\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x})) = (\mathbf{f}(\mathbf{x}),\mathbf{f}(\mathbf{x}),\mathbf{f}(\mathbf{x}),\mathbf{f}(\mathbf{x}))$. Therefore, $\mathbf{f}(\mathbf{x}) \in \mathrm{Encl}(\mathcal{F}((\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x})))$. Moreover, since both $(\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x})$ and $\mathcal{X}$ have nonempty enclosures and $(\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}) \preceq \mathcal{X}$, inclusion monotonicity implies that $\mathcal{F}((\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x})) \preceq \mathcal{F}(\mathcal{X})$, and hence $\mathrm{Encl}(\mathcal{F}((\mathbf{x},\mathbf{x},\mathbf{x},\mathbf{x}))) \subset \mathrm{Encl}(\mathcal{F}(\mathcal{X}))$. Therefore, $\mathbf{f}(\mathbf{x}) \in \mathrm{Encl}(\mathcal{F}(\mathcal{X}))$. □

Theorem 2 *Let $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ and let $\mathcal{F} : \mathcal{D} \subset \mathbb{MR}^n \to \mathbb{MR}^m$ be a McCormick extension of $\mathbf{f}$. If $\mathcal{F}$ is coherently concave on $\mathcal{D} \cap \mathbb{M}D$, coherently inclusion monotonic on $\mathcal{D} \cap \mathbb{M}D$, and inclusion monotonic on $\{\mathcal{X} \in \mathcal{D} \cap \mathbb{M}D : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$, then $\mathcal{F}$ is a relaxation function for $\mathbf{f}$ on $\mathcal{D} \cap \mathbb{M}D$.*

Proof The result follows immediately from Definition 9 and Lemma 1. □

The following lemma shows that all of the properties required by Theorem 2, as well as some others, are preserved when McCormick functions carrying these properties are composed.

Lemma 2 *Let $\mathcal{D}_1 \subset \mathbb{MR}^n$ and $\mathcal{D}_2 \subset \mathbb{MR}^m$ be closed under coherence and let $\mathcal{F}_1 : \mathcal{D}_1 \to \mathbb{MR}^m$ and $\mathcal{F}_2 : \mathcal{D}_2 \to \mathbb{MR}^k$. Let $\mathcal{D}_{12} \equiv \{\mathcal{X} \in \mathcal{D}_1 : \mathcal{F}_1(\mathcal{X}) \in \mathcal{D}_2\}$ and consider the composition $\mathcal{F}_2 \circ \mathcal{F}_1 : \mathcal{D}_{12} \to \mathbb{MR}^k$.*

1. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are coherent, then $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherent;*
2. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ preserve nonemptiness on $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively, then $\mathcal{F}_2 \circ \mathcal{F}_1$ preserves nonemptiness on $\mathcal{D}_{12}$;*

3. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are coherently concave on $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively, and $\mathcal{F}_2$ is coherently inclusion monotonic on $\mathcal{D}_2$, then $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherently concave on $\mathcal{D}_{12}$;*
4. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are coherently inclusion monotonic on $\mathcal{D}_1$ and $\mathcal{D}_2$, respectively, then $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherently inclusion monotonic on $\mathcal{D}_{12}$;*
5. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are inclusion monotonic on $\{\mathcal{X} \in \mathcal{D}_1 : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$ and $\{\mathcal{X} \in \mathcal{D}_2 : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$, respectively, and $\mathcal{F}_1$ preserves nonemptiness on $\mathcal{D}_1$, then $\mathcal{F}_2 \circ \mathcal{F}_1$ is inclusion monotonic on $\{\mathcal{X} \in \mathcal{D}_{12} : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$.*
6. *Let $\mathbf{f}_1 : D_1 \subset \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{f}_2 : D_2 \subset \mathbb{R}^m \to \mathbb{R}^k$, define $D_{12} \equiv \{\mathbf{x} \in D_1 : \mathbf{f}_1(\mathbf{x}) \in D_2\}$, and consider $\mathbf{f}_2 \circ \mathbf{f}_1 : D_{12} \to \mathbb{R}^k$. If $\mathcal{F}_1$ and $\mathcal{F}_2$ are McCormick extensions of $\mathbf{f}_1$ and $\mathbf{f}_2$, respectively, then $\mathcal{F}_2 \circ \mathcal{F}_1$ is a McCormick extension of $\mathbf{f}_2 \circ \mathbf{f}_1$.*

Proof

1. To show that $\mathcal{D}_{12}$ is closed under coherence, choose any $\mathcal{X} \in \mathcal{D}_{12}$ and let $\mathcal{Y} \in \mathbb{MR}^n$ be coherent to $\mathcal{X}$. Since $\mathcal{X}$ is in $\mathcal{D}_{12}$, it is also in $\mathcal{D}_1$, and since $\mathcal{D}_1$ is closed under coherence, $\mathcal{Y} \in \mathcal{D}_1$. Since $\mathcal{F}_1$ is coherent, it follows that $\mathcal{F}_1(\mathcal{X})$ and $\mathcal{F}_1(\mathcal{Y})$ are coherent. But $\mathcal{X} \in \mathcal{D}_{12}$ implies that $\mathcal{F}_1(\mathcal{X}) \in \mathcal{D}_2$, and since $\mathcal{D}_2$ is closed under coherence, this implies that $\mathcal{F}_1(\mathcal{Y}) \in \mathcal{D}_2$. Therefore, $\mathcal{Y} \in \mathcal{D}_{12}$, and since $\mathcal{Y}$ was chosen arbitrarily, $\mathcal{D}_{12}$ is closed under coherence.

   To show that $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherent, choose any coherent $\mathcal{X}, \mathcal{Y} \in \mathcal{D}_{12}$. Since $\mathcal{F}_1$ is coherent, it follows that $\mathcal{F}_1(\mathcal{X})$ and $\mathcal{F}_1(\mathcal{Y})$ are coherent. Then, since $\mathcal{F}_2$ is coherent, it follows that $\mathcal{F}_2 \circ \mathcal{F}_1(\mathcal{X})$ and $\mathcal{F}_2 \circ \mathcal{F}_1(\mathcal{Y})$ are coherent. Since $\mathcal{X}$ and $\mathcal{Y}$ were chosen arbitrarily, $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherent.

2. Choose any nonempty $\mathcal{X} \in \mathcal{D}_{12}$. Since $\mathcal{F}_1$ preserves nonemptiness on $\mathcal{D}_1$, $\mathcal{F}_1(\mathcal{X})$ is a nonempty element of $\mathcal{D}_2$, and since $\mathcal{F}_2$ preserves nonemptiness on $\mathcal{D}_2$, $\mathcal{F}_2(\mathcal{F}_1(\mathcal{X}))$ is nonempty as well.

3. It follows from Part 1 of the proof that $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherent. Choose any coherent $\mathcal{X}, \mathcal{Y} \in \mathcal{D}_{12}$ and any $\lambda \in [0, 1]$. Since $\mathcal{F}_1$ is coherently concave on $\mathcal{D}_1$,

$$\mathcal{F}_1(\mathrm{Conv}(\lambda, \mathcal{X}, \mathcal{Y})) \succeq \mathrm{Conv}(\lambda, \mathcal{F}_1(\mathcal{X}), \mathcal{F}_1(\mathcal{Y})) \tag{20}$$

   and $\mathcal{F}_1(\mathcal{X})$ and $\mathcal{F}_1(\mathcal{Y})$ are coherent. Since $\mathcal{F}_2$ is coherently concave on $\mathcal{D}_2$,

$$\mathcal{F}_2(\mathrm{Conv}(\lambda, \mathcal{F}_1(\mathcal{X}), \mathcal{F}_1(\mathcal{Y}))) \succeq \mathrm{Conv}(\lambda, \mathcal{F}_2(\mathcal{F}_1(\mathcal{X})), \mathcal{F}_2(\mathcal{F}_1(\mathcal{Y}))). \tag{21}$$

   To combine these, additionally note that $\mathcal{F}_1(\mathrm{Conv}(\lambda, \mathcal{X}, \mathcal{Y}))$ is coherent to $\mathcal{F}_1(\mathcal{X})$. Since $\mathcal{F}_1(\mathcal{X}) \in \mathcal{D}_2$ and $\mathcal{D}_2$ is closed under coherence, it follows that $\mathcal{F}_1(\mathrm{Conv}(\lambda, \mathcal{X}, \mathcal{Y})) \in \mathcal{D}_2$. Then, since $\mathcal{F}_2$ is coherently inclusion monotonic on $\mathcal{D}_2$, combining (20) and (21) yields

$$\mathcal{F}_2(\mathcal{F}_1(\mathrm{Conv}(\lambda, \mathcal{X}, \mathcal{Y}))) \succeq \mathrm{Conv}(\lambda, \mathcal{F}_2(\mathcal{F}_1(\mathcal{X})), \mathcal{F}_2(\mathcal{F}_1(\mathcal{Y}))), \tag{22}$$

   which shows that $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherently concave on $\mathcal{D}_{12}$.

4. It follows from Part 1 of the proof that $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherent. Choose any coherent $\mathcal{X}, \mathcal{Y} \in \mathcal{D}_{12}$ such that $\mathcal{X} \preceq \mathcal{Y}$. Coherent inclusion monotonicity of $\mathcal{F}_1$ on $\mathcal{D}_1$ gives $\mathcal{F}_1(\mathcal{X}) \preceq \mathcal{F}_1(\mathcal{Y})$. Moreover, $\mathcal{F}_1(\mathcal{X})$ and $\mathcal{F}_1(\mathcal{Y})$ are coherent elements of $\mathcal{D}_2$, so coherent inclusion monotonicity of $\mathcal{F}_2$ on $\mathcal{D}_2$ gives $\mathcal{F}_2(\mathcal{F}_1(\mathcal{X})) \preceq \mathcal{F}_2(\mathcal{F}_1(\mathcal{Y}))$. Therefore, $\mathcal{F}_2 \circ \mathcal{F}_1$ is coherently inclusion monotonic on $\mathcal{D}_{12}$.

5. Choose any $\mathcal{X}, \mathcal{Y} \in \{\mathcal{Z} \in \mathcal{D}_{12} : \mathrm{Encl}(\mathcal{Z}) \neq \emptyset\}$ such that $\mathcal{X} \preceq \mathcal{Y}$. Inclusion monotonicity of $\mathcal{F}_1$ on $\{\mathcal{Z} \in \mathcal{D}_1 : \mathrm{Encl}(\mathcal{Z}) \neq \emptyset\}$ gives $\mathcal{F}_1(\mathcal{X}) \preceq \mathcal{F}_1(\mathcal{Y})$. Moreover, since $\mathcal{F}_1$ preserves nonemptiness on $\mathcal{D}_1$, $\mathcal{F}_1(\mathcal{X})$ and $\mathcal{F}_1(\mathcal{Y})$ are elements of $\mathcal{D}_2$ with nonempty enclosures. Thus, inclusion monotonicity of $\mathcal{F}_2$ on $\{\mathcal{Z} \in \mathcal{D}_2 : \mathrm{Encl}(\mathcal{Z}) \neq \emptyset\}$ gives $\mathcal{F}_2(\mathcal{F}_1(\mathcal{X})) \preceq \mathcal{F}_2(\mathcal{F}_1(\mathcal{Y}))$. Thus, $\mathcal{F}_2 \circ \mathcal{F}_1$ is inclusion monotonic on $\{\mathcal{Z} \in \mathcal{D}_{12} : \mathrm{Encl}(\mathcal{Z}) \neq \emptyset\}$.

6. First it is shown that $\mathbf{x} \in D_{12}$ implies $(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) \in \mathcal{D}_{12}$. For any $\mathbf{x} \in D_{12}$, $\mathbf{x} \in D_1$ implies that $(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) \in \mathcal{D}_1$ because $\mathcal{D}_1$ is a McCormick extension of D_1. Because $\mathcal{F}_1$ is a McCormick extension of $\mathbf{f}_1$ on $\mathcal{D}_1$, we have $\mathcal{F}_1((\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x})) = (\mathbf{f}_1(\mathbf{x}), \mathbf{f}_1(\mathbf{x}), \mathbf{f}_1(\mathbf{x}), \mathbf{f}_1(\mathbf{x}))$. Since $\mathbf{x} \in D_{12}$, we have $\mathbf{f}_1(\mathbf{x}) \in D_2$, which implies that $\mathcal{F}_1((\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x})) \in \mathcal{D}_2$ because $\mathcal{D}_2$ is a McCormick extension of D_2. By definition, this implies $(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) \in \mathcal{D}_{12}$.

   To show that $\mathcal{F}_2 \circ \mathcal{F}_1$ is a McCormick extension of $\mathbf{f}_2 \circ \mathbf{f}_1$ on $\mathcal{D}_{12}$, choose any $\mathbf{x} \in D_{12}$. Since $\mathcal{D}_{12}$ is a McCormick extension of D_{12}, $(\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}) \in \mathcal{D}_{12}$. Then, since $\mathcal{F}_2$ is a McCormick extension of $\mathbf{f}_2$ on $\mathcal{D}_2$, we have

$$\begin{aligned}\mathcal{F}_2(\mathcal{F}_1((\mathbf{x}, \mathbf{x}, \mathbf{x}, \mathbf{x}))) &= \mathcal{F}_2((\mathbf{f}_1(\mathbf{x}), \mathbf{f}_1(\mathbf{x}), \mathbf{f}_1(\mathbf{x}), \mathbf{f}_1(\mathbf{x}))), \\ &= (\mathbf{f}_2(\mathbf{f}_1(\mathbf{x})), \mathbf{f}_2(\mathbf{f}_1(\mathbf{x})), \mathbf{f}_2(\mathbf{f}_1(\mathbf{x})), \mathbf{f}_2(\mathbf{f}_1(\mathbf{x}))). \quad \square\end{aligned}$$

3 Extended McCormick Rules for Elementary Operations

This section presents extended McCormick rules for propagating potentially empty McCormick objects through elementary functions. Specifically, we consider binary addition, binary multiplication, and composition with common univariate functions. In §4, these rules will be combined to construct relaxation functions for arbitrary factorable functions using Lemma 2 and Theorem 2. In preparation, several key properties are proven for each elementary function here, including coherent concavity, inclusion monotonicity, and the McCormick extension property.

3.1 The Cut Operation

We begin by establishing some properties of the Cut operation, which will be used in the definition of the extended McCormick rules. In the following definition and elsewhere, the min and max of vector arguments are taken elementwise.

Definition 11 Let $\mathrm{Cut} : \mathbb{MR}^n \to \mathbb{MR}^n$ be defined for every $\mathcal{X} \in \mathbb{MR}^n$ by

$$\mathrm{Cut}(\mathcal{X}) \equiv (\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \max(\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{cv}}), \min(\mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cc}})). \tag{23}$$

Lemma 3 *For any coherent* $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{MR}^n$ *with common interval part* $X = [\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}]$ *and any* $\lambda \in [0, 1]$,

$$\max(\mathbf{x}^{\mathrm{L}}, \lambda \mathbf{x}_1^{\mathrm{cv}} + (1-\lambda)\mathbf{x}_2^{\mathrm{cv}}) \leq \lambda \max(\mathbf{x}^{\mathrm{L}}, \mathbf{x}_1^{\mathrm{cv}}) + (1-\lambda)\max(\mathbf{x}^{\mathrm{L}}, \mathbf{x}_2^{\mathrm{cv}}), \tag{24}$$

$$\min(\mathbf{x}^{\mathrm{U}}, \lambda \mathbf{x}_1^{\mathrm{cc}} + (1-\lambda)\mathbf{x}_2^{\mathrm{cc}}) \geq \lambda \min(\mathbf{x}^{\mathrm{U}}, \mathbf{x}_1^{\mathrm{cc}}) + (1-\lambda)\min(\mathbf{x}^{\mathrm{U}}, \mathbf{x}_2^{\mathrm{cc}}). \tag{25}$$

Proof The proof follows directly from convexity of $\max(\mathbf{x}^{\mathrm{L}}, \cdot)$ and concavity of $\min(\mathbf{x}^{\mathrm{U}}, \cdot)$ on $\mathbb{R}$. □

Corollary 2 $\mathrm{Cut} : \mathbb{MR}^n \to \mathbb{MR}^n$ *is coherently concave on* $\mathbb{MR}^n$*.*

Proof By Definition 11, Cut is clearly coherent. Choose any $\lambda \in [0, 1]$ and any coherent $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{MR}^n$. It suffices to show that

$$\mathrm{Cut}(\mathrm{Conv}(\lambda, \mathcal{X}_1, \mathcal{X}_2)) \succeq \mathrm{Conv}(\lambda, \mathrm{Cut}(\mathcal{X}_1), \mathrm{Cut}(\mathcal{X}_2)). \tag{26}$$

This is an immediate consequence of (24) and (25). □

Theorem 3 $\mathrm{Cut} : \mathbb{MR}^n \to \mathbb{MR}^n$ *is inclusion monotonic on* $\mathbb{MR}^n$*.*

Proof Choose $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{MR}^n$ such that $\mathcal{X}_1 \preceq \mathcal{X}_2$. We must prove that $\mathrm{Cut}(\mathcal{X}_1) \preceq \mathrm{Cut}(\mathcal{X}_2)$. By Definition 11, it suffices to show that $\max(\mathbf{x}_1^{\mathrm{L}}, \mathbf{x}_1^{\mathrm{cv}}) \geq \max(\mathbf{x}_2^{\mathrm{L}}, \mathbf{x}_2^{\mathrm{cv}})$ and $\min(\mathbf{x}_1^{\mathrm{U}}, \mathbf{x}_1^{\mathrm{cc}}) \leq \min(\mathbf{x}_2^{\mathrm{U}}, \mathbf{x}_2^{\mathrm{cc}})$. Since $\mathcal{X}_1 \preceq \mathcal{X}_2$, we have $\mathbf{x}_1^{\mathrm{cv}} \geq \mathbf{x}_2^{\mathrm{cv}}$ and $\mathbf{x}_1^{\mathrm{L}} \geq \mathbf{x}_2^{\mathrm{L}}$. Since $\max(a, b) \leq \max(a', b')$ if $a \leq a'$ and $b \leq b'$, this implies that $\max(\mathbf{x}_1^{\mathrm{L}}, \mathbf{x}_1^{\mathrm{cv}}) \geq \max(\mathbf{x}_2^{\mathrm{L}}, \mathbf{x}_2^{\mathrm{cv}})$. Analogous arguments show that $\min(\mathbf{x}_1^{\mathrm{U}}, \mathbf{x}_1^{\mathrm{cc}}) \leq \min(\mathbf{x}_2^{\mathrm{U}}, \mathbf{x}_2^{\mathrm{cc}})$. □

Corollary 3 $\mathrm{Cut} : \mathbb{MR}^n \to \mathbb{MR}^n$ *is coherently inclusion monotonic on* $\mathbb{MR}^n$*.*

Proof Since $\mathrm{Cut} : \mathbb{MR}^{n_x} \to \mathbb{MR}^{n_x}$ is fully inclusion monotonic by Theorem 3, it only remains to argue that it is coherent, which follows from (23). □

Theorem 4 $\mathrm{Cut} : \mathbb{MR}^n \to \mathbb{MR}^n$ *preserves nonemptiness on* $\mathbb{MR}^n$*.*

Proof The proof follows immediately from Definitions 3 and 11. □

3.2 Binary Addition

We now introduce the extended McCormick rule for binary addition. In the remainder of the paper, we will refer to binary addition and a few other functions using the triplet notation $(+, \mathbb{R}^2, \mathbb{R})$ that explicitly specifies the domain $\mathbb{R}^2$ and codomain $\mathbb{R}$. This allows us to reuse the same symbol to denote the McCormick version as well without ambiguity, as in $(+, \mathbb{MR}^2, \mathbb{MR})$.

Definition 12 Define $(+, \mathbb{MR}^2, \mathbb{MR})$ by

$$+(\mathcal{X}, \mathcal{Y}) = \mathcal{X} + \mathcal{Y} \equiv (x^{\mathrm{L}} + y^{\mathrm{L}}, x^{\mathrm{U}} + y^{\mathrm{U}}, \bar{x}^{\mathrm{cv}} + \bar{y}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}} + \bar{y}^{\mathrm{cc}}), \tag{27}$$

where $(x^{\mathrm{L}}, x^{\mathrm{U}}, \bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}) = \mathrm{Cut}(\mathcal{X})$ and $(y^{\mathrm{L}}, y^{\mathrm{U}}, \bar{y}^{\mathrm{cv}}, \bar{y}^{\mathrm{cc}}) = \mathrm{Cut}(\mathcal{Y})$.

Aside from the use of Cut, which originates in [9], Definition 12 is the same as McCormick's original rule [4]. However, on account of Definition 2, the domain $\mathbb{MR}^2$ includes empty objects here, whereas it included only non-empty objects in prior work. Therefore, it is still necessary to prove the desired properties of the rule on this extended domain.

Theorem 5 $(+, \mathbb{MR}^2, \mathbb{MR})$ *is a McCormick extension of* $(+, \mathbb{R}^2, \mathbb{R})$ *on* $\mathbb{MR}^2$.

Proof For any $(x, y) \in \mathbb{R}^2$, the object $((x, y), (x, y), (x, y), (x, y))$ must be in $\mathbb{MR}^2$. Thus, $\mathbb{MR}^2$ is a McCormick extension of $\mathbb{R}^2$. Choose any $(\mathcal{X}, \mathcal{Y}) \in \mathbb{MR}^2$ such that $\mathcal{X} = (x, x, x, x)$ and $\mathcal{Y} = (y, y, y, y)$ and let $\mathcal{Z} = \mathcal{X} + \mathcal{Y}$. It suffices the show that $\mathcal{Z} = (x + y, x + y, x + y, x + y)$, which follows immediately from (27). $\square$

Theorem 6 $(+, \mathbb{MR}^2, \mathbb{MR})$ *is coherently concave on* $\mathbb{MR}^2$.

Proof Choose any coherent $(\mathcal{X}_1, \mathcal{Y}_1), (\mathcal{X}_2, \mathcal{Y}_2) \in \mathbb{MR}^2$ and denote their common interval part by $X \times Y = [x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}]$ without subscripts. Let $\mathcal{Z}_1 \equiv \mathcal{X}_1 + \mathcal{Y}_1$ and $\mathcal{Z}_2 \equiv \mathcal{X}_2 + \mathcal{Y}_2$. Furthermore, choose any $\lambda \in [0, 1]$ and let $\mathcal{X}_\lambda = \mathrm{Conv}(\lambda, \mathcal{X}_1, \mathcal{X}_2)$, $\mathcal{Y}_\lambda = \mathrm{Conv}(\lambda, \mathcal{Y}_1, \mathcal{Y}_2)$, and $\mathcal{Z}_\lambda = \mathcal{X}_\lambda + \mathcal{Y}_\lambda$. We must show that

$$\mathcal{Z}_\lambda \succeq \mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2). \tag{28}$$

Since $(\mathcal{X}_1, \mathcal{Y}_1)$ and $(\mathcal{X}_2, \mathcal{Y}_2)$ are coherent, $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are coherent as well, with common interval part $X + Y$. By Definition 7, this implies that the interval part of $\mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2)$ is also $X + Y$. Similarly, the convex combinations $\mathcal{X}_\lambda$ and $\mathcal{Y}_\lambda$ have interval parts X and Y, respectively, which implies that the interval part of $\mathcal{Z}_\lambda = \mathcal{X}_\lambda + \mathcal{Y}_\lambda$ is $X + Y$ as well. Therefore, the left-hand and right-hand sides of (28) have the same interval parts. Then, to prove (28), it only remains to show that $z_\lambda^{\mathrm{cv}} \le \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}$ and $z_\lambda^{\mathrm{cc}} \ge \lambda z_1^{\mathrm{cc}} + (1-\lambda) z_2^{\mathrm{cc}}$. Using Lemma 3, we have

$$\begin{aligned}
z_\lambda^{\mathrm{cv}} &= \max(x^{\mathrm{L}}, x_\lambda^{\mathrm{cv}}) + \max(y^{\mathrm{L}}, y_\lambda^{\mathrm{cv}}), \\
&= \max(x^{\mathrm{L}}, \lambda x_1^{\mathrm{cv}} + (1-\lambda) x_2^{\mathrm{cv}}) + \max(y^{\mathrm{L}}, \lambda y_1^{\mathrm{cv}} + (1-\lambda) y_2^{\mathrm{cv}}), \\
&\le \lambda \max(x^{\mathrm{L}}, x_1^{\mathrm{cv}}) + (1-\lambda) \max(x^{\mathrm{L}}, x_2^{\mathrm{cv}}) \\
&\quad + \lambda \max(y^{\mathrm{L}}, y_1^{\mathrm{cv}}) + (1-\lambda) \max(y^{\mathrm{L}}, y_2^{\mathrm{cv}}), \\
&= \lambda(\max(x^{\mathrm{L}}, x_1^{\mathrm{cv}}) + \max(y^{\mathrm{L}}, y_1^{\mathrm{cv}})) \\
&\qquad + (1-\lambda)(\max(x^{\mathrm{L}}, x_2^{\mathrm{cv}}) + \max(y^{\mathrm{L}}, y_2^{\mathrm{cv}})), \\
&= \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}.
\end{aligned}$$

This proves that $z_\lambda^{\mathrm{cv}} \le \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}$, and $z_\lambda^{\mathrm{cc}} \ge \lambda z_1^{\mathrm{cc}} + (1-\lambda) z_2^{\mathrm{cc}}$ can be proven analogously. $\square$

Theorem 7 $(+, \mathbb{MR}^2, \mathbb{MR})$ *is inclusion monotonic on* $\mathbb{MR}^2$.

Proof Choose any $(\mathcal{X}_1, \mathcal{Y}_1), (\mathcal{X}_2, \mathcal{Y}_2) \in \mathbb{MR}^2$ with $(\mathcal{X}_1, \mathcal{Y}_1) \preceq (\mathcal{X}_2, \mathcal{Y}_2)$. By Definition 4, it follows that $\mathcal{X}_1 \preceq \mathcal{X}_2$ and $\mathcal{Y}_1 \preceq \mathcal{Y}_2$. Let $\mathcal{Z}_1 = \mathcal{X}_1 + \mathcal{Y}_1$ and $\mathcal{Z}_2 = \mathcal{X}_2 + \mathcal{Y}_2$. We must prove that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$. Since the interval parts of $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are computed using standard interval arithmetic, $[z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] \subset [z_2^{\mathrm{L}}, z_2^{\mathrm{U}}]$ by Theorem 2.3.7 in [9]. It remains to prove that $z_1^{\mathrm{cv}} \ge z_2^{\mathrm{cv}}$ and $z_1^{\mathrm{cc}} \le z_2^{\mathrm{cc}}$. Since $\max(a, b) \le \max(a', b')$ if $a \le a'$ and $b \le b'$, the relations $\mathcal{X}_1 \preceq \mathcal{X}_2$ and $\mathcal{Y}_1 \preceq \mathcal{Y}_2$ imply that

$$\begin{aligned}
z_1^{\mathrm{cv}} &= \max(x_1^{\mathrm{L}}, x_1^{\mathrm{cv}}) + \max(y_1^{\mathrm{L}}, y_1^{\mathrm{cv}}), \\
&\ge \max(x_2^{\mathrm{L}}, x_2^{\mathrm{cv}}) + \max(y_2^{\mathrm{L}}, y_2^{\mathrm{cv}}), \\
&= z_2^{\mathrm{cv}}.
\end{aligned}$$

This proves that $z_1^{\mathrm{cv}} \ge z_2^{\mathrm{cv}}$, and $z_1^{\mathrm{cc}} \le z_2^{\mathrm{cc}}$ can be proven analogously. $\square$

Corollary 4 $(+, \mathbb{MR}^2, \mathbb{MR})$ *is coherently inclusion monotonic on* $\mathbb{MR}^2$*.*

Proof Since $(+, \mathbb{MR}^2, \mathbb{MR})$ is fully inclusion monotonic by Theorem 7, it only remains to argue that it is coherent, which follows immediately from (27). $\square$

Theorem 8 $(+, \mathbb{MR}^2, \mathbb{MR})$ *preserves nonemptiness on* $\mathbb{MR}^2$*.*

Proof The result follows from Lemma 1 with Theorems 5 and 7. $\square$

Theorem 9 $(+, \mathbb{MR}^2, \mathbb{MR})$ *is a relaxation function for* $(+, \mathbb{R}^2, \mathbb{R})$*.*

Proof The result follows from Theorem 2 with Corollary 4 and Theorems 5–7. $\square$

3.3 Binary Multiplication

The conventional McCormick multiplication rule is often written in two different ways. These two forms are equivalent for nonempty objects. However, we show below that they are not equivalent for empty objects, and only one of them retains the desired properties. Therefore, our extended McCormick rule for binary multiplication is a particular form of the conventional rule. The two forms are defined below as $(\times, \mathbb{MR}^2, \mathbb{MR})$ and $(\hat{\times}, \mathbb{MR}^2, \mathbb{MR})$. Subsequently, we show that $(\times, \mathbb{MR}^2, \mathbb{MR})$ satisfies all of the required properties on the extended domain $\mathbb{MR}^2$.

Definition 13 Define $\psi^{\mathrm{cv}}, \psi^{\mathrm{cc}}, \hat{\psi}^{\mathrm{cv}}, \hat{\psi}^{\mathrm{cc}} : \mathbb{R} \times \mathbb{MR} \to \mathbb{R}$ by

$$\psi^{\mathrm{cv}}(\alpha, \mathcal{X}) \equiv \begin{cases} \alpha x^{\mathrm{cv}} & \text{if } \alpha \geq 0 \\ \alpha x^{\mathrm{cc}} & \text{otherwise} \end{cases}, \qquad \psi^{\mathrm{cc}}(\alpha, \mathcal{X}) \equiv \begin{cases} \alpha x^{\mathrm{cc}} & \text{if } \alpha \geq 0 \\ \alpha x^{\mathrm{cv}} & \text{otherwise} \end{cases}, \tag{29}$$

$$\hat{\psi}^{\mathrm{cv}}(\alpha, \mathcal{X}) \equiv \min(\alpha x^{\mathrm{cv}}, \alpha x^{\mathrm{cc}}), \qquad \hat{\psi}^{\mathrm{cc}}(\alpha, \mathcal{X}) \equiv \max(\alpha x^{\mathrm{cv}}, \alpha x^{\mathrm{cc}}). \tag{30}$$

Definition 14 Define $(\times, \mathbb{MR}^2, \mathbb{MR})$ by

$$\times(\mathcal{X}, \mathcal{Y}) = \mathcal{X}\mathcal{Y} \equiv (z^{\mathrm{L}}, z^{\mathrm{U}}, z^{\mathrm{cv}}, z^{\mathrm{cc}}), \tag{31}$$

where, z^{L}, z^{U}, z^{cv}, and z^{cc} are defined as follows with $\bar{\mathcal{X}} = \mathrm{Cut}(\mathcal{X})$ and $\bar{\mathcal{Y}} = \mathrm{Cut}(\mathcal{Y})$:

$$z^{\mathrm{L}} = \min(x^{\mathrm{L}}y^{\mathrm{L}}, x^{\mathrm{L}}y^{\mathrm{U}}, x^{\mathrm{U}}y^{\mathrm{L}}, x^{\mathrm{U}}y^{\mathrm{U}}), \tag{32}$$

$$z^{\mathrm{U}} = \max(x^{\mathrm{L}}y^{\mathrm{L}}, x^{\mathrm{L}}y^{\mathrm{U}}, x^{\mathrm{U}}y^{\mathrm{L}}, x^{\mathrm{U}}y^{\mathrm{U}}), \tag{33}$$

$$z^{\mathrm{cv}} = \max(\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}) - x^{\mathrm{L}}y^{\mathrm{L}}, \; \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}) - x^{\mathrm{U}}y^{\mathrm{U}}), \tag{34}$$

$$z^{\mathrm{cc}} = \min(\psi^{\mathrm{cc}}(y^{\mathrm{L}}, \bar{\mathcal{X}}) + \psi^{\mathrm{cc}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}) - x^{\mathrm{U}}y^{\mathrm{L}}, \; \psi^{\mathrm{cc}}(y^{\mathrm{U}}, \bar{\mathcal{X}}) + \psi^{\mathrm{cc}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}) - x^{\mathrm{L}}y^{\mathrm{U}}). \tag{35}$$

Definition 15 Define $(\hat{\times}, \mathbb{MR}^2, \mathbb{MR})$ by

$$\hat{\times}(\mathcal{X}, \mathcal{Y}) \equiv (z^{\mathrm{L}}, z^{\mathrm{U}}, \hat{z}^{\mathrm{cv}}, \hat{z}^{\mathrm{cc}}), \tag{36}$$

where z^{L} and z^{U} are defined by (32)–(33) and $\hat{z}^{\mathrm{cv}}$ and $\hat{z}^{\mathrm{cc}}$ are defined as in (34)–(35) but with ψ^{cv} and ψ^{cc} replaced by $\hat{\psi}^{\mathrm{cv}}$ and $\hat{\psi}^{\mathrm{cc}}$, respectively.

For any nonempty $\mathcal{X} \in \mathbb{MR}$, we have $x^{\mathrm{cv}} \leq x^{\mathrm{cc}}$, and it follows that $\psi^{\mathrm{cv}}(\alpha, \mathcal{X}) = \hat{\psi}^{\mathrm{cv}}(\alpha, \mathcal{X})$ and $\psi^{\mathrm{cc}}(\alpha, \mathcal{X}) = \hat{\psi}^{\mathrm{cc}}(\alpha, \mathcal{X})$, $\forall \alpha \in \mathbb{R}$. Thus, $\times(\mathcal{X}, \mathcal{Y}) = \hat{\times}(\mathcal{X}, \mathcal{Y})$ for all nonempty $\mathcal{X}, \mathcal{Y} \in \mathbb{MR}$. Accordingly, prior literature uses Definitions 14 and 15 interchangeably, although Definition 15 is more common and is used in [9]. In contrast, if either $\mathcal{X}$ or $\mathcal{Y}$ is empty, then $\times(\mathcal{X}, \mathcal{Y})$ may not equal $\hat{\times}(\mathcal{X}, \mathcal{Y})$ (specifically, when $x^{\mathrm{cv}} > x^{\mathrm{cc}}$ or $y^{\mathrm{cv}} > y^{\mathrm{cc}}$). The results below show that $\times(\mathcal{X}, \mathcal{Y})$ still satisfies the appropriate properties in this case, while §3.3.1 shows by counterexample that $\hat{\times}(\mathcal{X}, \mathcal{Y})$ does not.

Theorem 10 $(\times, \mathbb{MR}^2, \mathbb{MR})$ *is a McCormick extension of* $(\times, \mathbb{R}^2, \mathbb{R})$ *on* $\mathbb{MR}^2$*.*

Proof For any $(x, y) \in \mathbb{R}^2$, the object $((x, y), (x, y), (x, y), (x, y))$ must be in $\mathbb{MR}^2$. Thus, $\mathbb{MR}^2$ is a McCormick extension of $\mathbb{R}^2$. Choose any $(\mathcal{X}, \mathcal{Y}) \in \mathbb{MR}^2$ such that $\mathcal{X} = (x, x, x, x)$ and $\mathcal{Y} = (y, y, y, y)$, and let $\mathcal{Z} = \mathcal{X}\mathcal{Y}$. It will be shown that $\mathcal{Z} = (xy, xy, xy, xy)$. Equations (32)–(33) clearly give $z^{\mathrm{L}} = z^{\mathrm{U}} = xy$, so it remains to prove that $z^{\mathrm{cv}} = z^{\mathrm{cc}} = xy$. The definition of Cut gives $\bar{\mathcal{X}} = \mathrm{Cut}(\mathcal{X}) = (x, x, x, x)$ and $\bar{\mathcal{Y}} = \mathrm{Cut}(\mathcal{Y}) = (y, y, y, y)$. Substituting these into (34) gives

$$\begin{aligned} z^{\mathrm{cv}} &= \max(\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}) - x^{\mathrm{L}}y^{\mathrm{L}}, \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}) - x^{\mathrm{U}}y^{\mathrm{U}}), \\ &= \max(yx + xy - xy, yx + xy - xy), \\ &= xy. \end{aligned}$$

This proves that $z^{\mathrm{cv}} = xy$, and $z^{\mathrm{cc}} = xy$ can be proven analogously. □

Lemma 4 will assist in proving the coherent concavity of $(\times, \mathbb{MR}^2, \mathbb{MR})$.

Lemma 4 *Choose any coherent* $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{MR}$ *and let* $X = [x^{\mathrm{L}}, x^{\mathrm{U}}]$ *without subscripts denote their common interval part. Choose any* $\lambda \in [0, 1]$ *and let* $\mathcal{X}_\lambda = \mathrm{Conv}(\lambda, \mathcal{X}_1, \mathcal{X}_2)$*. Let* $\bar{\mathcal{X}}_1 = \mathrm{Cut}(\mathcal{X}_1)$*,* $\bar{\mathcal{X}}_2 = \mathrm{Cut}(\mathcal{X}_2)$*, and* $\bar{\mathcal{X}}_\lambda = \mathrm{Cut}(\mathcal{X}_\lambda)$*. For any* $\alpha \in \mathbb{R}$*, we must have*

$$\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_\lambda) \leq \lambda\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_1) + (1 - \lambda)\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_2), \tag{37}$$

$$\psi^{\mathrm{cc}}(\alpha, \bar{\mathcal{X}}_\lambda) \geq \lambda\psi^{\mathrm{cc}}(\alpha, \bar{\mathcal{X}}_1) + (1 - \lambda)\psi^{\mathrm{cc}}(\alpha, \bar{\mathcal{X}}_2). \tag{38}$$

Proof If $\alpha \geq 0$, then

$$\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_\lambda) = \alpha\max(x^{\mathrm{L}}, \lambda x_1^{\mathrm{cv}} + (1 - \lambda)x_2^{\mathrm{cv}}). \tag{39}$$

Since $\alpha \geq 0$, $\alpha \max(x^{\mathrm{L}}, \cdot)$ is convex on $\mathbb{R}$, and hence

$$\begin{aligned} \psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_\lambda) &\leq \lambda[\alpha \max(x^{\mathrm{L}}, x_1^{\mathrm{cv}})] + (1 - \lambda)[\alpha \max(x^{\mathrm{L}}, x_2^{\mathrm{cv}})], \\ &= \lambda\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_1) + (1 - \lambda)\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_2). \end{aligned} \tag{40}$$

Similarly, if $\alpha < 0$, then

$$\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_\lambda) = \alpha\min(x^{\mathrm{U}}, \lambda x_1^{\mathrm{cc}} + (1 - \lambda)x_2^{\mathrm{cc}}). \tag{41}$$

Since $\alpha < 0$, $\alpha \min(x^{\mathrm{U}}, \cdot)$ is convex on $\mathbb{R}$, and hence

$$\begin{aligned} \psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_\lambda) &\leq \lambda[\alpha \min(x^{\mathrm{U}}, x_1^{\mathrm{cc}})] + (1 - \lambda)[\alpha \min(x^{\mathrm{U}}, x_2^{\mathrm{cc}})], \\ &= \lambda\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_1) + (1 - \lambda)\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}_2). \end{aligned} \tag{42}$$

This proves (37), and (38) can be proven analogously. □

Theorem 11 $(\times, \mathbb{MR}^2, \mathbb{MR})$ *is coherently concave on* $\mathbb{MR}^2$*.*

Proof Choose any coherent $(\mathcal{X}_1, \mathcal{Y}_1), (\mathcal{X}_2, \mathcal{Y}_2) \in \mathbb{MR}^2$ and let $X \times Y = [x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}]$ without subscripts denote their common interval part. Let $\mathcal{Z}_1 \equiv \mathcal{X}_1\mathcal{Y}_1$ and $\mathcal{Z}_2 \equiv \mathcal{X}_2\mathcal{Y}_2$. Furthermore, choose any $\lambda \in [0,1]$ and let $\mathcal{X}_\lambda = \mathrm{Conv}(\lambda, \mathcal{X}_1, \mathcal{X}_2)$, $\mathcal{Y}_\lambda = \mathrm{Conv}(\lambda, \mathcal{Y}_1, \mathcal{Y}_2)$, and $\mathcal{Z}_\lambda = \mathcal{X}_\lambda\mathcal{Y}_\lambda$. We must show that

$$\mathcal{Z}_\lambda \succeq \mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2). \tag{43}$$

Since $(\mathcal{X}_1, \mathcal{Y}_1)$ and $(\mathcal{X}_2, \mathcal{Y}_2)$ are coherent, $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are coherent as well with common interval part XY. By Definition 7, this implies that the interval part of $\mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2)$ is also XY. Similarly, the convex combinations $\mathcal{X}_\lambda$ and $\mathcal{Y}_\lambda$ have interval parts X and Y, respectively, which implies that the interval part of $\mathcal{Z}_\lambda = \mathcal{X}_\lambda\mathcal{Y}_\lambda$ is XY as well. Therefore, the left-hand and right-hand sides of (43) have the same interval parts. Then, to prove (43), it only remains to show that $z_\lambda^{\mathrm{cv}} \le \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}$ and $z_\lambda^{\mathrm{cc}} \ge \lambda z_1^{\mathrm{cc}} + (1-\lambda) z_2^{\mathrm{cc}}$.

Letting $\bar{\mathcal{X}} = \mathrm{Cut}(\mathcal{X})$ for all $\mathcal{X} \in \mathbb{MR}$, by Definition 14, we have

$$\begin{aligned} z_\lambda^{\mathrm{cv}} = \max(&\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_\lambda) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_\lambda) - x^{\mathrm{L}}y^{\mathrm{L}}, \\ &\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_\lambda) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_\lambda) - x^{\mathrm{U}}y^{\mathrm{U}}). \end{aligned} \tag{44}$$

Considering the first term in the max in (44) and applying Lemma 4 to $\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_\lambda)$ and $\psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_\lambda)$,

$$\begin{aligned} \psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_\lambda) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_\lambda) - x^{\mathrm{L}}y^{\mathrm{L}} \le \lambda &\left[\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{L}}y^{\mathrm{L}}\right] \\ +(1-\lambda) &\left[\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{L}}y^{\mathrm{L}}\right]. \end{aligned} \tag{45}$$

Similarly for the second term in (44), applying Lemma 4 to $\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_\lambda)$ and $\psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_\lambda)$ gives

$$\begin{aligned} &\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_\lambda) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_\lambda) - x^{\mathrm{U}}y^{\mathrm{U}} \\ \le&\lambda\left[\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{U}}y^{\mathrm{U}}\right] \\ &\quad + (1-\lambda)\left[\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{U}}y^{\mathrm{U}}\right]. \end{aligned} \tag{46}$$

Since $\max(a,b) \le \max(a',b')$ if $a \le a'$ and $b \le b'$, it follows that

$$\begin{aligned} z_\lambda^{\mathrm{cv}} \le \max(&\lambda\left[\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{L}}y^{\mathrm{L}}\right] \\ &\quad + (1-\lambda)\left[\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{L}}y^{\mathrm{L}}\right], \\ &\lambda\left[\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{U}}y^{\mathrm{U}}\right] \\ &\quad + (1-\lambda)\left[\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{U}}y^{\mathrm{U}}\right]). \end{aligned}$$

Since max is convex on $\mathbb{R}^2$, it follows that

$$\begin{aligned} z_\lambda^{\mathrm{cv}} \leq & \lambda \max(\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{L}} y^{\mathrm{L}}, \\ & \qquad \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{U}} y^{\mathrm{U}}) \\ & + (1-\lambda) \max(\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{L}} y^{\mathrm{L}}, \\ & \qquad \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{U}} y^{\mathrm{U}}). \end{aligned}$$

By Definition 14, the previous inequality is exactly $z_\lambda^{\mathrm{cv}} \leq \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}$. An analogous proof can be given for $z_\lambda^{\mathrm{cc}} \geq \lambda z_1^{\mathrm{cc}} + (1-\lambda) z_2^{\mathrm{cc}}$. □

Theorem 12 $(\times, \mathbb{MR}^2, \mathbb{MR})$ *is inclusion monotonic on the set* $\{(\mathcal{X}, \mathcal{Y}) \in \mathbb{MR}^2 : \mathrm{Encl}(\mathcal{X}) \neq \emptyset, \mathrm{Encl}(\mathcal{Y}) \neq \emptyset\}$.

Proof Choose any nonempty $(\mathcal{X}_1, \mathcal{Y}_1), (\mathcal{X}_2, \mathcal{Y}_2) \in \mathbb{MR}^2$ with $(\mathcal{X}_1, \mathcal{Y}_1) \preceq (\mathcal{X}_2, \mathcal{Y}_2)$. It follows that $\mathcal{X}_1 \preceq \mathcal{X}_2$ and $\mathcal{Y}_1 \preceq \mathcal{Y}_2$. Let $\mathcal{Z}_1 = \mathcal{X}_1 \mathcal{Y}_1$ and $\mathcal{Z}_2 = \mathcal{X}_2 \mathcal{Y}_2$. We must prove that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$.

Let $\bar{\mathcal{X}}_1 = \mathrm{Cut}(\mathcal{X}_1)$, $\bar{\mathcal{X}}_2 = \mathrm{Cut}(\mathcal{X}_2)$, $\bar{\mathcal{Y}}_1 = \mathrm{Cut}(\mathcal{Y}_1)$ and $\bar{\mathcal{Y}}_2 = \mathrm{Cut}(\mathcal{Y}_2)$. Theorems 3–4 imply $\bar{\mathcal{X}}_1 \preceq \bar{\mathcal{X}}_2$ and $\bar{\mathcal{Y}}_1 \preceq \bar{\mathcal{Y}}_2$ and that $\bar{\mathcal{X}}_1$, $\bar{\mathcal{X}}_2$, $\bar{\mathcal{Y}}_1$ and $\bar{\mathcal{Y}}_2$ are nonempty.

We now prove $\mathcal{Z}_1 \preceq \mathcal{Z}_2$ by applying Theorem 2.4.23 from [9], which establishes inclusion monotonicity of the alternate multiplication rule $(\hat{\times}, \mathbb{MR}^2, \mathbb{MR})$ on the set of nonempty McCormick objects. A little care is required because [9] uses a slightly different definition of *inclusion monotonicity*.

Let $\hat{\mathcal{Z}}_1 = \hat{\times}(\mathcal{X}_1, \mathcal{Y}_1)$ and $\hat{\mathcal{Z}}_2 = \hat{\times}(\mathcal{X}_2, \mathcal{Y}_2)$. By Definition 3, the fact that $\mathrm{Encl}(\mathcal{X}_1) \neq \emptyset$ implies that $x_1^{\mathrm{L}} \leq x_1^{\mathrm{U}}$, $x_1^{\mathrm{cv}} \leq x_1^{\mathrm{cc}}$ and $[x_1^{\mathrm{L}}, x_1^{\mathrm{U}}] \cap [x_1^{\mathrm{cv}}, x_1^{\mathrm{cc}}] \neq \emptyset$. Moreover, the interval $[x_1^{\mathrm{L}}, x_1^{\mathrm{U}}] \times [x_1^{\mathrm{cv}}, x_1^{\mathrm{cc}}]$ is a nonempty subset of $\mathbb{R}^2$. The same is clearly true of $\mathcal{Y}_1$, $\mathcal{X}_2$, and $\mathcal{Y}_2$. Using this subset notation, Theorem 2.4.23 in [9] establishes that $\hat{\times}$ is *inclusion monotonic* in the sense that

$$\begin{aligned} & [x_1^{\mathrm{L}}, x_1^{\mathrm{U}}] \times [x_1^{\mathrm{cv}}, x_1^{\mathrm{cc}}] \subset [x_2^{\mathrm{L}}, x_2^{\mathrm{U}}] \times [x_2^{\mathrm{cv}}, x_2^{\mathrm{cc}}] \text{ and} \qquad (47) \\ & [y_1^{\mathrm{L}}, y_1^{\mathrm{U}}] \times [y_1^{\mathrm{cv}}, y_1^{\mathrm{cc}}] \subset [y_2^{\mathrm{L}}, y_2^{\mathrm{U}}] \times [y_2^{\mathrm{cv}}, y_2^{\mathrm{cc}}] \\ \Longrightarrow & [\hat{z}_1^{\mathrm{L}}, \hat{z}_1^{\mathrm{U}}] \times [\hat{z}_1^{\mathrm{cv}}, \hat{z}_1^{\mathrm{cc}}] \subset [\hat{z}_2^{\mathrm{L}}, \hat{z}_2^{\mathrm{U}}] \times [\hat{z}_2^{\mathrm{cv}}, \hat{z}_2^{\mathrm{cc}}]. \qquad (48) \end{aligned}$$

To apply this result to $(\times, \mathbb{MR}^2, \mathbb{MR})$, we observe that

$$\hat{\times}(\mathcal{X}, \mathcal{Y}) = \times(\mathcal{X}, \mathcal{Y}), \quad \forall \mathcal{X}, \mathcal{Y} \in \mathbb{MR} \ \text{ s.t. } \ \mathrm{Encl}(\mathcal{X}) \neq \emptyset, \mathrm{Encl}(\mathcal{Y}) \neq \emptyset. \qquad (49)$$

Thus, $\hat{\mathcal{Z}}_1 = \mathcal{Z}_1$ and $\hat{\mathcal{Z}}_2 = \mathcal{Z}_2$, so (48) gives

$$[z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] \times [z_1^{\mathrm{cv}}, z_1^{\mathrm{cc}}] \subset [z_2^{\mathrm{L}}, z_2^{\mathrm{U}}] \times [z_2^{\mathrm{cv}}, z_2^{\mathrm{cc}}]. \qquad (50)$$

It follows immediately that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$, as desired. □

Theorem 13 $(\times, \mathbb{MR}^2, \mathbb{MR})$ *is coherently inclusion monotonic on* $\mathbb{MR}^2$.

Proof Choose any coherent $(\mathcal{X}_1, \mathcal{Y}_1), (\mathcal{X}_2, \mathcal{Y}_2) \in \mathbb{MR}^2$ with $(\mathcal{X}_1, \mathcal{Y}_1) \preceq (\mathcal{X}_2, \mathcal{Y}_2)$ and let $X \times Y = [x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}]$ without subscripts denote their common interval part. It follows that $\mathcal{X}_1 \preceq \mathcal{X}_2$ and $\mathcal{Y}_1 \preceq \mathcal{Y}_2$. Let $\mathcal{Z}_1 = \mathcal{X}_1\mathcal{Y}_1$ and $\mathcal{Z}_2 = \mathcal{X}_2\mathcal{Y}_2$. We must prove that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$. Since the interval parts of $\mathcal{Z}_1$ and $\mathcal{Z}_2$ depend only on $[x^{\mathrm{L}}, x^{\mathrm{U}}]$ and $[y^{\mathrm{L}}, y^{\mathrm{U}}]$, we must have $[z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] = [z_2^{\mathrm{L}}, z_2^{\mathrm{U}}]$. It remains to prove that $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$ and $z_1^{\mathrm{cc}} \leq z_2^{\mathrm{cc}}$.

Let $\bar{\mathcal{X}}_1 = \mathrm{Cut}(\mathcal{X}_1)$, $\bar{\mathcal{Y}}_1 = \mathrm{Cut}(\mathcal{Y}_1)$, $\bar{\mathcal{X}}_2 = \mathrm{Cut}(\mathcal{X}_2)$ and $\bar{\mathcal{Y}}_2 = \mathrm{Cut}(\mathcal{Y}_2)$. It will be shown that

$$\begin{aligned} z_2^{\mathrm{cv}} &= \max(\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{L}}y^{\mathrm{L}}, \\ &\qquad \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{U}}y^{\mathrm{U}}), \\ &\leq \max(\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{L}}y^{\mathrm{L}}, \\ &\qquad \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{U}}y^{\mathrm{U}}) = z_1^{\mathrm{cv}}. \end{aligned} \tag{51}$$

Since $\max(a, b) \leq \max(a', b')$ if $a \leq a'$ and $b \leq b'$, (51) holds provided that

$$\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{L}}y^{\mathrm{L}} \leq \psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{L}}y^{\mathrm{L}}, \tag{52}$$

$$\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_2) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_2) - x^{\mathrm{U}}y^{\mathrm{U}} \leq \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_1) + \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_1) - x^{\mathrm{U}}y^{\mathrm{U}}. \tag{53}$$

If $y^{\mathrm{L}} \geq 0$, then $\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) = y^{\mathrm{L}}\max(x^{\mathrm{L}}, x_2^{\mathrm{cv}})$ and $\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1) = y^{\mathrm{L}}\max(x^{\mathrm{L}}, x_1^{\mathrm{cv}})$. Therefore, $\mathcal{X}_1 \preceq \mathcal{X}_2$ implies that $\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) \leq \psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1)$. If $y^{\mathrm{L}} < 0$, then $\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) = y^{\mathrm{L}}\min(x^{\mathrm{U}}, x_2^{\mathrm{cc}})$ and $\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1) = y^{\mathrm{L}}\min(x^{\mathrm{U}}, x_1^{\mathrm{cc}})$. Therefore, $\mathcal{X}_1 \preceq \mathcal{X}_2$ still implies $\psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_2) \leq \psi^{\mathrm{cv}}(y^{\mathrm{L}}, \bar{\mathcal{X}}_1)$. By similar reasoning, it can be verified that $\psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_2) \leq \psi^{\mathrm{cv}}(x^{\mathrm{L}}, \bar{\mathcal{Y}}_1)$, $\psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_2) \leq \psi^{\mathrm{cv}}(y^{\mathrm{U}}, \bar{\mathcal{X}}_1)$, and $\psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_2) \leq \psi^{\mathrm{cv}}(x^{\mathrm{U}}, \bar{\mathcal{Y}}_1)$. Thus, (52) and (53) hold. This proves that $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$, and $z_1^{\mathrm{cc}} \leq z_2^{\mathrm{cc}}$ can be proven analogously. □

Theorem 14 $(\times, \mathbb{MR}^2, \mathbb{MR})$ *preserves nonemptiness on* $\mathbb{MR}^2$.

Proof The result follows from Lemma 1 with Theorems 10 and 12. □

Theorem 15 $(\times, \mathbb{MR}^2, \mathbb{MR})$ *is a relaxation function for* $(\times, \mathbb{R}^2, \mathbb{R})$.

Proof The result follows from Theorem 2 with Theorems 10–13. □

3.3.1 A Multiplication Example

Let $x : P \to \mathbb{R}$ be an arbitrary function on $P = [-2, 2]$. We consider computing relaxations of the composite function

$$f(p) = -2x(p) \tag{54}$$

on P given relaxation information for x on P in the form of a McCormick function $\mathcal{X} : P \to \mathbb{MR}$. When computing McCormick relaxations, a scalar multiplication like this is typically handled as a composition with a univariate function, not as a binary multiplication. However, it can be viewed as a special case of the binary multiplication $f(p) = y(p)x(p)$ with $y(p) = -2$, and this provides a simple example to illustrate the distinction between $\times$ and $\hat{\times}$. Accordingly, we consider the binary multiplications $\mathcal{F}(p) = \mathcal{Y}(p) \times \mathcal{X}(p)$ and $\hat{\mathcal{F}}(p) = \mathcal{Y}(p) \hat{\times} \mathcal{X}(p)$, respectively, with

$\mathcal{Y}(p) = (-2, -2, -2, -2)$ and $\mathcal{X}(p) = (x^{\mathrm{L}}, x^{\mathrm{U}}, x^{\mathrm{cv}}(p), x^{\mathrm{cc}}(p))$ defined as in the left panel of Figure 1. Note that $\mathcal{X}(p)$ is nonempty for all $p \in P^* \equiv [-0.87, 0.87]$ and empty otherwise. Even so, x^{cv} and x^{cc} are convex and concave, respectively, on all of P. As shown in the middle and right panels of Figure 1, the two multiplication rules agree for all p such that $\mathcal{X}(p)$ is nonempty. However, they clearly differ at infeasible p values. While $\hat{\mathcal{F}}(p)$ remains nonempty for all p, $\hat{f}^{\mathrm{cv}}$ and $\hat{f}^{\mathrm{cc}}$ are clearly not convex and concave. In contrast, $\mathcal{F}(p)$ takes empty values for some $p \notin P^*$ to ensure that f^{cv} and f^{cc} are convex and concave on all of P, which is the desired behavior. This difference can be traced back to the fact that $\psi^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}(p)) \neq \hat{\psi}^{\mathrm{cv}}(\alpha, \bar{\mathcal{X}}(p))$ and $\psi^{\mathrm{cc}}(\alpha, \bar{\mathcal{X}}(p)) \neq \hat{\psi}^{\mathrm{cc}}(\alpha, \bar{\mathcal{X}}(p))$ when $\alpha = -2$ and $\mathcal{X}$ is empty.

To reiterate the significance of this difference, suppose that $\mathcal{X}$ can be interpreted as a relaxation of x in the sense that it satisfies $x(p) \in \mathrm{Encl}(\mathcal{X}(p))$ for all p in some feasible set of interest, $P^{\mathrm{feas}} \subset P$. For example, P^{feas} may be defined as the subset of P for which some constraint $g(x(p), p) \leq 0$ holds, and $\mathcal{X}$ may be the result of a domain reduction procedure that begins with a nonempty initial relaxation $\mathcal{X}_0(p)$ and yields a tighter relaxation $\mathcal{X}(p) \preceq \mathcal{X}_0(p)$ such that $x(p) \in \mathrm{Encl}(\mathcal{X}(p))$ for all $p \in P^{\mathrm{feas}}$ (a simple example of such a procedure is given in §5). In such a situation, it follows that $P^{\mathrm{feas}} \subset P^*$, and the fact that f^{cv} is a valid underestimator for f on P^* implies that it is also a valid underestimator on P^{feas} (i.e., where it "matters"). At the same time, f^{cv} is well-defined and convex on all of P, meaning that it can be easily minimized to obtain a lower bound for f on P^{feas}. In contrast, the standard McCormick rule (as stated in [9]) yields an underestimator $\hat{f}^{\mathrm{cv}}$ that is valid on all of P, but nonconvex, which is much less useful. More generally, we will eventually show that $\mathcal{F}(p)$ has the right properties to be propagated through subsequent McCormick operations if needed, while $\hat{\mathcal{F}}(p)$ does not.

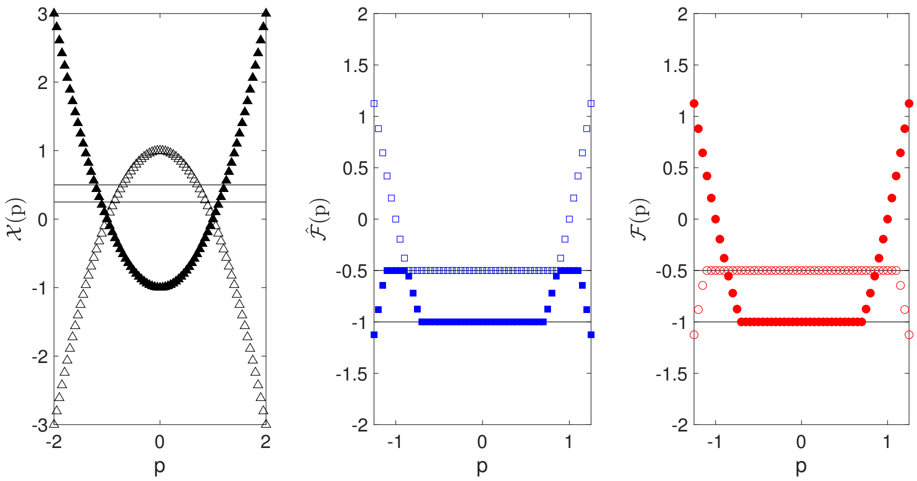

Fig. 1 Left: McCormick object $\mathcal{X}(p)$ corresponding to $x(p)$ in (54) defined by its bounds x^{L} and x^{U} (solid), convex component $x^{\mathrm{cv}}(p)$ (filled triangles), and concave component $x^{\mathrm{cc}}(p)$ (open triangles). Middle: McCormick object $\hat{\mathcal{F}}(p)$ corresponding to $f(p)$ in (54) defined by its bounds f^{L} and f^{U} (solid), convex component $\hat{f}^{\mathrm{cv}}(p)$ (filled squares), and concave component $\hat{f}^{\mathrm{cc}}(p)$ (open squares). Right: McCormick object $\mathcal{F}(p)$ corresponding to $f(p)$ in (54) defined by its bounds f^{L} and f^{U} (solid), convex component $f^{\mathrm{cv}}(p)$ (filled circles), and concave component $f^{\mathrm{cc}}(p)$ (open circles). For all $p \in [-0.87, 0.87]$, $\hat{f}^{\mathrm{cv}}(p)/\hat{f}^{\mathrm{cc}}(p)$ are identical to $f^{\mathrm{cv}}(p)/f^{\mathrm{cc}}(p)$.

3.4 Composition with Univariate Functions

This section presents an extension of McCormick's rule for relaxing compositions with univariate functions such as $\ln(x)$, $\exp(x)$, $\sin(x)$, x^n, $-x$, etc. Compared to binary addition and multiplication, the composition rule requires a much more significant modification to handle empty inputs. We first state the new rule and subsequently discuss its relation to the original rule from [4] as presented in [9].

Let $\mathcal{L}$ be a library of univariate functions $(u, B, \mathbb{R})$, where $B \subset \mathbb{R}$. To propogate McCormick relaxations through a function from this library, several pieces of information about the function must be available. These requirements are very similar to those of the original rule (see Remark 3) and are formalized in Assumption 1. Recall that $\mathbb{I}B$ denotes the set of all intervals contained in B and $\mathbb{M}B$ denotes the set of all McCormick objects with interval part contained in B.

Assumption 1 For every $(u, B, \mathbb{R}) \in \mathcal{L}$, functions $u : \mathbb{I}B \to \mathbb{I}\mathbb{R}$, $u^{\mathrm{cv}}, u^{\mathrm{cc}} : \mathbb{I}B \times \mathbb{R} \to \mathbb{R}$, and $x^{\min}, x^{\max} : \mathbb{I}B \to \mathbb{R} \cup \{-\infty, +\infty\}$ are known such that:

1. $(u, \mathbb{I}B, \mathbb{I}\mathbb{R})$ is an inclusion monotonic interval extension of $(u, B, \mathbb{R})$ on $\mathbb{I}B$.
2. For every $X \in \mathbb{I}B$, $u^{\mathrm{cv}}(X, \cdot)$ and $u^{\mathrm{cc}}(X, \cdot)$ are convex and concave on $\mathbb{R}$, respectively, and satisfy $u^{\mathrm{cv}}(X, x) \le u(x) \le u^{\mathrm{cc}}(X, x)$ for all $x \in X$.
3. For every $X \in \mathbb{I}B$, $x^{\min}(X)$ is a minimizer of $u^{\mathrm{cv}}(X, \cdot)$ on $\mathbb{R}$ and $x^{\max}(X)$ is a maximizer of $u^{\mathrm{cc}}(X, \cdot)$ on $\mathbb{R}$; i.e., $\lim_{x \to x^{\min}(X)}[u^{\mathrm{cv}}(X, x)] = \inf_{x \in \mathbb{R}}[u^{\mathrm{cv}}(X, x)]$ and $\lim_{x \to x^{\max}(X)}[u^{\mathrm{cc}}(X, x)] = \sup_{x \in \mathbb{R}}[u^{\mathrm{cc}}(X, x)]$.
4. For any $X_1, X_2 \in \mathbb{I}B$ with $X_1 \subset X_2$, we have $u^{\mathrm{cv}}(X_2, x) \le u^{\mathrm{cv}}(X_1, x)$ and $u^{\mathrm{cc}}(X_2, x) \ge u^{\mathrm{cc}}(X_1, x)$ for all $x \in X_1$.

5. $u^{\mathrm{cv}}([x,x],x) = u^{\mathrm{cc}}([x,x],x) = u(x)$ for every $x \in B$.

Remark 3 Assumption 1 is analogous to Assumptions 2.3.8 and 2.4.25 in [9] but includes two generalizations. First, $u^{\mathrm{cv}}(X,\cdot)$ and $u^{\mathrm{cc}}(X,\cdot)$ are defined on all of $\mathbb{R}$, and are convex and concave there, instead of only on X as in [9]. Second, $x^{\min}$ and $x^{\max}$ minimize $u^{\mathrm{cv}}(X,\cdot)$ and maximize $u^{\mathrm{cc}}(X,\cdot)$, respectively, on $\mathbb{R}$, rather than on X. These changes are necessary for handling empty McCormick objects without causing domain violations.

Remark 4 The Supplementary Information for this paper provides all of the information required by Assumption 1 for a variety of common univariate functions. For some, the provided definitions of u^{cv}, u^{cc}, $x^{\min}$, and $x^{\max}$ differ from those in standard libraries due to the differences highlighted in Remark 3.

The extended composition rule is defined as follows.

Definition 16 Let $(u, B, \mathbb{R}) \in \mathcal{L}$. Define $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$ by

$$u(\mathcal{X}) \equiv (z^{\mathrm{L}}, z^{\mathrm{U}}, z^{\mathrm{cv}}, z^{\mathrm{cc}}), \tag{55}$$

where $[z^{\mathrm{L}}, z^{\mathrm{U}}] = u(X)$ is the interval extension specified by Assumption 1.1 and

$$\begin{aligned} z^{\mathrm{cv}} &= u^{\mathrm{cv}}(X, \min(\bar{x}^{\mathrm{cc}}, x^{\min}(X))) + u^{\mathrm{cv}}(X, \max(\bar{x}^{\mathrm{cv}}, x^{\min}(X))) \qquad (56)\\ &\quad - u^{\mathrm{cv}}(X, x^{\min}(X)),\\ z^{\mathrm{cc}} &= u^{\mathrm{cc}}(X, \min(\bar{x}^{\mathrm{cc}}, x^{\max}(X))) + u^{\mathrm{cc}}(X, \max(\bar{x}^{\mathrm{cv}}, x^{\max}(X))) \qquad (57)\\ &\quad - u^{\mathrm{cc}}(X, x^{\max}(X)), \end{aligned}$$

where $\bar{\mathcal{X}} = \mathrm{Cut}(\mathcal{X})$. If $x^{\min}(X) = \pm\infty$, then (56) is replaced with

$$z^{\mathrm{cv}} = \begin{cases} u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cc}}) \text{ if } x^{\min}(X) = +\infty \\ u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cv}}) \text{ if } x^{\min}(X) = -\infty \end{cases}. \tag{58}$$

Similarly, if $x^{\max}(X) = \pm\infty$, then (57) is replaced with

$$z^{\mathrm{cc}} = \begin{cases} u^{\mathrm{cc}}(X, \bar{x}^{\mathrm{cc}}) \text{ if } x^{\max}(X) = +\infty \\ u^{\mathrm{cc}}(X, \bar{x}^{\mathrm{cv}}) \text{ if } x^{\max}(X) = -\infty \end{cases}. \tag{59}$$

Remark 5 All of the cases (56)–(59) can be represented by the single definitions

$$z^{\mathrm{cv}} = \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(\bar{x}^{\mathrm{cc}}, x)) + u^{\mathrm{cv}}(X, \max(\bar{x}^{\mathrm{cv}}, x)) - u^{\mathrm{cv}}(X, x)], \tag{60}$$

$$z^{\mathrm{cc}} = \lim_{x \to x^{\max}(X)} [u^{\mathrm{cc}}(X, \min(\bar{x}^{\mathrm{cc}}, x)) + u^{\mathrm{cc}}(X, \max(\bar{x}^{\mathrm{cv}}, x)) - u^{\mathrm{cc}}(X, x)]. \tag{61}$$

Since $u^{\mathrm{cv}}(X,\cdot)$ is convex on $\mathbb{R}$, it must be continuous, and therefore (60) is equivalent to (56) whenever $x^{\min}(X)$ is finite. Similarly, (61) is equivalent to (57) whenever $x^{\max}(X)$ is finite. If $x^{\min}(X) = +\infty$, then as x approaches $x^{\min}(X)$, the terms $u^{\mathrm{cv}}(X, \max(\bar{x}^{\mathrm{cv}}, x))$ and $u^{\mathrm{cv}}(X, \min(\bar{x}^{\mathrm{cc}}, x))$ in (60) eventually simplify to $u^{\mathrm{cv}}(X, x)$ and $u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cc}})$, respectively. Thus, we have from (60) that $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cc}})$, as in (58). Similarly, if $x^{\min}(X) = -\infty$, then analogous arguments show that (60) simplifies to $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cv}})$, as in (58). Analogous arguments also show that (61) simplifies to (59) if $x^{\max}(X) = \pm\infty$.

To explain the relationship between the definition of $(u, \mathbb{M}B, \mathbb{MR})$ above and the standard rule in [9], a few preliminary results and definitions are needed. The next proposition will also be used extensively in proving several properties of $(u, \mathbb{M}B, \mathbb{MR})$.

Proposition 1 *Let $g^{\mathrm{cv}}, g^{\mathrm{cc}} : \mathbb{R} \to \mathbb{R}$ and $x^{\min}, x^{\max} \in \mathbb{R} \cup \{-\infty, +\infty\}$. Assume that g^{cv} is convex and $g^{\mathrm{cv}}(x)$ approaches the infimum of g^{cv} on $\mathbb{R}$ as x approaches $x^{\min}$. Then, g^{cv} is non-increasing for all $x \in (-\infty, x^{\min})$ and is non-decreasing for all $x \in (x^{\min}, +\infty)$. Likewise, assume that g^{cc} is concave and $g^{\mathrm{cc}}(x)$ approaches the supremum of g^{cc} on $\mathbb{R}$ as x approaches $x^{\max}$. Then, g^{cc} is non-decreasing for all $x \in (-\infty, x^{\max})$ and is non-increasing for all $x \in (x^{\max}, +\infty)$.*

Proof We first prove that g^{cv} is non-decreasing for all $x \in (x^{\min}, +\infty)$. If $x^{\min} = +\infty$, then there is nothing to prove. Suppose $x^{\min} < +\infty$ and choose any $x_1, x_2 \in (x^{\min}, +\infty)$ such that $x_1 \le x_2$. Choose any $\epsilon > 0$. Since $g^{\mathrm{cv}}(x)$ approaches the infimum of g on $\mathbb{R}$ as x approaches $x^{\min}$, there must exist $z \in (x^{\min}, x_1]$ such that $g^{\mathrm{cv}}(z) \le g^{\mathrm{cv}}(x_2) + \epsilon$. Since $z \le x_1 \le x_2$, there must exist $\lambda \in [0, 1]$ such that $x_1 = \lambda z + (1 - \lambda)x_2$. By the convexity of g^{cv}, this implies that

$$\begin{aligned} g^{\mathrm{cv}}(x_1) &\le \lambda g^{\mathrm{cv}}(z) + (1 - \lambda) g^{\mathrm{cv}}(x_2), \\ &\le \lambda(g^{\mathrm{cv}}(x_2) + \epsilon) + (1 - \lambda) g^{\mathrm{cv}}(x_2), \\ &\le g^{\mathrm{cv}}(x_2) + \epsilon. \end{aligned} \tag{62}$$

Since (62) holds for any $\epsilon > 0$, we have $g^{\mathrm{cv}}(x_1) \le g^{\mathrm{cv}}(x_2)$. Finally, since this is true for any $x_1, x_2 \in (x^{\min}, +\infty)$, we have shown that g^{cv} is non-decreasing on $(x^{\min}, \infty)$. The facts that g^{cv} is non-increasing on $(-\infty, x^{\min})$, g^{cc} is non-decreasing on $(-\infty, x^{\max})$, and g^{cc} is non-increasing on $(x^{\max}, \infty)$ can all be proven analogously. □

We now define modified versions of $x^{\min}$ and $x^{\max}$, denoted $x^{\min'}$ and $x^{\max'}$, that bear the same meaning as $x^{\min}$ and $x^{\max}$ in [9]. This provides a critical link between the extended and standard composition rules. The necessary interpretation of $x^{\min'}$ and $x^{\max'}$ is then established in Lemma 5 using Proposition 1.

Definition 17 For all $(u, B, \mathbb{R}) \in \mathcal{L}$, define $x^{\min'}, x^{\max'} : \mathbb{I}B \to \mathbb{R}$ by

$$x^{\min'}(X) \equiv \mathrm{mid}(x^{\mathrm{L}}, x^{\mathrm{U}}, x^{\min}(X)), \tag{63}$$

$$x^{\max'}(X) \equiv \mathrm{mid}(x^{\mathrm{L}}, x^{\mathrm{U}}, x^{\max}(X)), \tag{64}$$

where mid returns the middle value of its three arguments.

Lemma 5 *For any $(u, B, \mathbb{R}) \in \mathcal{L}$ and any $X \in \mathbb{I}B$,*

1. *$u^{\mathrm{cv}}(X, \cdot)$ reaches its minimum on X at the point $x^{\min'}(X)$, and*
2. *$u^{\mathrm{cc}}(X, \cdot)$ reaches its maximum on X at the point $x^{\max'}(X)$.*

Proof Choose any $(u, B, \mathbb{R}) \in \mathcal{L}$ and any $X \in \mathbb{I}B$. To prove the first claim, first assume that $x^{\min}(X) < x^{\mathrm{L}} \le x^{\mathrm{U}}$. Since $u^{\mathrm{cv}}(X, x)$ approaches its infimum on $\mathbb{R}$ as x approaches $x^{\min}(X)$, Proposition 1 shows that $u^{\mathrm{cv}}(X, \cdot)$ is non-decreasing on X. Thus, $u^{\mathrm{cv}}(X, \cdot)$ must reach its minimum on X at $x^{\mathrm{L}} = \mathrm{mid}(x^{\mathrm{L}}, x^{\mathrm{U}}, x^{\min}(X)) = x^{\min'}(X)$. Next, assume that $x^{\mathrm{L}} \le x^{\mathrm{U}} < x^{\min}(X)$. In this case, Proposition 1 shows that $u^{\mathrm{cv}}(X,\cdot)$ is non-increasing on X. Therefore, $u^{\mathrm{cv}}(X,\cdot)$ must reach its minimum on X at $x^{\mathrm{U}} = \mathrm{mid}(x^{\mathrm{L}}, x^{\mathrm{U}}, x^{\min}(X)) = x^{\min'}(X)$. Finally, assume that $x^{\mathrm{L}} \leq x^{\min}(X) \leq x^{\mathrm{U}}$. In this case, $x^{\min'}(X) = x^{\min}(X)$ by definition, and since $u^{\mathrm{cv}}(X,x)$ approaches its infimum on $\mathbb{R}$ as x approaches $x^{\min}(X)$, it must also approach its infimum on X as x approaches $x^{\min}(X)$. This again shows that $u^{\mathrm{cv}}(X,\cdot)$ reaches its minimum on X at $x^{\min'}(X)$, and completes the proof of the first claim. The second claim can be proven analogously. □

Definition 18 now gives the standard composition rule [4] as presented in [9].

Definition 18 Let $(u, B, \mathbb{R}) \in \mathcal{L}$. Define $(\hat{u}, \mathbb{M}B, \mathbb{M}\mathbb{R})$ by

$$\hat{u}(\mathcal{X}) \equiv (z^{\mathrm{L}}, z^{\mathrm{U}}, \hat{z}^{\mathrm{cv}}, \hat{z}^{\mathrm{cc}}), \tag{65}$$

where $[z^{\mathrm{L}}, z^{\mathrm{U}}] = u(X)$ is the interval extension specified by Assumption 1.1 and

$$\hat{z}^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\min'}(X))), \tag{66}$$

$$\hat{z}^{\mathrm{cc}} = u^{\mathrm{cc}}(X, \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\max'}(X))), \tag{67}$$

where $\bar{\mathcal{X}} = \mathrm{Cut}(\mathcal{X})$.

Remark 6 Definition 18 uses $u(X)$, u^{cv}, u^{cc}, $x^{\min'}(X)$, and $x^{\max'}(X)$ from this paper. For this to be consistent with [9], these quantities need to satisfy Assumptions 2.3.8 and 2.4.25 in [9]. The required properties of $u(X)$, u^{cv}, and u^{cc} all follow from Conditions 1, 2, 4, and 5 of Assumption 1 in this paper, while the required properties of $x^{\min'}(X)$ and $x^{\max'}(X)$ were proven in Lemma 5.

Remark 7 Actually, in both [9] and McCormick's original paper [4], $\hat{z}^{\mathrm{cv}}$ and $\hat{z}^{\mathrm{cc}}$ are given as

$$\hat{z}^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \mathrm{mid}(x^{\mathrm{cv}}, x^{\mathrm{cc}}, x^{\min'}(X))), \tag{68}$$

$$\hat{z}^{\mathrm{cc}} = u^{\mathrm{cc}}(X, \mathrm{mid}(x^{\mathrm{cv}}, x^{\mathrm{cc}}, x^{\max'}(X))). \tag{69}$$

However, it is straightforward to show that (68)–(69) are equivalent to (66)–(67) for all nonempty $\mathcal{X} \in \mathbb{M}B$. We prefer (66)–(67) because they are easier to relate to Definition 16.

Lemma 6 establishes the connection between the extended and standard composition rules by showing that they agree for all nonempty inputs. The proof also sheds some light on the relationship between (56)–(57) and (66)–(67).

Lemma 6 *Let $(u, B, \mathbb{R}) \in \mathcal{L}$ and choose any $\mathcal{X} \in \mathbb{M}B$. If $\mathrm{Encl}(\mathcal{X}) \neq \emptyset$, then*

$$u(\mathcal{X}) = \hat{u}(\mathcal{X}). \tag{70}$$

Proof Choose any $\mathcal{X} \in \mathbb{M}B$ such that $\mathrm{Encl}(\mathcal{X}) \neq \emptyset$ and let $\bar{\mathcal{X}} = \mathrm{Cut}(\mathcal{X})$ and $\mathcal{Z} = u(\mathcal{X})$. By Definition 16 and Remark 5,

$$z^{\mathrm{cv}} = \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(\bar{x}^{\mathrm{cc}}, x)) + u^{\mathrm{cv}}(X, \max(\bar{x}^{\mathrm{cv}}, x)) - u^{\mathrm{cv}}(X, x)], \tag{71}$$

$$z^{\mathrm{cc}} = \lim_{x \to x^{\max}(X)} [u^{\mathrm{cc}}(X, \min(\bar{x}^{\mathrm{cc}}, x)) + u^{\mathrm{cc}}(X, \max(\bar{x}^{\mathrm{cv}}, x)) - u^{\mathrm{cc}}(X, x)]. \tag{72}$$

First, assume that $x^{\min}(X) \in X$. Since $\mathrm{Encl}(\mathcal{X}) \neq \emptyset$, we have $\bar{x}^{\mathrm{cv}} \leq \bar{x}^{\mathrm{cc}}$. Thus, one of the following sub-cases must occur:

1. $x^{\min}(X) < \bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}}$,
2. $\bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}} < x^{\min}(X)$,
3. $\bar{x}^{\mathrm{cv}} \le x^{\min}(X) \le \bar{x}^{\mathrm{cc}}$.

In the first sub-case, (71) gives

$$\begin{aligned} z^{\mathrm{cv}} &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(\bar{x}^{\mathrm{cc}}, x)) + u^{\mathrm{cv}}(X, \max(\bar{x}^{\mathrm{cv}}, x)) - u^{\mathrm{cv}}(X, x)], && (73) \\ &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, x) + u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cv}}) - u^{\mathrm{cv}}(X, x)], && (74) \\ &= u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cv}}). && (75) \end{aligned}$$

Since $x^{\min}(X) \in X$, Definition 17 implies that $x^{\min'}(X) = x^{\min}(X)$. Therefore, $\bar{x}^{\mathrm{cv}} = \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\min'}(X))$. Plugging this into (75) gives (66) as claimed.

In the second sub-case, (71) gives

$$\begin{aligned} z^{\mathrm{cv}} &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(\bar{x}^{\mathrm{cc}}, x)) + u^{\mathrm{cv}}(X, \max(\bar{x}^{\mathrm{cv}}, x)) - u^{\mathrm{cv}}(X, x)], && (76) \\ &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cc}}) + u^{\mathrm{cv}}(X, x) - u^{\mathrm{cv}}(X, x)], && (77) \\ &= u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cc}}). && (78) \end{aligned}$$

Since $x^{\min}(X) \in X$, we have $\bar{x}^{\mathrm{cc}} = \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\min'}(X))$. Plugging this into (78) gives (66) as claimed.

In the third sub-case, (71) gives

$$\begin{aligned} z^{\mathrm{cv}} &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(\bar{x}^{\mathrm{cc}}, x)) + u^{\mathrm{cv}}(X, \max(\bar{x}^{\mathrm{cv}}, x)) - u^{\mathrm{cv}}(X, x)], && (79) \\ &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, x) + u^{\mathrm{cv}}(X, x) - u^{\mathrm{cv}}(X, x)], && (80) \\ &= u^{\mathrm{cv}}(X, x^{\min}(X)). && (81) \end{aligned}$$

Since $x^{\min}(X) \in X$, we have $x^{\min}(X) = \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\min'}(X))$. Plugging this into (81) gives (66) as claimed.

Next, assume that $x^{\min}(X) \notin X$. One of the following sub-cases must occur:

1. $x^{\min}(X) < x^{\mathrm{L}} \le x^{\mathrm{U}}$,
2. $x^{\mathrm{L}} \le x^{\mathrm{U}} < x^{\min}(X)$.

Since $\mathrm{Encl}(\mathcal{X}) \neq \emptyset$, the definition of Cut (Definition 11) implies $x^{\mathrm{L}} \le \bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}} \le x^{\mathrm{U}}$. Thus, in the first sub-case, we have $x^{\min}(X) < \bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}}$. We have already shown in (73)–(75) that (71) gives $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cv}})$ if $x^{\min}(X) < \bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}}$. Since $x^{\mathrm{L}} \le \bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}} \le x^{\mathrm{U}}$, we have $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\mathrm{L}}))$. Moreover, Definition 17 implies that $x^{\min'}(X) = \mathrm{mid}(x^{\mathrm{L}}, x^{\mathrm{U}}, x^{\min}(X)) = x^{\mathrm{L}}$. Hence, it follows that $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\min'}(X)))$, which gives (66) as claimed.

In the second sub-case, it follows from $\mathrm{Encl}(\mathcal{X}) \neq \emptyset$ and from the definition of Cut (Definition 11) that we must have $\bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}} < x^{\min}(X)$. We have already shown in (76)–(78) that (71) gives $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \bar{x}^{\mathrm{cc}})$ if $\bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}} < x^{\min}(X)$. Since $x^{\mathrm{L}} \le \bar{x}^{\mathrm{cv}} \le \bar{x}^{\mathrm{cc}} \le x^{\mathrm{U}}$, we have $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\mathrm{U}}))$. Moreover, Definition 17 implies that $x^{\min'}(X) = \mathrm{mid}(x^{\mathrm{L}}, x^{\mathrm{U}}, x^{\min}(X)) = x^{\mathrm{U}}$. Hence, it follows that $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \mathrm{mid}(\bar{x}^{\mathrm{cv}}, \bar{x}^{\mathrm{cc}}, x^{\min'}(X)))$, which again gives (66) as claimed.

Since $\mathcal{X}$ was chosen arbitrarily, we have shown that (71) and (66) are equivalent for all $\mathcal{X} \in \mathbb{M}B$ satisfying $\mathrm{Encl}(\mathcal{X}) \neq \emptyset$. The equivalence of (72) and (67) under this condition can be proven analogously. This proves that $u(\mathcal{X}) = \hat{u}(\mathcal{X})$. $\square$

Remark 8 The equivalence (70) is only guaranteed when both the standard and extended rules use the same u^{cv} and u^{cc}. The properties of u^{cv} and u^{cc} required by Assumption 1 here are stronger than in [9] (see Remark 3). Thus, any u^{cv} and u^{cc} used in the extended rule can always be used in the standard rule and will satisfy (70). However, some choices of u^{cv} and u^{cc} commonly used in the standard rule cannot be used in the extended rule, and modifying these as needed may lead to violations of (70) in practice.

The last result shows that $u(\mathcal{X}) = \hat{u}(\mathcal{X})$ whenever $\mathcal{X}$ is nonempty. In contrast, if $\mathcal{X}$ is empty, then $u(\mathcal{X}) \neq \hat{u}(\mathcal{X})$. The results below show that $u(\mathcal{X})$ still satisfies the appropriate properties in this case. §3.4.1 shows by counterexample that $\hat{u}(\mathcal{X})$ does not.

Theorem 16 *For every $(u, B, \mathbb{R}) \in \mathcal{L}$, $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$ is a McCormick extension of $(u, B, \mathbb{R})$ on $\mathbb{M}B$.*

Proof Choose any $(u, B, \mathbb{R}) \in \mathcal{L}$. By the definition of $\mathbb{M}B$, for any $x \in B$, we must have $(x, x, x, x) \in \mathbb{M}B$. Hence, $\mathbb{M}B$ is a McCormick extension of B. Now, choose any $\mathcal{X} \in \mathbb{M}B$ with $\mathcal{X} = (x, x, x, x)$. Moreover, let $\mathcal{Z} = u(\mathcal{X})$. It will be shown that $\mathcal{Z} = (u(x), u(x), u(x), u(x))$. By Condition 1 of Assumption 1, we have $z^{\mathrm{L}} = z^{\mathrm{U}} = u(x)$. Thus, it remains to prove that $z^{\mathrm{cv}} = z^{\mathrm{cc}} = u(x)$.

By the definition of Cut (Definition 11), $\bar{\mathcal{X}} = \mathrm{Cut}(\mathcal{X}) = (x, x, x, x)$. Therefore, (60) yields

$$z^{\mathrm{cv}} = \lim_{t \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(x, t)) + u^{\mathrm{cv}}(X, \max(x, t)) - u^{\mathrm{cv}}(X, t)]. \tag{82}$$

If $x^{\min}(X) \leq x$, then

$$\begin{aligned} z^{\mathrm{cv}} &= \lim_{t \to x^{\min}(X)} [u^{\mathrm{cv}}(X, t) + u^{\mathrm{cv}}(X, x) - u^{\mathrm{cv}}(X, t)], \\ &= u^{\mathrm{cv}}(X, x). \end{aligned} \tag{83}$$

Alternatively, if $x^{\min}(X) > x$, then

$$\begin{aligned} z^{\mathrm{cv}} &= \lim_{t \to x^{\min}(X)} [u^{\mathrm{cv}}(X, x) + u^{\mathrm{cv}}(X, t) - u^{\mathrm{cv}}(X, t)], \\ &= u^{\mathrm{cv}}(X, x). \end{aligned} \tag{84}$$

Therefore, in either case, we have $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, x)$. But since $X = [x, x]$, Condition 5 of Assumption 1 ensures that $z^{\mathrm{cv}} = u^{\mathrm{cv}}(X, x) = u(x)$. It can be shown in a similar way that $z^{\mathrm{cc}} = u(x)$. Therefore, $z^{\mathrm{cv}} = z^{\mathrm{cc}} = u(x)$ as desired. □

Theorem 17 *For all $(u, B, \mathbb{R}) \in \mathcal{L}$, $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$ is coherently concave on $\mathbb{M}B$.*

Proof Choose any $(u, B, \mathbb{R}) \in \mathcal{L}$. By definition, $\mathbb{M}B$ is closed under coherence. Choose any coherent $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{M}B$ and let $X = [x^{\mathrm{L}}, x^{\mathrm{U}}]$ without subscripts denote their common interval part. Let $\mathcal{Z}_1 = u(\mathcal{X}_1)$ and $\mathcal{Z}_2 = u(\mathcal{X}_2)$. Furthermore, choose any $\lambda \in [0, 1]$ and let $\mathcal{X}_\lambda = \mathrm{Conv}(\lambda, \mathcal{X}_1, \mathcal{X}_2)$ and $\mathcal{Z}_\lambda = u(\mathcal{X}_\lambda)$. We must show that

$$\mathcal{Z}_\lambda \succeq \mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2). \tag{85}$$

Since $\mathcal{X}_1$ and $\mathcal{X}_2$ are coherent, $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are coherent as well with common interval part $u(X)$. By Definition 7, this implies that the interval part of $\mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2)$

is also $u(X)$. Similarly, the convex combination $\mathcal{X}_\lambda$ has interval part X, which implies that the interval part of $\mathcal{Z}_\lambda = u(\mathcal{X}_\lambda)$ is $u(X)$ as well. Therefore, the left-hand and right-hand sides of (85) have the same interval parts. Then, to prove (85), it suffices to show that $z_\lambda^{\mathrm{cv}} \le \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}$ and $z_\lambda^{\mathrm{cc}} \ge \lambda z_1^{\mathrm{cc}} + (1-\lambda) z_2^{\mathrm{cc}}$. Only the first inequality will be shown. The second is proven analogously.

Suppose first that $x^{\min}(X)$ is finite. By Definition 16, in this case z_1^{cv}, z_2^{cv}, and z_λ^{cv} are all given by (56). In particular,

$$\begin{aligned} z_\lambda^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \max(\bar{x}_\lambda^{\mathrm{cv}}, x^{\min}(X))) + u^{\mathrm{cv}}(X, \min(\bar{x}_\lambda^{\mathrm{cc}}, x^{\min}(X))) \\ - u^{\mathrm{cv}}(X, x^{\min}(X)), \end{aligned} \tag{86}$$

where $\bar{\mathcal{X}}_\lambda = \mathrm{Cut}(\mathcal{X}_\lambda)$. Using the definition of $\bar{x}_\lambda^{\mathrm{cv}}$ and the convexity of $\max(\cdot,\cdot)$ with respect to both arguments,

$$\begin{aligned} \max(\bar{x}_\lambda^{\mathrm{cv}}, x^{\min}(X)) &= \max(\max(x^{\mathrm{L}}, \lambda x_1^{\mathrm{cv}} + (1-\lambda) x_2^{\mathrm{cv}}), x^{\min}(X)), \\ &\le \max(\lambda \max(x^{\mathrm{L}}, x_1^{\mathrm{cv}}) + (1-\lambda)\max(x^{\mathrm{L}}, x_2^{\mathrm{cv}}), x^{\min}(X)), \\ &= \max(\lambda \bar{x}_1^{\mathrm{cv}} + (1-\lambda)\bar{x}_2^{\mathrm{cv}}, x^{\min}(X)), \\ &\le \lambda \max(\bar{x}_1^{\mathrm{cv}}, x^{\min}(X)) + (1-\lambda)\max(\bar{x}_2^{\mathrm{cv}}, x^{\min}(X)). \end{aligned} \tag{87}$$

Since $\max(\bar{x}_\lambda^{\mathrm{cv}}, x^{\min}(X))$ is to the right of $x^{\min}(X)$ and $u^{\mathrm{cv}}(X,\cdot)$ is non-decreasing to the right of $x^{\min}(X)$ by Proposition 1, Eq. (87) implies that

$$\begin{aligned} &u^{\mathrm{cv}}(X, \max(\bar{x}_\lambda^{\mathrm{cv}}, x^{\min}(X))) \\ \le\, &u^{\mathrm{cv}}(X, \lambda \max(\bar{x}_1^{\mathrm{cv}}, x^{\min}(X)) + (1-\lambda)\max(\bar{x}_2^{\mathrm{cv}}, x^{\min}(X))). \end{aligned} \tag{88}$$

By the convexity of $u^{\mathrm{cv}}(X,\cdot)$ on $\mathbb{R}$, (88) further implies that

$$\begin{aligned} u^{\mathrm{cv}}(X, \max(\bar{x}_\lambda^{\mathrm{cv}}, x^{\min}(X))) \le \lambda \left[u^{\mathrm{cv}}(X, \max(\bar{x}_1^{\mathrm{cv}}, x^{\min}(X))) \right] \\ + (1-\lambda) \left[u^{\mathrm{cv}}(X, \max(\bar{x}_2^{\mathrm{cv}}, x^{\min}(X))) \right]. \end{aligned} \tag{89}$$

By an analogous sequence of arguments using the facts that $\min(\cdot,\cdot)$ is concave with respect to both arguments, $\min(\bar{x}_\lambda^{\mathrm{cc}}, x^{\min}(X))$ is to the left of $x^{\min}(X)$, $u^{\mathrm{cv}}(X,\cdot)$ is non-increasing to the left of $x^{\min}(X)$, and $u^{\mathrm{cv}}(X,\cdot)$ is convex on $\mathbb{R}$, we can establish that

$$\begin{aligned} u^{\mathrm{cv}}(X, \min(\bar{x}_\lambda^{\mathrm{cc}}, x^{\min}(X))) \le \lambda \left[u^{\mathrm{cv}}(X, \min(\bar{x}_1^{\mathrm{cc}}, x^{\min}(X))) \right] \\ + (1-\lambda) \left[u^{\mathrm{cv}}(X, \min(\bar{x}_2^{\mathrm{cc}}, x^{\min}(X))) \right]. \end{aligned} \tag{90}$$

Combining (86), (89), and (90), we have

$$\begin{aligned} z_\lambda^{\mathrm{cv}} &= u^{\mathrm{cv}}(X, \max(\bar{x}_\lambda^{\mathrm{cv}}, x^{\min}(X))) \\ &\qquad + u^{\mathrm{cv}}(X, \min(\bar{x}_\lambda^{\mathrm{cc}}, x^{\min}(X))) - u^{\mathrm{cv}}(X, x^{\min}(X)), \\ &\le \lambda [u^{\mathrm{cv}}(X, \max(\bar{x}_1^{\mathrm{cv}}, x^{\min}(X))) \\ &\qquad + u^{\mathrm{cv}}(X, \min(\bar{x}_1^{\mathrm{cc}}, x^{\min}(X))) - u^{\mathrm{cv}}(X, x^{\min}(X))] \\ &\quad + (1-\lambda)[u^{\mathrm{cv}}(X, \max(\bar{x}_2^{\mathrm{cv}}, x^{\min}(X))) \\ &\qquad + u^{\mathrm{cv}}(X, \min(\bar{x}_2^{\mathrm{cc}}, x^{\min}(X))) - u^{\mathrm{cv}}(X, x^{\min}(X))], \\ &\le \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}, \end{aligned} \tag{91}$$

as desired.

Suppose now that $x^{\min}(X) = +\infty$. By Definition 16, in this case z_1^{cv}, z_2^{cv}, and z_λ^{cv} are all given by (58). Thus, beginning from (58) we have,

$$\begin{aligned} z_\lambda^{\mathrm{cv}} &= u^{\mathrm{cv}}(X, \bar{x}_\lambda^{\mathrm{cc}}), \\ &= u^{\mathrm{cv}}(X, \min(x^{\mathrm{U}}, x_\lambda^{\mathrm{cc}})), \\ &= u^{\mathrm{cv}}(X, \min(x^{\mathrm{U}}, \lambda x_1^{\mathrm{cc}} + (1-\lambda) x_2^{\mathrm{cc}})). \end{aligned} \tag{92}$$

Since $x^{\min}(X) = +\infty$, $u^{\mathrm{cv}}(X, \cdot)$ is non-increasing on all of $\mathbb{R}$. Using this along with the facts that $\min(x^{\mathrm{U}}, \cdot)$ is concave on $\mathbb{R}$ and $u^{\mathrm{cv}}(X, \cdot)$ is convex on $\mathbb{R}$, (92) implies that

$$\begin{aligned} z_\lambda^{\mathrm{cv}} &\le u^{\mathrm{cv}}(X, \lambda \min(x^{\mathrm{U}}, x_1^{\mathrm{cc}}) + (1-\lambda)\min(x^{\mathrm{U}}, x_2^{\mathrm{cc}})), \\ &\le \lambda u^{\mathrm{cv}}(X, \min(x^{\mathrm{U}}, x_1^{\mathrm{cc}})) + (1-\lambda) u^{\mathrm{cv}}(X, \min(x^{\mathrm{U}}, x_2^{\mathrm{cc}})), \\ &= \lambda u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cc}}) + (1-\lambda) u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cc}}), \\ &= \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}. \end{aligned} \tag{93}$$

This proves $z_\lambda^{\mathrm{cv}} \le \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}$ for the case where $x^{\min}(X) = +\infty$. The case where $x^{\min}(X) = -\infty$ can be proven analogously. □

Theorem 18 *For all* $(u, B, \mathbb{R}) \in \mathcal{L}$, $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$ *is inclusion monotonic on* $\{\mathcal{X} \in \mathbb{M}B : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$.

Proof Choose any $(u, B, \mathbb{R}) \in \mathcal{L}$ and any nonempty $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{M}B$ with $\mathcal{X}_1 \preceq \mathcal{X}_2$. Let $\mathcal{Z}_1 = u(\mathcal{X}_1)$ and $\mathcal{Z}_2 = u(\mathcal{X}_2)$. We must prove that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$.

We prove $\mathcal{Z}_1 \preceq \mathcal{Z}_2$ by applying Theorem 2.4.29 from [9], which establishes the inclusion monotonicity of $(\hat{u}, \mathbb{M}B, \mathbb{M}\mathbb{R})$, but using slightly different definitions of $\mathbb{M}B$ and *inclusion monotonicity* than are used here. Let $\hat{\mathcal{Z}}_1 = \hat{u}(\mathcal{X}_1)$ and $\hat{\mathcal{Z}}_2 = \hat{u}(\mathcal{X}_2)$. By Definition 3, the fact that $\mathrm{Encl}(\mathcal{X}_1) \neq \emptyset$ implies that $x_1^{\mathrm{L}} \le x_1^{\mathrm{U}}$, $x_1^{\mathrm{cv}} \le x_1^{\mathrm{cc}}$, and $[x_1^{\mathrm{L}}, x_1^{\mathrm{U}}] \cap [x_1^{\mathrm{cv}}, x_1^{\mathrm{cc}}] \neq \emptyset$. Moreover, the interval $[x_1^{\mathrm{L}}, x_1^{\mathrm{U}}] \times [x_1^{\mathrm{cv}}, x_1^{\mathrm{cc}}]$ is a nonempty subset of $B \times \mathbb{R}$. The same is clearly true of $\mathcal{X}_2$. Using this subset interpretation, Theorem 2.4.29 in [9] establishes that $\hat{u}$ is *inclusion monotonic* in the sense that

$$[x_1^{\mathrm{L}}, x_1^{\mathrm{U}}] \times [x_1^{\mathrm{cv}}, x_1^{\mathrm{cc}}] \subset [x_2^{\mathrm{L}}, x_2^{\mathrm{U}}] \times [x_2^{\mathrm{cv}}, x_2^{\mathrm{cc}}] \tag{94}$$

$$\Longrightarrow [\hat{z}_1^{\mathrm{L}}, \hat{z}_1^{\mathrm{U}}] \times [\hat{z}_1^{\mathrm{cv}}, \hat{z}_1^{\mathrm{cc}}] \subset [\hat{z}_2^{\mathrm{L}}, \hat{z}_2^{\mathrm{U}}] \times [\hat{z}_2^{\mathrm{cv}}, \hat{z}_2^{\mathrm{cc}}]. \tag{95}$$

Since (94) follows from the fact that $\mathcal{X}_1 \preceq \mathcal{X}_2$, (95) holds. To apply this result to $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$, we use Lemma 6, which shows that

$$\hat{u}(\mathcal{X}) = u(\mathcal{X}), \quad \forall \mathcal{X} \in \mathbb{M}B \ \text{ s.t. } \ \mathrm{Encl}(\mathcal{X}) \neq \emptyset. \tag{96}$$

Applying this relation to $\mathcal{X}_1$ and $\mathcal{X}_2$ shows that $\hat{\mathcal{Z}}_1 = \mathcal{Z}_1$ and $\hat{\mathcal{Z}}_2 = \mathcal{Z}_2$. Substituting these relations into (95) gives

$$[z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] \times [z_1^{\mathrm{cv}}, z_1^{\mathrm{cc}}] \subset [z_2^{\mathrm{L}}, z_2^{\mathrm{U}}] \times [z_2^{\mathrm{cv}}, z_2^{\mathrm{cc}}]. \tag{97}$$

From this, it follows immediately that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$, as desired. □

Theorem 19 *For all* $(u, B, \mathbb{R}) \in \mathcal{L}$*,* $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$ *is coherently inclusion monotonic on* $\mathbb{M}B$*.*

Proof Choose any $(u, B, \mathbb{R}) \in \mathcal{L}$ and any coherent $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{M}B$ with $\mathcal{X}_1 \preceq \mathcal{X}_2$ and let $X = [x^{\mathrm{L}}, x^{\mathrm{U}}]$ without subscripts denote their common interval part. Let $\mathcal{Z}_1 = u(\mathcal{X}_1)$ and $\mathcal{Z}_2 = u(\mathcal{X}_2)$. We must prove that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$. Since the interval parts of $\mathcal{Z}_1$ and $\mathcal{Z}_2$ depend only on $[x^{\mathrm{L}}, x^{\mathrm{U}}]$, we must have $[z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] = [z_2^{\mathrm{L}}, z_2^{\mathrm{U}}]$. It remains to prove that $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$ and $z_1^{\mathrm{cc}} \leq z_2^{\mathrm{cc}}$.

If both $\mathcal{X}_1$ and $\mathcal{X}_2$ are nonempty, then the result follows immediately from Theorem 18. Thus, we assume that at least one of $\mathcal{X}_1$ and $\mathcal{X}_2$ is empty. Since $\mathcal{X}_1 \preceq \mathcal{X}_2$, we cannot have $\mathcal{X}_1$ nonempty and $\mathcal{X}_2$ empty. Therefore, there are only two remaining cases: (i) both $\mathcal{X}_1$ and $\mathcal{X}_2$ are empty, and (ii) $\mathcal{X}_1$ is empty and $\mathcal{X}_2$ is nonempty. For these remaining cases, we prove that $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$. The proof that $z_1^{\mathrm{cc}} \leq z_2^{\mathrm{cc}}$ is analogous. Let $\bar{\mathcal{X}}_1 = \mathrm{Cut}(\mathcal{X}_1)$ and $\bar{\mathcal{X}}_2 = \mathrm{Cut}(\mathcal{X}_2)$ and note that $\bar{\mathcal{X}}_1 \preceq \bar{\mathcal{X}}_2$ by Theorem 3. Recall from Remark 5 that

$$
\begin{aligned}
z_1^{\mathrm{cv}} &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(\bar{x}_1^{\mathrm{cc}}, x)) + u^{\mathrm{cv}}(X, \max(\bar{x}_1^{\mathrm{cv}}, x)) - u^{\mathrm{cv}}(X, x)], \qquad (98)\\
z_2^{\mathrm{cv}} &= \lim_{x \to x^{\min}(X)} [u^{\mathrm{cv}}(X, \min(\bar{x}_2^{\mathrm{cc}}, x)) + u^{\mathrm{cv}}(X, \max(\bar{x}_2^{\mathrm{cv}}, x)) - u^{\mathrm{cv}}(X, x)].
\end{aligned}
$$

Case (i): $\mathrm{Encl}(\mathcal{X}_1) = \emptyset$*,* $\mathrm{Encl}(\mathcal{X}_2) = \emptyset$*.*

Since $\mathcal{X}_2$ is empty, $\bar{\mathcal{X}}_2$ must satisfy $\bar{x}_2^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cv}}$. Since $\bar{\mathcal{X}}_1 \preceq \bar{\mathcal{X}}_2$, we can additionally conclude that $\bar{x}_1^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cv}} \leq \bar{x}_1^{\mathrm{cv}}$. Thus, we consider five subcases based on where the value of $x^{\min}(X)$ falls within this chain of inequalities.

Case (i)(a): $x^{\min}(X) \leq \bar{x}_1^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cv}} \leq \bar{x}_1^{\mathrm{cv}}$. By (98), z_2^{cv} simplifies to $u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cv}})$ and z_1^{cv} simplifies to $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cv}})$. Since $u^{\mathrm{cv}}(X, \cdot)$ is convex, it is non-decreasing to the right of $x^{\min}(X)$ by Proposition 1. Since $x^{\min}(X) \leq \bar{x}_2^{\mathrm{cv}} \leq \bar{x}_1^{\mathrm{cv}}$, it follows that $z_1^{\mathrm{cv}} = u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cv}}) \geq u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cv}}) = z_2^{\mathrm{cv}}$, as desired.

Case (i)(b): $\bar{x}_1^{\mathrm{cc}} \leq x^{\min}(X) \leq \bar{x}_2^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cv}} \leq \bar{x}_1^{\mathrm{cv}}$. By (98), z_2^{cv} simplifies to $u^{\mathrm{cv}}(X, x^{\min}(X)) + u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cv}}) - u^{\mathrm{cv}}(X, x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cc}}) + u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cv}}) - u^{\mathrm{cv}}(X, x^{\min}(X))$. Comparing these term-wise, we will have $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$ as desired if $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cc}}) \geq u^{\mathrm{cv}}(X, x^{\min}(X))$ and $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cv}}) \geq u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cv}})$. The former follows from the definition of $x^{\min}(X)$, while the latter follows from the fact that $u^{\mathrm{cv}}(X, \cdot)$ is non-decreasing to the right of $x^{\min}(X)$ (Proposition 1) and $x^{\min}(X) \leq \bar{x}_2^{\mathrm{cv}} \leq \bar{x}_1^{\mathrm{cv}}$.

Case (i)(c): $\bar{x}_1^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cc}} \leq x^{\min}(X) \leq \bar{x}_2^{\mathrm{cv}} \leq \bar{x}_1^{\mathrm{cv}}$. By (98), z_2^{cv} simplifies to $u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cc}}) + u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cv}}) - u^{\mathrm{cv}}(X, x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cc}}) + u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cv}}) - u^{\mathrm{cv}}(X, x^{\min}(X))$. Comparing these term-wise, we will have $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$ as desired if $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cc}}) \geq u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cc}})$ and $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cv}}) \geq u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cv}})$. The former holds because $u^{\mathrm{cv}}(X, \cdot)$ is non-increasing to the left of $x^{\min}(X)$ (Proposition 1) and $\bar{x}_1^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cc}} \leq x^{\min}(X)$, while the latter holds because $u^{\mathrm{cv}}(X, \cdot)$ is non-decreasing to the right of $x^{\min}(X)$ and $x^{\min}(X) \leq \bar{x}_2^{\mathrm{cv}} \leq \bar{x}_1^{\mathrm{cv}}$.

Case (i)(d): $\bar{x}_1^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cc}} \leq \bar{x}_2^{\mathrm{cv}} \leq x^{\min}(X) \leq \bar{x}_1^{\mathrm{cv}}$. By (98), z_2^{cv} simplifies to $u^{\mathrm{cv}}(X, \bar{x}_2^{\mathrm{cc}}) + u^{\mathrm{cv}}(X, x^{\min}(X)) - u^{\mathrm{cv}}(X, x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\mathrm{cv}}(X, \bar{x}_1^{\mathrm{cc}})+$

$u^{\text{cv}}(X,\bar{x}_1^{\text{cv}})-u^{\text{cv}}(X,x^{\min}(X))$. Comparing these term-wise, we will have $z_1^{\text{cv}} \geq z_2^{\text{cv}}$ as desired if $u^{\text{cv}}(X,\bar{x}_1^{\text{cc}}) \geq u^{\text{cv}}(X,\bar{x}_2^{\text{cc}})$ and $u^{\text{cv}}(X,\bar{x}_1^{\text{cv}}) \geq u^{\text{cv}}(X,x^{\min}(X))$. The former holds because $u^{\text{cv}}(X,\cdot)$ is non-increasing to the left of $x^{\min}(X)$ (Proposition 1) and $\bar{x}_1^{\text{cc}} \leq \bar{x}_2^{\text{cc}} \leq x^{\min}(X)$, while the latter holds by the definition of $x^{\min}(X)$.

Case (i)(e): $\bar{x}_1^{\text{cc}} \leq \bar{x}_2^{\text{cc}} \leq \bar{x}_2^{\text{cv}} \leq \bar{x}_1^{\text{cv}} \leq x^{\min}(X)$. By (98), z_2^{cv} simplifies to $u^{\text{cv}}(X,\bar{x}_2^{\text{cc}})$ and z_1^{cv} simplifies to $u^{\text{cv}}(X,\bar{x}_1^{\text{cc}})$. Since $u^{\text{cv}}(X,\cdot)$ is non-increasing to the left of $x^{\min}(X)$ and $\bar{x}_1^{\text{cc}} \leq \bar{x}_2^{\text{cc}} \leq x^{\min}(X)$, it follows that $z_1^{\text{cv}} = u^{\text{cv}}(X,\bar{x}_1^{\text{cc}}) \geq u^{\text{cv}}(X,\bar{x}_2^{\text{cc}}) = z_2^{\text{cv}}$, as desired.

From the previous five subcases, we conclude that $z_1^{\text{cv}} \geq z_2^{\text{cv}}$ in Case (i).

Case (ii): $\text{Encl}(\mathcal{X}_1) = \emptyset$, $\text{Encl}(\mathcal{X}_2) \neq \emptyset$.

Since $\mathcal{X}_1$ is empty, $\mathcal{X}_2$ is nonempty, and $\mathcal{X}_1 \preceq \mathcal{X}_2$, the definition of Cut (Definition 11) and Theorems 3–4 imply that $\bar{\mathcal{X}}_1$ is empty, $\bar{\mathcal{X}}_2$ is nonempty, and $\bar{\mathcal{X}}_1 \preceq \bar{\mathcal{X}}_2$. In turn, these conditions imply the four requirements below:

$$\bar{x}_1^{\text{cc}} \leq \bar{x}_1^{\text{cv}}, \tag{99}$$

$$\bar{x}_2^{\text{cv}} \leq \bar{x}_2^{\text{cc}}, \tag{100}$$

$$\bar{x}_1^{\text{cc}} \leq \bar{x}_2^{\text{cc}}, \tag{101}$$

$$\bar{x}_2^{\text{cv}} \leq \bar{x}_1^{\text{cv}}. \tag{102}$$

Assign the four values $\bar{x}_1^{\text{cv}}$, $\bar{x}_1^{\text{cc}}$, $\bar{x}_2^{\text{cv}}$, and $\bar{x}_2^{\text{cc}}$ to the notations r_1, r_2, r_3, and r_4 in such a way that $r_1 \leq r_2 \leq r_3 \leq r_4$. We now consider five subcases based on where $x^{\min}(X)$ lies within this chain of inequalities.

Case (ii)(a): $x^{\min}(X) \leq r_1 \leq r_2 \leq r_3 \leq r_4$. By (98), z_2^{cv} simplifies to $u^{\text{cv}}(X,\bar{x}_2^{\text{cv}})$ and z_1^{cv} simplifies to $u^{\text{cv}}(X,\bar{x}_1^{\text{cv}})$. Since $u^{\text{cv}}(X,\cdot)$ is non-decreasing to the right of $x^{\min}(X)$ and $x^{\min}(X) \leq \bar{x}_2^{\text{cv}} \leq \bar{x}_1^{\text{cv}}$, it follows that $z_1^{\text{cv}} = u^{\text{cv}}(X,\bar{x}_1^{\text{cv}}) \geq u^{\text{cv}}(X,\bar{x}_2^{\text{cv}}) = z_2^{\text{cv}}$, as desired.

Case (ii)(b): $r_1 \leq x^{\min}(X) \leq r_2 \leq r_3 \leq r_4$. Considering (99)–(102), either $r_1 = \bar{x}_2^{\text{cv}}$ or $r_1 = \bar{x}_1^{\text{cc}}$.

Case (ii)(b)(i): $r_1 = \bar{x}_2^{\text{cv}}$. By (98), z_2^{cv} simplifies to $u^{\text{cv}}(X,x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\text{cv}}(X,\bar{x}_1^{\text{cv}})$. By the definition of $x^{\min}(X)$, it follows that $z_1^{\text{cv}} = u^{\text{cv}}(X,\bar{x}_1^{\text{cv}}) \geq u^{\text{cv}}(X,x^{\min}(X)) = z_2^{\text{cv}}$, as desired.

Case (ii)(b)(ii): $r_1 = \bar{x}_1^{\text{cc}}$. By (98), z_2^{cv} can be simplified to $u^{\text{cv}}(X,x^{\min}(X)) + u^{\text{cv}}(X,\bar{x}_2^{\text{cv}}) - u^{\text{cv}}(X,x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\text{cv}}(X,\bar{x}_1^{\text{cc}}) + u^{\text{cv}}(X,\bar{x}_1^{\text{cv}}) - u^{\text{cv}}(X,x^{\min}(X))$. Comparing these term-wise, we will have $z_1^{\text{cv}} \geq z_2^{\text{cv}}$ as desired provided that $u^{\text{cv}}(X,\bar{x}_1^{\text{cc}}) \geq u^{\text{cv}}(X,x^{\min}(X))$ and $u^{\text{cv}}(X,\bar{x}_1^{\text{cv}}) \geq u^{\text{cv}}(X,\bar{x}_2^{\text{cv}})$. The former follows from the definition of $x^{\min}(X)$, while the latter holds because $u^{\text{cv}}(X,\cdot)$ is non-decreasing to the right of $x^{\min}(X)$ and $x^{\min}(X) \leq \bar{x}_2^{\text{cv}} \leq \bar{x}_1^{\text{cv}}$.

Case (ii)(c): $r_1 \leq r_2 \leq x^{\min}(X) \leq r_3 \leq r_4$. The inequalities (99)–(102) imply that $\max(\bar{x}_2^{\text{cv}},\bar{x}_1^{\text{cc}}) \leq \min(\bar{x}_2^{\text{cc}},\bar{x}_1^{\text{cv}})$. Therefore, we must have $\bar{x}_2^{\text{cv}} \leq x^{\min}(X) \leq \bar{x}_2^{\text{cc}}$ and $\bar{x}_1^{\text{cc}} \leq x^{\min}(X) \leq \bar{x}_1^{\text{cv}}$. Then, by (98), z_2^{cv} simplifies to $u^{\text{cv}}(X,x^{\min}(X)) +$

$u^{\text{cv}}(X, x^{\min}(X)) - u^{\text{cv}}(X, x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\text{cv}}(X, \bar{x}_1^{\text{cc}}) + u^{\text{cv}}(X, \bar{x}_1^{\text{cv}}) - u^{\text{cv}}(X, x^{\min}(X))$. Comparing these term-wise, we will have $z_1^{\text{cv}} \geq z_2^{\text{cv}}$ as desired provided that $u^{\text{cv}}(X, \bar{x}_1^{\text{cc}}) \geq u^{\text{cv}}(X, x^{\min}(X))$ and $u^{\text{cv}}(X, \bar{x}_1^{\text{cv}}) \geq u^{\text{cv}}(X, x^{\min}(X))$. Both the former and the latter follow from the definition of $x^{\min}(X)$.

Case (ii)(d): $r_1 \leq r_2 \leq r_3 \leq x^{\min}(X) \leq r_4$. Considering (99)–(102), either $r_4 = \bar{x}_1^{\text{cv}}$ or $r_4 = \bar{x}_2^{\text{cc}}$.

Case (ii)(d)(i): $r_4 = \bar{x}_1^{\text{cv}}$. By (98), z_2^{cv} can be simplified to $u^{\text{cv}}(X, \bar{x}_2^{\text{cc}}) + u^{\text{cv}}(X, x^{\min}(X)) - u^{\text{cv}}(X, x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\text{cv}}(X, \bar{x}_1^{\text{cc}}) + u^{\text{cv}}(X, \bar{x}_1^{\text{cv}}) - u^{\text{cv}}(X, x^{\min}(X))$. Comparing these term-wise, we will have $z_1^{\text{cv}} \geq z_2^{\text{cv}}$ as desired provided that $u^{\text{cv}}(X, \bar{x}_1^{\text{cc}}) \geq u^{\text{cv}}(X, \bar{x}_2^{\text{cc}})$ and $u^{\text{cv}}(X, \bar{x}_1^{\text{cv}}) \geq u^{\text{cv}}(X, x^{\min}(X))$. The former holds because $u^{\text{cv}}(X, \cdot)$ is non-increasing to the left of $x^{\min}(X)$ and $\bar{x}_1^{\text{cc}} \leq \bar{x}_2^{\text{cc}} \leq x^{\min}(X)$, while the latter follows from the definition of $x^{\min}(X)$.

Case (ii)(d)(ii): $r_4 = \bar{x}_2^{\text{cc}}$. By (98), z_2^{cv} simplifies to $u^{\text{cv}}(X, x^{\min}(X))$ and z_1^{cv} simplifies to $u^{\text{cv}}(X, \bar{x}_1^{\text{cc}})$. By the definition of $x^{\min}(X)$, it follows that $z_1^{\text{cv}} = u^{\text{cv}}(X, \bar{x}_1^{\text{cc}}) \geq u^{\text{cv}}(X, x^{\min}(X)) = z_2^{\text{cv}}$, as desired.

Case (ii)(e): $r_1 \leq r_2 \leq r_3 \leq r_4 \leq x^{\min}(X)$. By (98), z_2^{cv} simplifies to $u^{\text{cv}}(X, \bar{x}_2^{\text{cc}})$ and z_1^{cv} simplifies to $u^{\text{cv}}(X, \bar{x}_1^{\text{cc}})$. Since $u^{\text{cv}}(X, \cdot)$ is non-increasing to the left of $x^{\min}(X)$ and $\bar{x}_1^{\text{cc}} \leq \bar{x}_2^{\text{cc}} \leq x^{\min}(X)$, it follows that $z_1^{\text{cv}} = u^{\text{cv}}(X, \bar{x}_1^{\text{cc}}) \geq u^{\text{cv}}(X, \bar{x}_2^{\text{cc}}) = z_2^{\text{cv}}$, as desired.

From the previous five subcases, we conclude that $z_1^{\text{cv}} \geq z_2^{\text{cv}}$ also holds in Case (ii). □

Theorem 20 $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$ *preserves nonemptiness on* $\mathbb{M}B$.

Proof The result follows from Lemma 1 with Theorems 16 and 18. □

Theorem 21 $(u, \mathbb{M}B, \mathbb{M}\mathbb{R})$ *is a relaxation function of* $(u, B, \mathbb{R})$.

Proof The result follows from Theorem 2 with Theorems 16–19. □

3.4.1 A Composition Example

Let $x : P \to \mathbb{R}$ be an arbitrary function on $P = [-2, 2]$. We consider computing relaxations of the composite function

$$f(p) = \exp(x(p)) \tag{103}$$

on P given relaxation information for x on P in the form of a McCormick function $\mathcal{X} : P \to \mathbb{M}\mathbb{R}$. Accordingly, let $u = \exp$, let $(u, \mathbb{M}\mathbb{R}, \mathbb{M}\mathbb{R})$ and $(\hat{u}, \mathbb{M}\mathbb{R}, \mathbb{M}\mathbb{R})$ denote the extended and standard relaxation functions for u defined above, and let $\mathcal{F}(p) = u(\mathcal{X}(p))$ and $\hat{\mathcal{F}}(p) = \hat{u}(\mathcal{X}(p))$ with $\mathcal{X}(p)$ defined as in the left panel of Figure 2. Note that $\mathcal{X}(p)$ is nonempty for all $p \in P^* \equiv [-0.87, 0.87]$ and empty otherwise. Even so, x^{cv} and x^{cc} are convex and concave, respectively, on all of P. As shown in the middle and right panels of Figure 2, the two composition rules agree for all p such that $\mathcal{X}(p)$ is nonempty. However, they clearly differ at infeasible p values.

While $\hat{\mathcal{F}}(p)$ remains nonempty for all p, $\hat{f}^{\mathrm{cc}}$ is clearly not concave. In contrast, $\mathcal{F}(p)$ takes empty values for some infeasible p to ensure that f^{cv} and f^{cc} are convex and concave on all of P, which is the desired behavior. Although the standard rule failed by producing a nonconcave $\hat{f}^{\mathrm{cc}}$ in this case, in general its behavior is unpredictable when $\mathcal{X}(p)$ is empty; $\hat{f}^{\mathrm{cv}}$ may be nonconvex, $\hat{\mathcal{F}}(p)$ may be empty, or $\hat{u}(\mathcal{X}(p))$ may be undefined due to a domain violation (this can occur when $u^{\mathrm{cv}}(X,\cdot)$ and $u^{\mathrm{cc}}(X,\cdot)$ are not defined on all of $\mathbb{R}$).

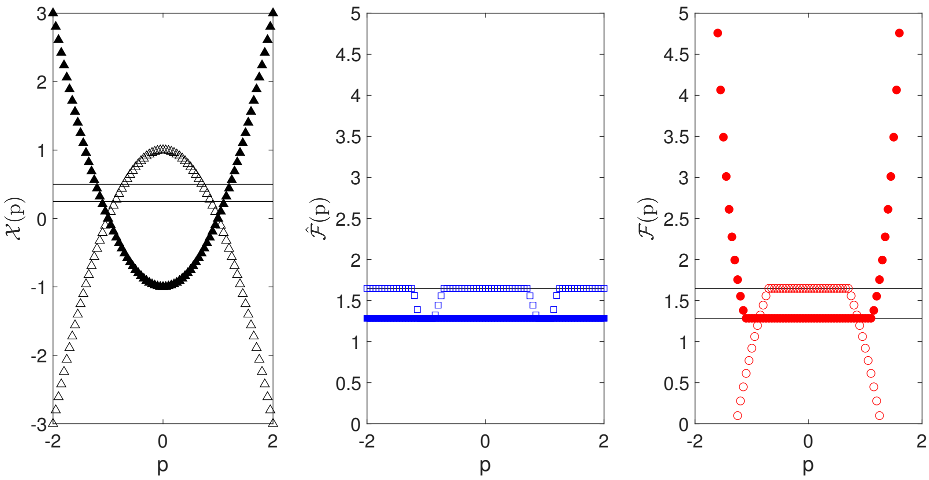

Fig. 2 Left: McCormick object $\mathcal{X}(p)$ corresponding to $x(p)$ in (103) defined by its bounds x^{L} and x^{U} (solid), convex component $x^{\mathrm{cv}}(p)$ (filled triangles), and concave component $x^{\mathrm{cc}}(p)$ (open triangles). Middle: McCormick object $\hat{\mathcal{F}}(p)$ corresponding to $f(p)$ in (103) defined by its bounds f^{L} and f^{U} (solid), convex component $\hat{f}^{\mathrm{cv}}(p)$ (filled squares), and concave component $\hat{f}^{\mathrm{cc}}(p)$ (open squares). Right: McCormick object $\mathcal{F}(p)$ corresponding to $f(p)$ in (103) defined by its bounds f^{L} and f^{U} (solid), convex component $f^{\mathrm{cv}}(p)$ (filled circles), and concave component $f^{\mathrm{cc}}(p)$ (open circles). For all $p \in [-0.87, 0.87]$, $\hat{f}^{\mathrm{cv}}(p)$ and $\hat{f}^{\mathrm{cc}}(p)$ are actually identical to $f^{\mathrm{cv}}(p)$ and $f^{\mathrm{cc}}(p)$, respectively.

4 Natural McCormick Extension

In this section, we define the *natural McCormick extension* of a factorable function, which is obtained by recursively applying the rules defined in the previous section to a sequence of elementary operations that defines the function. Using the composition results (Lemma 2) from §2, we then prove that this provides a valid relaxation function for the original function, which can be used to obtain convex and concave relaxations via Theorem 1 and Corollary 1.

To define factorable functions precisely, following [9] we first define the notion of a *computational sequence*, which is essentially an ordered list of elementary operations along with some indexing maps that connect the inputs of each operation to the outputs of earlier operations. Next, we define the functions created by applying the operations in this list recursively up to a generic step k (termed *factors*), which ultimately produces the complete function encoded by the list (termed the *natural function*). The notation below is heavier than is usually necessary to introduce these concepts, but it offers the precision needed to clearly define the natural McCormick extension and establish its properties.

Definition 19 Let $n_i, n_o, n_f \in \mathbb{N}$. An *$\mathcal{L}$-computational sequence* with n_i inputs, n_o outputs, and n_f factors is a pair $(\mathcal{S}, \pi_o)$ where:

1. $\mathcal{S}$ is a finite sequence of pairs $\{((o_k, B_k, \mathbb{R}), (\pi_k, \mathbb{R}^{k-1}, \mathbb{R}^{d_k}))\}_{k=n_i+1}^{n_f}$, where $(o_k, B_k, \mathbb{R})$ is an elementary operation and $(\pi_k, \mathbb{R}^{k-1}, \mathbb{R}^{d_k})$ is an input selection map. Specifically, every element of $\mathcal{S}$ is defined by one of the following options:
   (a) $(o_k, B_k, \mathbb{R})$ is either $(+, \mathbb{R}^2, \mathbb{R})$ or $(\times, \mathbb{R}^2, \mathbb{R})$ and $\pi_k : \mathbb{R}^{k-1} \to \mathbb{R}^2$ is defined by $\pi_k(\mathbf{v}) = (v_i, v_j)$ for some integers $i, j \in \{1, \ldots, k-1\}$.
   (b) $(o_k, B_k, \mathbb{R}) \in \mathcal{L}$ where $\mathcal{L}$ is a library of univariate functions (see §3.4) and $\pi_k : \mathbb{R}^{k-1} \to \mathbb{R}$ is defined by $\pi_k(\mathbf{v}) = v_i$ for some integer $i \in \{1, \ldots, k-1\}$.
2. $\pi_o : \mathbb{R}^{n_f} \to \mathbb{R}^{n_o}$ is an output selection map defined by $\pi_o(\mathbf{v}) = (v_{i(1)}, \ldots, v_{i(n_o)})$ for some integers $i(1), \ldots, i(n_o) \in \{1, \ldots, n_f\}$.

Definition 20 Let $(\mathcal{S}, \pi_o)$ be an $\mathcal{L}$-computational sequence with n_i inputs and n_o outputs. Define the *sequence of factors* $\{(v_k, D_k, \mathbb{R})\}_{k=1}^{n_f}$ as follows:

1. For $k = 1, \ldots, n_i$, $D_k = \mathbb{R}^{n_i}$ and $v_k(\mathbf{x}) = x_k$, $\forall \mathbf{x} \in D_k$,
2. For $k = n_i + 1, \ldots, n_f$, $D_k = \{\mathbf{x} \in D_{k-1} : \pi_k(v_1(\mathbf{x}), \ldots, v_{k-1}(\mathbf{x})) \in B_k\}$ and $v_k(\mathbf{x}) = o_k \circ \pi_k \circ (v_1(\mathbf{x}), \ldots, v_{k-1}(\mathbf{x}))$, $\forall \mathbf{x} \in D_k$.

The set $D_{\mathcal{S}} \equiv D_{n_f}$ is called the *natural domain* of $(\mathcal{S}, \pi_o)$, and the *natural function* $(\mathbf{f}_{\mathcal{S}}, D_{\mathcal{S}}, \mathbb{R}^{n_o})$ is defined by $\mathbf{f}_{\mathcal{S}}(\mathbf{x}) = \pi_o \circ (v_1(\mathbf{x}), \ldots, v_{n_f}(\mathbf{x}))$, $\forall \mathbf{x} \in D_{\mathcal{S}}$.

Definition 21 A function $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ is called *$\mathcal{L}$-factorable* if there exists an $\mathcal{L}$-computational sequence $(\mathcal{S}, \pi_o)$ with n inputs and m outputs such that the natural function $(\mathbf{f}_{\mathcal{S}}, D_{\mathcal{S}}, \mathbb{R}^{n_o})$ satisfies $D \subset D_{\mathcal{S}}$ and $\mathbf{f} = \mathbf{f}_{\mathcal{S}}|_D$.

We are now prepared to define the natural McCormick extension of an $\mathcal{L}$-computational sequence (and hence one possible natural McCormick extension of an $\mathcal{L}$-factorable function) using the extended McCormick rules from §3. Since each $(o_k, B_k, \mathbb{R})$ in the sequence is one of the elementary operations considered in §3, McCormick extensions $(o_k, \mathbb{M}B_k, \mathbb{MR})$ satisfying all of the required properties have already been established. We also require McCormick versions of the input selections maps, $(\pi_k, \mathbb{MR}^{k-1}, \mathbb{MR}^{d_k})$, which simply select the same element(s) from the input vector as the corresponding real-valued version.

Definition 22 Let $(\mathcal{S}, \pi_o)$ be an $\mathcal{L}$-computational sequence with n_i inputs and n_o outputs. Define the *sequence of relaxation factors* $\{(\mathcal{V}_k, \mathcal{D}_k, \mathbb{MR})\}_{k=1}^{n_f}$, with $\mathcal{D}_k \subset \mathbb{MR}^{n_i}$, as follows:

1. For all $k = 1, \ldots, n_i$, $\mathcal{D}_k = \mathbb{MR}^{n_i}$ and $\mathcal{V}_k(\mathcal{X}) = \mathcal{X}_k$, $\forall \mathcal{X} \in \mathcal{D}_k$,
2. For all $k = n_i + 1, \ldots, n_f$, $\mathcal{D}_k = \{\mathcal{X} \in \mathcal{D}_{k-1} : \pi_k \circ (\mathcal{V}_1(\mathcal{X}), \ldots, \mathcal{V}_{k-1}(\mathcal{X})) \in \mathbb{M}B_k\}$ and $\mathcal{V}_k(\mathcal{X}) = o_k \circ \pi_k \circ (\mathcal{V}_1(\mathcal{X}), \ldots, \mathcal{V}_{k-1}(\mathcal{X}))$, $\forall \mathcal{X} \in \mathcal{D}_k$.

The *natural McCormick extension* of $(\mathcal{S}, \pi_o)$ is the function $(\mathcal{F}_{\mathcal{S}}, \mathcal{D}_{\mathcal{S}}, \mathbb{MR}^{n_o})$ defined by $\mathcal{D}_{\mathcal{S}} \equiv \mathcal{D}_{n_f}$ and $\mathcal{F}_{\mathcal{S}}(\mathcal{X}) = \pi_o \circ (\mathcal{V}_1(\mathcal{X}), \ldots, \mathcal{V}_{n_f}(\mathcal{X}))$, $\forall \mathcal{X} \in \mathcal{D}_{\mathcal{S}}$.

Definition 23 Let $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ be an $\mathcal{L}$-factorable function. Then, for any $\mathcal{L}$-computational sequence describing $\mathbf{f}$, the natural McCormick extension $(\mathcal{F}_{\mathcal{S}}, \mathcal{D}_{\mathcal{S}}, \mathbb{MR}^m)$ is called a natural McCormick extension of $\mathbf{f}$.

The next two results establish the properties of the natural McCormick extension, ultimately concluding that it is a relaxation function, as desired. In the proof by induction for establishing Theorem 22, the fact that $(o_K, \mathbb{M}B_K, \mathbb{MR})$ satisfies the required conditions is evident if o_K corresponds to either $(+, \mathbb{MR}^2, \mathbb{MR})$, $(\times, \mathbb{MR}^2, \mathbb{MR})$, or $(u, \mathbb{M}B, \mathbb{MR})$, as defined in §3.

Theorem 22 *Let $(\mathcal{S}, \pi_o)$ be an $\mathcal{L}$-computational sequence with natural function $(\mathbf{f}_\mathcal{S}, D_\mathcal{S}, \mathbb{R}^{n_o})$. The natural McCormick extension $(\mathcal{F}_\mathcal{S}, \mathcal{D}_\mathcal{S}, \mathbb{MR}^{n_o})$ is a McCormick extension of $(\mathbf{f}_\mathcal{S}, D_\mathcal{S}, \mathbb{R}^{n_o})$ on $\mathcal{D}_\mathcal{S}$. Moreover, it is coherently concave and coherently inclusion monotonic on $\mathcal{D}_\mathcal{S}$, is fully inclusion monotonic on $\{\mathcal{X} \in \mathcal{D}_\mathcal{S} : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$, and preserves nonemptiness on $\mathcal{D}_\mathcal{S}$.*

Proof Consider the sequence of factors $\{(v_k, D_k, \mathbb{R})\}_{k=1}^{n_f}$ and the sequence of relaxation factors $\{(\mathcal{V}_k, \mathcal{D}_k, \mathbb{MR})\}_{k=1}^{n_f}$. To set up an inductive argument, choose any $K \in \{n_i+1, \ldots, n_f\}$ and assume that, for all $k < K$, $(\mathcal{V}_k, \mathcal{D}_k, \mathbb{MR})$ is a McCormick extension of $(v_k, D_k, \mathbb{R})$ on $\mathcal{D}_k$, is coherently concave and coherently inclusion monotonic on $\mathcal{D}_k$, is fully inclusion monotonic on $\{\mathcal{X} \in \mathcal{D}_k : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$, and preserves nonemptiness on $\mathcal{D}_k$. This assumption is straightforward to verify for $K = n_i + 1$ because, for every $k < n_i + 1$, we have $v_k(\mathbf{x}) = x_k$ and $\mathcal{V}_k(\mathcal{X}) = \mathcal{X}_k$. Assuming it holds for an arbitrary K, we now prove that $\mathcal{V}_K$ satisfies these conditions as well. Recall that

$$\begin{aligned} v_K(\mathbf{x}) &= o_K \circ \pi_K \circ (v_1, \ldots, v_{K-1}), \\ \mathcal{V}_K(\mathcal{X}) &= o_K \circ \pi_K \circ (\mathcal{V}_1, \ldots, \mathcal{V}_{K-1}). \end{aligned}$$

Since π_K merely selects one or two of its arguments, the inductive hypothesis implies that $\pi_K \circ (\mathcal{V}_1, \ldots, \mathcal{V}_{K-1})$ is a McCormick extension of $\pi_K \circ (v_1, \ldots, v_{K-1})$ on $\mathcal{D}_{K-1}$, is coherently concave and coherently inclusion monotonic on $\mathcal{D}_{K-1}$, is fully inclusion monotonic on $\{\mathcal{X} \in \mathcal{D}_{K-1} : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$, and preserves nonemptiness on $\mathcal{D}_{K-1}$. Moreover, regardless of the identity of the operation $(o_K, B_K, \mathbb{R})$, the results in §3 ensure that $(o_K, \mathbb{M}B_K, \mathbb{MR})$ is a McCormick extension of $(o_K, B_K, \mathbb{R})$ on $\mathbb{M}B_K$ (Thm. 5, 10, 16), is coherently concave on $\mathbb{M}B_K$ (Thm. 6, 11, 17), is coherently inclusion monotonic on $\mathbb{M}B_K$ (Cor. 4, Thm. 13, 19), is fully inclusion monotonic on $\{\mathcal{X} \in \mathbb{M}B_K : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$ (Thm. 7, 12, 18), and preserves nonemptiness on $\mathbb{M}B_K$ (Thm. 8, 14, 20). Given these facts, Parts 2–6 of Lemma 2 show that $(\mathcal{V}_K, \mathcal{D}_K, \mathbb{MR})$ is a McCormick extension of $(v_K, D_K, \mathbb{R})$ on $\mathcal{D}_K$, is coherently concave and coherently inclusion monotonic on $\mathcal{D}_K$, is fully inclusion monotonic on $\{\mathcal{X} \in \mathcal{D}_K : \mathrm{Encl}(\mathcal{X}) \neq \emptyset\}$, and preserves nonemptiness on $\mathcal{D}_K$. By induction, this holds for every $K \in \{1, \ldots, n_f\}$, and the result then follows from the definition of $(\mathcal{F}_\mathcal{S}, \mathcal{D}_\mathcal{S}, \mathbb{MR}^{n_o})$. □

Corollary 5 *The natural McCormick extension $(\mathcal{F}_\mathcal{S}, \mathcal{D}_\mathcal{S}, \mathbb{MR}^{n_o})$ is a relaxation function of the natural function $(\mathbf{f}_\mathcal{S}, D_\mathcal{S}, \mathbb{R}^{n_o})$ on $\mathcal{D}_\mathcal{S}$.*

Proof The result follows immediately from Theorems 2 and 22. □

5 A Relaxation Refinement Operation

This section describes a constraint-based relaxation refinement algorithm that can be used to tighten the relaxations used in reduced space global optimization algorithms (see discussion in §1). This algorithm is very simple and only applies to linear constraints. It is included here primarily as a concrete example of how empty McCormick objects can be generated in the first place, and hence why we have emphasized the ability to compute with such objects throughout this paper. We begin by defining the notion of a *relaxation refinement operator* for a given set of constraints, which is analogous to the definition of a relaxation function for a given real function. Theorem 23 then shows that relaxation refinement operators have the desired behavior for use in reduced space global optimization; i.e., they produce convex and concave relaxations of the function of interest that are potentially tighter but still valid over the feasible domain. This use case is explored further by example in §5.1.

Definition 24 Let $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^{n_c}$. A function $(I, \mathbb{MR}^n, \mathbb{MR}^n)$ is a *relaxation refinement operator* for the constraints $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$ if it is coherently concave, coherently inclusion monotonic, and satisfies

$$\mathrm{Encl}(I(\mathcal{X})) \supset \{\mathbf{x} \in \mathrm{Encl}(\mathcal{X}) : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}, \quad \forall \mathcal{X} \in \mathbb{MR}^n.$$

Remark 9 For a relaxation refinement operator to be useful, $\mathrm{Encl}(I(\mathcal{X}))$ should be a strict subset of $\mathrm{Encl}(\mathcal{X})$ for at least some $\mathcal{X}$, but this is not required by the definition.

Theorem 23 *Let $\mathbf{x} : P \subset \mathbb{R}^{n_p} \to \mathbb{R}^{n_x}$ and $\mathbf{g} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_c}$. Let $\mathcal{X} : P \to \mathbb{MR}^{n_x}$ be a McCormick function with components $\mathcal{X}(\mathbf{p}) = (\mathbf{x}^{\mathrm{L}}, \mathbf{x}^{\mathrm{U}}, \mathbf{x}^{\mathrm{cv}}(\mathbf{p}), \mathbf{x}^{\mathrm{cc}}(\mathbf{p}))$ such that $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$ are convex and concave on P, respectively. If $(I, \mathbb{MR}^{n_x}, \mathbb{MR}^{n_x})$ is a relaxation refinement operator for the constraints $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$, then the object defined by*

$$\mathcal{X}^*(\mathbf{p}) \equiv I(\mathcal{X}(\mathbf{p})), \quad \forall \mathbf{p} \in P,$$

satisfies

1. *$\mathbf{x}^{*,\mathrm{cv}}$ and $\mathbf{x}^{*,\mathrm{cc}}$ are convex and concave on P, respectively, and*
2. *$\mathbf{x}^{*,\mathrm{L}} \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{*,\mathrm{U}}$ and $\mathbf{x}^{*,\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{*,\mathrm{cc}}(\mathbf{p})$ for all $\mathbf{p} \in P^{\mathrm{feas}}$, where*

$$P^{\mathrm{feas}} \equiv \left\{\mathbf{p} \in P : \mathbf{x}^{\mathrm{L}} \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{\mathrm{U}},\ \mathbf{x}^{\mathrm{cv}}(\mathbf{p}) \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{\mathrm{cc}}(\mathbf{p}),\ \mathbf{g}(\mathbf{x}(\mathbf{p})) \leq \mathbf{0}\right\}.$$

Proof Let $\mathbf{x}$, $\mathbf{g}$, $\mathcal{X}$, and $\mathcal{X}^*$ be defined as in the theorem statement and assume I is a relaxation refinement operator for $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$. To prove the first conclusion, choose any $\mathbf{p}_1, \mathbf{p}_2 \in P$ and any $\lambda \in [0, 1]$ and define $\mathbf{p}_\lambda = \lambda \mathbf{p}_1 + (1 - \lambda)\mathbf{p}_2$. Since $\mathbf{x}^{\mathrm{cv}}$ and $\mathbf{x}^{\mathrm{cc}}$ are convex and concave, it is straightforward to show that

$$\mathcal{X}(\mathbf{p}_\lambda) \succeq \mathrm{Conv}(\lambda, \mathcal{X}(\mathbf{p}_1), \mathcal{X}(\mathbf{p}_2)). \tag{104}$$

Since I is coherently inclusion monotonic, it follows that

$$I[\mathcal{X}(\mathbf{p}_\lambda)] \succeq I[\mathrm{Conv}(\lambda, \mathcal{X}(\mathbf{p}_1), \mathcal{X}(\mathbf{p}_2))]. \tag{105}$$

Since I is coherently concave, this implies that

$$I[\mathcal{X}(\mathbf{p}_\lambda)] \succeq \mathrm{Conv}(\lambda, I[\mathcal{X}(\mathbf{p}_1)], I[\mathcal{X}(\mathbf{p}_2)]). \tag{106}$$

By definition, this is equivalent to

$$\mathcal{X}^*(\mathbf{p}_\lambda) \succeq \text{Conv}(\lambda, \mathcal{X}^*(\mathbf{p}_1), \mathcal{X}^*(\mathbf{p}_2)). \tag{107}$$

By Definition 4, this implies that

$$\mathbf{x}^{*,\text{cv}}(\mathbf{p}_\lambda) \leq \lambda \mathbf{x}^{*,\text{cv}}(\mathbf{p}_1) + (1-\lambda)\mathbf{x}^{*,\text{cv}}(\mathbf{p}_2), \tag{108}$$
$$\mathbf{x}^{*,\text{cc}}(\mathbf{p}_\lambda) \geq \lambda \mathbf{x}^{*,\text{cc}}(\mathbf{p}_1) + (1-\lambda)\mathbf{x}^{*,\text{cc}}(\mathbf{p}_2), \tag{109}$$

which establishes convexity and concavity of $\mathbf{x}^{*,\text{cv}}$ and $\mathbf{x}^{*,\text{cc}}$ on P, respectively.

To prove the second conclusion, choose any $\mathbf{p} \in P^{\text{feas}}$. By the definition of P^{feas}, we must have $\mathbf{x}(\mathbf{p}) \in \text{Encl}(\mathcal{X}(\mathbf{p}))$ and $\mathbf{g}(\mathbf{x}(\mathbf{p})) \leq \mathbf{0}$. By Definition 24, this implies that $\mathbf{x}(\mathbf{p}) \in \text{Encl}(I(\mathcal{X}(\mathbf{p}))) = \text{Encl}(\mathcal{X}^*(\mathbf{p}))$. Therefore, $\mathbf{x}^{*,\text{L}} \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{*,\text{U}}$ and $\mathbf{x}^{*,\text{cv}}(\mathbf{p}) \leq \mathbf{x}(\mathbf{p}) \leq \mathbf{x}^{*,\text{cc}}(\mathbf{p})$, as desired. □

The goal in the remainder of this section is to define a specific relaxation refinement operator for linear constraints. To do so, we require a notion of intersection between two McCormick objects. This is defined next and several useful properties are subsequently proven.

Definition 25 Define $(\cap, \mathbb{MR}^2, \mathbb{MR})$ by

$$\mathcal{X} \cap \mathcal{Y} \equiv (\max(x^{\text{L}}, y^{\text{L}}), \min(x^{\text{U}}, y^{\text{U}}), \max(x^{\text{cv}}, y^{\text{cv}}), \min(x^{\text{cc}}, y^{\text{cc}})) \tag{110}$$

if $\max(x^{\text{L}}, y^{\text{L}}) \leq \min(x^{\text{U}}, y^{\text{U}})$, and otherwise by

$$\mathcal{X} \cap \mathcal{Y} \equiv (\min(x^{\text{U}}, y^{\text{U}}), \max(x^{\text{L}}, y^{\text{L}}), \max(x^{\text{L}}, y^{\text{L}}), \min(x^{\text{U}}, y^{\text{U}})). \tag{111}$$

Remark 10 When $\max(x^{\text{L}}, y^{\text{L}}) > \min(x^{\text{U}}, y^{\text{U}})$, the interval parts of $\mathcal{X}$ and $\mathcal{Y}$ do not overlap. The value we assign to $\mathcal{X} \cap \mathcal{Y}$ in this case is arbitrary because there is no reason to do any further computations with such an intersection in the applications we have in mind (see Remark 11). However, it is useful to have $\mathcal{X} \cap \mathcal{Y}$ produce an element of $\mathbb{MR}$ even in this case so that concepts such as coherent concavity and inclusion monotonicity, which were defined for functions mapping into $\mathbb{MR}$, can be applied to $\cap$ without caveat. Thus, we use (111) in this case, which defines $\mathcal{X} \cap \mathcal{Y}$ as an empty element of $\mathbb{MR}$, rather than (110), which does not produce an element of $\mathbb{MR}$.

Lemma 7 *For any $\mathcal{X}, \mathcal{Y} \in \mathbb{MR}$,*

$$\text{Encl}(\mathcal{X}) \cap \text{Encl}(\mathcal{Y}) = \text{Encl}(\mathcal{X} \cap \mathcal{Y}). \tag{112}$$

Proof Choose any $\mathcal{X}, \mathcal{Y} \in \mathbb{MR}$ and any $z \in \mathbb{R}$. The point z is in $\text{Encl}(\mathcal{X}) \cap \text{Encl}(\mathcal{Y})$ if and only if

$$\begin{aligned} \max(x^{\text{L}}, x^{\text{cv}}) &\leq z \leq \min(x^{\text{U}}, x^{\text{cc}}), \\ \max(y^{\text{L}}, y^{\text{cv}}) &\leq z \leq \min(y^{\text{U}}, y^{\text{cc}}). \end{aligned} \tag{113}$$

On the other hand, z is in $\text{Encl}(\mathcal{X} \cap \mathcal{Y})$ if and only if

$$\begin{aligned} \max(x^{\text{L}}, y^{\text{L}}) &\leq z \leq \min(x^{\text{U}}, y^{\text{U}}), \\ \max(x^{\text{cv}}, y^{\text{cv}}) &\leq z \leq \min(x^{\text{cc}}, y^{\text{cc}}). \end{aligned} \tag{114}$$

The result follows from the fact that (113) implies (114) and vice versa. □

Theorem 24 $(\cap, \mathbb{MR}^2, \mathbb{MR})$ *is coherently concave on* $\mathbb{MR}^2$*.*

Proof Choose any coherent $(\mathcal{X}_1, \mathcal{Y}_1), (\mathcal{X}_2, \mathcal{Y}_2) \in \mathbb{MR}^2$ and let $X \times Y = [x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}]$ without subscripts denote their common interval part. Let $\mathcal{Z}_1 \equiv \mathcal{X}_1 \cap \mathcal{Y}_1$ and $\mathcal{Z}_2 \equiv \mathcal{X}_2 \cap \mathcal{Y}_2$. Furthermore, choose any $\lambda \in [0, 1]$ and let $\mathcal{X}_\lambda = \mathrm{Conv}(\lambda, \mathcal{X}_1, \mathcal{X}_2)$, $\mathcal{Y}_\lambda = \mathrm{Conv}(\lambda, \mathcal{Y}_1, \mathcal{Y}_2)$, and $\mathcal{Z}_\lambda = \mathcal{X}_\lambda \cap \mathcal{Y}_\lambda$. We must show that

$$\mathcal{Z}_\lambda \succeq \mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2). \tag{115}$$

Since $(\mathcal{X}_1, \mathcal{Y}_1)$ and $(\mathcal{X}_2, \mathcal{Y}_2)$ are coherent, $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are coherent as well. By Definition 7, $\mathrm{Conv}(\lambda, \mathcal{Z}_1, \mathcal{Z}_2)$ must also be coherent to $\mathcal{Z}_1$. Similarly, the convex combinations $\mathcal{X}_\lambda$ and $\mathcal{Y}_\lambda$ are coherent to $\mathcal{X}_1$ and $\mathcal{Y}_1$, respectively, which implies that $\mathcal{Z}_\lambda = \mathcal{X}_\lambda \cap \mathcal{Y}_\lambda$ is coherent to $\mathcal{Z}_1$ as well. Therefore, the left-hand and right-hand sides of (115) have the same interval parts. Then, to prove (115), it only remains to show that $z_\lambda^{\mathrm{cv}} \leq \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}$ and $z_\lambda^{\mathrm{cc}} \geq \lambda z_1^{\mathrm{cc}} + (1-\lambda) z_2^{\mathrm{cc}}$. We prove the former and note that the latter follows from analogous arguments. If $\max(x^{\mathrm{L}}, y^{\mathrm{L}}) \leq \min(x^{\mathrm{U}}, y^{\mathrm{U}})$, then the intersections defining $\mathcal{Z}_1$, $\mathcal{Z}_2$, and $\mathcal{Z}_\lambda$ are all defined by (110). From (110), we have

$$\begin{aligned} z_\lambda^{\mathrm{cv}} &= \max(\lambda x_1^{\mathrm{cv}} + (1-\lambda) x_2^{\mathrm{cv}}, \lambda y_1^{\mathrm{cv}} + (1-\lambda) y_2^{\mathrm{cv}}) \\ &\leq \lambda \max(x_1^{\mathrm{cv}}, y_1^{\mathrm{cv}}) + (1-\lambda) \max(x_2^{\mathrm{cv}}, y_2^{\mathrm{cv}}), \\ &= \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}, \end{aligned} \tag{116}$$

where the inequality follows from convexity of max on $\mathbb{R}^2$. Alternatively, if $\max(x^{\mathrm{L}}, y^{\mathrm{L}}) > \min(x^{\mathrm{U}}, y^{\mathrm{U}})$, then the intersections defining $\mathcal{Z}_1$, $\mathcal{Z}_2$, and $\mathcal{Z}_\lambda$ are all defined by (111). From (111), we have

$$\begin{aligned} z_\lambda^{\mathrm{cv}} &= \max(x^{\mathrm{L}}, y^{\mathrm{L}}) \\ &= \lambda \max(x_1^{\mathrm{L}}, y_1^{\mathrm{L}}) + (1-\lambda) \max(x_2^{\mathrm{L}}, y_2^{\mathrm{L}}), \\ &= \lambda z_1^{\mathrm{cv}} + (1-\lambda) z_2^{\mathrm{cv}}. \end{aligned} \tag{117}$$

□

Theorem 25 $(\cap, \mathbb{MR}^2, \mathbb{MR})$ *is coherently inclusion monotonic on* $\mathbb{MR}^2$*.*

Proof Choose any coherent $(\mathcal{X}_1, \mathcal{Y}_1), (\mathcal{X}_2, \mathcal{Y}_2) \in \mathbb{MR}^2$ with $(\mathcal{X}_1, \mathcal{Y}_1) \preceq (\mathcal{X}_2, \mathcal{Y}_2)$, and let $X \times Y = [x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}]$ without subscripts denote their common interval part. By Definition 4, it follows that $\mathcal{X}_1 \preceq \mathcal{X}_2$ and $\mathcal{Y}_1 \preceq \mathcal{Y}_2$. Let $\mathcal{Z}_1 = \mathcal{X}_1 \cap \mathcal{Y}_1$ and $\mathcal{Z}_2 = \mathcal{X}_2 \cap \mathcal{Y}_2$. We must prove that $\mathcal{Z}_1 \preceq \mathcal{Z}_2$. Since the interval parts of $\mathcal{Z}_1$ and $\mathcal{Z}_2$ depend only on $[x^{\mathrm{L}}, x^{\mathrm{U}}]$ and $[y^{\mathrm{L}}, y^{\mathrm{U}}]$, we must have $[z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] = [z_2^{\mathrm{L}}, z_2^{\mathrm{U}}]$. It remains to prove that $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$ and $z_1^{\mathrm{cc}} \leq z_2^{\mathrm{cc}}$.

Suppose first that $\max(x^{\mathrm{L}}, y^{\mathrm{L}}) \leq \min(x^{\mathrm{U}}, y^{\mathrm{U}})$, so that the intersections defining $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are given by (110). In this case, $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$ and $z_1^{\mathrm{cc}} \leq z_2^{\mathrm{cc}}$ hold if

$$\max(x_1^{\mathrm{cv}}, y_1^{\mathrm{cv}}) \geq \max(x_2^{\mathrm{cv}}, y_2^{\mathrm{cv}}), \tag{118}$$

$$\min(x_1^{\mathrm{cc}}, y_1^{\mathrm{cc}}) \leq \min(x_2^{\mathrm{cc}}, y_2^{\mathrm{cc}}). \tag{119}$$

Since $\mathcal{X}_1 \preceq \mathcal{X}_2$ and $\mathcal{Y}_1 \preceq \mathcal{Y}_2$, Definition 4 implies that $x_1^{\mathrm{cv}} \geq x_2^{\mathrm{cv}}$, $x_1^{\mathrm{cc}} \leq x_2^{\mathrm{cc}}$, $y_1^{\mathrm{cv}} \geq y_2^{\mathrm{cv}}$, and $y_1^{\mathrm{cc}} \leq y_2^{\mathrm{cc}}$. Thus, the inequalities (118)–(119) follow from the fact that $\max(a', b') \geq \max(a, b)$ and $\min(a, b) \leq \min(a', b')$ when $a \leq a'$ and $b \leq b'$.

Next, suppose that $\max(x^{\mathrm{L}}, y^{\mathrm{L}}) > \min(x^{\mathrm{U}}, y^{\mathrm{U}})$, so that the intersections defining $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are given by (111). In this case, $\mathcal{Z}_1 = \mathcal{Z}_2$ because (111) only depends on the input bounds. Thus, we trivially have $z_1^{\mathrm{cv}} \geq z_2^{\mathrm{cv}}$ and $z_1^{\mathrm{cc}} \leq z_2^{\mathrm{cc}}$. $\square$

We now apply the McCormick intersection to define a valid relaxation refinement operator for constraints of the form $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{A} \in \mathbb{R}^{n_c \times n}$ and $\mathbf{b} \in \mathbb{R}^{n_c}$. The basic idea is to consider all possible rearrangements of the individual constraints into the form

$$x_k = \frac{b_i}{A_{ik}} + \sum_{j \neq k} \frac{-A_{ij}}{A_{ik}} x_j. \tag{120}$$

For each one, we then take the natural McCormick extension of the right-hand side using the extended McCormick rules from §3 and intersect the result with the original enclosure $\mathcal{X}_k$. The complete refinement operator based on this approach is defined in Algorithm 1.

Algorithm 1 A relaxation refinement operator for $\mathbf{Ax} = \mathbf{b}$

1: **function** RelaxRefine($\mathcal{X}, \mathbf{A}, \mathbf{b}, \mathrm{tol}$)
2:     $\mathcal{X} \leftarrow \mathrm{Cut}(\mathcal{X})$
3:     **for** $i \leftarrow 1$ to n_c **do**
4:         **for** $k \leftarrow 1$ to n **do**
5:             **if** $|A_{ik}| > \mathrm{tol}$ **then**
6:                 $\mathcal{X}_k^{\mathrm{updated}} \leftarrow \frac{b_i}{A_{ik}} + \sum_{j \neq k} \left(\frac{-A_{ij}}{A_{ik}} \right) \mathcal{X}_j$
7:                 $\mathcal{X}_k \leftarrow \mathcal{X}_k \cap \mathcal{X}_k^{\mathrm{updated}}$
8:             **end if**
9:         **end for**
10:     **end for**
11:     **return** $\mathcal{X}$
12: **end function**

The next several results prove that Algorithm 1 defines a valid relaxation refinement operator.

Theorem 26 *The function $(I, \mathbb{MR}^n, \mathbb{MR}^n)$ defined by Algorithm 1 satisfies*

$$\mathrm{Encl}(I(\mathcal{X})) \supset \{\mathbf{x} \in \mathrm{Encl}(\mathcal{X}) : \mathbf{Ax} = \mathbf{b}\}, \quad \forall \mathcal{X} \in \mathbb{MR}^n.$$

Proof Choose any $\mathcal{X} \in \mathbb{MR}^n$ and any $\mathbf{x} \in \mathrm{Encl}(\mathcal{X})$ such that $\mathbf{Ax} = \mathbf{b}$. Suppose $\mathcal{X}$ is given as input to Algorithm 1. Clearly, we have $\mathbf{x} \in \mathrm{Encl}(\mathcal{X})$ upon entry to the algorithm, and it follows that $\mathbf{x} \in \mathrm{Encl}(\mathcal{X})$ when line 6 is reached for the first time. To set up an inductive argument, assume that $\mathbf{x} \in \mathrm{Encl}(\mathcal{X})$ when line 6 is reached for ℓ^{th} time for some $\ell \geq 1$. Whatever the values of k and i may be upon this visit to line 6, the fact that $\mathbf{Ax} = \mathbf{b}$ ensures that (120) holds. Since line 6 defines $\mathcal{X}_k^{\mathrm{updated}}$ as the natural McCormick extension of the right-hand side of (120), it follows that $x_k \in \mathrm{Encl}(\mathcal{X}_k^{\mathrm{updated}})$ immediately after line 6. Since we also have $x_k \in \mathrm{Encl}(\mathcal{X}_k)$ at this point, Lemma 7 ensures that $x_k \in \mathrm{Encl}(\mathcal{X}_k \cap \mathcal{X}_k^{\mathrm{updated}})$. Therefore, $\mathbf{x} \in \mathrm{Encl}(\mathcal{X})$ still holds after line 7. This recovers the inductive hypothesis and ensures that $\mathbf{x} \in \mathrm{Encl}(\mathcal{X})$ at all times during the execution of Algorithm 1. Thus, $\mathbf{x} \in \mathrm{Encl}(I(\mathcal{X}))$. $\square$

Theorem 27 *The function $(I, \mathbb{MR}^n, \mathbb{MR}^n)$ defined by Algorithm 1 is coherently concave and coherently inclusion monotonic on $\mathbb{MR}^n$.*

Proof For every $i \in \{1, \dots, n_c\}$ and $k \in \{1, \dots, n\}$, let $G_{i,k} : \mathbb{MR}^n \to \mathbb{MR}^n$ denote the update of $\mathcal{X}$ that occurs on lines 6–7 of Algorithm 1; i.e.,

$$[G_{i,k}(\mathcal{X})]_j \equiv \mathcal{X}_j, \quad \forall j \neq k, \tag{121}$$

$$[G_{i,k}(\mathcal{X})]_k \equiv \mathcal{X}_k \cap \left[\frac{b_i}{A_{ik}} + \sum_{j \neq k} \left(\frac{-A_{ij}}{A_{ik}} \right) \mathcal{X}_j \right]. \tag{122}$$

Algorithm 1 defines I as a finite composition of these functions. Specifically,

$$I(\mathcal{X}) \equiv G_{n_c,n} \circ G_{n_c,n-1} \circ \dots \circ G_{1,2} \circ G_{1,1} \circ \mathrm{Cut}(\mathcal{X}), \quad \forall \mathcal{X} \in \mathbb{MR}^n. \tag{123}$$

Parts 3 and 4 of Lemma 2 state that, if two functions are both coherently concave and coherently inclusion monotonic on $\mathbb{MR}^n$, then their composition is also coherently concave and coherently inclusion monotonic on $\mathbb{MR}^n$. The Cut operator has these two properties by Corollaries 2 and 3. If each $G_{i,k}$ has these properties, then it follows by induction that I does as well. Thus, it suffices to establish these properties for an arbitrary $G_{i,k}$.

Each $G_{i,k}$ is itself a composition of intersections, binary additions, and the univariate operations of adding and multiplying by a constant. Thus, by another application of Lemma 2, it suffices to show that these four basic operations are coherently concave and coherently inclusion monotonic on $\mathbb{MR}$ and $\mathbb{MR}^2$. The intersection has these properties by Theorems 24 and 25, binary addition has them by Theorems 6 and 7, and both the addition of and multiplication by a constant has them by Theorems 17 and 19 (noting in the last case that $B = \mathbb{R}$ for the univariate functions $u(x) = a + x$ and $u(x) = ax$). □

Theorem 28 *The function $(I, \mathbb{MR}^n, \mathbb{MR}^n)$ defined by Algorithm 1 is a relaxation refinement operator for the constraints $\mathbf{Ax} = \mathbf{b}$.*

Proof The result follows immediately from Theorems 26 and 27. □

Remark 11 Consider the set up of Theorem 23 with I defined by Algorithm 1. Suppose that, during the evaluation of $I(\mathcal{X}(\mathbf{p}))$ for some $\mathbf{p} \in P$, the intersection in line 7 of Algorithm 1 triggers the case defined by (111). Since $I(\mathcal{X}(\mathbf{p}))$ will be empty, it follows from Theorem 26 that this $\mathbf{p}$ is not in P^{feas}. Moreover, since the condition leading to (111) only depends on the bound components of the two operands of $\cap$, it is guaranteed that (111) will again be used when evaluating $I(\mathcal{X}(\mathbf{p}))$ for any other $\mathbf{p} \in P$. Thus, if (111) is ever encountered within Algorithm 1, we can conclude immediately that $P^{\mathrm{feas}} = \emptyset$. In the context of reduced-space optimization, this is grounds to fathom P without further calculation. This is why is was said in Remark 10 that, if (111) occurs in the applications we have in mind, then there is no reason to do any further computations with the result of the intersection, and hence the definition in (111) is somewhat arbitrary.

5.1 A Relaxation Refinement Example

Let $P \equiv [-3, 3]$ and let $x_1, x_2 : P \to \mathbb{R}$ be defined as $x_1(p) = p^2$ and $x_2(p) = e^p$, $\forall p \in P$. Furthermore, let $\mathcal{X}_1, \mathcal{X}_2 : P \to \mathbb{R}$ be the McCormick functions shown in Figures 3 and 4 (red), which are defined componentwise by

$$\mathcal{X}_1(p) = \left(0, 9, p^2, 9\right), \quad \mathcal{X}_2(p) = \left(e^{-3}, e^3, e^p, \frac{e^3 - e^{-3}}{6}(p+3) + e^{-3}\right).$$

Consider a situation where $\mathcal{X}_1$ and $\mathcal{X}_2$ represent initial relaxations of x_1 and x_2, providing valid enclosures on all of P, and it is desirable to refine them based on the constraint

$$x_1(p) + x_2(p) = 5. \tag{124}$$

Specifically, we are interested in tighter relaxations that remain convex and concave on all of P, but need only provide a valid enclosure on the subset of P for which (124) holds, P^{feas}. In this case, P^{feas} contains only two isolated points marked by the black dots in Figures 3 and 4. To accomplish this, we apply Algorithm 1.

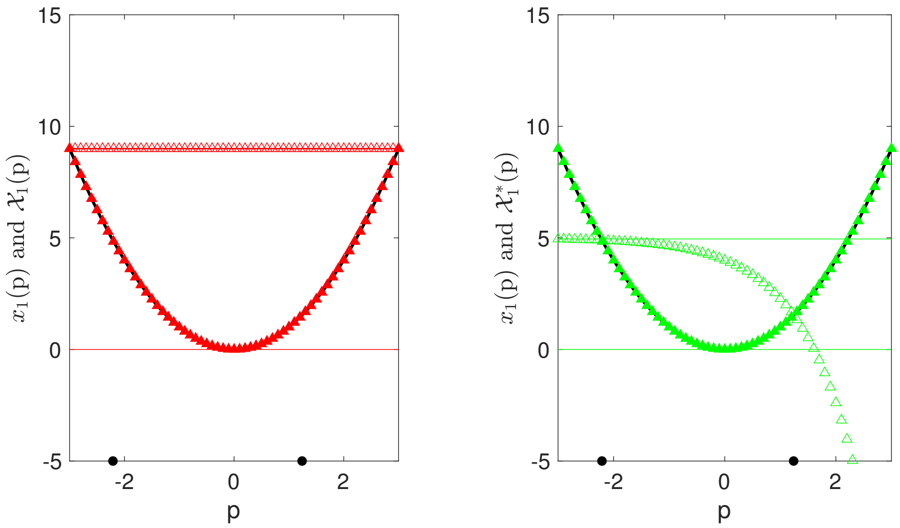

Fig. 3 Relaxations of x_1 before (left) and after (right) refinement with (124). The left panel shows the original function x_1 (solid black) and the McCormick object $\mathcal{X}_1(p)$ defined by its bounds x_1^{L} and x_1^{U} (solid red) and its relaxations x_1^{cv} and x_1^{cc} (filled and open red triangles, resp.) prior to refinement. The right panel shows the original function x_1 (solid black) and the McCormick object $\mathcal{X}_1^*(p)$ defined by its bounds $x_1^{\mathrm{L},*}$ and $x_1^{\mathrm{U},*}$ (solid green) and its relaxations $x_1^{\mathrm{cv},*}$ and $x_1^{\mathrm{cc},*}$ (filled and open green triangles, resp.) after refinement. Black dots represent the values of p that are feasible in (124).

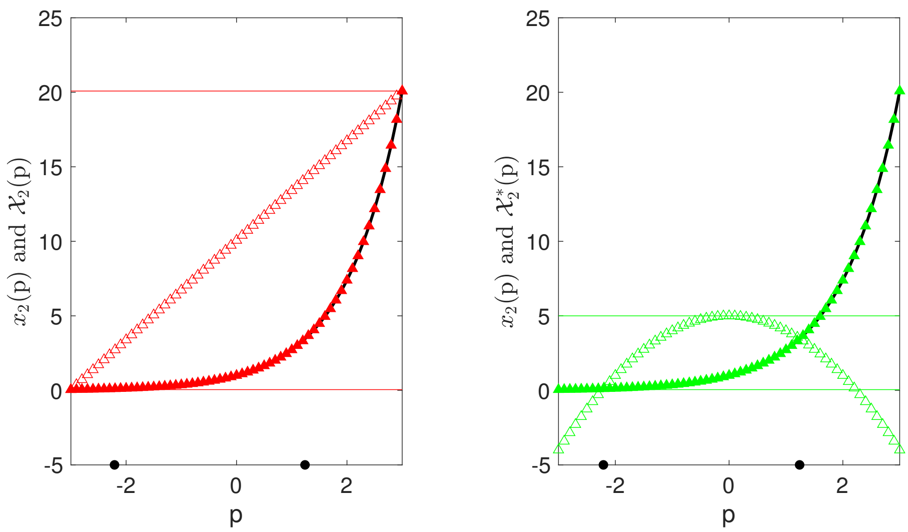

Fig. 4 Relaxations of x_2 before (left) and after (right) refinement with (124). The left panel shows the original function x_2 (solid black) and the McCormick object $\mathcal{X}_2(p)$ defined by its bounds x_2^{L} and x_2^{U} (solid red) and its relaxations x_2^{cv} and x_2^{cc} (filled and open red triangles, resp.) prior to refinement. The right panel shows the original function x_2 (solid black) and the McCormick object $\mathcal{X}_2^*(p)$ defined by its bounds $x_2^{\mathrm{L},*}$ and $x_2^{\mathrm{U},*}$ (solid green) and its relaxations $x_2^{\mathrm{cv},*}$ and $x_2^{\mathrm{cc},*}$ (filled and open green triangles, resp.) after refinement. Black dots represent the values of p that are feasible in (124).

Letting I denote the refinement operator defined by Algorithm 1, the refined McCormick objects defined by $\mathcal{X}^*(p) \equiv I(\mathcal{X}(p))$ are shown in Figures 3 and 4 (green). Both the bound and relaxation components of the refined relaxations are significantly tighter than before refinement. Even so, the refined relaxations remain convex and concave. Moreover, although the objects $\mathcal{X}_1^*(p)$ and $\mathcal{X}_2^*(p)$ are empty for many p, they are nonempty on P^{feas}, as desired. In fact, they are nonempty on the convex set $P^* = [-2.2, 1.2]$, which contains P^{feas}.

To demonstrate the utility of these refined relaxations, next suppose that the constraint (124) is part of the reduced-space optimization problem

$$\begin{aligned} \min_{p \in P} \quad & -x_1(p)x_2(p) \\ \text{s.t.} \quad & x_1(p) + x_2(p) = 5 \end{aligned} \tag{125}$$

Recall that, in the reduced-space formulations of interest, $x_1(p)$ and $x_2(p)$ would be defined as the (explicit or implicit) solutions of some system of equations not shown, and the initial relaxations $\mathcal{X}_1$ and $\mathcal{X}_2$ would be computed by specialized methods [11, 18, 20]. In this context, we are interested in formulating a convex lower bounding problem on P by constructing a relaxation of the objective $f(p) = -x_1(p)x_2(p)$. The standard approach is to relax both f and the constraint using $\mathcal{X}_1$ and $\mathcal{X}_2$. Specifically, the relaxation of f is obtained by natural McCormick extension as $\mathcal{F}(p) = -[\mathcal{X}_1(p)\mathcal{X}_2(p)]$. However, since our extended McCormick rules enable computations with empty objects, we can instead consider the relaxation

$\mathcal{F}^*(p) = -[\mathcal{X}_1^*(p)\mathcal{X}_2^*(p)]$, which implicitly carries information about the feasible domain that can lead to less underestimation of f.

These two approaches are compared in Figure 5. Clearly, $\mathcal{F}^*$ provides a much tighter convex relaxation while still underestimating f at all feasible points. Notably, $f^{*,\mathrm{cv}}$ is not just tighter on the infeasible space, but dominates f^{cv} even on the convex hull of the feasible set, P^{feas}. It follows that minimizing $f^{*,\mathrm{cv}}$ on P will produce a tighter lower bound than the usual approach of minimizing f^{cv} over a convex relaxation of P^{feas}.

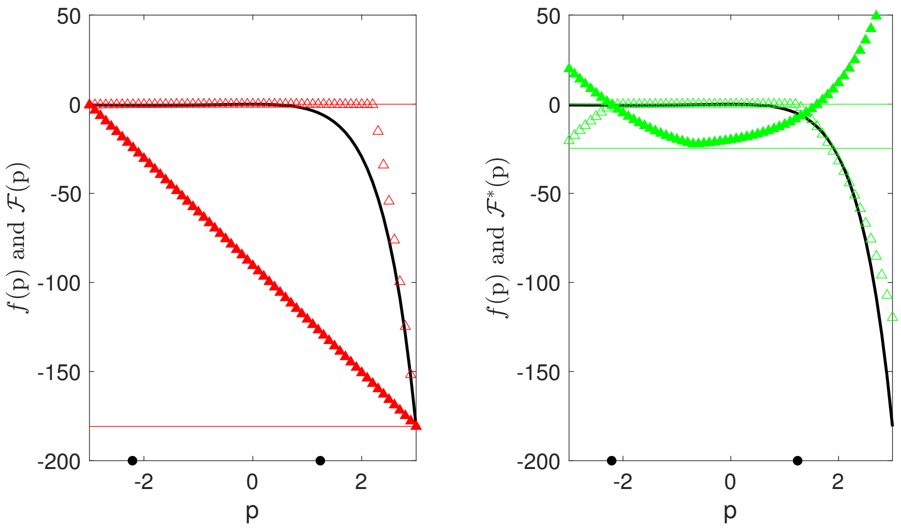

Fig. 5 Relaxations of f before (left) and after (right) refinement with (124). The left panel shows the original function f (solid black) and the McCormick object $\mathcal{F}(p)$ defined by its bounds f^{L} and f^{U} (solid red) and its relaxations f^{cv} and f^{cc} (filled and open red triangles, resp.) prior to such refinement. The right panel shows the original function f (solid black) and the McCormick object $\mathcal{F}^*(p)$ defined by its bounds $f^{\mathrm{L},*}$ and $f^{\mathrm{U},*}$ (solid green) and its relaxations $f^{\mathrm{cv},*}$ and $f^{\mathrm{cc},*}$ (filled and open green triangles, resp.) after such refinement. Black dots represent the values of p that are feasible in (124).

Figure 6 shows that the refined relaxation $f^{*,\mathrm{cv}}$ cannot be achieved without the extended McCormick rules developed in this paper. Specifically, the figure compares $\mathcal{F}^*(p)$ (green) with the result of the same procedure using the standard McCormick rules (blue); namely, $\hat{\mathcal{F}}^*(p) = \hat{-}[\mathcal{X}_1^*(p)\hat{\times}\mathcal{X}_2^*(p)]$, where the hat over the minus sign indicates that the standard univariate composition rule is used for the multiplication by -1. While the resulting relaxation $\hat{f}^{*,\mathrm{cv}}$ is effectively just as tight as $f^{*,\mathrm{cv}}$, it is clearly nonconvex. Therefore, it does not produce a convex lower-bounding problem for (125). Similarly, although the McCormick-based refinement and relaxation methods in [20] could also be used to compute versions of $\mathcal{X}_1^*$, $\mathcal{X}_2^*$, and $f^{*,\mathrm{cv}}$, that approach would yield $f^{*,\mathrm{cv}}(p) = \mathrm{NaN}$ for many values of p, which again fails to produce a computationally useful lower-bounding problem.

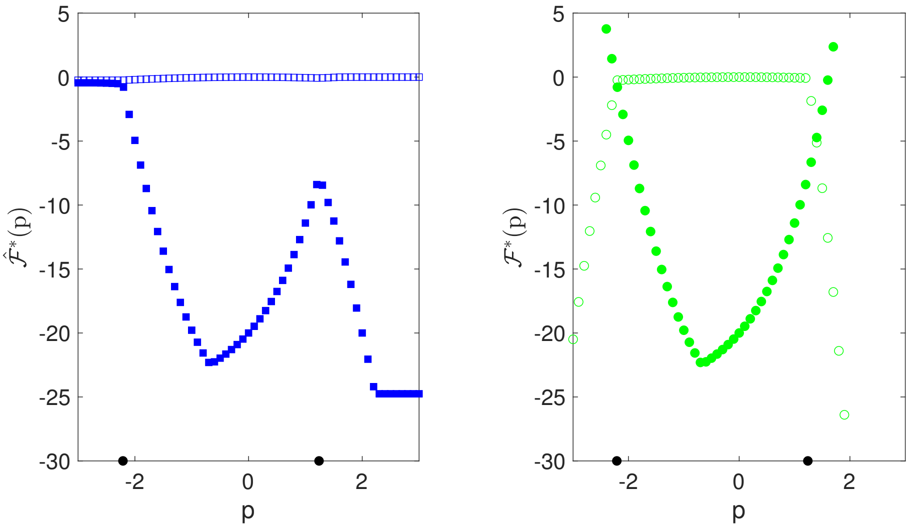

Fig. 6 Refined relaxations of f computed using the original McCormick's rules [4] (left) and using the extended McCormick's rules in this paper (right). The left panel shows the refined relaxations of f, $\hat{f}^{\mathrm{cv},*}$ and $\hat{f}^{\mathrm{cc},*}$ (filled and open blue squares, resp.), computed using the original McCormick's rules. The right panel shows the refined relaxations of f, $f^{\mathrm{cv},*}$ and $f^{\mathrm{cc},*}$ (filled and open green circles, resp.), computed using the extended McCormick's rules. Black dots represent the values of p that are feasible in (124).

6 Conclusion

In this paper, we extended the notion of a McCormick object – the basic computational object in McCormick relaxation routines – to include *empty* objects where either $x^{\mathrm{cv}} > x^{\mathrm{cc}}$ or $[x^{\mathrm{cv}}, x^{\mathrm{cc}}] \cap [x^{\mathrm{L}}, x^{\mathrm{U}}] = \emptyset$. We then generalized McCormick's relaxation rules for binary addition, binary multiplication, and univariate composition to be well-defined and preserve their essential properties on this extended domain of objects. Empty McCormick objects provide a natural way to represent infeasibility in reduced-space global optimization formulations and can be readily generated by domain reduction procedures in that context. We showed in §5 that allowing emptiness enables a very natural intersection operation between McCormick objects that preserves desirable convexity and concavity properties, and that this intersection can further be used to develop constraint-based refinement procedures in a straightforward way. Our extended McCormick relaxation rules then enable subsequent calculations to be done with the possibly-empty refined objects. For example, they can be used to compute objective function relaxations that are tighter on the feasible parts of the domain without compromising convexity (and hence ease of minimization) anywhere on the domain. We hope that these capabilities will significantly ease the development of improved McCormick-based algorithms for reduced-space global optimization, global dynamic optimization, and domain reduction for nonconvex NLPs.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Bompadre, A., Mitsos, A.: Convergence rate of McCormick relaxations. Journal of Global Optimization **52**, 1–28 (2012). DOI https://doi.org/10.1007/s10898-011-9685-2
2. Khan, K.A., Watson, H.A.J., Barton, P.I.: Differentiable McCormick relaxations. Journal of Global Optimization **67**, 687–729 (2017). DOI 10.1007/s10898-016-0440-6
3. Land, A.H., Doig, A.G.: An automatic method of solving discrete programming problems. Econometrica **28**(3), 497–520 (1960). DOI https://doi.org/10.2307/1910129. URL https://www.jstor.org/stable/1910129
4. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: Part I - convex underestimating problems. Mathematical Programming **10**(1), 147–175 (1976). DOI https://doi.org/10.1007/BF01580665. URL https://doi.org/10.1080/10556788.2014.924514
5. Mitsos, A., Chachuat, B., Barton, P.I.: McCormick-based relaxations of algorithms. Society for Industrial and Applied Mathematics **20**(2), 573–601 (2009). DOI 10.1137/080717341
6. Moore, R.E.: Interval Analysis. Prentice-Hall, Englewood Cliffs, NJ (1966)
7. Najman, J., Mitsos, A.: Tighter McCormick relaxations through subgradient propagation. Journal of Global Optimization **75**, 565–593 (2019). DOI https://doi.org/10.1007/s10898-019-00791-0
8. Schweidtmann, A.M., Mitsos, A.: Deterministic global optimization with artificial neural networks embedded. Journal of Optimization Theory and Applications **180**, 925–948 (2019). DOI https://doi.org/10.1007/s10957-018-1396-0
9. Scott, J.K.: Reachability analysis and deterministic global optimization of differential-algebraic systems. phdthesis, Massachusetts Institute of Technology (2012)
10. Scott, J.K., Barton, P.I.: Convex and concave relaxations for the parametric solutions of semi-explicit index-one differential-algebraic equations. Journal of Optimization Theory and Applications **156**(3), 617–649 (2013). DOI https://doi.org/10.1007/s10957-012-0149-8
11. Scott, J.K., Barton, P.I.: Improved relaxations for the parametric solutions of ODEs using differential inequalities. Journal of Global Optimization **57**, 143–176 (2013). DOI 10.1007/s10898-012-9909-0
12. Scott, J.K., Chachuat, B., Barton, P.I.: Nonlinear convex and concave relaxations for the solutions of parametric ODEs. Optimal Control Applications and Methods **34**(2), 145–163 (2013). DOI https://doi.org/10.1002/oca.2014
13. Scott, J.K., Stuber, M.D., Barton, P.I.: Generalized McCormick relaxations. Journal of Global Optimization **51**(4), 569–606 (2011). DOI https://doi.org/10.1007/s10898-011-9664-7. URL https://link.springer.com/article/10.1007/s10898-011-9664-7
14. Shao, Y., Scott, J.K.: Convex relaxations for global optimization under uncertainty described by continuous random variables. AIChE Journal **64**(8), 3023–3033 (2018). DOI https://doi.org/10.1002/aic.16064
15. Shen, K., Scott, J.K.: Rapid and accurate reachability analysis for nonlinear dynamic systems by exploiting model redundancy. Computers and Chemical Engineering **106**, 596–608 (2017). DOI https://doi.org/10.1016/j.compchemeng.2017.08.001
16. Shen, K., Scott, J.K.: Exploiting nonlinear invariants and path constraints to achieve tighter reachable set enclosures using differential inequalities. Mathematics of Control, Signals, and Systems **32**, 101–127 (2020). DOI https://doi.org/10.1007/s00498-020-00254-y

17. Stuber, M.D., Barton, P.I.: Robust simulation and design using semi-infinite programs with implicit functions. International Journal of Reliability and Safety **5**(3/4), 378–397 (2011). DOI https://doi.org/10.1504/IJRS.2011.041186
18. Stuber, M.D., Scott, J.K., Barton, P.I.: Convex and concave relaxations of implicit functions. Optimization Methods and Software **30**(3), 424–460 (2015). DOI https://doi.org/10.1080/10556788.2014.924514
19. Tsoukalas, A., Mitsos, A.: Multivariate McCormick relaxations. Journal of Global Optimization **59**, 633–662 (2014). DOI https://doi.org/10.1007/s10898-014-0176-0
20. Wechsung, A.: Global optimization in reduced space. phdthesis, Massachusetts Institute of Technology (2014)
21. Wechsung, A., Barton, P.I.: Global optimization of bounded factorable functions with discontinuities. Journal of Global Optimization **58**, 1–30 (2014). DOI https://doi.org/10.1007/s10898-013-0060-3
22. Wechsung, A., Scott, J.K., Watson, H.A.J., Barton, P.I.: Reverse propagation of McCormick relaxations. Journal of Global Optimization **63**, 1–36 (2015). DOI 10.1007/s10898-015-0303-6

Extended McCormick Relaxation Rules for Handling Empty Arguments Representing Infeasibility: Supplementary Information*

Jason Ye and Joseph K. Scott†

July 19, 2022

This document provides some suitable choices of the data required by Assumption 1 in the main paper for some common univariate functions. Specifically, it provides choices for the data $(u, \mathbb{I}B, \mathbb{IR})$, u^{cv}, u^{cc}, $x^{\min}$, and $x^{\max}$, as defined in Assumption 1.

I Data Satisfying Assumption 1 for Common Univariate Functions

Addition of a constant

$$u(x) = x + c, \quad B = \mathbb{R}$$
$$u^{\mathrm{L}}(X) = x^{\mathrm{L}} + c, \quad u^{\mathrm{U}}(X) = x^{\mathrm{U}} + c$$
$$u^{\mathrm{cv}}(X, x) = x + c, \quad x^{\min}(X) = -\infty$$
$$u^{\mathrm{cc}}(X, x) = x + c, \quad x^{\max}(X) = +\infty$$

Multiplication by a non-negative constant

$$u(x) = cx,\ c \geq 0, \quad B = \mathbb{R}$$
$$u^{\mathrm{L}}(X) = cx^{\mathrm{L}}, \quad u^{\mathrm{U}}(X) = cx^{\mathrm{U}}$$
$$u^{\mathrm{cv}}(X, x) = cx, \quad x^{\min}(X) = -\infty$$
$$u^{\mathrm{cc}}(X, x) = cx, \quad x^{\max}(X) = +\infty$$

Multiplication by a negative constant

$$u(x) = cx,\ c < 0, \quad B = \mathbb{R}$$
$$u^{\mathrm{L}}(X) = cx^{\mathrm{U}}, \quad u^{\mathrm{U}}(X) = cx^{\mathrm{L}}$$
$$u^{\mathrm{cv}}(X, x) = cx, \quad x^{\min}(X) = +\infty$$
$$u^{\mathrm{cc}}(X, x) = cx, \quad x^{\max}(X) = -\infty$$

*Electronic supplementary material for "Extended McCormick Relaxation Rules for Handling Empty Arguments Representing Infeasibility," by Jason Ye and Joseph K. Scott, submitted to the *Journal of Global Optimization*, July 2022

†Corresponding author: Joseph K. Scott, Dept. of Chemical and Biomolecular Engineering, Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332, email: joseph.scott@chbe.gatech.edu

Reciprocal

$$u(x) = \frac{1}{x}, \quad B = \mathbb{R} - \{0\}$$
$$u^{\mathrm{L}}(X) = \frac{1}{x^{\mathrm{U}}}, \quad u^{\mathrm{U}}(X) = \frac{1}{x^{\mathrm{L}}}$$

The definitions of $u^{\mathrm{cv}}(X, x)$ and $u^{\mathrm{cc}}(X, x)$ are separated into cases where $X > 0$ or $X < 0$. Since B does not include 0 for this operation, any $X \in \mathbb{I}B$ is either all positive or all negative.

When $X > 0$, $\frac{1}{x}$ is convex on X. For standard McCormick relaxations, $u^{\mathrm{cv}}(X, x)$ is typically defined as $\frac{1}{x}$ and $u^{\mathrm{cc}}(X, x)$ is defined as the secant of $\frac{1}{x}$ on X. However, this choice of $u^{\mathrm{cv}}(X, x)$ violates Assumption 1.2 because it is not defined on all of $\mathbb{R}$. To correct this, $u^{\mathrm{cv}}(X, x)$ is defined here as $\frac{1}{x}$ for x above a threshold δ, and as a linearization of $\frac{1}{x}$ at δ for all $x < \delta$.

$$u^{\mathrm{cv}}(X, x) = \begin{cases} \frac{1}{x} & \text{if } x \geq \delta \\ \frac{-1}{\delta^2}x + \frac{2}{\delta} & \text{if } x < \delta \end{cases}.$$
$$u^{\mathrm{cc}}(X, x) = \frac{1}{x^{\mathrm{L}}} + \frac{1/x^{\mathrm{U}} - 1/x^{\mathrm{L}}}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}}).$$
$$x^{\min}(X) = +\infty, \quad x^{\max}(X) = -\infty$$

Note that δ may depend on X. A reasonable choice is $\delta = \min(10^{-3}, x^{\mathrm{L}})$.

When $X < 0$, an analogous strategy is used with $\delta < 0$.

$$u^{\mathrm{cv}}(X, x) = \frac{1}{x^{\mathrm{L}}} + \frac{1/x^{\mathrm{U}} - 1/x^{\mathrm{L}}}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}}).$$
$$u^{\mathrm{cc}}(X, x) = \begin{cases} \frac{1}{x} & \text{if } x \leq \delta \\ \frac{-1}{\delta^2}x + \frac{2}{\delta} & \text{if } x > \delta \end{cases}.$$
$$x^{\min}(X) = +\infty, \quad x^{\max}(X) = -\infty$$

A reasonable choice for δ in this case is $\delta = \max(-10^{-3}, x^{\mathrm{U}})$.

Exponential

$$u(x) = \exp(x), \quad B = \mathbb{R}$$
$$u^{\mathrm{L}}(X) = \exp(x^{\mathrm{L}}), \quad u^{\mathrm{U}}(X) = \exp(x^{\mathrm{U}})$$
$$u^{\mathrm{cv}}(X, x) = \exp(x)$$
$$u^{\mathrm{cc}}(X, x) = \exp(x^{\mathrm{L}}) + \frac{(\exp(x^{\mathrm{U}}) - \exp(x^{\mathrm{L}}))}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}})$$
$$x^{\min}(X) = -\infty, \quad x^{\max}(X) = +\infty$$

Natural log

$$u(x) = \ln(x), \quad B = (0, +\infty)$$
$$u^{\mathrm{L}}(X) = \ln(x^{\mathrm{L}}), \quad u^{\mathrm{U}}(X) = \ln(x^{\mathrm{U}})$$
$$u^{\mathrm{cv}}(X, x) = \ln(x^{\mathrm{L}}) + \frac{\ln(x^{\mathrm{U}}) - \ln(x^{\mathrm{L}})}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}})$$
$$x^{\min}(X) = -\infty$$

The usual definition $u^{\mathrm{cc}}(X, x) = \ln(x)$ needs to be extended onto all of $\mathbb{R}$ to satisfy Assumption 1.2. This is accomplished by following a linearization of $\ln(x)$ below a threshold $\delta > 0$.

$$u^{\mathrm{cc}}(X, x) = \begin{cases} \ln(x) & \text{if } x \geq \delta \\ \frac{1}{\delta}x + [(\ln\delta) - 1] & \text{if } x < \delta \end{cases}$$
$$x^{\max}(X) = +\infty$$

Note that δ may depend on X. A reasonable choice is $\delta = \min(10^{-3}, x^{\mathrm{L}})$.

xln(x)

$$u(x) = x\ln(x), \quad B = (0, +\infty)$$

This function is convex and takes a minimum value of $-\exp(-1)$ at $x = \exp(-1)$.

$$u^{\mathrm{L}}(X) = \begin{cases} -\exp(-1) & \text{if} & \exp(-1) \in [x^{\mathrm{L}}, x^{\mathrm{U}}] \\ \min(x^{\mathrm{L}}\ln x^{\mathrm{L}}, x^{\mathrm{U}}\ln x^{\mathrm{U}}) & \text{otherwise} & \end{cases}$$
$$u^{\mathrm{U}}(X) = \max(x^{\mathrm{L}}\ln x^{\mathrm{L}}, x^{\mathrm{U}}\ln x^{\mathrm{U}})$$
$$u^{\mathrm{cc}}(X, x) = x^{\mathrm{L}}\ln(x^{\mathrm{L}}) + \frac{x^{\mathrm{U}}\ln(x^{\mathrm{U}}) - x^{\mathrm{L}}\ln(x^{\mathrm{L}})}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}})$$
$$x^{\max}(X) = \begin{cases} -\infty & \text{if} & x^{\mathrm{L}}\ln x^{\mathrm{L}} \geq x^{\mathrm{U}}\ln x^{\mathrm{U}} \\ +\infty & \text{otherwise} & \end{cases}$$

The usual definition $u^{\mathrm{cc}}(X, x) = x\ln x$ needs to be extended onto all of $\mathbb{R}$ to satisfy Assumption 1.2. This is accomplished by following a linearization of $x\ln x$ below a threshold $\delta > 0$. The value of δ is chosen to be less than $\exp(-1)$, which is where $x\ln x$ takes its minimum.

$$u^{\mathrm{cv}}(X, x) = \begin{cases} x\ln(x) & \text{if } x \geq \delta \\ (1 + \ln(\delta))x - \delta & \text{if } x < \delta \end{cases}$$
$$x^{\min}(X) = \exp(-1)$$

Note that δ may depend on X. A reasonable choice is $\delta = \min(10^{-3}, x^{\mathrm{L}})$.

Square root

$$u(x)=\sqrt{x},\quad B=[0,+\infty)$$
$$u^{\mathrm{L}}(X)=\sqrt{x^{\mathrm{L}}},\quad u^{\mathrm{U}}(X)=\sqrt{x^{\mathrm{U}}}$$
$$u^{\mathrm{cv}}(X,x)=\sqrt{x^{\mathrm{L}}}+\frac{\sqrt{x^{\mathrm{U}}}-\sqrt{x^{\mathrm{L}}}}{x^{\mathrm{U}}-x^{\mathrm{L}}}(x-x^{\mathrm{L}})$$
$$x^{\min}(X)=-\infty$$

The usual definition $u^{\mathrm{cc}}(X,x)=\sqrt{x}$ needs to be extended onto all of $\mathbb{R}$ to satisfy Assumption 1.2. This is accomplished by following a linearization of $\sqrt{x}$ below a threshold $\delta>0$.

$$u^{\mathrm{cc}}(X,x)=\begin{cases}\sqrt{x} & \text{if } x\geq\delta\\ \frac{1}{2\sqrt{\delta}}x+\frac{\sqrt{\delta}}{2} & \text{if } x<\delta\end{cases}$$
$$x^{\max}(X)=+\infty$$

Note that δ may depend on X. A reasonable choice is $\delta=\min(10^{-3},x^{\mathrm{L}})$.

Even integer powers

$$u(x)=x^n,\ n=2,4,6,\ldots,\quad B=\mathbb{R}$$
$$u^{\mathrm{L}}(X)=\begin{cases}0 & \text{if} & 0\in[x^{\mathrm{L}},x^{\mathrm{U}}]\\ \min((x^{\mathrm{L}})^n,(x^{\mathrm{U}})^n) & \text{otherwise}\end{cases}$$
$$u^{\mathrm{U}}(X)=\max((x^{\mathrm{L}})^n,(x^{\mathrm{U}})^n)$$
$$u^{\mathrm{cv}}(X,x)=x^n$$
$$u^{\mathrm{cc}}(X,x)=(x^{\mathrm{L}})^n+\frac{(x^{\mathrm{U}})^n-(x^{\mathrm{L}})^n}{x^{\mathrm{U}}-x^{\mathrm{L}}}(x-x^{\mathrm{L}})$$
$$x^{\min}(X)=0$$
$$x^{\max}(X)=\begin{cases}-\infty & \text{if} & (x^{\mathrm{L}})^n\geq(x^{\mathrm{U}})^n\\ +\infty & \text{otherwise}\end{cases}$$

Odd integer powers

$$u(x)=x^n,\ n=3,5,7,\ldots,\quad B=\mathbb{R}$$
$$u^{\mathrm{L}}(X)=(x^{\mathrm{L}})^n$$
$$u^{\mathrm{U}}(X)=(x^{\mathrm{U}})^n$$

If $x^{\mathrm{U}}\leq 0$:

$$u^{\mathrm{cv}}(X,x)=(x^{\mathrm{L}})^n+\frac{(x^{\mathrm{U}})^n-(x^{\mathrm{L}})^n}{x^{\mathrm{U}}-x^{\mathrm{L}}}(x-x^{\mathrm{L}}).$$
$$u^{\mathrm{cc}}(X,x)=\begin{cases}x^n & \text{if } x\leq 0\\ 0 & \text{if } x>0\end{cases}.$$

If $x^{\mathrm{L}}\geq 0$:

$$u^{\mathrm{cv}}(X,x)=\begin{cases}0 & \text{if } x<0\\ x^n & \text{if } x\geq 0\end{cases}.$$

$$u^{\mathrm{cc}}(X,x) = (x^{\mathrm{L}})^n + \frac{(x^{\mathrm{U}})^n - (x^{\mathrm{L}})^n}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}}).$$

If $0 \in [x^{\mathrm{L}}, x^{\mathrm{U}}]$:

Let x' and x'' be the solutions of

$$(n-1)(x')^n - nx^{\mathrm{L}}(x')^{n-1} + (x^{\mathrm{L}})^n = 0,$$

$$(n-1)(x'')^n - nx^{\mathrm{U}}(x'')^{n-1} + (x^{\mathrm{U}})^n = 0.$$

If $x' > x^{\mathrm{U}}$,

$$u^{\mathrm{cv}}(X,x) = (x^{\mathrm{L}})^n + \frac{(x^{\mathrm{U}})^n - (x^{\mathrm{L}})^n}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}}).$$

Else

$$u^{\mathrm{cv}}(X,x) = \begin{cases} (x^{\mathrm{L}})^n + \frac{(x')^n - (x^{\mathrm{L}})^n}{x' - x^{\mathrm{L}}}(x - x^{\mathrm{L}}) & \text{if } x \leq x' \\ x^n & \text{if } x > x' \end{cases}.$$

If $x'' < x^{\mathrm{L}}$,

$$u^{\mathrm{cc}}(X,x) = (x^{\mathrm{L}})^n + \frac{(x^{\mathrm{U}})^n - (x^{\mathrm{L}})^n}{x^{\mathrm{U}} - x^{\mathrm{L}}}(x - x^{\mathrm{L}}).$$

Else

$$u^{\mathrm{cc}}(X,x) = \begin{cases} (x'')^n + \frac{(x^{\mathrm{U}})^n - (x'')^n}{x^{\mathrm{U}} - x''}(x - x'') & \text{if } x \geq x'' \\ x^n & \text{if } x < x'' \end{cases}.$$

For all cases, $x^{\min}(X) = -\infty$ and $x^{\max}(X) = +\infty$.

Conflict of interest

The authors declare that they have no conflict of interest.