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

A twice-continuously differentiable version of McCormick's relaxations

Kamil A. Khan · **Paul I. Barton**

Received: date / Accepted: date

Abstract McCormick's classical relaxation technique evaluates convex and concave relaxations of compositions of simple functions and arithmetic operations. These relaxations have several properties which make them useful for lower bounding problems in global optimization: they can be evaluated accurately and computationally inexpensively, and they converge rapidly to the relaxed function as the underlying domain is reduced in size. However, McCormick's product rule and composition rule can introduce nondifferentiability into these relaxations, which can create theoretical and computational obstacles. To address this problem, this article develops a variant of McCormick's relaxation scheme in which the produced relaxations are twice-continuously differentiable. This modified scheme proceeds by weakening McCormick's relaxations of products and compositions to yield twice-continuous differentiability, without sacrificing the useful theoretical and computational properties of McCormick's classical technique. Gradients of the modified relaxations may be computed efficiently using the standard forward or reverse modes of automatic differentiation. A C++ implementation based on the library MC++ is described and applied for illustration.

Keywords Nonconvex optimization · Convex underestimators · McCormick relaxations · Interval analysis

Mathematics Subject Classification (2000) 49M20 · 90C26 · 65G40 · 26B25

This material was supported by Novartis Pharmaceuticals as part of the Novartis-MIT Center for Continuous Manufacturing, and was based in part on work supported by the U.S. Department of Energy, Office of Science, under contract DE-AC02-06CH11357.

K.A. Khan
Process Systems Engineering Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA.
Current address: Mathematics and Computer Science Division, Argonne National Laboratory.
E-mail: kamil@alum.mit.edu

P.I. Barton
Process Systems Engineering Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA.
Tel.: +1 617 253 6526, Fax: +1 617 258 5042
E-mail: pib@mit.edu

1 Introduction

Branch-and-bound methods for deterministic global optimization [18] require the ability to evaluate a lower bound on a nonconvex function on particular classes of subdomains. This bounding information may be generated using a relaxation scheme by McCormick [23], which evaluates convex underestimators of a nonconvex objective function on interval subdomains. McCormick's relaxation method assumes that the objective function can be expressed as a finite, known composition of simple functions and arithmetic operations. Subgradients may be computed for these underestimators using dedicated variants [24, 3] of automatic differentiation [14, 26]. Using this information, a lower bound on a nonconvex objective function on an interval may be supplied by minimizing the corresponding convex McCormick underestimator using a local optimization solver. Other methods for global optimization, such as nonconvex outer approximation [19] and nonconvex generalized Benders decomposition [21], also require the construction and minimization of convex underestimators.

McCormick's relaxation method has several useful properties. Firstly, accurate evaluation of a convex underestimator and a corresponding subgradient is computationally inexpensive and automatable; the C++ library MC++ [8, 24] uses operator overloading to compute these quantities for well-defined user-supplied compositions of the basic arithmetic operations and functions such as sin / cos and exp / log. Secondly, as the width of the interval on which a McCormick relaxation is constructed is reduced to zero, the relaxation approaches the objective function sufficiently rapidly [6] to mitigate a phenomenon called the *cluster effect* [9, 41], in which a branch-and-bound method will branch many times on intervals that either contain or are near a global minimum. By extending McCormick's method in an intuitive manner, *generalized McCormick relaxations* [37, 33] have been developed to handle compositions of functions in a more systematic manner, and to handle various extensions of McCormick's theory to implicit functions [42, 38, 34].

However, as the following example shows, McCormick's relaxations can be nondifferentiable.

Example 1 Let a function $\mathrm{mid} : \mathbb{R}^3 \to \mathbb{R}$ map to the median of its three scalar arguments, consider the smooth composite function $g : \mathbb{R} \to \mathbb{R} : z \mapsto \exp(z^3)$, and set $z^* := -1 + \sqrt{3}$. As shown in [24, Example 2.1], the function $g^{\mathrm{cv}} : [-1, 1] \to \mathbb{R}$ for which

$$g^{\mathrm{cv}} : z \mapsto \exp(\mathrm{mid}(z^3 + 3z^2 - 3, z^3 - 3z^2 + 3, -1)) = \begin{cases} \exp(-1), & \text{if } z \le z^*, \\ \exp(z^3 + 3z^2 - 3), & \text{if } z > z^*, \end{cases}$$

can be generated from g according to McCormick's rule [24, Section 3] for constructing convex relaxations of a composite function. (In this application of McCormick's rule, αBB relaxations [1] of the inner function $z \mapsto z^3$ have been employed.) Indeed, g^{cv} is convex on $[-1, 1]$, and $g^{\mathrm{cv}}(z) \le g(z)$ for each $z \in [-1, 1]$. However, even though g^{cv} satisfies McCormick's proposed sufficient condition for differentiability of a convex relaxation [23, p. 151], it is in fact nondifferentiable at z^*.

Several factors can introduce failure of continuous or twice-continuous differentiability of McCormick's relaxations. Firstly, as illustrated by the above example, the median function used in defining McCormick's composition rule is itself nondifferentiable. Secondly, any nondifferentiability in supplied relaxations of composed functions can propagate to yield nondifferentiability in constructed relaxations of composite functions. (Whether the composed functions are themselves smooth is irrelevant.) Thirdly, as presented in [24, Proposition 2.6], McCormick's rule for generating relaxations of products introduces nondifferentiability, due to its use of bivariate max and min functions.

A relaxation scheme preserving continuous or twice-continuous differentiability would be desirable for a number of reasons. In general, minimization of nondifferentiable convex objective functions requires dedicated numerical methods for nondifferentiable problems such as bundle methods [20, 17], which lack the strong convergence rate results of their smooth counterparts. On the other hand, continuously differentiable convex relaxations may be minimized using gradient-based algorithms for local optimization, which typically exhibit Q-linear convergence. Twice-continuously differentiable relaxations can be minimized by Newton's method (discussed in [28]), which exhibits Q-quadratic convergence under certain nonsingularity assumptions on the Hessian matrix. Computation of the required Hessian or Hessian-vector products can be avoided by using a secant-based quasi-Newton method [7], which exhibits Q-superlinear convergence under the assumptions of Newton's method.

Furthermore, an automatable and computationally inexpensive method for generating continuously differentiable relaxations would yield theoretical and numerical benefits when used in established methods for generating convex and concave relaxations of solutions of parametric ordinary differential equations (ODEs). If continuously differentiable relaxations were available for the right-hand side function of such an ODE, then the relaxation-generating ODE described in [36] would have a continuously differentiable right-hand side function. The corresponding relaxation of the ODE solution would then be differentiable with respect to the ODE parameter, permitting computation of the corresponding parametric derivatives according to classical ODE theory [15], thus overcoming theoretical hurdles concerning subgradient evaluation. Similarly, incorporation of continuously differentiable relaxations of an ODE right-hand side function into the relaxation method of [35] would yield ODEs whose parametric sensitivities are decribed by the hybrid system sensitivity results of [12].

Thus, the goal of this article is to present a variant of McCormick's relaxation scheme which produces continuously or twice-continuously differentiable relaxations, while retaining the various theoretical and computational benefits of McCormick's original method. To achieve this, variants of McCormick's product rule are introduced in Definition 13, in which the original product rule is further relaxed in a particular manner. An additional assumption (Assumption 1) is imposed on user-supplied relaxations of composed *univariate intrinsic functions*, so as to enforce differentiability in McCormick's composition rule. This assumption is readily satisfied for standard arithmetic operations and functions. Under these modifications, the aforementioned sources of nonsmoothness in McCormick's relaxation scheme are circumvented. For broader applicability, the relaxation theory developed in this article is presented in the framework of generalized McCormick relaxations [33]. To construct twice-continuously differentiable relaxations rather than once-continuously differentiable relaxations according to the methods in this article, more stringent (yet readily satisfied) assumptions are required on the supplied relaxations of univariate intrinsic functions, and the employed product rule must be relaxed further. Gradients of the developed relaxations can be evaluated efficiently using the standard forward or reverse modes of automatic differentiation [14, 26].

The product rule variants developed in this article make use of certain smoothing approximations. Smooth approximations of simple nonsmooth functions have previously been considered [4], particularly in the context of complementarity problems [11, 10, 29]. The smoothing approach taken in this article is similar in spirit, but is modified so as to accommodate our requirement that the posited convex/concave relaxations are well-defined, are indeed convex or concave, are valid bounds on the underlying function, and are either once- or twice-continuously differentiable, as desired.

Observe that the αBB relaxation scheme [1] represents an alternative to McCormick's scheme, and shares several of the features of McCormick's method outlined above. Moreover, αBB relaxations of twice-continuously differentiable functions are themselves twice-continuously differentiable. This article instead focuses on variants of McCormick's method, due to the ability of McCormick's theory to handle more general compositions of functions, and due to its extensions to relaxations of implicit functions and solutions of differential-algebraic equations.

This article is structured as follows. Section 2 summarizes and extends established definitions and properties concerning differentiability on intervals, interval analysis, and McCormick's relaxation technique. Section 3 develops the smoothing constructions used in the remainder of the article, and uses these to construct a variant of McCormick's multiplication operation. Section 4 develops variants of McCormick's overall relaxation technique and presents the main theorem of the article, in which these variants are asserted to be once- or twice-continuously differentiable as desired, and to have the various desirable properties of McCormick's original relaxation scheme. Section 5 describes a C++ implementation of the methods in this article, and presents examples of its application for illustration.

2 Background and preliminaries

This section summarizes and extends established definitions and properties concerning differentiability on closed sets, interval analysis, McCormick's relaxation scheme, and convergence analysis of relaxation schemes.

2.1 Differentiability on open and closed sets

Let $\|\cdot\|$ denote the Euclidean norm on $\mathbb{R}^n$. Given an open set $X \subset \mathbb{R}^n$, a function $f : X \to \mathbb{R}^m$ is *(Fréchet) differentiable* at $x \in X$ if there exists a matrix $A \in \mathbb{R}^{m\times n}$ for which

$$0 = \lim_{h\to 0} \frac{f(x+h) - (f(x) + Ah)}{\|h\|}.$$

In this case, the above equation defines A uniquely, and A is called the *Jacobian* $Jf(x)$ of f at x. If $m = 1$, in which case f is scalar-valued, then the *gradient* of f at x is the column vector $\nabla f(x) := (Jf(x))^{\mathrm{T}} \in \mathbb{R}^m$.

Given an open set $X \subset \mathbb{R}^n$, a function $f : X \to \mathbb{R}^n$ is *continuously differentiable* ($\mathscr{C}^1$) on X if it is differentiable on X and the Jacobian mapping $x \mapsto Jf(x)$ is continuous on X. Equivalently, f is $\mathscr{C}^1$ on X if its first-order partial derivatives each exist on X and are continuous. If $m = 1$, in which case f is scalar-valued, then f is *twice-continuously differentiable* ($\mathscr{C}^2$) on X if f is $\mathscr{C}^1$ on X and there exists a continuous *Hessian* mapping $x \mapsto \nabla^2 f(x) \in \mathbb{R}^{n\times n}$ for which

$$0 = \lim_{h\to 0} \frac{f(x+h) - (f(x) + \nabla f(x)^{\mathrm{T}} h + \frac{1}{2} h^{\mathrm{T}} \nabla^2 f(x)\, h)}{\|h\|^2}, \qquad \forall x \in X.$$

Equivalently [30], f is $\mathscr{C}^2$ on X if its second-order partial derivatives each exist on X and are continuous. A vector-valued function f is $\mathscr{C}^2$ if each of its component functions is $\mathscr{C}^2$.

By specializing a classical result by Whitney [43], differentiability on closed sets such as intervals can be defined in a manner that is consistent with the classical chain rule of differentiation, as follows.

Definition 1 (adapted from [43]) Given a closed set $B \subset \mathbb{R}^n$ and some $i \in \{1,2\}$, a function $f : B \to \mathbb{R}^m$ is $\mathscr{C}^i$ on B if there exist an open set $X \subset \mathbb{R}^n$ and a function $\hat{f} : X \to \mathbb{R}^m$ such that $B \subset X$, $\hat{f}(x) = f(x)$ for each $x \in B$, and $\hat{f}$ is $\mathscr{C}^i$ (in the classical sense) on X. Given any point x in the boundary of B, define $Jf(x) := J\hat{f}(x)$. If $m = 1$, in which case f is scalar-valued, then define $\nabla f(x) := Jf(x)^{\mathrm{T}}$.

Remark 1 When x lies in the boundary of B, it is possible that $Jf(x)$ is not uniquely specified by the above definition, since $\hat{f}$ might not be specified uniquely. For example, if B comprises a single point $\{x_0\} \subset \mathbb{R}^n$, then $Jf(x_0)$ may be chosen to be any element of $\mathbb{R}^{m\times n}$, since $\hat{f}$ may be chosen to be any $\mathscr{C}^i$ function for which $\hat{f}(x_0) = f(x_0)$.

Despite the possible nonuniqueness implied by the previous remark, the following propositions show that the classical chain rule continues to hold. Both propositions are immediate corollaries of Theorem 1 in [43].

Proposition 1 *Consider B, i, and f as in Definition 1, and any point x in the boundary of B. If there exists any sequence $\{x_{(k)}\}_{k\in\mathbb{N}} \to x$ in $B\backslash\{x\}$, then any Jacobian $Jf(x)$ satisfies*

$$0 = \lim_{\substack{h\to 0\\(x+h)\in B}} \frac{f(x+h) - (f(x) + Jf(x)\,h)}{\|h\|}.$$

Proposition 2 *Consider nonempty sets $B \subset \mathbb{R}^n$ and $D \subset \mathbb{R}^m$ such that B is either closed, open, or both, and such that D is either closed, open, or both. For any fixed $i \in \{1,2\}$, given $\mathscr{C}^i$ functions $g : B \to D$ and $f : D \to \mathbb{R}^p$, the composite function $h \equiv f \circ g : B \to \mathbb{R}^p$ is well-defined and $\mathscr{C}^i$ on B.*

Moreover, for each $x \in B$, $Jh(x) = Jf(g(x))\,Jg(x)$. (If B is closed and x lies in the boundary of B, then this construction of $Jh(x)$ satisfies both Definition 1 and Proposition 1 for some valid choice of $\hat{h}$.)

Corollary 1 *Given a closed convex set $B \subset \mathbb{R}^n$ and a convex $\mathscr{C}^1$ function $f : B \to \mathbb{R}$, for each $x \in B$, $\nabla f(x)$ is a* subgradient *of f at x in that*

$$f(y) \geq f(x) + \nabla f(x)^{\mathrm{T}}(y-x), \qquad \forall y \in B.$$

2.2 Interval analysis

An *interval* $\boldsymbol{x} \equiv [\underline{x}, \overline{x}]$ is a nonempty compact set $\{z \in \mathbb{R} : \underline{x} \leq z \leq \overline{x}\} \subset \mathbb{R}$; the set of all such intervals is denoted $\mathbb{IR}$. Intervals and vectors of intervals are denoted in this article as boldfaced, italicized, lowercase letters (e.g., $\boldsymbol{y}$). Given a set $B \subset \mathbb{R}^n$, the set of intervals (or vectors of intervals) that are subsets of B will be denoted as $\mathbb{I}B$. If B is nonempty, then $\mathbb{I}B$ is necessarily nonempty. An interval vector $\boldsymbol{y} \equiv (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \in \mathbb{IR}^n$ will be represented equivalently as $[\underline{y}, \overline{y}]$, where $\underline{y} := (\underline{y}_1, \ldots, \underline{y}_n) \in \mathbb{R}^n$ and $\overline{y} := (\overline{y}_1, \ldots, \overline{y}_n) \in \mathbb{R}^n$. An interval $\boldsymbol{x} \in \mathbb{IR}$ has a *width* of $\mathrm{wid}\,\boldsymbol{x} := \overline{x} - \underline{x}$, and an interval vector $\boldsymbol{y} \in \mathbb{IR}^n$ has a width of $\mathrm{wid}\,\boldsymbol{y} := \max_{k\in\{1,\ldots,n\}} \mathrm{wid}\,\boldsymbol{y}_k$.

This article makes use of standard definitions and results from interval analysis concerning operations, *inclusion monotonicity*, *interval extensions*, and *interval hulls*; these are summarized in Appendix A.1. For further details, the reader is directed to the introductory sources [27,25,2].

Definition 2 (adapted from [25]) Consider a nonempty set $B \subset \mathbb{R}^n$. An interval function $\boldsymbol{f} : \mathbb{I}B \to \mathbb{IR}^m$ is *locally Lipschitz continuous* if for each $\boldsymbol{q} \in \mathbb{I}B$, there exists $k \geq 0$ for which

$$\text{wid}\,(\boldsymbol{f}(\boldsymbol{x})) \leq k\,\text{wid}\,\boldsymbol{x}, \qquad \forall \boldsymbol{x} \in \mathbb{I}\boldsymbol{q}.$$

A locally Lipschitz continuous, inclusion-monotonic interval extension of a function f will be called a *tight interval extension* of f.

The results in this article apply to finite compositions of the functions formalized by the following definition.

Definition 3 (adapted from [33]) Given an open set $B \subset \mathbb{R}$, a function $u : B \to \mathbb{R}$ is a *univariate intrinsic function (UIF)* if there exists a known tight interval extension $\tilde{\boldsymbol{u}} : \mathbb{I}B \to \mathbb{IR}$ of u, and if, with $\bar{B} := \{(\boldsymbol{x}, z) \in \mathbb{I}B \times B : z \in \boldsymbol{x}\}$, there exist known functions $u^{\text{cv}}, u^{\text{cc}} : \bar{B} \to \mathbb{R}$ and $\zeta_u^{\min}, \zeta_u^{\max} : \mathbb{I}B \to \mathbb{R}$ satisfying all of the following conditions.

- For each $\boldsymbol{x} \in \mathbb{I}B$, $u^{\text{cv}}(\boldsymbol{x}, \cdot)$ is convex on $\boldsymbol{x}$, $u^{\text{cc}}(\boldsymbol{x}, \cdot)$ is concave on $\boldsymbol{x}$, and $u^{\text{cv}}(\boldsymbol{x}, z) \leq u(z) \leq u^{\text{cc}}(\boldsymbol{x}, z)$ for each $z \in \boldsymbol{x}$.
- For each $\boldsymbol{x} \in \mathbb{I}B$, $\zeta_u^{\min}(\boldsymbol{x}) \in \arg\min\{u^{\text{cv}}(\boldsymbol{x}, z) : z \in \boldsymbol{x}\}$ and $\zeta_u^{\max}(\boldsymbol{x}) \in \arg\max\{u^{\text{cc}}(\boldsymbol{x}, z) : z \in \boldsymbol{x}\}$.
- For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{I}B$ with $\boldsymbol{x} \subset \boldsymbol{y}$, and for any $z \in \boldsymbol{x}$, $u^{\text{cv}}(\boldsymbol{y}, z) \leq u^{\text{cv}}(\boldsymbol{x}, z)$ and $u^{\text{cc}}(\boldsymbol{y}, z) \geq u^{\text{cc}}(\boldsymbol{x}, z)$.
- For each $z \in B$, $u^{\text{cv}}([z, z], z) = u^{\text{cc}}([z, z], z) = u(z)$.

For any $z \in \boldsymbol{x} \in \mathbb{I}B$, define

$$u_I^{\text{cv}}(\boldsymbol{x}, z) := u^{\text{cv}}(\max\{z, \zeta_u^{\min}(\boldsymbol{x})\}), \qquad u_D^{\text{cv}}(\boldsymbol{x}, z) := u^{\text{cv}}(\min\{z, \zeta_u^{\min}(\boldsymbol{x})\}),$$
$$u_I^{\text{cc}}(\boldsymbol{x}, z) := u^{\text{cc}}(\min\{z, \zeta_u^{\max}(\boldsymbol{x})\}), \quad \text{and} \quad u_D^{\text{cc}}(\boldsymbol{x}, z) := u^{\text{cc}}(\max\{z, \zeta_u^{\max}(\boldsymbol{x})\}).$$

The interval hull of a locally Lipschitz continuous function is clearly a tight interval extension of the function. The interval operations in Definition 15 are interval hulls of the corresponding operations on real numbers. Tight interval extensions are provided for a number of UIFs in Table 1; these interval extensions are also interval hulls. Appropriate constructions of the functions u^{cv} and u^{cc} are also provided for these UIFs in Table 2. By inspection, these particular constructions all satisfy the properties:

$$\min_{z \in \boldsymbol{x}} u^{\text{cv}}(\boldsymbol{x}, z) = \min_{z \in \boldsymbol{x}} u(z), \qquad \text{and} \qquad \max_{z \in \boldsymbol{x}} u^{\text{cc}}(\boldsymbol{x}, z) = \max_{z \in \boldsymbol{x}} u(z);$$

in general, a weaker version of these properties will be required in Assumption 1 below.

Definition 4 (adapted from [23,24]) Given a nonempty set $B \subset \mathbb{R}^n$, a function $f : B \to \mathbb{R}^m$ is *factorable* if each of the following conditions is satisfied:

- f can be expressed on B as a finite composition (in some order) of addition, multiplication, and UIFs, and
- a well-defined *natural interval extension* $\tilde{\boldsymbol{f}} : \mathbb{I}B \to \mathbb{IR}^n$ of f can be constructed by replacing each addition/multipication/UIF by its corresponding tight interval extension, without introducing any domain violations.

The natural interval extension of a factorable function is a tight interval extension of the function [25, Section 3.3].

Table 1 Tight interval extensions for various UIFs u.

B	$u(z)$ for $z \in B$	$\tilde{\boldsymbol{u}}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{I}B$
$\mathbb{R}$	cz for fixed $c \in \mathbb{R}$	$c\boldsymbol{x}$
$\mathbb{R}$	$\exp z$	$[\exp \underline{x}, \exp \overline{x}]$
$(0,+\infty)$	$\ln z$	$[\ln \underline{x}, \ln \overline{x}]$
$\mathbb{R}$	z^{2k} for fixed $k \in \mathbb{N}$	$[(\mathrm{mid}(0,\underline{x},\overline{x}))^{2k}, \max\{\underline{x}^{2k}, \overline{x}^{2k}\}]$
$\mathbb{R}$	z^{2k+1} for fixed $k \in \mathbb{N}$	$[\underline{x}^{2k+1}, \overline{x}^{2k+1}]$
$(0,+\infty)$	$\sqrt{z}$	$[\sqrt{\underline{x}}, \sqrt{\overline{x}}]$
$\mathbb{R}$	$\lvert z\rvert$	$[\lvert\mathrm{mid}(0,\underline{x},\overline{x})\rvert, \max\{\lvert\underline{x}\rvert, \lvert\overline{x}\rvert\}]$
$(0,+\infty)$	$\frac{1}{z^k}$ for fixed $k \in \mathbb{N}$	$\left[\frac{1}{\overline{x}^k}, \frac{1}{\underline{x}^k}\right]$
$(-\infty,0)$	$\frac{1}{z^{2k}}$ for fixed $k \in \mathbb{N}$	$\left[\frac{1}{\underline{x}^{2k}}, \frac{1}{\overline{x}^{2k}}\right]$
$(-\infty,0)$	$\frac{1}{z^{2k-1}}$ for fixed $k \in \mathbb{N}$	$\left[\frac{1}{\overline{x}^{2k-1}}, \frac{1}{\underline{x}^{2k-1}}\right]$

2.3 McCormick objects and relaxations

This section presents and extends definitions and properties concerning the *generalized McCormick* framework [42,33], which expresses the classical development of McCormick's relaxation technique [23] in terms of the abstract objects containing bounding and relaxing information that are propagated by MC++ [8] in order to carry out McCormick's scheme in practice. As described in Section 1, the main results in the current article are presented in the generalized McCormick framework. Notation from [42] is employed.

Definition 5 (from [33]) The set of *McCormick objects* of n variables is defined as $\mathbb{MR}^n := \{(\boldsymbol{z}^{\mathrm{B}}, \boldsymbol{z}^{\mathrm{C}}) \in \mathbb{IR}^n \times \mathbb{IR}^n : \boldsymbol{z}^{\mathrm{B}} \cap \boldsymbol{z}^{\mathrm{C}} \neq \emptyset\}$. For any $\mathscr{X} \in \mathbb{MR}^n$, $\mathscr{X}$ will be represented equivalently as

$$\mathscr{X} \equiv (\boldsymbol{x}^{\mathrm{B}}, \boldsymbol{x}^{\mathrm{C}}) \equiv ([\underline{x}^{\mathrm{B}}, \overline{x}^{\mathrm{B}}], [\underline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{C}}]).$$

Given $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}^n$, $\mathscr{X} \subset \mathscr{Y}$ if and only if both $\boldsymbol{x}^{\mathrm{B}} \subset \boldsymbol{y}^{\mathrm{B}}$ and $\boldsymbol{x}^{\mathrm{C}} \subset \boldsymbol{y}^{\mathrm{C}}$. The set of *proper McCormick objects* of n variables is $\mathbb{MR}^n_{\mathrm{prop}} := \{(\boldsymbol{z}^{\mathrm{B}}, \boldsymbol{z}^{\mathrm{C}}) \in \mathbb{MR}^n : \boldsymbol{z}^{\mathrm{C}} \subset \boldsymbol{z}^{\mathrm{B}}\}$. Given a set $B \subset \mathbb{R}^n$, define $\mathbb{M}B := \{\mathscr{X} \in \mathbb{MR}^n : \boldsymbol{x}^{\mathrm{B}} \in \mathbb{I}B\}$, and $\mathbb{M}B_{\mathrm{prop}} := \{\mathscr{X} \in \mathbb{MR}^n : \boldsymbol{x}^{\mathrm{C}} \subset \boldsymbol{x}^{\mathrm{B}} \in \mathbb{I}B\} \subset \mathbb{MR}^n_{\mathrm{prop}}$.

Roughly, the $\boldsymbol{x}^{\mathrm{B}}$-component of a McCormick object $\mathscr{X}$ contains interval bounding information, and the $\boldsymbol{x}^{\mathrm{C}}$-component contains information used to construct convex underestimators and concave overestimators. These notions will be formalized by the following results.

Notions of *McCormick extensions*, *inclusion monotonicity*, and *coherent concavity* for functions of McCormick objects were developed in [42]. In this article, these notions have been altered slightly to permit restrictions to proper McCormick objects, and are presented in Appendix A.2. The following definition is stricter than in its analog in [42], and combines these properties.

Definition 6 Given a function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$, a mapping $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\mathrm{prop}}) \to \mathbb{MR}^m$ is a *relaxation function* for f if it is coherently concave, inclusion monotonic, and a McCormick extension of f.

The following two propositions demonstrate the utility of relaxation functions: they are closed under composition, and effectively define convex underestimators and concave overestimators of the underlying functions they relax.

Proposition 3 (Lemmas 2.4.15 and 2.4.17 in [33]) *Consider functions $f : B \subset \mathbb{R}^n \to D \subset \mathbb{R}^m$ and $g : D \to \mathbb{R}^k$, a relaxation function $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) $\to \mathbb{M}\mathbb{R}^m$ for f, and a relaxation function $\mathscr{G} : \mathbb{M}D$ (or $\mathbb{M}D_{\text{prop}}$) $\to \mathbb{M}\mathbb{R}^k$ for g. Define $B_0 := \{\mathscr{X} \in \mathbb{M}B \text{ (or } \mathbb{M}B_{\text{prop}}) : \mathscr{F}(\mathscr{X}) \in \mathbb{M}D\}$. If there are no domain violations in constructing the composition $\mathscr{G} \circ \mathscr{F} : B_0 \to \mathbb{M}\mathbb{R}^k$, then $\mathscr{G} \circ \mathscr{F}$ is a relaxation function for $g \circ f : B \to \mathbb{R}^k$.*

Proposition 4 (Lemma 2.4.11 in [33]) *Given a function $f : B \subset \mathbb{R}^n \to \mathbb{R}$, a relaxation function $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) $\to \mathbb{M}\mathbb{R}$ for f on B, and some $\boldsymbol{x} \in \mathbb{I}B$, define functions $\phi_{f,\boldsymbol{x}}, \psi_{f,\boldsymbol{x}} : \boldsymbol{x} \to \mathbb{R}$ such that:*

$$\phi_{f,\boldsymbol{x}}(z) = \underline{f}^{\mathrm{C}}(\boldsymbol{x}, [z,z]), \qquad \text{and} \qquad \psi_{f,\boldsymbol{x}}(z) = \overline{f}^{\mathrm{C}}(\boldsymbol{x}, [z,z]), \qquad \forall z \in \boldsymbol{x}.$$

Then $\phi_{f,\boldsymbol{x}}$ is convex on $\boldsymbol{x}$, $\psi_{f,\boldsymbol{x}}$ is concave on $\boldsymbol{x}$, and $\phi_{f,\boldsymbol{x}}(z) \leq f(z) \leq \psi_{f,\boldsymbol{x}}(z)$ for each $z \in \boldsymbol{x}$.

The goal of this article is to obtain $\mathscr{C}^i$ relaxations of a factorable function, for any particular $i \in \{1,2\}$. Achieving this will require appending the following nonstandard assumption to Definition 3 for each employed UIF. This assumption will be invoked explicitly whenever it is required.

Assumption 1 *For particular $i \in \{1,2\}$, given a UIF $u : B \subset \mathbb{R} \to \mathbb{R}$, assume for each $\boldsymbol{x} \in \mathbb{I}B$ that the functions $u^{\text{cv}}(\boldsymbol{x}, \cdot)$ and $u^{\text{cc}}(\boldsymbol{x}, \cdot)$ are each $\mathscr{C}^i$ on $\boldsymbol{x}$, and that $\underline{\tilde{u}}(\boldsymbol{x}) \leq u^{\text{cv}}(\boldsymbol{x}, \zeta_u^{\min}(\boldsymbol{x}))$ and $\overline{\tilde{u}}(\boldsymbol{x}) \geq u^{\text{cc}}(\boldsymbol{x}, \zeta_u^{\max}(\boldsymbol{x}))$. If $i = 2$, assume additionally that:*

- *if $\zeta_u^{\min}(\boldsymbol{x}) \in \operatorname{int}(\boldsymbol{x})$, then the second derivative of $u^{\text{cv}}(\boldsymbol{x}, \cdot)$ is zero at $\zeta_u^{\min}(\boldsymbol{x})$, and*
- *if $\zeta_u^{\max}(\boldsymbol{x}) \in \operatorname{int}(\boldsymbol{x})$, then the second derivative of $u^{\text{cc}}(\boldsymbol{x}, \cdot)$ is zero at $\zeta_u^{\max}(\boldsymbol{x})$.*

Observe that the above assumption does not require u itself to be $\mathscr{C}^i$. Indeed, the $\mathscr{C}^i$ relaxations obtained in this article will remain valid even when nondifferentiable UIFs are employed. However, [6] shows that nondifferentiable UIFs cannot satisfy Assumption 2 below, which will be required in this article to ensure sufficiently rapid convergence of the obtained relaxations to the original function as the width of $\boldsymbol{x}$ approaches zero.

Remark 2 Lemmata 12 and 13 in Appendix B show that, if $i = 1$ and a function $u : B \to \mathbb{R}$ is $\mathscr{C}^1$, then Assumption 1 and the conditions of Definition 3 are satisfied when $u^{\text{cv}}(\boldsymbol{x}, \cdot)$ and $u^{\text{cc}}(\boldsymbol{x}, \cdot)$ are chosen to be the convex and concave envelopes of u on $\boldsymbol{x}$, respectively, and when $\zeta_u^{\min}(\boldsymbol{x})$ and $\zeta_u^{\max}(\boldsymbol{x})$ are chosen according to Definition 3.

Remark 3 Consider a univariate function u that is $\mathscr{C}^2$, is either convex or concave, and is either monotonically increasing or monontonically decreasing on its domain. Moreover, note that the concave envelope of a univariate convex function on an interval is a secant, as is the convex envelope of a univariate concave function on an interval. Thus, even if $i = 2$, Assumption 1 and the conditions of Definition 3 are satisfied when $u^{\text{cv}}(\boldsymbol{x}, \cdot)$ and $u^{\text{cc}}(\boldsymbol{x}, \cdot)$ are chosen to be the convex and concave envelopes of u on $\boldsymbol{x}$, respectively, and when $\zeta_u^{\min}(\boldsymbol{x})$ and $\zeta_u^{\max}(\boldsymbol{x})$ are chosen according to Definition 3.

Remark 4 The condition in Assumption 1 that both $\underline{\tilde{u}}(\boldsymbol{x}) \leq u^{\text{cv}}(\boldsymbol{x}, \zeta_u^{\min}(\boldsymbol{x}))$ and $\overline{\tilde{u}}(\boldsymbol{x}) \geq u^{\text{cv}}(\boldsymbol{x}, \zeta_u^{\max}(\boldsymbol{x}))$ can be imposed without loss of generality, as detailed in Remark 5 below.

Any univariate function $u : B \subset \mathbb{R} \to \mathbb{R}$ on an open set can be considered to be a UIF, provided that the functions $\tilde{\boldsymbol{u}}$, u^{cv}, and u^{cc} are known or can be constructed. Table 2 presents

Table 2 Functions $u^{\mathrm{cv}}, u^{\mathrm{cc}}$ that satisfy the conditions of Definition 3 and Assumption 1 for various UIFs u. These functions also satisfy Assumption 2, except when $u : z \mapsto |z|$. Examples 7–9 are in Appendix B.

B	$u(z)$ for $z \in B$	$u^{\mathrm{cv}}(\boldsymbol{x}, z)$ for $\boldsymbol{x} \in \mathbb{I}B, z \in B$	$u^{\mathrm{cc}}(\boldsymbol{x}, z)$ for $\boldsymbol{x} \in \mathbb{I}B, z \in B$
$\mathbb{R}$	cz for fixed $c \in \mathbb{R}$	cz	cz
$\mathbb{R}$	$\exp z$	$\exp z$	$\exp \underline{x} + (\exp \bar{x} - \exp \underline{x}) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$
$(0, +\infty)$	$\ln z$	$\ln \underline{x} + (\ln \bar{x} - \ln \underline{x}) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$	$\ln z$
$\mathbb{R}$	z^2	See Example 7	$\underline{x}^2 + (\bar{x}^2 - \underline{x}^2) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$
$\mathbb{R}$	z^{2k+2} for fixed $k \in \mathbb{N}$	z^{2k+2}	$\underline{x}^{2k+2} + (\bar{x}^{2k+2} - \underline{x}^{2k+2}) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$
$\mathbb{R}$	z^{2k+1} for fixed $k \in \mathbb{N}$	See Example 9	See Example 9
$(0, +\infty)$	$\sqrt{z}$	$\sqrt{\underline{x}} + (\sqrt{\bar{x}} - \sqrt{\underline{x}}) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$	$\sqrt{z}$
$\mathbb{R}$	$\lvert z \rvert$	See Example 8	$\lvert \underline{x} \rvert + (\lvert \bar{x} \rvert - \lvert \underline{x} \rvert) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$
$(0, +\infty)$	$\frac{1}{z^k}$ for fixed $k \in \mathbb{N}$	$\frac{1}{z^k}$	$\frac{1}{\underline{x}^k} + (\frac{1}{\bar{x}^k} - \frac{1}{\underline{x}^k}) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$
$(-\infty, 0)$	$\frac{1}{z^{2k}}$ for fixed $k \in \mathbb{N}$	$\frac{1}{z^{2k}}$	$\frac{1}{\underline{x}^{2k}} + (\frac{1}{\bar{x}^{2k}} - \frac{1}{\underline{x}^{2k}}) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$
$(-\infty, 0)$	$\frac{1}{z^{2k-1}}$ for fixed $k \in \mathbb{N}$	$\frac{1}{\underline{x}^{2k-1}} + (\frac{1}{\bar{x}^{2k-1}} - \frac{1}{\underline{x}^{2k-1}}) \left(\frac{z - \underline{x}}{\bar{x} - \underline{x}} \right)$	$\frac{1}{z^{2k-1}}$

functions $u^{\mathrm{cv}}, u^{\mathrm{cc}}$ which satisfy the conditions of Definition 3 and Assumption 1 for the UIFs u considered in Table 1.

Within this framework, McCormick's classical relaxations [23] can be restated as the convex/concave relaxations implied by Propositions 3 and 4 for a factorable function, when each addition/multiplication/UI operation is replaced by a relaxation function of the operation. Such relaxation functions were described in [33], and are presented as Definitions 21–23 in Appendix A.2. These relaxation functions suggest the construction of an analog of a natural interval extension for a factorable function, using McCormick objects instead of intervals. This notion is formalized in the following definition, which is motivated by the subsequent theorem.

Definition 7 (adapted from [33]) Given a factorable function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$, a *natural McCormick extension* $\mathscr{F} : \mathbb{M}B \to \mathbb{M}\mathbb{R}^m$ of f is defined by replacing each addition operation, multiplication operation, and UIF in the construction of f with its McCormick counterpart described by Definitions 21–23 in Appendix A.2, provided that there are no domain violations in the introduced McCormick arithmetic.

Theorem 1 (Theorem 2.4.32 in [33]) *Given a factorable function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$ with a well-defined natural McCormick extension $\mathscr{F}$, $\mathscr{F}$ is a relaxation function for f.*

The classical McCormick relaxations of a factorable function are the convex/concave relaxations implied by the above theorem and by Proposition 4. These relaxations may be nonsmooth; the central goal of this article is to develop $\mathscr{C}^1$ and $\mathscr{C}^2$ variants of these relaxations.

2.4 Convergence analysis

Intuitively, to be useful, a scheme for constructing convex and concave relaxations of a scalar-valued function on an interval should converge rapidly to the underlying function as the width of interval is reduced to zero. Appopriate notions of convergence were formalized by Bompadre and Mitsos [6], and were extended to McCormick objects by Schaber [31]; these notions are summarized here.

Definition 8 (adapted from [6]) Given a continuous function $f : B \subset \mathbb{R}^n \to \mathbb{R}$, functions $\{f^{\mathrm{cv}}(\boldsymbol{x},\cdot), f^{\mathrm{cc}}(\boldsymbol{x},\cdot) : \boldsymbol{x} \to \mathbb{R}\}_{\boldsymbol{x}\in\mathbb{I}B}$ comprise a *scheme of estimators* for f if, for each $\boldsymbol{x} \in \mathbb{I}B$, $f^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ is convex on $\boldsymbol{x}$, $f^{\mathrm{cc}}(\boldsymbol{x},\cdot)$ is concave on $\boldsymbol{x}$, and

$$f^{\mathrm{cv}}(\boldsymbol{x},z) \le f(z) \le f^{\mathrm{cc}}(\boldsymbol{x},z), \qquad \forall z \in \boldsymbol{x}.$$

Such a scheme is *pointwise convergent of order* t_0 if, for each $\boldsymbol{q} \in \mathbb{I}B$, there exists $a_0 > 0$ such that

$$\begin{aligned} &\sup_{z\in\boldsymbol{x}}(f(z) - f^{\mathrm{cv}}(\boldsymbol{x},z)) \le a_0(\mathrm{wid}\,\boldsymbol{x})^{t_0}, \qquad \forall \boldsymbol{x} \in \mathbb{I}\boldsymbol{q}, \\ \text{and}\qquad &\sup_{z\in\boldsymbol{x}}(f^{\mathrm{cc}}(\boldsymbol{x},z) - f(z)) \le a_0(\mathrm{wid}\,\boldsymbol{x})^{t_0}, \qquad \forall \boldsymbol{x} \in \mathbb{I}\boldsymbol{q}. \end{aligned}$$

The following example motivates the incorporation of the interval $\boldsymbol{q}$ into this definition.

Example 2 Consider a function $f : \mathbb{R} \to \mathbb{R}$ and a scheme of estimators $\{f^{\mathrm{cv}}(\boldsymbol{x},\cdot), f^{\mathrm{cc}}(\boldsymbol{x},\cdot)\}_{\boldsymbol{x}\in\mathbb{IR}}$ for f, for which, for each $\boldsymbol{x} \in \mathbb{IR}$,

$$\sup_{z\in\boldsymbol{x}}(f(z) - f^{\mathrm{cv}}(\boldsymbol{x},z)) = \sup_{z\in\boldsymbol{x}}(f^{\mathrm{cc}}(\boldsymbol{x},z) - f(z)) = \begin{cases} (\mathrm{wid}\,\boldsymbol{x})^2, & \text{if } \mathrm{wid}\,\boldsymbol{x} \le 1, \\ (\mathrm{wid}\,\boldsymbol{x})^3, & \text{if } \mathrm{wid}\,\boldsymbol{x} > 1. \end{cases}$$

According to the above definition, this scheme is pointwise convergent of order 2. In the original definition [6], however, this scheme is not pointwise convergent of order 2, since, for each $a_0 > 0$, there exists a sufficiently large interval $\boldsymbol{x} \in \mathbb{IR}$ for which

$$a_0(\mathrm{wid}\,\boldsymbol{x})^2 < (\mathrm{wid}\,\boldsymbol{x})^3.$$

In fact, according to the definition in [6], this scheme is not pointwise convergent of any order. Since applications of pointwise convergence in [6] are only concerned with sufficiently small intervals, the interval $\boldsymbol{q}$ was added to the definition above so that the constants a_0 and t_0 need not apply to arbitrarily large intervals in $\mathbb{I}B$.

By Theorem 2 in [6], if f is nonaffine and twice-continuously differentiable, then there does not exist any scheme of estimators for f with pointwise convergence of order greater than 2. Given a factorable function expressed as a composition of twice-continuously differentiable functions, the classical McCormick relaxations of this function are pointwise convergent of order 2 [6], as are the αBB relaxations [1,6]. A scheme of estimators with second-order pointwise convergence is typically necessary to mitigate clustering when carrying out a branch-and-bound method for global optimization [9]. Certain optimization problems with nondifferentiable objective functions, however, are not subject to this requirement [40].

Consider a factorable function f that is a composition only of locally Lipschitz continuous functions. Given a natural interval extension $\tilde{\boldsymbol{f}}$ of f, the constant mappings $\{z \mapsto \underline{\tilde{f}}(\boldsymbol{x}), z \mapsto \overline{\tilde{f}}(\boldsymbol{x})\}_{\boldsymbol{x}\in\mathbb{I}B}$ comprise a scheme of estimators for f that is pointwise convergent of order 1 [32].

The following definition formalizes a notion of *width* of a McCormick object, and a corresponding notion of convergence of a function of McCormick objects, as the width of the argument tends to zero.

Definition 9 (adapted from [31]) A McCormick object $\mathscr{X} \in \mathbb{MR}$ has a *width* of

$$\text{wid}_{\mathscr{M}}\,\mathscr{X} := \text{wid}\,(\boldsymbol{x}^{\text{B}} \cap \boldsymbol{x}^{\text{C}}) = \min\{\bar{x}^{\text{C}}, \bar{x}^{\text{B}}\} - \max\{\underline{x}^{\text{C}}, \underline{x}^{\text{B}}\}.$$

A vector $\mathscr{Y} \in \mathbb{MR}^n$ of McCormick objects has a width of

$$\text{wid}_{\mathscr{M}}\,\mathscr{Y} \equiv \text{wid}_{\mathscr{M}}\,(\mathscr{Y}_1, \ldots, \mathscr{Y}_n) := \max_{k \in \{1,\ldots,n\}} \text{wid}_{\mathscr{M}}\,\mathscr{Y}_k.$$

A function $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}) \subset \mathbb{MR}^n \to \mathbb{MR}^m$ is (t_1, t_2)-*convergent* on $\mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) if, for each $\boldsymbol{q} \in \mathbb{I}B$, there exist $a_1, a_2 > 0$ such that

$$\text{wid}_{\mathscr{M}}\,(\mathscr{F}(\mathscr{X})) \leq a_1(\text{wid}_{\mathscr{M}}\,\mathscr{X})^{t_1} + a_2(\text{wid}\,\boldsymbol{x}^{\text{B}})^{t_2}, \qquad \forall \mathscr{X} \in \mathbb{M}\boldsymbol{q} \text{ (or } \mathbb{M}\boldsymbol{q}_{\text{prop}}).$$

Again, the interval $\boldsymbol{q}$ has been added to this definition to prevent the fixed constants a_1, a_2 from having to be applicable to every choice of $\boldsymbol{x}^{\text{B}} \in \mathbb{I}B$.

As described in Section 3.2 of [31], given a (t_1, t_2)-convergent relaxation function $\mathscr{F}$ for a function f, the corresponding convex/concave relaxations of f described by Proposition 4 exhibit pointwise convergence of order t_2. Moreover, as described in Section 3.9.7 of [31], a well-defined composition of $(1,2)$-convergent McCormick-valued functions is itself $(1,2)$-convergent. This notion motivates the following assumption, which will be appended frequently to Definition 3.

Assumption 2 *Given a UIF $u : B \subset \mathbb{R} \to \mathbb{R}$, assume that $\{u^{\text{cv}}(\boldsymbol{x}, \cdot), u^{\text{cc}}(\boldsymbol{x}, \cdot)\}_{\boldsymbol{x} \in \mathbb{I}B}$ comprises a scheme of estimators for u on B that is pointwise convergent of order 2.*

It will be shown in this article that the above assumption yields $(1,2)$-convergent relaxation functions for UIFs. This assumption is satisfied by the functions $u^{\text{cv}}, u^{\text{cc}}$ described in Table 2, except when u is the absolute value function $z \mapsto |z|$. This is demonstrated in [6] for each u other than $z \mapsto z^2$ and $z \mapsto z^{2k+1}$ for $k \in \mathbb{N}$, which are considered in Lemmata 14 and 15 in Appendix B.

3 Smoothing constructions

This section establishes basic properties of certain $\mathscr{C}^1$ and $\mathscr{C}^2$ relaxations of simple nonsmooth functions such as $z \mapsto \max\{z, 0\}$ and $(x,y) \mapsto \max\{x,y\}$, and uses these to construct variants of McCormick's multiplication rule. These rules will be shown in subsequent sections to have various desirable properties.

3.1 Relaxing simple nonsmooth functions

Definition 10 Define functions $\mu_1, \mu_2 : \mathbb{R} \to \mathbb{R}$ as follows:

$$\mu_1 : y \mapsto \begin{cases} 0, & \text{if } y \leq 0, \\ \frac{1}{4}y^2, & \text{if } 0 < y < 2, \\ y - 1, & \text{if } 2 \leq y, \end{cases} \qquad \mu_2 : y \mapsto \begin{cases} 0, & \text{if } y \leq 0, \\ \frac{1}{16}y^3(4-y), & \text{if } 0 < y < 2, \\ y - 1, & \text{if } 2 \leq y. \end{cases}$$

For each $i \in \{1,2\}$, define functions $\gamma_i, \sigma_i : \mathbb{R} \times \mathbb{R} \times [0, +\infty) \to \mathbb{R}$ as follows:

$$\gamma_i : (z, a, p) \mapsto \begin{cases} \max\{z, a\}, & \text{if } p = 0, \\ a + p\,\mu_i(\frac{z-a}{p}), & \text{if } p > 0, \end{cases} \qquad \sigma_i : (z, b, p) \mapsto \begin{cases} \min\{z, b\}, & \text{if } p = 0, \\ b - p\,\mu_i(\frac{b-z}{p}), & \text{if } p > 0, \end{cases}$$

and define functions $\nu_i, \lambda_i : \mathbb{R} \times \mathbb{R} \times [0, +\infty) \to \mathbb{R}$ as follows:

$$\nu_i : (x,y,p) \mapsto \tfrac{1}{2}(\gamma_i(x,y,p) + \gamma_i(y,x,p)),$$
$$\lambda_i : (x,y,p) \mapsto \tfrac{1}{2}(\sigma_i(x,y,p) + \sigma_i(y,x,p)).$$

Observe that μ_1 is a member of the family of functions considered in [10, Example 11.8.11(c)]. In this article, in the spirit of [4], [10, Section 11.8] and [5, Section 1.10], μ_1 and μ_2 essentially serve as analogs of the mapping $y \mapsto \max\{y, 0\}$ which exhibit several useful properties. Ultimately, μ_1, γ_1, and ν_1 will be used to construct $\mathscr{C}^1$ analogs of McCormick relaxations, while μ_2, γ_2, and ν_2 will be used to construct $\mathscr{C}^2$ relaxations. By inspection, μ_1 and μ_2 are each $\mathscr{C}^1$, with

$$\nabla\mu_1 : y \mapsto \begin{cases} 0, & \text{if } y \le 0, \\ \frac{1}{2}y, & \text{if } 0 < y < 2, \\ 1, & \text{if } 2 \le y, \end{cases} \qquad \nabla\mu_2 : y \mapsto \begin{cases} 0, & \text{if } y \le 0, \\ \frac{1}{4}y^2(3-y), & \text{if } 0 < y < 2, \\ 1, & \text{if } 2 \le y. \end{cases} \tag{1}$$

The above expressions show that $\mu_1(y)$, $\mu_2(y)$, $\nabla\mu_1(y)$, and $\nabla\mu_2(y)$ are each nonnegative for each $y \in \mathbb{R}$, noting that $3 - y > 0$ when $0 < y < 2$. Thus, μ_1 and μ_2 are increasing on $\mathbb{R}$. Moreover, μ_2 is $\mathscr{C}^2$, with

$$\nabla^2\mu_2 : y \mapsto \begin{cases} 0, & \text{if } y \le 0, \\ \frac{3}{4}y(2-y), & \text{if } 0 < y < 2, \\ 0, & \text{if } 2 \le y. \end{cases} \tag{2}$$

The following lemmata summarize basic properties of the functions μ_i, γ_i, σ_i, ν_i, and λ_i for each $i \in \{1,2\}$; each of these properties can be demonstrated readily. First, basic properties of μ_i are presented.

Lemma 1 *For each $i \in \{1,2\}$ and each $y \in \mathbb{R}$,* $\max\{y-1, 0\} \le \mu_i(y) \le \max\{y, 0\}$.

Lemma 2 *The functions μ_1 and μ_2 are convex on $\mathbb{R}$.*

Next, useful properties of γ_i, σ_i, ν_i, and λ_i will be established. Intuitively, throughout this article, $\gamma_i(z,a,p)$ plays a similar role to $\max\{z,a\}$ for fixed a, $\sigma_i(z,b,p)$ is analogous to $\min\{z,b\}$ for fixed b, $\nu_i(x,y,p)$ is analogous to $\max\{x,y\}$ for varying x and y, and $\lambda_i(x,y,p)$ is analogous to $\min\{x,y\}$ for varying x and y. Roughly, the parameter p quantifies the extent to which γ_i and σ_i are relaxed to yield a differentiable underestimator of $\max\{\cdot, a\}$ and a differentiable overestimator of $\min\{\cdot, b\}$, as is formalized in the following lemma.

Lemma 3 *Consider any fixed $i \in \{1,2\}$, $a, b \in \mathbb{R}$, and $p \ge 0$. The mapping $\gamma_i(\cdot, a, p)$ is convex and increasing. Moreover,*

$$a \le \max\{z-p, a\} \le \gamma_i(z,a,p) \le \max\{z,a\}, \qquad \forall z \in \mathbb{R}.$$

Similarly, the mapping $\sigma_i(\cdot, b, p)$ is concave and increasing, with

$$\min\{z,b\} \le \sigma_i(z,b,p) \le \min\{z+p, b\} \le b, \qquad \forall z \in \mathbb{R}.$$

If $p > 0$, then $\gamma_i(\cdot, a, p)$ and $\sigma_i(\cdot, b, p)$ are both $\mathscr{C}^i$.

Proposition 5 *Consider any fixed $i \in \{1,2\}$, $a, b \in \mathbb{R}$ and $p > 0$. Gradients of the mappings $z \mapsto \gamma_i(z,a,p)$ and $z \mapsto \sigma_i(z,b,p)$ at some $z_0 \in \mathbb{R}$ may be computed using* (1) *as follows.*

$$\frac{\partial \gamma_i}{\partial z}(z_0, a, p) = \nabla\mu_i\left(\tfrac{z_0 - a}{p}\right), \qquad \frac{\partial \sigma_i}{\partial z}(z_0, b, p) = \nabla\mu_i\left(\tfrac{b - z_0}{p}\right).$$

Lemma 4 *Given triples* $(z_1,a_1,p_1),(z_2,a_2,p_2)\in\mathbb{R}\times\mathbb{R}\times[0,+\infty)$*, suppose that* $z_1\leq z_2$*,* $a_1\leq a_2$*, and* $p_1\geq p_2$*. Then, for each* $i\in\{1,2\}$*,* $\gamma_i(z_1,a_1,p_1)\leq\gamma_i(z_2,a_2,p_2)$*.*

Similarly, given triples $(z_1,b_1,p_1),(z_2,b_2,p_2)\in\mathbb{R}\times\mathbb{R}\times[0,+\infty)$*, suppose that* $z_1\geq z_2$*,* $b_1\geq b_2$*, and* $p_1\geq p_2$*. Then, for each* $i\in\{1,2\}$*,* $\sigma_i(z_1,b_1,p_1)\geq\sigma_i(z_2,b_2,p_2)$*.*

Lemma 5 *For any fixed* $p>0$ *and* $i\in\{1,2\}$*, the mappings* $(x,y)\mapsto\nu_i(x,y,p)$ *and* $(x,y)\mapsto\lambda_i(x,y,p)$ *are each* $\mathscr{C}^i$ *on* $\mathbb{R}^2$*.*

Proposition 6 *Consider any fixed* $i\in\{1,2\}$ *and* $p>0$*. Partial derivatives of the mappings* $(x,y)\mapsto\nu_i(x,y,p)$ *and* $(x,y)\mapsto\lambda_i(x,y,p)$ *at some* $x_0,y_0\in\mathbb{R}$ *may be computed using* (1) *as follows.*

$$\begin{aligned}\frac{\partial\nu_i}{\partial x}(x_0,y_0,p)=\frac{\partial\lambda_i}{\partial y}(x_0,y_0,p)&=\frac{1}{2}\left(1+\nabla\mu_i(\tfrac{x_0-y_0}{p})-\nabla\mu_i(\tfrac{y_0-x_0}{p})\right),\\ \frac{\partial\nu_i}{\partial y}(x_0,y_0,p)=\frac{\partial\lambda_i}{\partial x}(x_0,y_0,p)&=\frac{1}{2}\left(1-\nabla\mu_i(\tfrac{x_0-y_0}{p})+\nabla\mu_i(\tfrac{y_0-x_0}{p})\right).\end{aligned}$$

Lemma 6 *Given triples* $(x_1,y_1,p_1),(x_2,y_2,p_2)\in\mathbb{R}\times\mathbb{R}\times[0,+\infty)$*, suppose that* $x_1\leq x_2$*,* $y_1\leq y_2$*, and* $p_1\geq p_2$*. Then, for each* $i\in\{1,2\}$*,* $\nu_i(x_1,y_1,p_1)\leq\nu_i(x_2,y_2,p_2)$*.*

Similarly, given triples $(x_3,y_3,p_3),(x_4,y_4,p_4)\in\mathbb{R}\times\mathbb{R}\times[0,+\infty)$*, suppose that* $x_3\geq x_4$*,* $y_3\geq y_4$*, and* $p_3\geq p_4$*. Then, for each* $i\in\{1,2\}$*,* $\lambda_i(x_3,y_3,p_3)\geq\lambda_i(x_4,y_4,p_4)$*.*

Lemma 7 *Given* $p\geq 0$ *and* $i\in\{1,2\}$*, the mapping* $(x,y)\mapsto\nu_i(x,y,p)$ *is convex on* $\mathbb{R}^2$*, and the mapping* $(x,y)\mapsto\lambda_i(x,y,p)$ *is concave on* $\mathbb{R}^2$*. Moreover,*

$$\begin{aligned}&\tfrac{1}{2}(x+y)\leq\tfrac{1}{2}(\max\{x-p,y\}+\max\{x,y-p\})\leq\nu_i(x,y,p)\leq\max\{x,y\},\quad\forall x,y\in\mathbb{R},\\ \textit{and}\quad&\min\{x,y\}\leq\lambda_i(x,y,p)\leq\tfrac{1}{2}(\min\{x+p,y\}+\min\{x,y+p\})\leq\tfrac{1}{2}(x+y),\quad\forall x,y\in\mathbb{R}.\end{aligned}$$

Definition 11 Define a function $p:\mathbb{IR}\to[0,+\infty)$ such that for some constant $a_p>0$, $p(\boldsymbol{x}):=a_p(\mathrm{wid}\,\boldsymbol{x})^2$ for each $\boldsymbol{x}\in\mathbb{IR}$. Denote $p(\boldsymbol{x})$ as $p_{\boldsymbol{x}}$.

In Definition 11, the particular quadratic expression for p is irrelevant to the results developed in Sections C.1 and C.2 below; the results in these sections remain valid if p is redefined so that $p(\boldsymbol{x}):=\pi(\mathrm{wid}\,\boldsymbol{x})$, where $\pi:[0,+\infty)\to[0,+\infty)$ is any particular strictly-increasing function for which $\pi(0)=0$. Defining $\pi:z\mapsto a_pz^2$, however, yields the convergence results obtained in Section C.3. The particular choice of the constant a_p does not affect the theoretical results developed in this article; appropriate choices of a_p will be discussed in Section 5.1 from a numerical standpoint.

Remark 5 As claimed earlier, the condition in Assumption 1 that both $\underline{\tilde{u}}(\boldsymbol{x})\leq u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))$ and $\overline{\tilde{u}}(\boldsymbol{x})\geq u^{\mathrm{cc}}(\boldsymbol{x},\zeta_u^{\max}(\boldsymbol{x}))$ can be imposed without loss of generality. If this condition either fails or is not known to be true, then, for each $\boldsymbol{x}\in\mathbb{I}B$, $u^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ can be replaced with the mapping $z\mapsto\gamma_i(u^{\mathrm{cv}}(\boldsymbol{x},z),\underline{\tilde{u}}(\boldsymbol{x}),p_{\boldsymbol{x}})$, and $u^{\mathrm{cc}}(\boldsymbol{x},\cdot)$ can be replaced with the mapping $z\mapsto\sigma_i(u^{\mathrm{cc}}(\boldsymbol{x},z),\overline{\tilde{u}}(\boldsymbol{x}),p_{\boldsymbol{x}})$; these replacements now satisfy the condition. The established properties of γ_i and σ_i ensure that the other conditions required of u^{cv} and u^{cc} by Definition 3 and Assumption 1 continue to hold under these replacements.

3.2 Relaxing intersections of bounds and relaxations

Roughly, for each $i \in \{1,2\}$, the $\mathscr{S}qu_i$ and $\boldsymbol{belt}_i$ operations introduced in this section are $\mathscr{C}^i$ relaxations of the "Cut" and "Enc" operations presented in Definitions 2.4.3 and 2.4.5 of [33], and serve analogous roles. It will be shown in this section that $\mathscr{S}qu_i$ is a relaxation function of the identity function on $\mathbb{R}$. Moreover, Lemma 26 in Appendix C shows that $\mathscr{S}qu_i$ is $(1,2)$-convergent. Intuitively, $\mathscr{S}qu_i$ also inherits the $\mathscr{C}^i$ nature of γ_i and σ_i. These properties will be exploited in Section 3.3 when constructing a $\mathscr{C}^i$ variant of McCormick's multiplication operation.

Definition 12 For each $\mathscr{X} \in \mathbb{MR}$ and each $i \in \{1,2\}$, define a *belt operation* $\boldsymbol{belt}_i(\mathscr{X}) \in \mathbb{IR}$ as follows:

$$\boldsymbol{belt}_i(\mathscr{X}) := \begin{cases} [x,x] & \text{if } \underline{x}^{\mathrm{B}} = \overline{x}^{\mathrm{B}} =: x, \\ [\gamma_i(\underline{x}^{\mathrm{C}}, \underline{x}^{\mathrm{B}}, p_{\boldsymbol{x}^{\mathrm{B}}}), \sigma_i(\overline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{B}}, p_{\boldsymbol{x}^{\mathrm{B}}})] & \text{if } \underline{x}^{\mathrm{B}} < \overline{x}^{\mathrm{B}}. \end{cases}$$

Define a *squashing operation* $\mathscr{S}qu_i(\mathscr{X}) := (\boldsymbol{x}^{\mathrm{B}}, \boldsymbol{belt}_i(\mathscr{X})) \in \mathbb{IR}^2$. Given a vector $\mathscr{Y} \in \mathbb{MR}^n$, define

$$\mathscr{S}qu_i(\mathscr{Y}) := \begin{bmatrix} \mathscr{S}qu_i(\mathscr{Y}_1) \\ \vdots \\ \mathscr{S}qu_i(\mathscr{Y}_n) \end{bmatrix} \in (\mathbb{IR}^2)^n.$$

For any $\mathscr{X} \in \mathbb{MR}$, Lemma 3 implies that $\boldsymbol{x}^{\mathrm{B}} \cap \boldsymbol{x}^{\mathrm{C}} \subset \boldsymbol{belt}_i(\mathscr{X}) \subset \boldsymbol{x}^{\mathrm{B}}$, which in turn yields $\mathscr{S}qu_i(\mathscr{X}) \in \mathbb{MR}_{\mathrm{prop}}$. Thus, $\mathscr{S}qu_i(\mathscr{Y}) \in \mathbb{MR}^n_{\mathrm{prop}}$ for any $\mathscr{Y} \in \mathbb{MR}^n$. Furthermore, observe that $\boldsymbol{belt}_i(([x,x],[x,x])) = [x,x]$ for each $x \in \mathbb{R}$.

Lemma 8 *For each $i \in \{1,2\}$, for each coherent pair $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}$ and each $\ell \in [0,1]$,*

$$\boldsymbol{belt}_i(\mathscr{C}onv(\ell, \mathscr{X}, \mathscr{Y})) \supset \ell\, \boldsymbol{belt}_i(\mathscr{X}) + (1-\ell)\, \boldsymbol{belt}_i(\mathscr{Y}).$$

Moreover, $\mathscr{S}qu_i$ is coherently concave for each $i \in \{1,2\}$.

Proof Since $\mathscr{X}$ and $\mathscr{Y}$ are coherent, define $\boldsymbol{z} := \boldsymbol{x}^{\mathrm{B}} = \boldsymbol{y}^{\mathrm{B}}$. The convexity of $\gamma_i(\cdot, \underline{z}^{\mathrm{B}}, p_{\boldsymbol{z}})$ and the concavity of $\sigma_i(\cdot, \overline{z}^{\mathrm{B}}, p_{\boldsymbol{z}})$ yield the following pair of inequalities:

$$\begin{aligned} \gamma_i(\ell \underline{x}^{\mathrm{C}} + (1-\ell)\underline{y}^{\mathrm{C}}, \underline{z}^{\mathrm{B}}, p_{\boldsymbol{z}}) &\le \ell\, \gamma_i(\underline{x}^{\mathrm{C}}, \underline{z}^{\mathrm{B}}, p_{\boldsymbol{z}}) + (1-\ell)\, \gamma_i(\underline{y}^{\mathrm{C}}, \underline{z}^{\mathrm{B}}, p_{\boldsymbol{z}}), \\ \sigma_i(\ell \overline{x}^{\mathrm{C}} + (1-\ell)\overline{y}^{\mathrm{C}}, \overline{z}^{\mathrm{B}}, p_{\boldsymbol{z}}) &\ge \ell\, \sigma_i(\overline{x}^{\mathrm{C}}, \overline{z}^{\mathrm{B}}, p_{\boldsymbol{z}}) + (1-\ell)\, \sigma_i(\overline{y}^{\mathrm{C}}, \overline{z}^{\mathrm{B}}, p_{\boldsymbol{z}}), \end{aligned}$$

which are equivalent to the required inclusion. Moreover, since $\mathscr{X}$, $\mathscr{Y}$, and ℓ were chosen arbitrarily, it follows immediately that $\mathscr{S}qu_i$ is coherently concave. □

Lemma 9 *For each $i \in \{1,2\}$, $\boldsymbol{belt}_i$ and $\mathscr{S}qu_i$ are inclusion monotonic.*

Proof Consider any $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}$ for which $\mathscr{X} \subset \mathscr{Y}$. If $\underline{x}^{\mathrm{B}} = \overline{x}^{\mathrm{B}} =: x$, then $\boldsymbol{x}^{\mathrm{B}} \cap \boldsymbol{x}^{\mathrm{C}} \neq \emptyset$ implies $x \in \boldsymbol{x}^{\mathrm{C}}$. Thus, $\mathscr{X} \subset \mathscr{Y}$ implies

$$\boldsymbol{belt}_i(\mathscr{X}) = [x,x] = \boldsymbol{x}^{\mathrm{B}} = \boldsymbol{x}^{\mathrm{B}} \cap \boldsymbol{x}^{\mathrm{C}} \subset \boldsymbol{y}^{\mathrm{B}} \cap \boldsymbol{y}^{\mathrm{C}} \subset \boldsymbol{belt}_i(\mathscr{Y}),$$

as required. If $\underline{x}^{\mathrm{B}} < \overline{x}^{\mathrm{B}}$, then since $\underline{x}^{\mathrm{C}} \ge \underline{y}^{\mathrm{C}}$, $\underline{x}^{\mathrm{B}} \ge \underline{y}^{\mathrm{B}}$, and $p_{\boldsymbol{x}^{\mathrm{B}}} \le p_{\boldsymbol{y}^{\mathrm{B}}}$, Lemma 4 implies that $\gamma_i(\underline{x}^{\mathrm{C}}, \underline{x}^{\mathrm{B}}, p_{\boldsymbol{x}^{\mathrm{B}}}) \ge \gamma_i(\underline{y}^{\mathrm{C}}, \underline{y}^{\mathrm{B}}, p_{\boldsymbol{y}^{\mathrm{B}}})$. A similar argument shows that $\sigma_i(\overline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{B}}, p_{\boldsymbol{x}^{\mathrm{B}}}) \le \sigma_i(\overline{y}^{\mathrm{C}}, \overline{y}^{\mathrm{B}}, p_{\boldsymbol{y}^{\mathrm{B}}})$, and so $\boldsymbol{belt}_i(\mathscr{X}) \subset \boldsymbol{belt}_i(\mathscr{Y})$. The inclusion $\mathscr{S}qu_i(\mathscr{X}) \subset \mathscr{S}qu_i(\mathscr{Y})$ follows immediately. □

Lemma 10 *For each fixed $i \in \{1,2\}$ and $\boldsymbol{x}^{\mathrm{B}} := [\underline{x}^{\mathrm{B}}, \overline{x}^{\mathrm{B}}] \in \mathbb{IR}$, consider the interval-valued mapping $\boldsymbol{y} : (\underline{\xi}, \overline{\xi}) \mapsto \boldsymbol{belt}_i((\boldsymbol{x}^{\mathrm{B}}, [\underline{\xi}, \overline{\xi}]))$. The mappings $\underline{y}$ and $\overline{y}$ are both $\mathscr{C}^i$ on $\mathbb{R}^2$.*

Proof If $\underline{x}^{\mathrm{B}} = \overline{x}^{\mathrm{B}}$, then the mapping $\boldsymbol{belt}_i((\boldsymbol{x}^{\mathrm{B}}, \cdot))$ is constant, and is therefore $\mathscr{C}^i$. Otherwise, if $\underline{x}^{\mathrm{B}} < \overline{x}^{\mathrm{B}}$, then the required result follows immediately from Lemma 3 and Definition 11. □

Remark 6 It follows from the above definitions and lemmata that, for each $i \in \{1,2\}$, $\mathscr{S}qu_i$ is a relaxation function for the identity mapping on $\mathbb{R}^n$.

3.3 Relaxing multiplication

Let the value of $i \in \{1,2\}$ be fixed throughout this section. Ultimately, setting $i = 1$ will yield $\mathscr{C}^1$ relaxations. Setting $i = 2$ instead will yield $\mathscr{C}^2$ relaxations, but will place stricter requirements on the UIFs considered, as demanded by Assumption 1.

The following definition replaces Definition 22 in Appendix A.2; it will be shown in Section 4 that this replacement weakens McCormick's classical multiplication operation to yield an alternative that is $\mathscr{C}^i$, while maintaining $(1,2)$-convergence. This modified multiplication operation depends on i, but this dependence will not be reflected in its "$\mathscr{X}\mathscr{Y}$" notation.

Definition 13 Define a multiplication operation $\times_i : \mathbb{MR}^2_{\mathrm{prop}} \to \mathbb{MR}$ so that, for each $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}$,

$$\times_i(\mathscr{X}, \mathscr{Y}) \equiv \mathscr{X}\mathscr{Y} := \mathscr{S}qu_i((\boldsymbol{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{B}}, \boldsymbol{z})),$$

where $\boldsymbol{z} \equiv [\underline{z}, \overline{z}] \in \mathbb{IR}$ is defined so that:

$$\underline{z} := \nu_i\left(\underline{(\underline{y}^{\mathrm{B}}\boldsymbol{x}^{\mathrm{C}})} + \underline{(\underline{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{C}})} - \underline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \underline{(\overline{y}^{\mathrm{B}}\boldsymbol{x}^{\mathrm{C}})} + \underline{(\overline{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{C}})} - \overline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}},\, p_{\boldsymbol{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{B}}}\right),$$
$$\overline{z} := \lambda_i\left(\overline{(\underline{y}^{\mathrm{B}}\boldsymbol{x}^{\mathrm{C}})} + \overline{(\overline{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{C}})} - \overline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \overline{(\overline{y}^{\mathrm{B}}\boldsymbol{x}^{\mathrm{C}})} + \overline{(\underline{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{C}})} - \underline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}},\, p_{\boldsymbol{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{B}}}\right).$$

3.4 Restrictions to proper McCormick objects

Again, let the value of $i \in \{1,2\}$ be fixed throughout this section. The following result shows that the codomains of $+ : \mathbb{MR}^2_{\mathrm{prop}} \to \mathbb{MR}$ (cf. Definition 21) and $\times_i : \mathbb{MR}^2_{\mathrm{prop}} \to \mathbb{MR}$ may be restricted to $\mathbb{MR}_{\mathrm{prop}}$ without loss of generality.

Proposition 7 *Consider any $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}_{\mathrm{prop}}$, and define $\mathscr{S} := \mathscr{X} + \mathscr{Y}$ and $\mathscr{P} := \mathscr{X}\mathscr{Y}$ for some $i \in \{1,2\}$. Then, $\mathscr{S}, \mathscr{P} \in \mathbb{MR}_{\mathrm{prop}}$.*

Proof Firstly, to show that $\mathscr{S} \in \mathbb{MR}_{\mathrm{prop}}$, observe that

$$\boldsymbol{s}^{\mathrm{C}} = [\underline{s}^{\mathrm{C}}, \overline{s}^{\mathrm{C}}] = [\underline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{C}}] + [\underline{y}^{\mathrm{C}}, \overline{y}^{\mathrm{C}}] \subset [\underline{x}^{\mathrm{B}}, \overline{x}^{\mathrm{B}}] + [\underline{y}^{\mathrm{B}}, \overline{y}^{\mathrm{B}}] = \boldsymbol{x}^{\mathrm{B}} + \boldsymbol{y}^{\mathrm{B}} = \boldsymbol{s}^{\mathrm{B}}.$$

Secondly, $\mathscr{P} = \mathscr{S}qu_i((\boldsymbol{p}^{\mathrm{B}}, \boldsymbol{z}))$, with $\boldsymbol{z} \in \mathbb{IR}$ given as in Definition 13. Define $\boldsymbol{v} := \boldsymbol{x}^{\mathrm{B}} \cap \boldsymbol{x}^{\mathrm{C}}$ and $\boldsymbol{w} := \boldsymbol{y}^{\mathrm{B}} \cap \boldsymbol{y}^{\mathrm{C}}$. Since $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}_{\mathrm{prop}}$, it follows that $\boldsymbol{v} = \boldsymbol{x}^{\mathrm{C}}$ and $\boldsymbol{w} = \boldsymbol{y}^{\mathrm{C}}$. Now, making use of Lemma 7, it follows that

$$\underline{z} \leq \max\left(\underline{(\underline{y}^{\mathrm{B}}\boldsymbol{v})} + \underline{(\underline{x}^{\mathrm{B}}\boldsymbol{w})} - \underline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \underline{(\overline{y}^{\mathrm{B}}\boldsymbol{v})} + \underline{(\overline{x}^{\mathrm{B}}\boldsymbol{w})} - \overline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}}\right)$$

$$\text{and} \qquad \overline{z} \geq \min\left(\overline{(\underline{y}^{\mathrm{B}}\boldsymbol{v})} + \overline{(\overline{x}^{\mathrm{B}}\boldsymbol{w})} - \overline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \overline{(\overline{y}^{\mathrm{B}}\boldsymbol{v})} + \overline{(\underline{x}^{\mathrm{B}}\boldsymbol{w})} - \underline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}}\right).$$

Defining

$$\underline{q} := \max\left(\underline{(\underline{y}^{\mathrm{B}}\boldsymbol{v})} + \underline{(\underline{x}^{\mathrm{B}}\boldsymbol{w})} - \underline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \underline{(\overline{y}^{\mathrm{B}}\boldsymbol{v})} + \underline{(\overline{x}^{\mathrm{B}}\boldsymbol{w})} - \overline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}}\right),$$

$$\text{and}\qquad \overline{q} := \min\left(\overline{(\underline{y}^{\mathrm{B}}\boldsymbol{v})} + \overline{(\overline{x}^{\mathrm{B}}\boldsymbol{w})} - \overline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \overline{(\overline{y}^{\mathrm{B}}\boldsymbol{v})} + \overline{(\underline{x}^{\mathrm{B}}\boldsymbol{w})} - \underline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}}\right),$$

it is argued on [33, Page 69] that $[\underline{q},\overline{q}] \cap \boldsymbol{p}^{\mathrm{B}} \neq \emptyset$. Thus,

$$\boldsymbol{z} \cap \boldsymbol{p}^{\mathrm{B}} \supset [\underline{q},\overline{q}] \cap \boldsymbol{p}^{\mathrm{B}} \neq \emptyset.$$

This shows that $(\boldsymbol{p}^{\mathrm{B}}, \boldsymbol{z}) \in \mathbb{MR}$, which implies that $\mathscr{P} = \mathscr{S}qu_i((\boldsymbol{p}^{\mathrm{B}}, \boldsymbol{z})) \in \mathbb{MR}_{\mathrm{prop}}$. □

The following result considers UIFs in the same manner as the above result, and shows that the codomains of their McCormick analogs may be restricted to $\mathbb{MR}_{\mathrm{prop}}$ without loss of generality.

Proposition 8 *Consider a UIF $u : B \subset \mathbb{R} \to \mathbb{R}$ that satisfies Assumption 1. With $\mathscr{U}$ described by Definition 23, $\mathscr{U}(\mathscr{X}) \in \mathbb{MR}_{\mathrm{prop}}$ for each $\mathscr{X} \in \mathbb{M}B_{\mathrm{prop}}$.*

Proof By construction, $\zeta_u^{\min}(\boldsymbol{x}^{\mathrm{B}}) \in \boldsymbol{x}^{\mathrm{B}}$ and $\zeta_u^{\max}(\boldsymbol{x}^{\mathrm{B}}) \in \boldsymbol{x}^{\mathrm{B}}$. Moreover, since $\mathscr{X} \in \mathbb{M}B_{\mathrm{prop}}$, $\boldsymbol{x}^{\mathrm{C}} \subset \boldsymbol{x}^{\mathrm{B}}$. It follows that $\mathrm{mid}(\zeta_u^{\min}(\boldsymbol{x}^{\mathrm{B}}), \underline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{C}}) \in \boldsymbol{x}^{\mathrm{B}}$ and $\mathrm{mid}(\zeta_u^{\max}(\boldsymbol{x}^{\mathrm{B}}), \underline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{C}}) \in \boldsymbol{x}^{\mathrm{B}}$. It therefore follows from the bounds on u^{cv} and u^{cc} in Assumption 1 that $[\underline{u}^{\mathrm{C}}(\mathscr{X}), \overline{u}^{\mathrm{C}}(\mathscr{X})] \subset \tilde{\boldsymbol{u}}(\boldsymbol{x}^{\mathrm{B}})$, which implies that $\mathscr{U}(\mathscr{X}) \in \mathbb{MR}_{\mathrm{prop}}$. □

4 Main results

The following definition is a variant of Definition 7. Recall that all UIFs listed in Table 2 satisfy Assumption 1 for each $i \in \{1,2\}$, and that all of these functions except for the absolute-value function satisfy Assumption 2.

Definition 14 Given some $i^* \in \{1,2\}$ and a factorable function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$ whose composed UIFs each satisfy Assumption 1 with $i := i^*$, a *natural $\mathscr{C}^{i^*}$ McCormick extension* $\mathscr{F} : \mathbb{M}B_{\mathrm{prop}} \to \mathbb{MR}^m$ of f is defined by replacing each addition operation in the construction of f with its McCormick counterpart described in Definition 21 in Appendix A.2, each multiplication operation with its counterpart in Definition 13 with $i := i^*$, and each UIF with its counterpart in Definition 23 in Appendix A.2.

Define an *unconstrained $\mathscr{C}^{i^*}$ McCormick extension* of f as $\mathscr{F}_{\mathrm{unc}} := \mathscr{F} \circ \mathscr{S}qu_{i^*} : \mathbb{M}B \to \mathbb{MR}^m$.

The following theorem is the main theorem of this article. This theorem shows that the convex/concave relaxations obtained from a natural $\mathscr{C}^i$ McCormick extension of a given factorable function according to Proposition 4 are indeed $\mathscr{C}^i$, and that they satisfy the various useful properties of McCormick's original relaxation method.

Theorem 2 *Given some $i^* \in \{1,2\}$ and a factorable function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$ whose composed UIFs each satisfy Assumption 1 with $i := i^*$, there are no domain violations in the construction of a natural $\mathscr{C}^{i^*}$ McCormick extension $\mathscr{F} : \mathbb{M}B_{\mathrm{prop}} \to \mathbb{MR}^m$ of f on B. The function $\mathscr{F}$ is a relaxation function for f on B. Additionally, if each UIF describing f satisfies Assumption 2, then $\mathscr{F}$ is $(1,2)$-convergent.*

Moreover, if $m = 1$, in which case f is scalar-valued, then the convex/concave relaxations $\phi_{f,\boldsymbol{x}}, \psi_{f,\boldsymbol{x}}$ defined by Proposition 4 in terms of $\mathscr{F}$ for each $\boldsymbol{x} \in \mathbb{I}B$ are each $\mathscr{C}^{i^*}$ on $\boldsymbol{x}$.

Proof Since f is factorable, it has a well-defined natural interval extension. Thus, Proposition 7, Proposition 8, and Assumption 1 imply that there are no domain violations in the construction of $\mathscr{F}$. The remaining claims of the theorem are proved separately as Theorems 5, 6, and 7 in Appendix C. □

Roughly, an unconstrained $\mathscr{C}^{i^*}$ McCormick extension of a function $f : B \subset \mathbb{R}^n \to \mathbb{R}$ yields convex/concave relaxations of f that are weaker than those described by a natural $\mathscr{C}^{i^*}$ McCormick extension, yet are well-defined on all of $\mathbb{R}^n$ rather than particular interval subsets, and satisfy the following corollary. As a result, natural $\mathscr{C}^{i^*}$ McCormick extensions are preferable to unconstrained $\mathscr{C}^{i^*}$ McCormick extensions in general, since the former generate tighter relaxations. Unconstrained $\mathscr{C}^{i^*}$ McCormick extensions are useful in two particular situations: firstly, if the problem $\min_{z \in \boldsymbol{x}} \phi_{f,\boldsymbol{x}}(z)$ is solved using a constrained convex optimization method that visits infeasible points, and secondly, if generalized McCormick relaxations [37,42] are employed in a manner that permits inputs $\mathscr{X} \equiv (\boldsymbol{x}^{\mathrm{B}}, \boldsymbol{x}^{\mathrm{C}})$ for which $\boldsymbol{x}^{\mathrm{C}} \nsubseteq \boldsymbol{x}^{\mathrm{B}}$.

Corollary 2 *Given some $i^* \in \{1,2\}$ and a factorable function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$ whose composed UIFs each satisfy Assumption 1 with $i := i^*$, an unconstrained $\mathscr{C}^{i^*}$ McCormick extension $\mathscr{F}_{\mathrm{unc}} : \mathbb{M}B \to \mathbb{M}\mathbb{R}^m$ of f is a relaxation function for f. Additionally, if each UIF describing f satisfies Assumption 2, then $\mathscr{F}_{\mathrm{unc}}$ is $(1,2)$-convergent.*

Moreover, if $m = 1$, in which case f is scalar-valued, then the convex/concave relaxations $\phi_{f,\boldsymbol{x}}, \psi_{f,\boldsymbol{x}}$ defined by Proposition 4 in terms of $\mathscr{F}_{\mathrm{unc}}$ for each $\boldsymbol{x} \in \mathbb{I}B$ are each $\mathscr{C}^{i^*}$ on $\mathbb{R}^n$.

4.1 Gradient propagation

Using the obtained differentiability results, the standard forward or reverse modes of automatic differentiation [14,26] can be used to evaluate derivatives of the convex/concave relaxations obtained for natural or unconstrained $\mathscr{C}^i$ McCormick extensions, provided that gradients can be evaluated for the composed addition, multiplication, and UI operations. The obtained gradients are subgradients of the corresponding relaxations.

To evaluate derivatives for $\mathscr{C}^i$ McCormick extensions, addition and UI composition can be treated exactly as in Proposition 2.9 and Theorem 3.2 in [24], with all subgradients mentioned in these results replaced by the corresponding gradients. For multiplication, repeated application of the chain rule to Definition 13 yields the following, which makes use of the partial derivatives of ν_i and λ_i provided by Proposition 6.

Theorem 3 *Consider functions $f, g : D \subset \mathbb{R}^n \to \mathbb{R}$, and relaxation functions $\mathscr{F}, \mathscr{G} : \mathbb{M}D$ (or $\mathbb{M}D_{\mathrm{prop}}) \to \mathbb{M}\mathbb{R}$ for f and g on D, such that the mappings $\mathscr{X} \mapsto \boldsymbol{f}^{\mathrm{B}}(\mathscr{X})$ and $\mathscr{X} \mapsto \boldsymbol{g}^{\mathrm{B}}(\mathscr{X})$ are each independent of their $\boldsymbol{x}^{\mathrm{C}}$ argument. Consider the product function $h : D \to \mathbb{R} : z \mapsto f(z)\,g(z)$, and the corresponding product relaxation function $\mathscr{H} : \mathbb{M}D$ (or $\mathbb{M}D_{\mathrm{prop}}) \to \mathbb{M}\mathbb{R} : \mathscr{X} \mapsto \times_i(\mathscr{F}(\mathscr{X}), \mathscr{G}(\mathscr{X}))$ with $i \in \{1,2\}$. As in Proposition 4, for some fixed $\boldsymbol{y} \in \mathbb{I}D$, construct the convex/concave relaxations $\phi_{h,\boldsymbol{y}} : z \mapsto \underline{h}^{\mathrm{C}}((\boldsymbol{y}, [z,z]))$ and $\psi_{h,\boldsymbol{y}} : z \mapsto \overline{h}^{\mathrm{C}}((\boldsymbol{y}, [z,z]))$ of h on $\boldsymbol{y}$, and construct the analogous relaxations $\phi_{f,\boldsymbol{y}}/\psi_{f,\boldsymbol{y}}$ of f and $\phi_{g,\boldsymbol{y}}/\psi_{g,\boldsymbol{y}}$ of g. Gradients of $\phi_{h,\boldsymbol{y}}$ and $\psi_{h,\boldsymbol{y}}$ at some particular $x \in \boldsymbol{y}$ may be computed as follows, with $\mathscr{Y} := (\boldsymbol{y}, [x,x]) \in \mathbb{M}D$ (or $\mathbb{M}D_{\mathrm{prop}}$). For notational simplicity, the $\mathscr{Y}$ arguments of $\boldsymbol{f}^{\mathrm{B}}(\mathscr{Y}) \equiv [\underline{f}^{\mathrm{B}}(\mathscr{Y}), \overline{f}^{\mathrm{B}}(\mathscr{Y})]$, $\boldsymbol{g}^{\mathrm{B}}(\mathscr{Y}) \equiv [\underline{g}^{\mathrm{B}}(\mathscr{Y}), \overline{g}^{\mathrm{B}}(\mathscr{Y})]$, and $\boldsymbol{h}^{\mathrm{B}}(\mathscr{Y}) \equiv [\underline{h}^{\mathrm{B}}(\mathscr{Y}), \overline{h}^{\mathrm{B}}(\mathscr{Y})]$ will be omitted.*

If $\overline{h}^{\mathrm{B}} = \underline{h}^{\mathrm{B}}$*, then* $\nabla \phi_{h,\boldsymbol{y}}(x) = \nabla \psi_{h,\boldsymbol{y}}(x) = 0$*. Otherwise, if* $\overline{h}^{\mathrm{B}} > \underline{h}^{\mathrm{B}}$*, then define intermediate scalar quantities:*

$$\begin{aligned}
n_1(x) &:= \underline{(\underline{g}^{\mathrm{B}}\,\boldsymbol{f}^{\mathrm{C}}(\mathscr{Y}))} + \underline{(\underline{f}^{\mathrm{B}}\,\boldsymbol{g}^{\mathrm{C}}(\mathscr{Y}))} - \underline{f}^{\mathrm{B}}\underline{g}^{\mathrm{B}},\\
n_2(x) &:= \underline{(\overline{g}^{\mathrm{B}}\,\boldsymbol{f}^{\mathrm{C}}(\mathscr{Y}))} + \underline{(\overline{f}^{\mathrm{B}}\,\boldsymbol{g}^{\mathrm{C}}(\mathscr{Y}))} - \overline{f}^{\mathrm{B}}\overline{g}^{\mathrm{B}},\\
n_3(x) &:= \overline{(\underline{g}^{\mathrm{B}}\,\boldsymbol{f}^{\mathrm{C}}(\mathscr{Y}))} + \overline{(\overline{f}^{\mathrm{B}}\,\boldsymbol{g}^{\mathrm{C}}(\mathscr{Y}))} - \overline{f}^{\mathrm{B}}\underline{g}^{\mathrm{B}},\\
n_4(x) &:= \overline{(\overline{g}^{\mathrm{B}}\,\boldsymbol{f}^{\mathrm{C}}(\mathscr{Y}))} + \overline{(\underline{f}^{\mathrm{B}}\,\boldsymbol{g}^{\mathrm{C}}(\mathscr{Y}))} - \underline{f}^{\mathrm{B}}\overline{g}^{\mathrm{B}},
\end{aligned}$$

If $\overline{f}^{\mathrm{B}} = \underline{f}^{\mathrm{B}}$*, then define intermediate scalar quantities* $b_1(x) = b_2(x) = b_3(x) = b_4(x) := 0$*. Otherwise, if* $\overline{f}^{\mathrm{B}} > \underline{f}^{\mathrm{B}}$*, then define:*

$$b_1(x) := \begin{cases} \underline{g}^{\mathrm{B}}\,\nabla\phi_{f,\boldsymbol{y}}(x), & \text{if } \underline{g}^{\mathrm{B}} \geq 0,\\ \underline{g}^{\mathrm{B}}\,\nabla\psi_{f,\boldsymbol{y}}(x), & \text{if } \underline{g}^{\mathrm{B}} < 0,\end{cases} \qquad b_2(x) := \begin{cases} \overline{g}^{\mathrm{B}}\,\nabla\phi_{f,\boldsymbol{y}}(x), & \text{if } \overline{g}^{\mathrm{B}} \geq 0,\\ \overline{g}^{\mathrm{B}}\,\nabla\psi_{f,\boldsymbol{y}}(x), & \text{if } \overline{g}^{\mathrm{B}} < 0,\end{cases}$$

$$b_3(x) := \begin{cases} \underline{g}^{\mathrm{B}}\,\nabla\psi_{f,\boldsymbol{y}}(x), & \text{if } \underline{g}^{\mathrm{B}} \geq 0,\\ \underline{g}^{\mathrm{B}}\,\nabla\phi_{f,\boldsymbol{y}}(x), & \text{if } \underline{g}^{\mathrm{B}} < 0,\end{cases} \qquad b_4(x) := \begin{cases} \overline{g}^{\mathrm{B}}\,\nabla\psi_{f,\boldsymbol{y}}(x), & \text{if } \overline{g}^{\mathrm{B}} \geq 0,\\ \overline{g}^{\mathrm{B}}\,\nabla\phi_{f,\boldsymbol{y}}(x), & \text{if } \overline{g}^{\mathrm{B}} < 0.\end{cases}$$

If $\overline{g}^{\mathrm{B}} = \underline{g}^{\mathrm{B}}$*, then define intermediate scalar quantities* $b_5(x) = b_6(x) = b_7(x) = b_8(x) := 0$*. Otherwise, if* $\overline{g}^{\mathrm{B}} > \underline{g}^{\mathrm{B}}$*, then define:*

$$b_5(x) := \begin{cases} \underline{f}^{\mathrm{B}}\,\nabla\phi_{g,\boldsymbol{y}}(x), & \text{if } \underline{f}^{\mathrm{B}} \geq 0,\\ \underline{f}^{\mathrm{B}}\,\nabla\psi_{g,\boldsymbol{y}}(x), & \text{if } \underline{f}^{\mathrm{B}} < 0,\end{cases} \qquad b_6(x) := \begin{cases} \overline{f}^{\mathrm{B}}\,\nabla\phi_{g,\boldsymbol{y}}(x), & \text{if } \overline{f}^{\mathrm{B}} \geq 0,\\ \overline{f}^{\mathrm{B}}\,\nabla\psi_{g,\boldsymbol{y}}(x), & \text{if } \overline{f}^{\mathrm{B}} < 0,\end{cases}$$

$$b_7(x) := \begin{cases} \overline{f}^{\mathrm{B}}\,\nabla\psi_{g,\boldsymbol{y}}(x), & \text{if } \overline{f}^{\mathrm{B}} \geq 0,\\ \overline{f}^{\mathrm{B}}\,\nabla\phi_{g,\boldsymbol{y}}(x), & \text{if } \overline{f}^{\mathrm{B}} < 0,\end{cases} \qquad b_8(x) := \begin{cases} \underline{f}^{\mathrm{B}}\,\nabla\psi_{g,\boldsymbol{y}}(x), & \text{if } \underline{f}^{\mathrm{B}} \geq 0,\\ \underline{f}^{\mathrm{B}}\,\nabla\phi_{g,\boldsymbol{y}}(x), & \text{if } \underline{f}^{\mathrm{B}} < 0.\end{cases}$$

Next, define the following intermediate scalar quantities:

$$\begin{aligned}
a_1(x) &:= \frac{\partial \nu_i}{\partial x}(n_1(x), n_2(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_1(x) + \frac{\partial \nu_i}{\partial y}(n_1(x), n_2(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_2(x),\\
a_2(x) &:= \frac{\partial \nu_i}{\partial x}(n_1(x), n_2(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_5(x) + \frac{\partial \nu_i}{\partial y}(n_1(x), n_2(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_6(x),\\
a_3(x) &:= \frac{\partial \lambda_i}{\partial x}(n_3(x), n_4(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_3(x) + \frac{\partial \lambda_i}{\partial y}(n_3(x), n_4(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_4(x),\\
a_4(x) &:= \frac{\partial \lambda_i}{\partial x}(n_3(x), n_4(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_7(x) + \frac{\partial \lambda_i}{\partial y}(n_3(x), n_4(x), p_{\boldsymbol{h}^{\mathrm{B}}})\,b_8(x).
\end{aligned}$$

Then,

$$\begin{aligned}
\nabla\phi_{h,\boldsymbol{y}}(x) &= \frac{\partial \gamma_i}{\partial z}(\underline{h}^{\mathrm{C}}(\mathscr{Y}), \underline{h}^{\mathrm{B}}, p_{\boldsymbol{h}^{\mathrm{B}}})\;(a_1(x) + a_2(x)),\\
\textit{and} \qquad \nabla\psi_{h,\boldsymbol{y}}(x) &= \frac{\partial \sigma_i}{\partial z}(\overline{h}^{\mathrm{C}}(\mathscr{Y}), \overline{h}^{\mathrm{B}}, p_{\boldsymbol{h}^{\mathrm{B}}})\;(a_3(x) + a_4(x)).
\end{aligned}$$

Proof This result follows immediately from Definition 13 and the chain rule in Proposition 2. Observe that, in light of Remark 1, if a composed function is defined only at a single point, then its derivative at this point may be set to 0 without affecting the validity of this chain rule. □

When constructing unconstrained $\mathscr{C}^i$ McCormick relaxations, the following gradient propagation result can be used to handle the initial squashing operation.

Proposition 9 *For fixed $\boldsymbol{y} \in \mathbb{IR}$ and $i \in \{1,2\}$, consider the functions $\underline{s}^{\mathrm{C}}_{\boldsymbol{y}}, \overline{s}^{\mathrm{C}}_{\boldsymbol{y}} : \mathbb{R}^2 \to \mathbb{R}$ defined so that $\mathscr{S}qu_i((\boldsymbol{y},\boldsymbol{z})) = (\boldsymbol{y}, [\underline{s}^{\mathrm{C}}_{\boldsymbol{y}}(\underline{z},\overline{z}), \overline{s}^{\mathrm{C}}_{\boldsymbol{y}}(\underline{z},\overline{z})])$ for each $\boldsymbol{z} \in \mathbb{IR}$. Then*

$$\nabla \underline{s}^{\mathrm{C}}_{\boldsymbol{y}}(\underline{z},\overline{z}) = \begin{bmatrix} \frac{\partial \gamma_i}{\partial z}(\underline{z},\underline{y},p_{\boldsymbol{y}}) & 0 \end{bmatrix}, \qquad \textit{and} \qquad \nabla \overline{s}^{\mathrm{C}}_{\boldsymbol{y}}(\underline{z},\overline{z}) = \begin{bmatrix} 0 & \frac{\partial \sigma_i}{\partial z}(\overline{z},\overline{y},p_{\boldsymbol{y}}) \end{bmatrix}.$$

Proof This result follows immediately from the definition of the squashing operation. □

5 Implementation and examples

This section first discusses how to choose the parameter a_p in Definition 11 in accordance with numerical considerations. A C++ implementation of the relaxation theory in this article is then described, and is subsequently applied to various example problems for illustration.

5.1 Choosing the parameter a_p

When constructing a $\mathscr{C}^i$ McCormick extension of a function f, the parameter a_p in Definition 11 is only used if either f is described in terms of at least one product function, an unconstrained $\mathscr{C}^i$ McCormick extension is desired, or the constructions described in Remark 5 for pathological UIFs are required. If none of these circumstances apply, then there is no need to choose a_p.

Although the established $(1,2)$-convergence of $\mathscr{C}^i$ McCormick extensions is independent of a_p, larger values of a_p ultimately yield weaker relaxations $\phi_{f,\boldsymbol{x}}/\psi_{f,\boldsymbol{x}}$ when $\mathrm{wid}\,\boldsymbol{x}$ is large, making fathoming by value dominance less likely at the early stages of a branch-and-bound procedure for nonconvex optimization. On the other hand, smaller values of a_p yield relaxations that are theoretically $\mathscr{C}^i$, yet may differ (with respect to the L^2-norm) only marginally from a nondifferentiable function when $\mathrm{wid}\,\boldsymbol{x}$ is reduced.

Moreover, observe that in the results established in this article, there is no need for the same value of a_p to be used each time the function $p : \mathbb{IR} \to [0,+\infty)$ is invoked during construction of a particular $\mathscr{C}^i$ McCormick extension of a function. This notion provides a degree of freedom which can be exploited to ensure that the values of a_p employed are neither too great or too small, in accordance with the previous paragraph.

Now, it follows from Lemma 3 that, for any $\mathscr{X} \equiv (\boldsymbol{x}^{\mathrm{B}}, \boldsymbol{x}^{\mathrm{C}}) \in \mathbb{MR}_{\mathrm{prop}}$ and $i \in \{1,2\}$,

$$0 \leq \frac{\mathrm{wid}\,(\boldsymbol{belt}_i(\mathscr{X}))}{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}}} - \frac{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{C}}}{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}}} \leq \min\left\{\frac{2p_{\boldsymbol{x}^{\mathrm{B}}}}{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}}}, 1\right\} = \min\{2a_p\,\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}}, 1\}.$$

So, the belt operation increases the ratio $\frac{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{C}}}{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}}}$ by at most $(2a_p\,\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}})$. Observe that, if $\frac{\mathrm{wid}\,(\boldsymbol{belt}_i(\mathscr{X}))}{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}}} \approx \frac{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{C}}}{\mathrm{wid}\,\boldsymbol{x}^{\mathrm{B}}}$, then there is little numerical difference between the $\mathscr{C}^i$ McCormick relaxations and the classical McCormick relaxations.

In light of the above discussion, suppose that during execution of a branch-and-bound procedure, whenever any interval subdomain $\boldsymbol{x}$ is visited, the $\mathscr{C}^i$ McCormick extension of a function demands evaluation of $p_{\boldsymbol{y}^{\mathrm{B}}(\boldsymbol{x})}$, where the interval-valued function $\boldsymbol{y}^{\mathrm{B}}$ is defined by the natural interval extension of the factorable objective function. Due to inclusion monotonicity of natural interval extensions, $\mathrm{wid}\,(\boldsymbol{y}^{\mathrm{B}}(\boldsymbol{x}))$ decreases as $\mathrm{wid}\,\boldsymbol{x}$ decreases. Now, if $\boldsymbol{x}_0$

denotes the interval domain considered at the root node of the branch-and-bound procedure, the above discussion suggests setting

$$a_p \leftarrow \frac{b_p}{2\,\mathrm{wid}\,\boldsymbol{y}^{\mathrm{B}}(\boldsymbol{x}_0)} \tag{3}$$

for some constant b_p in the range $[0.01, 0.2]$. With this choice, the $\mathscr{C}^i$ McCormick extensions are not relaxed too much relative to the corresponding original natural McCormick extensions, and yet $(2a_p\,\mathrm{wid}\,\boldsymbol{y}^{\mathrm{B}}(\boldsymbol{x}))$ remains significantly greater than 0 (relative to a computer's typical numerical precision) even after several successive branches in the branch-and-bound procedure.

Lastly, note that $(2a_p\,\mathrm{wid}\,\boldsymbol{y}^{\mathrm{B}}(\boldsymbol{x})) \to 0^+$ in the limit $(\mathrm{wid}\,\boldsymbol{x}) \to 0^+$. If $(2a_p\,\mathrm{wid}\,\boldsymbol{y}^{\mathrm{B}}(\boldsymbol{x}))$ falls below some small tolerance $\varepsilon > 0$, then affine relaxations defined either by the subgradients of the classical natural McCormick extensions or the gradients of $\mathscr{C}^i$ McCormick extensions may be preferable to the McCormick extensions themselves.

5.2 Implementation

A C++ implementation of $\mathscr{C}^i$ McCormick extension evaluation was developed by modifying version 1.0 of the header library MC++ [8], to carry out the methods in this article using operator overloading. This new implementation describes McCormick objects using a template class `mc::smoothMcC<T>`, which is a modified version of the class `mc::McCormick<T>` defined by MC++. As in MC++, the templated argument `T` refers to the interval objects used by an employed interval arithmetic library. The specific modifications used to construct the class `mc::smoothMcC<T>` from `mc::McCormick<T>` are as follows.

Firstly, static member variables `_MCbp` and `_MCi` were added to the class, so as to hold the values of the parameters b_p and $i \in \{1,2\}$, respectively. These parameters can be set or retrieved using static member functions `setBp`, `getBp`, `setI`, and `getI`. Further static member functions evaluate the functions p, μ_i, $\nabla\mu_i$, γ_i, σ_i, v_i, $\frac{\partial v_i}{\partial x}$, $\frac{\partial v_i}{\partial y}$, λ_i, $\frac{\partial \lambda_i}{\partial x}$, and $\frac{\partial \lambda_i}{\partial y}$. The execution of `MCp`, the static member function evaluating p, is detailed in the next paragraph. In the following description, let `mcX` denote an arbitrary `mc::smoothMcC<T>` object representing a McCormick object $\mathscr{X}$. Member functions `squash` and `p` were added to the class, so that `mcX.squash()` replaces its calling member $\mathscr{X}$ with $\mathscr{S}qu_i(\mathscr{X})$, and so that `mcX.p()` invokes `MCp` to return the value $p_{\boldsymbol{x}^{\mathrm{B}}}$. Using these constructions, McCormick-McCormick multiplication was implemented according to Definition 13, via a `friend` function that overloads `operator*`, with gradients propagated according to Theorem 3. The relaxations described in Examples 7, 8 and 9 were implemented by modifying the overloaded operations `fabs` and `pow` appropriately, along with a squaring function `sqr` that was implemented in MC++.

To implement evaluation of p via `MCp` according to the discussion in Section 5.1, a static member `enum` variable `_apMode` was added to the `mc::smoothMcC<T>` class, to describe whether the parameters a_p should be evaluated as if the root node in a branch-and-bound process is being visited, or whether a child node is being visited instead. If `_apMode=SET_AP`, then whenever p is evaluated, the parameter a_p is evaluated in the root-node mode described in Section 5.1, and the value of a_p is pushed onto the end of a static member `std::vector<double>` named `_apList`. To handle child nodes in a branch-and-bound process, when values of a_p have already been stored in `_apList`, `_apMode` is set to `GET_AP`. In this mode, whenever p is evaluated, the appropriate value of a_p is retrieved from _apList; the appropriate component of _apList to be retrieved is tracked using a static member `std::vector<double>::const_iterator`.

Ultimately, given a user-supplied template subroutine `f` that is written as though its inputs and outputs are `doubles` or `double` arrays, the implementation described above permits natural $\mathscr{C}^i$ McCormick extensions of `f` to be evaluated using operator overloading, along with directional derivatives that are evaluated using the forward mode of automatic differentiation. To obtain unconstrained $\mathscr{C}^i$ McCormick extensions instead, the `squash` operation should first be applied to each `mc::smoothMcC<T>` input to `f`. The UIFs and operations described in Table 2 are all supported in this implementation.

5.3 Complexity analysis

Roughly, denote the computational cost of evaluating a factorable function $f : X \subset \mathbb{R}^n \to \mathbb{R}$ using its factored representation as "cost(f)". Observe that, when constructing the convex or concave relaxation suggested by a natural $\mathscr{C}^2$ McCormick extension for f, each addition, multiplication, and UIF in the factored representation of f is replaced with its $\mathscr{C}^2$ McCormick counterpart. Thus, there exists $\gamma_c > 0$ for which the computational cost of evaluating a $\mathscr{C}^2$ convex or concave relaxation of f is no greater than $\gamma_c \operatorname{cost}(f)$. The parameter γ_c is independent of f, but depends on the library of UIFs considered.

Similarly, using standard complexity results for automatic differentiation [14], it follows that there exist similar library-dependent constants $\gamma_a, \gamma_t > 0$, satisfying the following claim. If the reverse mode of automatic differentiation is used to evaluate a subgradient of such a relaxation, then the cost of doing so is bounded above by $\gamma_a \operatorname{cost}(f)$; if the forward mode is used instead, then the cost of evaluating this subgradient is bounded above by $n\gamma_t \operatorname{cost}(f)$, where n denotes the domain dimension of f.

5.4 Examples

In this section, the implementation of $\mathscr{C}^2$ McCormick relaxation described in Section 5.2 is applied to various example problems for illustration.

Example 3 To illustrate the modified multiplication rule provided by Definition 13, consider the function $f : \mathbb{R}^2 \to \mathbb{R} : (x,y) \mapsto y(x^2 - 1)$, which is plotted in Figure 1(a). The function f is (real-)analytic but nonconvex on $\boldsymbol{z} := [-4,4]^2 \subset \mathbb{R}^2$.

Using MC++ [8], the classical McCormick convex relaxation of f was constructed on $\boldsymbol{z}$, and is plotted in Figure 1(b). This relaxation is not differentiable everywhere; this nondifferentiability is introduced via McCormick's rule for relaxing the product of terms whose signs change on the interval of interest. A natural $\mathscr{C}^2$ McCormick relaxation of f on $\boldsymbol{z}$ was constructed using the implementation described in Sections 5.1 and 5.2, with $b_p := 0.2$; this relaxation is plotted in Figure 1(c). Observe that this relaxation is visibly differentiable (and is, in fact, $\mathscr{C}^2$), but is otherwise qualitatively similar to the classical McCormick relaxation. The classical McCormick relaxation dominates its $\mathscr{C}^2$ counterpart on $\boldsymbol{z}$.

For comparison, the αBB relaxation of f on $\boldsymbol{z}$ with a nonuniform diagonal shift matrix that minimizes maximum separation distance [1] was computed directly to be:

$$f^\alpha : (x,y) \mapsto f(x,y) + 8(x^2 - 16) + 4(y^2 - 16),$$

and is plotted in Figure 1(d). The obtained αBB relaxation is analytic, and has a minimum at $(x^*, y^*) := (0, 0.125)$. Observe that $f^\alpha(x^*, y^*) = -192.0625$, which is less than the lower

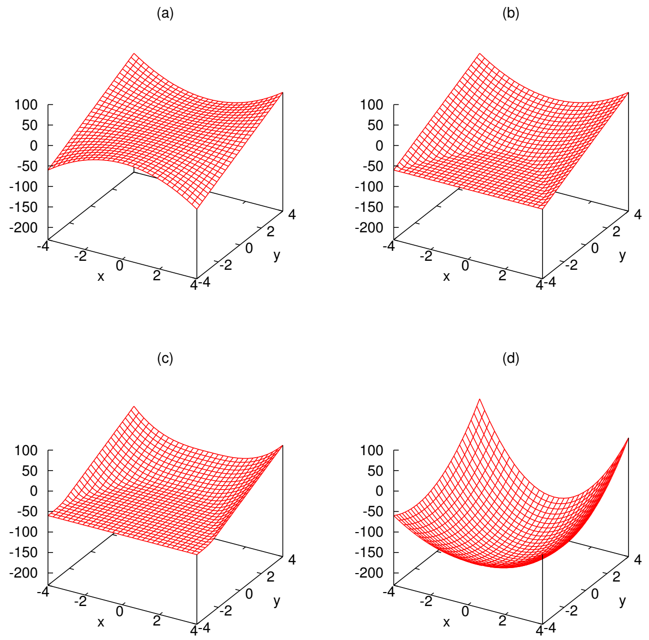

Fig. 1 The function $f:(x,y)\mapsto y(x^2-1)$ and its convex relaxations on $[-4,4]^2$: (a) the function f, (b) the classical McCormick relaxation of f, (c) a $\mathscr{C}^2$ McCormick relaxation of f, and (d) the αBB relaxation of f that minimizes maximum separation distance.

bound $\underline{\tilde{f}}(\boldsymbol{z})=-60$ provided by the natural interval extension of f on $\boldsymbol{z}$. This interval lower bound coincides with $\min_{(x,y)\in\boldsymbol{z}} f(x,y)$, and is dominated on $\boldsymbol{z}$ by both the constructed classical McCormick relaxation and the constructed $\mathscr{C}^2$ McCormick relaxation.

Example 4 To illustrate the handling of the absolute-value function according to Example 8, consider the function

$$g:\mathbb{R}^2\to\mathbb{R}:(x,y)\mapsto|x+1|+|x-1|-|x+y-1|-|x-y+1|, \tag{4}$$

which is plotted in Figure 2(a). The function g is piecewise affine, and is nonconvex on $\boldsymbol{z}:=[-2,2]^2\subset\mathbb{R}^2$.

As in the previous example, the classical McCormick convex relaxation of g on $\boldsymbol{z}$ was constructed using MC++, and is plotted in Figure 2(b); this relaxation is readily verified to be piecewise affine. The $\mathscr{C}^2$ McCormick relaxation of g on $\boldsymbol{z}$ was evaluated using the implementation described in Section 5.2, and is plotted in Figure 2(c). Since there does not exist a scheme of estimators satisfying Assumption 2 for the absolute-value function, the generated McCormick and $\mathscr{C}^2$ McCormick relaxations are not guaranteed to be pointwise convergent of order 2.

Example 5 To illustrate the handling of the squaring function $z\mapsto z^2$ according to Example 7, consider the function

$$h:\mathbb{R}^2\to\mathbb{R}:(x,y)\mapsto(xy-1)^2, \tag{5}$$

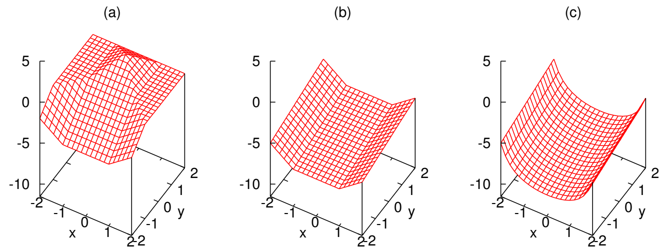

Fig. 2 The function g described in (4) and its convex relaxations on $[-2,2]^2$: (a) the function g, (b) the classical McCormick relaxation of g, and (c) a $\mathscr{C}^2$ McCormick relaxation of g.

which is plotted in Figure 3(a). The function h is analytic and nonconvex on $\boldsymbol{z} := [-2,2]^2 \subset \mathbb{R}^2$.

The classical McCormick convex relaxation h^{cv} of h on $\boldsymbol{z}$ was evaluated using MC++, along with a subgradient at each point. This relaxation h^{cv} is plotted in Figure 3(b), and x-components of the evaluated subgradients of h^{cv} are plotted in Figure 3(c). As a function of (x,y), the evaluated subgradient is evidently not differentiable everywhere; it follows that h^{cv} is not twice-differentiable, let alone $\mathscr{C}^2$. This example illustrates that, even though the squaring function is convex and smooth, considering the squaring UIF as its own convex relaxation can yield failures of twice-continuous differentiability. This observation motivates Assumption 1 and Example 7.

A natural $\mathscr{C}^2$ McCormick relaxation $\tilde{h}^{\mathrm{cv}}$ of h on $\boldsymbol{z}$ was constructed using the implementation described in Sections 5.1 and 5.2, with $b_p := 0.2$; this relaxation is plotted in Figure 1(d). Gradients of $\tilde{h}^{\mathrm{cv}}$ were also evaluated using the described implementation; the partial derivatives $\frac{\partial \tilde{h}^{\mathrm{cv}}}{\partial x}$ and $\frac{\partial \tilde{h}^{\mathrm{cv}}}{\partial y}$ are plotted in Figures 3(e) and 3(f), respectively. These partial derivatives appear to be differentiable, and are indeed $\mathscr{C}^1$.

Example 6 This example illustrates the second-order pointwise convergence of the $\mathscr{C}^2$ McCormick relaxations presented in this article. As in [6, Example 7], consider the function

$$f : \mathbb{R}_+ \to \mathbb{R} : x \mapsto (x - x^2)(\log x + e^{-x})$$

on intervals of the form $[0.5-\varepsilon, 0.5+\varepsilon]$ for $\varepsilon \in (0, 0.2]$. The function f is plotted in Figure 4, together with a series of $\mathscr{C}^2$ relaxations $\psi_{\boldsymbol{x}(\varepsilon)}$ of f constructed using the implementation described in Sections 5.1 and 5.2, on intervals $\boldsymbol{x} \in \{[0.5-\varepsilon, 0.5+\varepsilon] : \varepsilon = 0.4(2^k), k \in \{1,\ldots,20\}\}$, with the parameters in (3) set to $\boldsymbol{x}_0 := [0.3, 0.7]$ and $b_p := 0.2$.

For the considered values of ε, Figure 4(b) plots $\sup_{x \in \boldsymbol{x}(\varepsilon)}(f(x) - \psi_{\boldsymbol{x}(\varepsilon)}(x))$ against $\mathrm{wid}\,\boldsymbol{x}(\varepsilon)$ on a logarithmic scale; the slope of this plot suggests second-order pointwise convergence of the convex relaxation $\psi_{\boldsymbol{x}}(\varepsilon)$ to f as $\varepsilon \to 0^+$.

6 Conclusions

A variant of McCormick's relaxation scheme has been presented, which produces $\mathscr{C}^2$ convex and concave relaxations of a provided factorable function, while retaining the computational benefits of McCormick's method. Gradients are readily evaluated for the provided

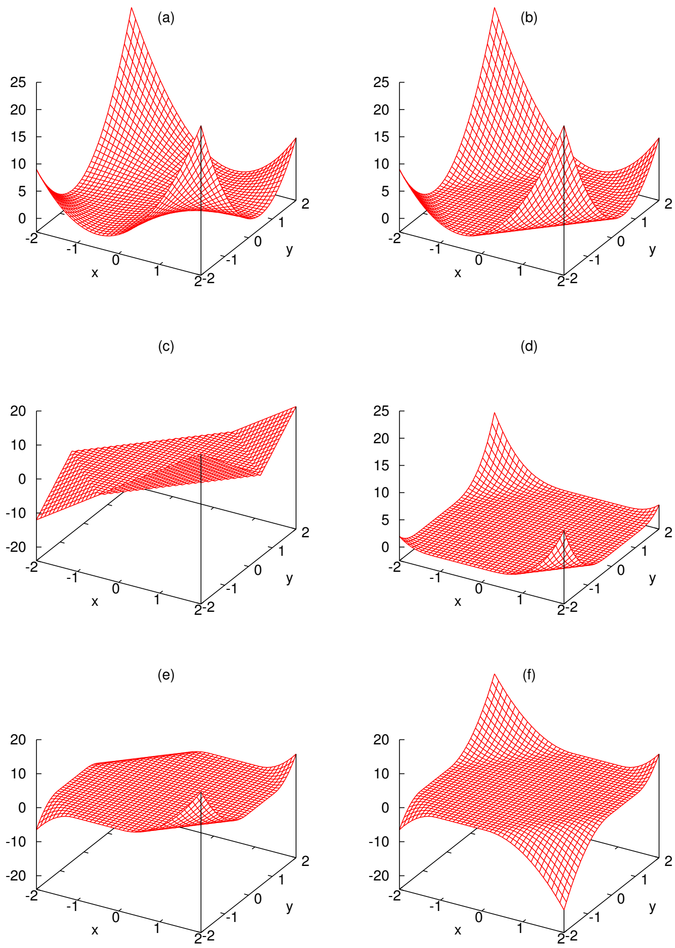

Fig. 3 The function h described in (5) and its convex relaxations and associated subgradients on $[-2,2]^2$: (a) the function h, (b) the classical McCormick relaxation h^{cv} of h, (c) the x-component of some subgradient of h^{cv}, (d) a $\mathscr{C}^2$ McCormick relaxation $\tilde{h}^{\mathrm{cv}}$ of h, (e) the partial derivative $\frac{\partial \tilde{h}^{\mathrm{cv}}}{\partial x}$, and (f) the partial derivative $\frac{\partial \tilde{h}^{\mathrm{cv}}}{\partial y}$.

relaxations using standard automatic differentiation methods. As an avenue for possible future work, we expect that the methods in this article are compatible with an established scheme for reverse propagation of McCormick relaxations [42], and could yield a scheme for constructing $\mathscr{C}^2$ relaxations for implicit functions.

As an open problem, observe that the methods in this article do not extend immediately to the multivariate relaxations described by Tsoukalas and Mitsos [39]. Such an extension would be desirable, since the multivariate product relaxations are tighter than the classical McCormick product relaxation described in Definition 22.

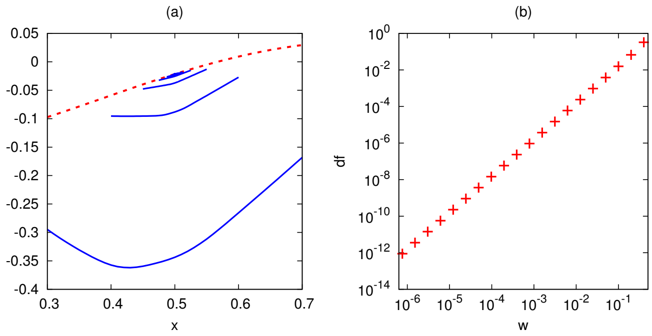

Fig. 4 (a) The function f described in Example 6 (dashed) and its $\mathscr{C}^2$ convex relaxations $\psi_{\boldsymbol{x}(\varepsilon)}$ of f on intervals $\boldsymbol{x}(\varepsilon) := [0.5-\varepsilon, 0.5+\varepsilon]$ for various $\varepsilon > 0$ (solid), and (b) a plot of $df := \sup_{x\in\boldsymbol{x}(\varepsilon)}(f(x) - \psi_{\boldsymbol{x}(\varepsilon)}(x))$ vs. $w := \mathrm{wid}\,\boldsymbol{x}(\varepsilon) = 2\varepsilon$.

Acknowledgements The authors would like to thank Achim Wechsung and Spencer Schaber for several helpful discussions.

Appendices

A Established definitions and results

A.1 Interval analysis

This appendix summarizes relevant standard definitions and results concerning interval analysis.

Definition 15 (from [2]) For each $c \in \mathbb{R}$, define scalar-interval multiplication so that for each $\boldsymbol{x} \in \mathbb{IR}$,

$$c\boldsymbol{x} := \begin{cases} [c\underline{x}, c\overline{x}], & \text{if } c \geq 0, \\ [c\overline{x}, c\underline{x}], & \text{if } c < 0. \end{cases}$$

Setting $c \leftarrow -1$ corresponds to a *negative* operation. Define interval operations $+, -, \times : \mathbb{IR} \times \mathbb{IR} \to \mathbb{IR}$ such that

$$+(\boldsymbol{x}, \boldsymbol{y}) \equiv \boldsymbol{x} + \boldsymbol{y} := [\underline{x} + \underline{y}, \overline{x} + \overline{y}], \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}.$$
$$-(\boldsymbol{x}, \boldsymbol{y}) \equiv \boldsymbol{x} - \boldsymbol{y} := [\underline{x} - \overline{y}, \overline{x} - \underline{y}], \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}.$$
$$\times(\boldsymbol{x}, \boldsymbol{y}) \equiv \boldsymbol{x}\boldsymbol{y} := [\min\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}, \max\{\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}\}], \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}.$$

It is readily verified that for any interval operation $\circ \in \{+, -, \times\}$, $\boldsymbol{x} \circ \boldsymbol{y} = \{v \circ w : v \in \boldsymbol{x}, w \in \boldsymbol{y}\}$ for any intervals $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}$, and $[v, v] \circ [w, w] = [v \circ w, v \circ w]$ for any $v, w \in \mathbb{R}$.

Lemma 11 *Consider an interval $\boldsymbol{x} \in \mathbb{IR}$ and scalars $a, b \in \mathbb{R}$ for which $a \leq b$. If $\underline{x} \geq 0$, then $\underline{(a\boldsymbol{x})} \leq \underline{(b\boldsymbol{x})}$. If $\overline{x} \leq 0$, then $\underline{(a\boldsymbol{x})} \geq \underline{(b\boldsymbol{x})}$. Similarly, if $\underline{x} \geq 0$, then $\overline{(a\boldsymbol{x})} \leq \overline{(b\boldsymbol{x})}$. If $\overline{x} \leq 0$, then $\overline{(a\boldsymbol{x})} \geq \overline{(b\boldsymbol{x})}$.*

Proof For any $c \in \mathbb{R}$,

$$\underline{(c\boldsymbol{x})} = \begin{cases} c\underline{x}, & \text{if } c \geq 0, \\ c\overline{x}, & \text{if } c < 0 \end{cases} = \overline{x}\min\{c, 0\} + \underline{x}\max\{c, 0\}. \tag{6}$$

If $\underline{x} \geq 0$, then each term in the final expression above is evidently increasing with respect to c, yielding the first required inequality. If, instead, $\overline{x} \leq 0$, then each term in the final expression is decreasing with respect to c, which yields the second required inequality. Next, for any $c \in \mathbb{R}$,

$$\overline{(c\boldsymbol{x})} = \begin{cases} c\overline{x}, & \text{if } c \geq 0, \\ c\underline{x}, & \text{if } c < 0 \end{cases} = \underline{x}\min\{c, 0\} + \overline{x}\max\{c, 0\}. \tag{7}$$

Using this result, a similar argument to the previous case yields the remaining inequalities. $\square$

Definition 16 (from [25]) Consider a nonempty set $B \subset \mathbb{R}^n$. An interval-valued function $\boldsymbol{f} : \mathbb{I}B \to \mathbb{IR}^m$ is *inclusion monotonic* if $\boldsymbol{f}(\boldsymbol{x}) \subset \boldsymbol{f}(\boldsymbol{y})$ for any pair $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{I}B$ for which $\boldsymbol{x} \subset \boldsymbol{y}$.

Given a function $g : B \to \mathbb{R}^m$, an interval-valued function $\tilde{\boldsymbol{g}} : \mathbb{I}B \to \mathbb{IR}^m$ is an *interval extension* of g if $\tilde{\boldsymbol{g}}([x,x]) = [g(x), g(x)]$ for each $x \in B$.

The following result from [25] motivates the above definition.

Theorem 4 (Theorem 3.1 in [25]) *Consider a function $g : B \subset \mathbb{R}^n \to \mathbb{R}^m$. If a function $\tilde{\boldsymbol{g}} : \mathbb{I}B \to \mathbb{IR}^m$ is inclusion monotonic and is an interval extension of g, then $g(\boldsymbol{x}) := \{g(z) : z \in \boldsymbol{x}\} \subset \tilde{\boldsymbol{g}}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{I}B$.*

Definition 17 Given a set $D \subset \mathbb{R}^n$, define the *interval hull* $\square D$ of D as the intersection of all intervals in $\mathbb{IR}^n$ that are supersets of D. Given a function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$, define the interval hull $\square f : \mathbb{I}B \to \mathbb{IR}^m$ so that $\square f(\boldsymbol{x}) = \square\{y \in \mathbb{R}^m : \exists z \in \boldsymbol{x} \text{ s.t. } y = f(z)\}$.

A.2 McCormick objects

This appendix summarizes definitions from [33,42] concerning generalized McCormick objects, and adapts these to permit restrictions to proper McCormick objects.

Definition 18 (adapted from [42]) Given a function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$, a mapping $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) $\to \mathbb{MR}^m$ is a *McCormick extension* of f if

$$\mathscr{F}([x,x],[x,x]) = ([f(x), f(x)], [f(x), f(x)]), \qquad x \in B.$$

Definition 19 (adapted from [42]) Given a set $B \subset \mathbb{R}^n$, a function $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) $\to \mathbb{MR}^m$ is *inclusion monotonic* if $\mathscr{F}(\mathscr{X}) \subset \mathscr{F}(\mathscr{Y})$ for each pair $\mathscr{X}, \mathscr{Y} \in \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) such that $\mathscr{X} \subset \mathscr{Y}$.

Definition 20 (adapted from [42]) A pair $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}^n$ is *coherent* if $\boldsymbol{x}^{\text{B}} = \boldsymbol{y}^{\text{B}}$. Given coherent $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}^n$, for each $\lambda \in [0,1]$, define:

$$\mathscr{C}onv(\lambda, \mathscr{X}, \mathscr{Y}) := (\boldsymbol{x}^{\text{B}}, \lambda \boldsymbol{x}^{\text{C}} + (1-\lambda)\boldsymbol{y}^{\text{C}}) \in \mathbb{MR}^n.$$

Given a set $B \subset \mathbb{R}^n$, a function $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) $\to \mathbb{MR}^m$ is *coherent* if $\mathscr{F}(\mathscr{X})$ is coherent to $\mathscr{F}(\mathscr{Y})$ for every coherent $\mathscr{X}, \mathscr{Y} \in \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$). A function $\mathscr{F} : \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$) $\to \mathbb{MR}^m$ is *coherently concave* if it is coherent, and, for every $\mathscr{X}, \mathscr{Y} \in \mathbb{M}B$ (or $\mathbb{M}B_{\text{prop}}$),

$$\mathscr{F}(\mathscr{C}onv(\lambda, \mathscr{X}, \mathscr{Y})) \supset \mathscr{C}onv(\lambda, \mathscr{F}(\mathscr{X}), \mathscr{F}(\mathscr{Y})), \qquad \forall \lambda \in [0,1].$$

The following analogs of addition, multiplication, and UIFs for McCormick objects were developed in [33], and are presented here in the notation of [42]. As demonstrated in [33], these McCormick operations are indeed relaxation functinos of the corresponding operations on real numbers.

Definition 21 (from [33]) Define an addition operation $+ : \mathbb{MR}^2 \to \mathbb{MR}$ so that, for each $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}$,

$$+(\mathscr{X}, \mathscr{Y}) \equiv \mathscr{X} + \mathscr{Y} := (\boldsymbol{x}^{\text{B}} + \boldsymbol{y}^{\text{B}}, \boldsymbol{x}^{\text{C}} + \boldsymbol{y}^{\text{C}}),$$

and define $+ : \mathbb{MR}^2_{\text{prop}} \to \mathbb{MR}$ as the restriction of $+ : \mathbb{MR}^2 \to \mathbb{MR}$ to the domain $\mathbb{MR}^2_{\text{prop}}$.

The following definition of a McCormick multiplication operation is adapted from McCormick's original presentation [23] and [24], and is essentially replaced in this article by Definition 13 in Section 3.3. The following operation will be denoted by the symbol "$\bullet$"; the usual notation for multiplication is reserved for Definition 13. Note that multiplication of a scalar and a McCormick object was treated as a UIF in Tables 1 and 2; the following definition instead concerns multiplication of two McCormick objects.

Definition 22 (from [33]) Define the *classical McCormick multiplication* operation $\bullet : \mathbb{MR}^2 \to \mathbb{MR}$ so that, for each $\mathscr{X}, \mathscr{Y} \in \mathbb{MR}$,

$$\bullet(\mathscr{X}, \mathscr{Y}) \equiv \mathscr{X} \bullet \mathscr{Y} := (\boldsymbol{x}^{\text{B}} \boldsymbol{y}^{\text{B}}, \boldsymbol{z}),$$

where $\boldsymbol{z} \equiv [\underline{z}, \overline{z}] \in \mathbb{IR}$ is defined in terms of the intermediate quantities $\boldsymbol{v} := \boldsymbol{x}^{\text{B}} \cap \boldsymbol{x}^{\text{C}}$ and $\boldsymbol{w} := \boldsymbol{y}^{\text{B}} \cap \boldsymbol{y}^{\text{C}}$ as follows:

$$\underline{z} := \max\left(\underline{(\underline{y}^{\text{B}}\boldsymbol{v})} + \underline{(\underline{x}^{\text{B}}\boldsymbol{w})} - \underline{x}^{\text{B}}\underline{y}^{\text{B}}, \underline{(\overline{y}^{\text{B}}\boldsymbol{v})} + \underline{(\overline{x}^{\text{B}}\boldsymbol{w})} - \overline{x}^{\text{B}}\overline{y}^{\text{B}}\right),$$
$$\overline{z} := \min\left(\overline{(\underline{y}^{\text{B}}\boldsymbol{v})} + \overline{(\overline{x}^{\text{B}}\boldsymbol{w})} - \overline{x}^{\text{B}}\underline{y}^{\text{B}}, \overline{(\overline{y}^{\text{B}}\boldsymbol{v})} + \overline{(\underline{x}^{\text{B}}\boldsymbol{w})} - \underline{x}^{\text{B}}\overline{y}^{\text{B}}\right).$$

Definition 23 (from [33]) Define a function mid : $\mathbb{R}^3 \to \mathbb{R}$ as mapping to the median of its three scalar arguments. Given a UIF $u : B \subset \mathbb{R} \to \mathbb{R}$ that satisfies Assumption 1, define $\mathscr{U} : \mathbb{M}B \to \mathbb{MR}$ so that for each $\mathscr{X} \in \mathbb{M}B$, $\mathscr{U}(\mathscr{X}) := (\tilde{\boldsymbol{u}}(\boldsymbol{x}^{\mathrm{B}}), \boldsymbol{z})$, where

$$\boldsymbol{z} := [u^{\mathrm{cv}}(\boldsymbol{x}^{\mathrm{B}}, \mathrm{mid}(\zeta_u^{\min}(\boldsymbol{x}^{\mathrm{B}}), \underline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{C}})), u^{\mathrm{cc}}(\boldsymbol{x}^{\mathrm{B}}, \mathrm{mid}(\zeta_u^{\max}(\boldsymbol{x}^{\mathrm{B}}), \underline{x}^{\mathrm{C}}, \overline{x}^{\mathrm{C}}))].$$

B Relaxations for UIFs

This appendix presents various results concerning the UIF relaxations u^{cv} and u^{cc} described in Definition 3. Satisfaction of Assumptions 1 and 2 is also discussed.

Lemma 12 *Consider an interval $\boldsymbol{x} \in \mathbb{IR}$, a Lipschitz continuous function $u : \boldsymbol{x} \to \mathbb{R}$, and the convex envelope $u^{\mathrm{cv}} : \boldsymbol{x} \to \mathbb{R}$ of u on $\boldsymbol{x}$. Then, $u^{\mathrm{cv}}(\underline{x}) = u(\underline{x})$ and $u^{\mathrm{cv}}(\overline{x}) = u(\overline{x})$. Moreover, u^{cv} is Lipschitz continuous on $\boldsymbol{x}$, with the same Lipschitz constant as u.*

Proof The required result is trivial if $\underline{x} = \overline{x}$, so assume that $\underline{x} < \overline{x}$. Let k_u denote a Lipschitz constant for u on $\boldsymbol{x}$. Applying the definition of the convex envelope,

$$u(y) \geq u^{\mathrm{cv}}(y) \geq u(\underline{x}) - k_u(y - \underline{x}), \qquad \forall y \in \boldsymbol{x}; \tag{8}$$

the first inequality above is due to u dominating u^{cv}, and the second inequality is due to u^{cv} dominating each convex underestimator of u on $\boldsymbol{x}$. Setting y to $\underline{x}$ in the above inequality chain yields $u^{\mathrm{cv}}(\underline{x}) = u(\underline{x})$.

A similar argument yields:

$$u(y) \geq u^{\mathrm{cv}}(y) \geq u(\overline{x}) + k_u(y - \overline{x}), \qquad \forall y \in \boldsymbol{x}; \tag{9}$$

setting y to $\overline{x}$ yields $u^{\mathrm{cv}}(\overline{x}) = u(\overline{x})$.

Thus, (8) and (9) become:

$$u^{\mathrm{cv}}(y) - u^{\mathrm{cv}}(\underline{x}) \geq -k_u(y - \underline{x}), \qquad \forall y \in \boldsymbol{x},$$
$$u^{\mathrm{cv}}(y) - u^{\mathrm{cv}}(\overline{x}) \geq k_u(y - \overline{x}), \qquad \forall y \in \boldsymbol{x}.$$

Defining D_+u^{cv} and D_-u^{cv} as the right-derivative and left-derivative of u^{cv} described in [16, Part I, Theorem 4.1.1], it follows from [16, Part I, Proposition 4.1.3] that $D_+u^{\mathrm{cv}}(\underline{x})$ and $D_-u^{\mathrm{cv}}(\overline{x})$ both exist, are finite, and satisfy

$$D_+u^{\mathrm{cv}}(\underline{x}) \geq -k_u, \qquad \text{and} \qquad D_-u^{\mathrm{cv}}(\overline{x}) \leq k_u.$$

Thus, u^{cv} is continuous at $\underline{x}$ and $\overline{x}$. Moreover, [16, Part I, Theorem 4.2.1] implies that for each $y \in \mathrm{int}(\boldsymbol{x})$, each subgradient of u^{cv} at y is an element of $[-k_u, k_u]$. This result, combined with the mean-value theorem [16, Part I, Theorem 4.2.4], shows that u^{cv} is Lipschitz continuous on $\boldsymbol{x}$, with a Lipschitz constant of k_u. □

Lemma 13 *Consider an interval $\boldsymbol{x} \in \mathbb{IR}$, and a $\mathscr{C}^1$ function $u : \boldsymbol{x} \to \mathbb{R}$. The convex envelope $u^{\mathrm{cv}} : \boldsymbol{x} \to \mathbb{R}$ of u on $\boldsymbol{x}$ is also $\mathscr{C}^1$ on $\boldsymbol{x}$.*

Proof The required result is trivial if $\underline{x} = \overline{x}$, so assume that $\underline{x} < \overline{x}$. Theorem 3.2 in [13] implies that u^{cv} is $\mathscr{C}^1$ on $(\underline{x}, \overline{x}) = \mathrm{int}(\boldsymbol{x})$; it remains to be shown that u^{cv} is also $\mathscr{C}^1$ at $\underline{x}$ and $\overline{x}$. Noting that u is Lipschitz continuous on $\boldsymbol{x}$, construct the right-derivative D_+u^{cv} and the left-derivative D_-u^{cv} as in the proof of Lemma 12. As in the proof of Lemma 12, $D_+u^{\mathrm{cv}}(\underline{x})$ and $D_-u^{\mathrm{cv}}(\overline{x})$ each exist and are finite. Define the following function, which extends the domain of u^{cv} to $\mathbb{R}$:

$$\psi : \mathbb{R} \to \mathbb{R} : y \mapsto \begin{cases} u^{\mathrm{cv}}(\underline{x}) + (D_+u^{\mathrm{cv}}(\underline{x}))(y - \underline{x}), & \text{if } y < \underline{x}, \\ u^{\mathrm{cv}}(y), & \text{if } y \in \boldsymbol{x}, \\ u^{\mathrm{cv}}(\overline{x}) + (D_-u^{\mathrm{cv}}(\overline{x}))(y - \overline{x}), & \text{if } \overline{x} < y. \end{cases}$$

The function ψ is evidently continuous, and is $\mathscr{C}^1$ at each $y \in \mathbb{R} \backslash \{\underline{x}, \overline{x}\}$. Applying the definitions of D_+u^{cv} and D_-u^{cv}, it follows that ψ is differentiable at $\underline{x}$ and $\overline{x}$ as well; thus,

$$\nabla\psi(y) = \begin{cases} D_+u^{\mathrm{cv}}(\underline{x}), & \text{if } y \leq \underline{x}, \\ \nabla u^{\mathrm{cv}}(y), & \text{if } y \in \mathrm{int}(\boldsymbol{x}), \\ D_-u^{\mathrm{cv}}(\overline{x}), & \text{if } \overline{x} \leq y. \end{cases}$$

This equation, together with [16, Part I, Theorem 4.2.1(iii)], shows that ψ is $\mathscr{C}^1$ even at $\underline{x}$ and $\overline{x}$, and is therefore $\mathscr{C}^1$ on $\mathbb{R}$. Hence, u^{cv} is $\mathscr{C}^1$ on $\boldsymbol{x}$. □

Example 7 Suppose the function $u : \mathbb{R} \to \mathbb{R} : z \mapsto z^2$ is considered as a UIF. Since u is convex, it is its own convex envelope on any subinterval of $\mathbb{R}$. In line with Remark 2, if $i = 1$, then setting $u^{\mathrm{cv}}(\boldsymbol{x}, \cdot) \equiv u$ for each $\boldsymbol{x} \in \mathbb{IR}$ is consistent with Definition 3 and Assumption 1. However, on an interval $\boldsymbol{y}$ with $\underline{y} < 0 < \overline{y}$, $\zeta_u^{\min}(\boldsymbol{y}) = 0 \in \mathrm{int}(\boldsymbol{y})$, but $\nabla^2 u(0) = 2$, so setting $u^{\mathrm{cv}}(\boldsymbol{y}, \cdot) \equiv u$ is inconsistent with Assumption 1 when $i = 2$. Nevertheless, it is readily shown that the following choice of u^{cv} is consistent with Definition 3 and Assumption 1 when $i = 2$:

$$u^{\mathrm{cv}}(\boldsymbol{x}, z) := \begin{cases} z^2, & \text{if } 0 \notin (\underline{x}, \overline{x}), \\ \frac{z^3}{(\overline{x})}, & \text{if } \underline{x} < 0 < \overline{x} \text{ and } 0 \leq z, \\ \frac{z^3}{(\underline{x})}, & \text{if } \underline{x} < 0 < \overline{x} \text{ and } z < 0. \end{cases}$$

Observe that setting $u^{\mathrm{cc}}(\boldsymbol{x}, \cdot)$ to be the affine concave envelope of u on $\boldsymbol{x}$ is consistent with Assumption 1, since, in this case, either $\underline{x}$ or $\overline{x}$ will be a valid choice of $\zeta_u^{\max}(\boldsymbol{x})$. Lemma 14 below shows that the relaxations described in this example also satisfy Assumption 2.

Example 8 Suppose the function $u : \mathbb{R} \to \mathbb{R} : z \mapsto |z|$ is considered as a UIF. In the spirit of the previous example, it is readily confirmed that the following choice of u^{cv} is consistent with Definition 3 and Assumption 1 for each $i \in \{1, 2\}$:

$$u^{\mathrm{cv}}(\boldsymbol{x}, z) := \begin{cases} |z|, & \text{if } 0 \notin (\underline{x}, \overline{x}), \\ \frac{z^{2+i}}{\overline{x}^{1+i}}, & \text{if } \underline{x} < 0 < \overline{x} \text{ and } 0 \leq z, \\ \left| \frac{z^{2+i}}{\underline{x}^{1+i}} \right|, & \text{if } \underline{x} < 0 < \overline{x} \text{ and } z < 0. \end{cases}$$

As in the previous example, observe that setting $u^{\mathrm{cc}}(\boldsymbol{x}, \cdot)$ to be the affine concave envelope of u on $\boldsymbol{x}$ is consistent with Assumption 1.

Example 9 For some fixed $k \in \mathbb{N}$, suppose the function $u : \mathbb{R} \to \mathbb{R} : z \mapsto z^{2k+1}$ is considered as a UIF. In line with Remark 2, if $i = 1$, setting $u^{\mathrm{cv}}(\boldsymbol{x}, \cdot)$ and $u^{\mathrm{cc}}(\boldsymbol{x}, \cdot)$ to be the convex/concave envelopes of u described in [22] is consistent with Definition 3 and Assumption 1. If $i = 2$, then it is readily verified that the following choices of u^{cv} and u^{cc} are consistent with Definition 3 and Assumption 1:

$$u^{\mathrm{cv}}(\boldsymbol{x}, z) := \begin{cases} \underline{x}^{2k+1} + (\overline{x}^{2k+1} - \underline{x}^{2k+1}) \left(\frac{z - \underline{x}}{\overline{x} - \underline{x}} \right), & \text{if } \overline{x} \leq 0, \\ \underline{x}^{2k+1} \left(\frac{\overline{x} - z}{\overline{x} - \underline{x}} \right) + (\max\{0, z\})^{2k+1}, & \text{if } \underline{x} < 0 < \overline{x}, \\ z^{2k+1}, & \text{if } 0 \leq \underline{x}, \end{cases}$$

$$u^{\mathrm{cc}}(\boldsymbol{x}, z) := \begin{cases} z^{2k+1}, & \text{if } \overline{x} \leq 0, \\ \overline{x}^{2k+1} \left(\frac{z - \underline{x}}{\overline{x} - \underline{x}} \right) + (\min\{0, z\})^{2k+1}, & \text{if } \underline{x} < 0 < \overline{x}, \\ \underline{x}^{2k+1} + (\overline{x}^{2k+1} - \underline{x}^{2k+1}) \left(\frac{z - \underline{x}}{\overline{x} - \underline{x}} \right), & \text{if } 0 \leq \underline{x}. \end{cases}$$

The functions $u^{\mathrm{cv}}(\boldsymbol{x}, \cdot)$ and $u^{\mathrm{cc}}(\boldsymbol{x}, \cdot)$ described above are evidently strictly increasing on $\boldsymbol{x}$ for each $\boldsymbol{x} \in \mathbb{IR}$. Thus, setting $\zeta_u^{\min}(\boldsymbol{x}) := \underline{x}$ and $\zeta_u^{\max}(\boldsymbol{x}) := \overline{x}$ is consistent with Definition 3. Lemma 15 below shows that the relaxations described in this example also satisfy Assumption 2.

Lemma 14 *Consider the relaxation schemes $\{u^{\mathrm{cv}}(\boldsymbol{x}, \cdot), u^{\mathrm{cc}}(\boldsymbol{x}, \cdot)\}_{\boldsymbol{x} \in \mathbb{IR}}$ for $u : z \mapsto z^2$ on $\mathbb{R}$ described in Example 7. These schemes satisfy Assumption 2.*

Proof By [6, Theorem 10], the concave relaxations of u described in Example 7 are pointwise convergent of order 2, as are the convex relaxations described in Example 7 when $i = 1$. It remains to consider only the convex relaxations of u when $i = 2$.

Now, if $\boldsymbol{x} \in \mathbb{IR}$ but $0 \notin \mathrm{int}(\boldsymbol{x})$, then $u(z) - u^{\mathrm{cv}}(\boldsymbol{x}, z) = 0$ for all $z \in \boldsymbol{x}$. If, instead, $\boldsymbol{x} \in \mathbb{IR}$ and $0 \in \mathrm{int}(\boldsymbol{x})$, then

$$\begin{aligned} \sup_{z \in \boldsymbol{x}} (u(z) - u^{\mathrm{cv}}(\boldsymbol{x}, z)) &= \max \left\{ \sup_{z \in [\underline{x}, 0]} (u(z) - u^{\mathrm{cv}}(\boldsymbol{x}, z)), \sup_{z \in [0, \overline{x}]} (u(z) - u^{\mathrm{cv}}(\boldsymbol{x}, z)) \right\}, \\ &= \max \left\{ \sup_{z \in [\underline{x}, 0]} \left(z^2 - \frac{z^3}{(\underline{x})} \right), \sup_{z \in [0, \overline{x}]} \left(z^2 - \frac{z^3}{(\overline{x})} \right) \right\}, \\ &= \max \left\{ \tfrac{4}{27} \underline{x}^2, \tfrac{4}{27} \overline{x}^2 \right\}, \\ &\leq \tfrac{4}{27} (\mathrm{wid}\, \boldsymbol{x})^2. \end{aligned}$$

Combining the above cases,

$$\sup_{z\in\boldsymbol{x}}(u(z)-u^{\mathrm{cv}}(\boldsymbol{x},z)) \le \tfrac{4}{27}(\operatorname{wid}\boldsymbol{x})^2, \qquad \forall \boldsymbol{x}\in\mathbb{IR},$$

as required. □

Lemma 15 *For fixed $k\in\mathbb{N}$, consider the relaxation schemes $\{u^{\mathrm{cv}}(\boldsymbol{x},\cdot),u^{\mathrm{cc}}(\boldsymbol{x},\cdot)\}_{\boldsymbol{x}\in\mathbb{IR}}$ for $u:z\mapsto z^{2k+1}$ on $\mathbb{R}$ described in Example 9. These schemes satisfy Assumption 2.*

Proof Again, by [6, Theorem 10], the relaxations of u described in Example 9 when $i=1$ are pointwise convergent of order 2; it remains only to demonstrate the $i=2$ case. It will be shown that the convex relaxations $u^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ of u are pointwise convergent of order 2; a similar argument applies to the concave relaxations $u^{\mathrm{cc}}(\boldsymbol{x},\cdot)$.

Consider any fixed interval $\boldsymbol{q}\in\mathbb{IR}$, and any $\boldsymbol{x}\in\mathbb{I}\boldsymbol{q}$. If $i=1$ or $0\notin\boldsymbol{x}$, then $u^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ is the convex envelope of u on $\boldsymbol{x}$, which, by [6, Theorem 10], is pointwise convergent of order 2 with respect to $\boldsymbol{x}$.

If $i=2$ and $0\in\boldsymbol{x}$, then, noting that $u^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ is increasing, we obtain:

$$\begin{aligned}\sup_{z\in\boldsymbol{x}}(u(z)-u^{\mathrm{cv}}(\boldsymbol{x},z)) &\le \sup_{z\in\boldsymbol{x}}(u(z)-u^{\mathrm{cv}}(\boldsymbol{x},\underline{x}))\\ &= \sup_{z\in\boldsymbol{x}}(z^{2k+1}-\underline{x}^{2k+1}) \le (\bar{x}-\underline{x})^{2k+1}+(\bar{x}-\underline{x})^{2k+1} \le 2(\operatorname{wid}\boldsymbol{q})^{2k-1}(\bar{x}-\underline{x})^2.\end{aligned}$$

The above results together show that $u^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ is pointwise convergent of order 2 to u with respect to $\boldsymbol{x}$, as required. □

C Intermediate results

This appendix establishes key intermediate results that are used in the proof of Theorem 2.

C.1 Establishing elemental relaxation functions

The following theorem shows that Definitions 21, 13, and 23 provide relaxation functions of addition, multiplication, and UIFs; the remainder of this section is concerned with proving this theorem.

Theorem 5 *The functions $+:\mathbb{MR}^2_{\mathrm{prop}}\to\mathbb{MR}_{\mathrm{prop}}$, $\times_i:\mathbb{MR}^2_{\mathrm{prop}}\to\mathbb{MR}_{\mathrm{prop}}$, and $\mathscr{U}:\mathbb{M}B\to\mathbb{MR}$ described in Definitions 21, 13, and 23 are relaxation functions for $+:\mathbb{R}^2\to\mathbb{R}$, $\times:\mathbb{R}^2\to\mathbb{R}$, and $u:B\to\mathbb{R}$, respectively.*

Proof This theorem combines Lemmata 16–20 below. □

Lemma 16 *The function $+:\mathbb{MR}^2_{\mathrm{prop}}\to\mathbb{MR}_{\mathrm{prop}}$ is coherently concave, inclusion monotonic, and a McCormick extension of $+:\mathbb{R}^2\to\mathbb{R}$.*

Proof This result follows from Theorem 2.4.20 in [33], noting that for any choice of $\mathscr{X},\mathscr{Y}\in\mathbb{MR}_{\mathrm{prop}}$, $\boldsymbol{x}^{\mathrm{B}}\cap\boldsymbol{x}^{\mathrm{C}}=\boldsymbol{x}^{\mathrm{C}}$ and $\boldsymbol{y}^{\mathrm{B}}\cap\boldsymbol{y}^{\mathrm{C}}=\boldsymbol{y}^{\mathrm{C}}$. □

Lemma 17 *The function $\times_i:\mathbb{MR}^2_{\mathrm{prop}}\to\mathbb{MR}_{\mathrm{prop}}$ is coherently concave.*

Proof Consider a coherent pair $(\mathscr{X}_1,\mathscr{Y}_1),(\mathscr{X}_2,\mathscr{Y}_2)\in\mathbb{MR}^2_{\mathrm{prop}}$, and a scalar $\ell\in[0,1]$. Since this pair is coherent, define $\boldsymbol{x}^{\mathrm{B}}:=\boldsymbol{x}^{\mathrm{B}}_1=\boldsymbol{x}^{\mathrm{B}}_2$ and $\boldsymbol{y}^{\mathrm{B}}:=\boldsymbol{y}^{\mathrm{B}}_1=\boldsymbol{y}^{\mathrm{B}}_2$. Define $\mathscr{Q}_1:=\mathscr{X}_1\mathscr{Y}_1$ and $\mathscr{Q}_2:=\mathscr{X}_2\mathscr{Y}_2$. Using the definition of the squashing operation, it follows that $\boldsymbol{q}^{\mathrm{B}}_1=\boldsymbol{q}^{\mathrm{B}}_2=\boldsymbol{x}^{\mathrm{B}}\boldsymbol{y}^{\mathrm{B}}=:\boldsymbol{q}^{\mathrm{B}}$, and so $\mathscr{Q}_1$ and $\mathscr{Q}_2$ are coherent. Define $\mathscr{X}_0:=\mathscr{C}onv(\ell,\mathscr{X}_1,\mathscr{X}_2)$, and define $\mathscr{Y}_0$ and $\mathscr{Q}_0$ analogously. To obtain the required result, it suffices to show that $\mathscr{X}_0\mathscr{Y}_0\supset\mathscr{Q}_0$.

If $\underline{q}^{\mathrm{B}}=\overline{q}^{\mathrm{B}}=:q$, then the definition of the squashing operation implies that $\mathscr{X}_0\mathscr{Y}_0=[q,q]=\mathscr{Q}_0$, as required. Thus, it will be assumed throughout the rest of this proof that $\underline{q}^{\mathrm{B}}<\overline{q}^{\mathrm{B}}$.

For each $j \in \{0,1,2\}$, define $\mathbf{z}_j \equiv [\underline{z}_j, \overline{z}_j] \in \mathbb{IR}$ such that

$$\underline{z}_j := \nu_i\left(\underline{(y^{\mathrm{B}}\mathbf{x}_j^{\mathrm{C}})} + \underline{(x^{\mathrm{B}}\mathbf{y}_j^{\mathrm{C}})} - \underline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \underline{(\overline{y}^{\mathrm{B}}\mathbf{x}_j^{\mathrm{C}})} + \underline{(\overline{x}^{\mathrm{B}}\mathbf{y}_j^{\mathrm{C}})} - \overline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}},\, p_{\mathbf{q}^{\mathrm{B}}}\right),$$

$$\text{and}\qquad \overline{z}_j := \lambda_i\left(\overline{(\underline{y}^{\mathrm{B}}\mathbf{x}_j^{\mathrm{C}})} + \overline{(\overline{x}^{\mathrm{B}}\mathbf{y}_j^{\mathrm{C}})} - \overline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}},\, \overline{(\overline{y}^{\mathrm{B}}\mathbf{x}_j^{\mathrm{C}})} + \overline{(\underline{x}^{\mathrm{B}}\mathbf{y}_j^{\mathrm{C}})} - \underline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}},\, p_{\mathbf{q}^{\mathrm{B}}}\right).$$

Since $\mathbf{q}_1^{\mathrm{B}} = \mathbf{q}_2^{\mathrm{B}} = \mathbf{x}^{\mathrm{B}}\mathbf{y}^{\mathrm{B}} = \mathbf{q}^{\mathrm{B}}$, the required inclusion, $\mathscr{X}_0\mathscr{Y}_0 \supset \mathscr{C}onv(\ell, \mathscr{Q}_1, \mathscr{Q}_2)$, is equivalent to the inclusion:

$$\mathscr{S}qu_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_0)) \supset \mathscr{C}onv(\ell, \mathscr{S}qu_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_1)), \mathscr{S}qu_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_2))),$$

which is in turn equivalent to the inclusion:

$$\mathbf{belt}_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_0)) \supset \ell\,\mathbf{belt}_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_1)) + (1-\ell)\,\mathbf{belt}_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_2)).$$

Thus, due to Lemma 8, it suffices to demonstrate the following inclusion:

$$\mathbf{belt}_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_0)) \supset \mathbf{belt}_i(\mathscr{C}onv(\ell, (\mathbf{q}^{\mathrm{B}}, \mathbf{z}_1), (\mathbf{q}^{\mathrm{B}}, \mathbf{z}_2))),$$

which can be rewritten as:

$$\mathbf{belt}_i((\mathbf{q}^{\mathrm{B}}, \mathbf{z}_0)) \supset \mathbf{belt}_i((\mathbf{q}^{\mathrm{B}}, [\ell\underline{z}_1 + (1-\ell)\underline{z}_2, \ell\overline{z}_1 + (1-\ell)\overline{z}_2])).$$

Since $\mathbf{belt}_i$ is inclusion monotonic, it thus suffices to demonstrate the inequalities:

$$\underline{z}_0 \le \ell\underline{z}_1 + (1-\ell)\underline{z}_2, \qquad \text{and} \qquad \overline{z}_0 \ge \ell\overline{z}_1 + (1-\ell)\overline{z}_2.$$

The first of these inequalities will be demonstrated here; the second can be shown to hold by an analogous argument. For each $j \in \{0,1,2\}$, define:

$$\alpha_j := \underline{(y^{\mathrm{B}}\mathbf{x}_j^{\mathrm{C}})} + \underline{(x^{\mathrm{B}}\mathbf{y}_j^{\mathrm{C}})} - \underline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}}, \qquad \text{and} \qquad \beta_j := \underline{(\overline{y}^{\mathrm{B}}\mathbf{x}_j^{\mathrm{C}})} + \underline{(\overline{x}^{\mathrm{B}}\mathbf{y}_j^{\mathrm{C}})} - \overline{x}^{\mathrm{B}}\overline{y}^{\mathrm{B}}.$$

Now, for each $j \in \{0,1,2\}$,

$$\underline{(\underline{y}^{\mathrm{B}}\mathbf{x}_j^{\mathrm{C}})} = \begin{cases} \underline{y}^{\mathrm{B}}\underline{x}_j^{\mathrm{C}} & \text{if } \underline{y}^{\mathrm{B}} \ge 0, \\ \underline{y}^{\mathrm{B}}\overline{x}_j^{\mathrm{C}} & \text{if } \underline{y}^{\mathrm{B}} < 0. \end{cases}$$

Moreover, by definition of the $\mathscr{C}onv$ operation,

$$\underline{x}_0^{\mathrm{C}} = \ell\underline{x}_1^{\mathrm{C}} + (1-\ell)\underline{x}_2^{\mathrm{C}}, \qquad \text{and} \qquad \overline{x}_0^{\mathrm{C}} = \ell\overline{x}_1^{\mathrm{C}} + (1-\ell)\overline{x}_2^{\mathrm{C}}.$$

Combining the above results, it follows that

$$\underline{(\underline{y}^{\mathrm{B}}\mathbf{x}_0^{\mathrm{C}})} = \ell\underline{(\underline{y}^{\mathrm{B}}\mathbf{x}_1^{\mathrm{C}})} + (1-\ell)\underline{(\underline{y}^{\mathrm{B}}\mathbf{x}_2^{\mathrm{C}})};$$

an analogous argument shows that

$$\underline{(\underline{x}^{\mathrm{B}}\mathbf{y}_0^{\mathrm{C}})} = \ell\underline{(\underline{x}^{\mathrm{B}}\mathbf{y}_1^{\mathrm{C}})} + (1-\ell)\underline{(\underline{x}^{\mathrm{B}}\mathbf{y}_2^{\mathrm{C}})}.$$

Adding these two equations and subtracting the constant term $\underline{x}^{\mathrm{B}}\underline{y}^{\mathrm{B}}$, it follows that

$$\alpha_0 = \ell\alpha_1 + (1-\ell)\alpha_2;$$

an analogous argument shows that

$$\beta_0 = \ell\beta_1 + (1-\ell)\beta_2.$$

Thus,

$$\nu_i(\alpha_0, \beta_0, p_{\mathbf{q}^{\mathrm{B}}}) = \nu_i(\ell\alpha_1 + (1-\ell)\alpha_2, \ell\beta_1 + (1-\ell)\beta_2, p_{\mathbf{q}^{\mathrm{B}}}),$$

which, by Lemma 7, implies that

$$\nu_i(\alpha_0, \beta_0, p_{\mathbf{q}^{\mathrm{B}}}) \le \ell\,\nu_i(\alpha_1, \beta_1, p_{\mathbf{q}^{\mathrm{B}}}) + (1-\ell)\,\nu_i(\alpha_2, \beta_2, p_{\mathbf{q}^{\mathrm{B}}}).$$

Comparing this inequality with the definitions of α_j, β_j, and $\underline{z}_j$ for each $j \in \{0,1,2\}$, it follows immediately that

$$\underline{z}_0 \le \ell\underline{z}_1 + (1-\ell)\underline{z}_2,$$

as required. □

Lemma 18 *The function* $\times_i : \mathbb{MR}^2_{\text{prop}} \to \mathbb{MR}_{\text{prop}}$ *is inclusion monotonic.*

Proof Consider any $\mathscr{X}_1, \mathscr{X}_2, \mathscr{Y}_1, \mathscr{Y}_2 \in \mathbb{MR}_{\text{prop}}$ such that $\mathscr{X}_2 \subset \mathscr{X}_1$ and $\mathscr{Y}_2 \subset \mathscr{Y}_1$. It will be shown that $\mathscr{X}_2\mathscr{Y}_2 \subset \mathscr{X}_1\mathscr{Y}_1$. Since $\times : \mathbb{IR}^2 \to \mathbb{IR}$ is inclusion monotonic, it follows that $\boldsymbol{x}_2^{\text{B}}\boldsymbol{y}_2^{\text{B}} \subset \boldsymbol{x}_1^{\text{B}}\boldsymbol{y}_1^{\text{B}}$, and so $p_{\boldsymbol{x}_2^{\text{B}}\boldsymbol{y}_2^{\text{B}}} \leq p_{\boldsymbol{x}_1^{\text{B}}\boldsymbol{y}_1^{\text{B}}}$. By construction $\boldsymbol{x}_2^{\text{C}} \subset \boldsymbol{x}_1^{\text{C}}$ and $\boldsymbol{y}_2^{\text{C}} \subset \boldsymbol{y}_1^{\text{C}}$. Define $\boldsymbol{z}_1 \in \mathbb{IR}$ as in Definition 13 for the product $\mathscr{X}_1\mathscr{Y}_1$, and define $\boldsymbol{z}_2 \in \mathbb{IR}$ analogously for the product $\mathscr{X}_2\mathscr{Y}_2$.

Due to Lemma 9 and the inclusion $\boldsymbol{x}_2^{\text{B}}\boldsymbol{y}_2^{\text{B}} \subset \boldsymbol{x}_1^{\text{B}}\boldsymbol{y}_1^{\text{B}}$, it suffices to show that $\boldsymbol{z}_2 \subset \boldsymbol{z}_1$. It will be shown that $\underline{z}_2 \geq \underline{z}_1$; an analogous argument shows that $\bar{z}_2 \leq \bar{z}_1$. In turn, due to Lemma 6 and the inequality $p_{\boldsymbol{x}_2^{\text{B}}\boldsymbol{y}_2^{\text{B}}} \leq p_{\boldsymbol{x}_1^{\text{B}}\boldsymbol{y}_1^{\text{B}}}$, it suffices to demonstrate the inequalities:

$$\begin{aligned}
&\underline{(\underline{y}_1^{\text{B}}\boldsymbol{x}_1^{\text{C}})} + \underline{(\underline{x}_1^{\text{B}}\boldsymbol{y}_1^{\text{C}})} - \underline{x}_1^{\text{B}}\underline{y}_1^{\text{B}} \leq \underline{(\underline{y}_2^{\text{B}}\boldsymbol{x}_2^{\text{C}})} + \underline{(\underline{x}_2^{\text{B}}\boldsymbol{y}_2^{\text{C}})} - \underline{x}_2^{\text{B}}\underline{y}_2^{\text{B}}, \\
\text{and} \quad &\underline{(\bar{y}_1^{\text{B}}\boldsymbol{x}_1^{\text{C}})} + \underline{(\bar{x}_1^{\text{B}}\boldsymbol{y}_1^{\text{C}})} - \bar{x}_1^{\text{B}}\bar{y}_1^{\text{B}} \leq \underline{(\bar{y}_2^{\text{B}}\boldsymbol{x}_2^{\text{C}})} + \underline{(\bar{x}_2^{\text{B}}\boldsymbol{y}_2^{\text{C}})} - \bar{x}_2^{\text{B}}\bar{y}_2^{\text{B}}.
\end{aligned}$$

Noting that $\boldsymbol{x}_j^{\text{C}} = \boldsymbol{x}_j^{\text{B}} \cap \boldsymbol{x}_j^{\text{C}}$ and $\boldsymbol{y}_j^{\text{C}} = \boldsymbol{y}_j^{\text{B}} \cap \boldsymbol{y}_j^{\text{C}}$ for each $j \in \{1,2\}$ by construction, the proof of [33, Theorem 2.4.23] demonstrates the above inequalities. $\square$

Lemma 19 *The function* $\times_i : \mathbb{MR}^2_{\text{prop}} \to \mathbb{MR}_{\text{prop}}$ *is a McCormick extension of* $\times : \mathbb{R}^2 \to \mathbb{R}$.

Proof Choose $x, y \in \mathbb{R}$, and set $\mathscr{X}_0 := ([x,x],[x,x]) \in \mathbb{MR}_{\text{prop}}$ and $\mathscr{Y}_0 := ([y,y],[y,y]) \in \mathbb{MR}_{\text{prop}}$. Observe that $[x,x][y,y] = [xy,xy]$, and that $p_{[x,x]} = p_{[y,y]} = p_{[x,x][y,y]} = 0$. Thus, according to Definition 13 and the established properties of γ_i and σ_i, it follows that:

$$\mathscr{X}_0\mathscr{Y}_0 = \mathscr{S}qu_i(([x,x][y,y],\boldsymbol{z})),$$

where $\boldsymbol{z} \equiv [\underline{z}, \bar{z}]$ is defined as follows:

$$\begin{aligned}
\underline{z} &:= \nu_i\left(\underline{(y[x,x])} + \underline{(x[y,y])} - xy,\, \underline{(y[x,x])} + \underline{(x[y,y])} - xy,\, 0\right) = \nu_i(xy,xy,0) = \gamma_i(xy,xy,0) = xy, \\
\bar{z} &:= \lambda_i\left(\overline{(y[x,x])} + \overline{(x[y,y])} - xy,\, \overline{(y[x,x])} + \overline{(x[y,y])} - xy,\, 0\right) = \lambda_i(xy,xy,0) = \sigma_i(xy,xy,0) = xy.
\end{aligned}$$

Thus, $\mathscr{X}_0\mathscr{Y}_0 = \mathscr{S}qu_i(([xy,xy],[xy,xy])) = ([xy,xy],[xy,xy])$. $\square$

Lemma 20 **(Theorems 2.4.27, 2.4.29, and 2.4.30 in [33])** *For any UIF* $u : B \subset \mathbb{R} \to \mathbb{R}$*, the function* $\mathscr{U} : \mathbb{M}B \to \mathbb{MR}$ *is coherently concave, inclusion monotonic, and a McCormick extension of* u.

C.2 Establishing continuous and twice-continuous differentiability

The results in this section show that for any natural or unconstrained $\mathscr{C}^i$ McCormick extension of a factorable function, the convex/concave relaxations suggested by Proposition 4 are indeed $\mathscr{C}^i$ on their interval domains, and their gradients may be evaluated using the standard forward or reverse modes of automatic differentiation [14].

In particular, Lemmata 24 and 25 effectively correct McCormick's proposed sufficient condition for differentiability of relaxations of composite functions [23, p. 151], by applying Assumption 1.

Theorem 6 *Consider any* $i^* \in \{1,2\}$*. For fixed intervals* $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{IR}$*, the mappings*

$$\begin{aligned}
&(\underline{y}_1, \bar{y}_1, \underline{y}_2, \bar{y}_2) \mapsto \underline{+}^{\text{C}}\left((\boldsymbol{x}_1, [\underline{y}_1, \bar{y}_1]), (\boldsymbol{x}_2, [\underline{y}_2, \bar{y}_2])\right), \\
&(\underline{y}_1, \bar{y}_1, \underline{y}_2, \bar{y}_2) \mapsto \overline{+}^{\text{C}}\left((\boldsymbol{x}_1, [\underline{y}_1, \bar{y}_1]), (\boldsymbol{x}_2, [\underline{y}_2, \bar{y}_2])\right), \\
&(\underline{y}_1, \bar{y}_1, \underline{y}_2, \bar{y}_2) \mapsto \underline{\times}_i^{\text{C}}\left((\boldsymbol{x}_1, [\underline{y}_1, \bar{y}_1]), (\boldsymbol{x}_2, [\underline{y}_2, \bar{y}_2])\right), \\
\textit{and} \quad &(\underline{y}_1, \bar{y}_1, \underline{y}_2, \bar{y}_2) \mapsto \overline{\times}_i^{\text{C}}\left((\boldsymbol{x}_1, [\underline{y}_1, \bar{y}_1]), (\boldsymbol{x}_2, [\underline{y}_2, \bar{y}_2])\right)
\end{aligned}$$

described in Definitions 21 and 13 are each $\mathscr{C}^{i^*}$ *on* $\{(\underline{y}_1, \bar{y}_1, \underline{y}_2, \bar{y}_2) \in \mathbb{R}^4 : \underline{y}_1 \leq \bar{y}_1, \underline{y}_2 \leq \bar{y}_2, [\underline{y}_1, \bar{y}_1] \in \boldsymbol{x}_1, [\underline{y}_2, \bar{y}_2] \in \boldsymbol{x}_2\}$.

Next, consider a UIF $u : B \subset \mathbb{R} \to \mathbb{R}$ that satisfies Assumption 1 with $i := i^*$, and choose any fixed interval $\boldsymbol{x} \in \mathbb{I}B$. The mappings

$$(\underline{y},\overline{y}) \mapsto \underline{u}^{\mathrm{C}}((\boldsymbol{x},[\underline{y},\overline{y}])) \qquad \text{and} \qquad (\underline{y},\overline{y}) \mapsto \overline{u}^{\mathrm{C}}((\boldsymbol{x},[\underline{y},\overline{y}])),$$

described in Definition 23, are each $\mathscr{C}^{i^*}$ on $\{(\underline{y},\overline{y}) \in \mathbb{R}^2 : \underline{y} \leq \overline{y}, [\underline{y},\overline{y}] \subset \boldsymbol{x}\}$.

Proof This theorem collects the results of Lemmata 21–25 below. □

Lemma 21 *For fixed intervals $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{IR}$, the mappings*

$$(\underline{y}_1,\overline{y}_1,\underline{y}_2,\overline{y}_2) \mapsto \underline{\pm}^{\mathrm{C}}\left((\boldsymbol{x}_1,[\underline{y}_1,\overline{y}_1]),(\boldsymbol{x}_2,[\underline{y}_2,\overline{y}_2])\right),$$
$$\text{and} \qquad (\underline{y}_1,\overline{y}_1,\underline{y}_2,\overline{y}_2) \mapsto \overline{\mp}^{\mathrm{C}}\left((\boldsymbol{x}_1,[\underline{y}_1,\overline{y}_1]),(\boldsymbol{x}_2,[\underline{y}_2,\overline{y}_2])\right)$$

are each $\mathscr{C}^2$ on $\{(\underline{y}_1,\overline{y}_1,\underline{y}_2,\overline{y}_2) \in \mathbb{R}^4 : \underline{y}_1 \leq \overline{y}_1, \underline{y}_2 \leq \overline{y}_2, [\underline{y}_1,\overline{y}_1] \in \boldsymbol{x}_1, [\underline{y}_2,\overline{y}_2] \in \boldsymbol{x}_2\}$.

Proof The definition of $+ : \mathbb{MR}^2_{\mathrm{prop}} \mapsto \mathbb{MR}_{\mathrm{prop}}$ implies that the mappings in question are linear, and are therefore $\mathscr{C}^2$. □

Lemma 22 *Suppose that scalars $a,b,c,d \in \mathbb{R}$ are such that $ab = ad = cb = cd$. At least one of the following conditions must hold:*

- *both $a = c$ and $b = d$ hold simultaneously,*
- *$a = c = 0$,*
- *$b = d = 0$.*

Proof Suppose that the first condition does not hold; it will be shown that either the second or third condition must hold in this case. Thus, suppose that either $a \neq c$ or $b \neq d$. If $a \neq c$, then the equations $(a-c)b = 0 = (a-c)d$ imply that $b = d = 0$, as required. Otherwise, if $b \neq d$, then the equations $a(b-d) = 0 = c(b-d)$ imply that $a = c = 0$, as required. □

Lemma 23 *For each $i \in \{1,2\}$, given fixed intervals $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{IR}$, the mappings*

$$(\underline{y}_1,\overline{y}_1,\underline{y}_2,\overline{y}_2) \mapsto \underline{\times}_i^{\mathrm{C}}\left((\boldsymbol{x}_1,[\underline{y}_1,\overline{y}_1]),(\boldsymbol{x}_2,[\underline{y}_2,\overline{y}_2])\right),$$
$$\text{and} \qquad (\underline{y}_1,\overline{y}_1,\underline{y}_2,\overline{y}_2) \mapsto \overline{\times}_i^{\mathrm{C}}\left((\boldsymbol{x}_1,[\underline{y}_1,\overline{y}_1]),(\boldsymbol{x}_2,[\underline{y}_2,\overline{y}_2])\right)$$

are each $\mathscr{C}^i$ on $\{(\underline{y}_1,\overline{y}_1,\underline{y}_2,\overline{y}_2) \in \mathbb{R}^4 : \underline{y}_1 \leq \overline{y}_1, \underline{y}_2 \leq \overline{y}_2, [\underline{y}_1,\overline{y}_1] \in \boldsymbol{x}_1, [\underline{y}_2,\overline{y}_2] \in \boldsymbol{x}_2\}$.

Proof The cases in which $\mathrm{wid}(\boldsymbol{x}_1\boldsymbol{x}_2) > 0$ and $\mathrm{wid}(\boldsymbol{x}_1\boldsymbol{x}_2) = 0$ will be considered separately.

Firstly, suppose that $\mathrm{wid}(\boldsymbol{x}_1\boldsymbol{x}_2) > 0$. For any $c \in \mathbb{R}$, (6) and (7) imply that the mappings $(v,w) \mapsto \underline{(c[v,w])}$ and $(v,w) \mapsto \overline{(c[v,w])}$ are both linear on $\{(v,w) \in \mathbb{R}^2 : v \leq w\}$, and are therefore $\mathscr{C}^i$. This observation, together with Lemma 5, Lemma 10, and Definition 13, implies that the required result holds.

Secondly, suppose that $\mathrm{wid}(\boldsymbol{x}_1\boldsymbol{x}_2) = 0$, in which case $\underline{x}_1\underline{x}_2 = \underline{x}_1\overline{x}_2 = \overline{x}_1\underline{x}_2 = \overline{x}_1\overline{x}_2$. Applying Lemma 22, it suffices to consider separately the cases in which $\underline{x}_1 = \overline{x}_1 = 0$, $\underline{x}_2 = \overline{x}_2 = 0$, and both $\underline{x}_1 = \overline{x}_1$ and $\underline{x}_2 = \overline{x}_2$.

If $\underline{x}_1 = \overline{x}_1 = 0$, then $\boldsymbol{x}_1\boldsymbol{x}_2 = [0,0]$, in which case the outer squashing operation in Definition 13 implies that $(\boldsymbol{x}_1,[\underline{y}_1,\overline{y}_1]) \times (\boldsymbol{x}_2,[\underline{y}_2,\overline{y}_2]) = ([0,0],[0,0])$. Thus, each of the two mappings in the statement of the lemma is the zero mapping, which is trivially $\mathscr{C}^i$. The case in which $\underline{x}_2 = \overline{x}_2 = 0$ is analogous.

Lastly, if both $\underline{x}_1 = \overline{x}_1 =: x_1$ and $\underline{x}_2 = \overline{x}_2 =: x_2$, then $\boldsymbol{x}_1\boldsymbol{x}_2 = [x_1x_2, x_1x_2]$, in which case the outer squashing operation in Definition 13 implies that $(\boldsymbol{x}_1,[\underline{y}_1,\overline{y}_1]) \times (\boldsymbol{x}_2,[\underline{y}_2,\overline{y}_2]) = ([x_1x_2,x_1x_2],[x_1x_2,x_1x_2])$. Thus, each of the two mappings in the statement of the lemma is a constant mapping, which, again, is trivially $\mathscr{C}^i$. □

The following two lemmata essentially show that McCormick's proposed sufficient condition for differentiable relaxations of composite functions [23, p. 151] becomes valid when Assumption 1 is applied.

Lemma 24 *Consider a UIF $u : B \subset \mathbb{R} \to \mathbb{R}$ that satisfies Assumption 1. For any intervals $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{I}B$ for which $\boldsymbol{y} \subset \boldsymbol{x}$,*

$$u^{\mathrm{cv}}(\boldsymbol{x}, \mathrm{mid}(\zeta_u^{\min}(\boldsymbol{x}),\underline{y},\overline{y})) = u_I^{\mathrm{cv}}(\boldsymbol{x},\underline{y}) + u_D^{\mathrm{cv}}(\boldsymbol{x},\overline{y}) - u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x})),$$
$$\text{and} \qquad u^{\mathrm{cc}}(\boldsymbol{x}, \mathrm{mid}(\zeta_u^{\max}(\boldsymbol{x}),\underline{y},\overline{y})) = u_I^{\mathrm{cc}}(\boldsymbol{x},\underline{y}) + u_D^{\mathrm{cc}}(\boldsymbol{x},\overline{y}) - u^{\mathrm{cc}}(\boldsymbol{x},\zeta_u^{\max}(\boldsymbol{x})).$$

Proof The first required equation will be shown to hold; the second can be demonstrated analogously. By construction,

$$u_I^{\mathrm{cv}}(\boldsymbol{x},\underline{y})+u_D^{\mathrm{cv}}(\boldsymbol{x},\bar{y})-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x})) \\ = u^{\mathrm{cv}}(\boldsymbol{x},\max\{\underline{y},\zeta_u^{\min}(\boldsymbol{x})\})+u^{\mathrm{cv}}(\boldsymbol{x},\min\{\bar{y},\zeta_u^{\min}(\boldsymbol{x})\})-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x})). \tag{10}$$

Since $\underline{y}\leq\bar{y}$, at least one of the following three cases must apply: $\zeta_u^{\min}(\boldsymbol{x})\leq\underline{y}\leq\bar{y}$, $\underline{y}\leq\zeta_u^{\min}(\boldsymbol{x})\leq\bar{y}$, or $\underline{y}\leq\bar{y}\leq\zeta_u^{\min}(\boldsymbol{x})$. These cases will be considered separately.

If $\zeta_u^{\min}(\boldsymbol{x})\leq\underline{y}\leq\bar{y}$, then $\underline{y}=\mathrm{mid}(\zeta_u^{\min}(\boldsymbol{x}),\underline{y},\bar{y})$, and (10) becomes

$$u_I^{\mathrm{cv}}(\boldsymbol{x},\underline{y})+u_D^{\mathrm{cv}}(\boldsymbol{x},\bar{y})-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))=u^{\mathrm{cv}}(\boldsymbol{x},\underline{y})+u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))=u^{\mathrm{cv}}(\boldsymbol{x},\underline{y}).$$

If $\underline{y}\leq\zeta_u^{\min}(\boldsymbol{x})\leq\bar{y}$, then $\zeta_u^{\min}(\boldsymbol{x})=\mathrm{mid}(\zeta_u^{\min}(\boldsymbol{x}),\underline{y},\bar{y})$, and (10) becomes

$$u_I^{\mathrm{cv}}(\boldsymbol{x},\underline{y})+u_D^{\mathrm{cv}}(\boldsymbol{x},\bar{y})-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x})) \\ = u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))+u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))=u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x})).$$

If $\underline{y}\leq\bar{y}\leq\zeta_u^{\min}(\boldsymbol{x})$, then $\bar{y}=\mathrm{mid}(\zeta_u^{\min}(\boldsymbol{x}),\underline{y},\bar{y})$, and (10) becomes

$$u_I^{\mathrm{cv}}(\boldsymbol{x},\underline{y})+u_D^{\mathrm{cv}}(\boldsymbol{x},\bar{y})-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))=u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))+u^{\mathrm{cv}}(\boldsymbol{x},\bar{y})-u^{\mathrm{cv}}(\boldsymbol{x},\zeta_u^{\min}(\boldsymbol{x}))=u^{\mathrm{cv}}(\boldsymbol{x},\bar{y}).$$

In each case, the required result is satisfied. □

Lemma 25 *Consider a UIF $u:B\subset\mathbb{R}\to\mathbb{R}$ that satisfies Assumption 1, and an interval $\boldsymbol{x}\in\mathbb{I}B$. The functions $u_I^{\mathrm{cv}}(\boldsymbol{x},\cdot)$, $u_D^{\mathrm{cv}}(\boldsymbol{x},\cdot)$, $u_I^{\mathrm{cc}}(\boldsymbol{x},\cdot)$, and $u_D^{\mathrm{cc}}(\boldsymbol{x},\cdot)$ are each $\mathscr{C}^i$ on $\boldsymbol{x}$.*

Proof It will be shown that $u_I^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ and $u_D^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ are $\mathscr{C}^i$; the remaining results can be demonstrated analogously. The cases in which $\underline{x}<\zeta_u^{\min}(\boldsymbol{x})<\bar{x}$, $\zeta_u^{\min}(\boldsymbol{x})=\underline{x}$, or $\zeta_u^{\min}(\boldsymbol{x})=\bar{x}$ will be considered separately.

Suppose first that $\underline{x}<\zeta_u^{\min}(\boldsymbol{x})<\bar{x}$. Since the mapping $\phi:=u^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ is $\mathscr{C}^1$ on $\boldsymbol{x}$, regardless of the value of $i\in\{1,2\}$, it follows that $\nabla\phi(\zeta_u^{\min}(\boldsymbol{x}))=0$. Using this result, it is readily verified that $\phi_I:=u_I^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ and $\phi_D:=u_D^{\mathrm{cv}}(\boldsymbol{x},\cdot)$ are $\mathscr{C}^1$ on $\boldsymbol{x}$, with

$$\nabla\phi_I(z)=\begin{cases}0 & \text{if } z\leq\zeta_u^{\min}(\boldsymbol{x}),\\ \nabla\phi(z) & \text{if } z>\zeta_u^{\min}(\boldsymbol{x}),\end{cases}\qquad\text{and}\qquad \nabla\phi_D(z)=\begin{cases}\nabla\phi(z) & \text{if } z<\zeta_u^{\min}(\boldsymbol{x}),\\ 0 & \text{if } z\geq\zeta_u^{\min}(\boldsymbol{x}).\end{cases} \tag{11}$$

Furthermore, if $i=2$, then Assumption 1 implies that ϕ is $\mathscr{C}^2$ on $\boldsymbol{x}$, and that $\nabla^2\phi(\zeta_u^{\min}(\boldsymbol{x}))=0$. Using this result, it is readily verified that ϕ_D and ϕ_I are $\mathscr{C}^2$ on $\boldsymbol{x}$, with

$$\nabla^2\phi_I(z)=\begin{cases}0 & \text{if } z\leq\zeta_u^{\min}(\boldsymbol{x}),\\ \nabla^2\phi(z) & \text{if } z>\zeta_u^{\min}(\boldsymbol{x}),\end{cases}\qquad\text{and}\qquad \nabla^2\phi_D(z)=\begin{cases}\nabla^2\phi(z) & \text{if } z<\zeta_u^{\min}(\boldsymbol{x}),\\ 0 & \text{if } z\geq\zeta_u^{\min}(\boldsymbol{x}).\end{cases}$$

Next, suppose that either $\zeta_u^{\min}(\boldsymbol{x})=\underline{x}$ or $\zeta_u^{\min}(\boldsymbol{x})=\bar{x}$. In these cases, the functions ϕ_I and ϕ_D are each equivalent on $\boldsymbol{x}$ to either ϕ or to the constant mapping $\phi^*:z\mapsto\phi(\zeta_u^{\min}(\boldsymbol{x}))$, and are therefore $\mathscr{C}^i$ on $\boldsymbol{x}$. □

C.3 Establishing convergence order

This section shows that both natural and unconstrained $\mathscr{C}^i$ McCormick extensions are $(1,2)$-convergent, provided that each employed UIF satisfies Assumptions 1 and 2. Thus, convex/concave relaxations based on these McCormick extensions exhibit second-order pointwise convergence. Each univariate function in Table 2 satisfies Assumption 2 except the absolute-value function; the nonsmoothness of the absolute-value function prevents second-order pointwise convergence from being achievable [6, Example 5].

Lemma 26 *The squashing operation is $(1,2)$-convergent for each fixed $i\in\{1,2\}$.*

Proof Choose any $\mathscr{X} \in \mathbb{MR}$. If $\text{wid}\,\boldsymbol{x}^{\text{B}} = 0$, then

$$\text{wid}_{\mathscr{M}}\,(\mathscr{S}qu_i(\mathscr{X})) = 0 = \text{wid}_{\mathscr{M}}\,\mathscr{X} + 2a_p(\text{wid}\,\boldsymbol{x}^{\text{B}})^2.$$

If $\text{wid}\,\boldsymbol{x}^{\text{B}} > 0$, then, using Lemma 3, and noting that $\mathscr{S}qu_i(\mathscr{X}) \in \mathbb{MR}_{\text{prop}}$, it follows that:

$$\begin{aligned}
\text{wid}_{\mathscr{M}}\,(\mathscr{S}qu_i(\mathscr{X})) &= \text{wid}\,(\boldsymbol{belt}_i(\mathscr{X})) \\
&= \sigma_i(\overline{x}^{\text{C}}, \overline{x}^{\text{B}}, p_{\boldsymbol{x}^{\text{B}}}) - \gamma_i(\underline{x}^{\text{C}}, \underline{x}^{\text{B}}, p_{\boldsymbol{x}^{\text{B}}}) \\
&\leq \min\{\overline{x}^{\text{C}} + p_{\boldsymbol{x}^{\text{B}}}, \overline{x}^{\text{B}}\} - \max\{\underline{x}^{\text{C}} - p_{\boldsymbol{x}^{\text{B}}}, \underline{x}^{\text{B}}\} \\
&\leq \min\{\overline{x}^{\text{C}} + p_{\boldsymbol{x}^{\text{B}}}, \overline{x}^{\text{B}} + p_{\boldsymbol{x}^{\text{B}}}\} - \max\{\underline{x}^{\text{C}} - p_{\boldsymbol{x}^{\text{B}}}, \underline{x}^{\text{B}} - p_{\boldsymbol{x}^{\text{B}}}\} \\
&= \min\{\overline{x}^{\text{C}}, \overline{x}^{\text{B}}\} - \max\{\underline{x}^{\text{C}}, \underline{x}^{\text{B}}\} + 2p_{\boldsymbol{x}^{\text{B}}} \\
&= \text{wid}_{\mathscr{M}}\,\mathscr{X} + 2a_p(\text{wid}\,\boldsymbol{x}^{\text{B}})^2. \qquad (12)
\end{aligned}$$

Noting that $\mathscr{X}$ was chosen arbitrarily, the required result follows. □

Lemma 27 *The multiplication operation described in Definition 13 is* $(1,2)$*-convergent for each fixed* $i \in \{1,2\}$.

Proof Choose any $q \equiv (\boldsymbol{q}_1, \boldsymbol{q}_2) \in \mathbb{IR}^2$, and any $\mathscr{X} \in (\mathbb{M}\boldsymbol{q}_1)_{\text{prop}}, \mathscr{Y} \in (\mathbb{M}\boldsymbol{q}_2)_{\text{prop}}$, in which case $\text{wid}_{\mathscr{M}}\,\mathscr{X} \leq \text{wid}\,\boldsymbol{x}^{\text{B}} \leq \text{wid}\,\boldsymbol{q}_1$, and $\text{wid}_{\mathscr{M}}\,\mathscr{Y} \leq \text{wid}\,\boldsymbol{y}^{\text{B}} \leq \text{wid}\,\boldsymbol{q}_1$. Construct the interval $\boldsymbol{z} \in \mathbb{IR}$ described in Definition 13. Define $\mathscr{Z} := (\boldsymbol{x}^{\text{B}}\boldsymbol{y}^{\text{B}}, \boldsymbol{z}) \in \mathbb{MR}_{\text{prop}}$ and $p := p_{\boldsymbol{x}^{\text{B}}\boldsymbol{y}^{\text{B}}}$ for notational convenience.

Applying Lemma 3.9.19 in [31], and noting that $\boldsymbol{x}_1^{\text{B}} = \boldsymbol{x}^{\text{B}}$ and $\boldsymbol{y}_1^{\text{B}} = \boldsymbol{y}^{\text{B}}$ by construction, there exist $a_1, a_2 > 0$ (which may depend on $\boldsymbol{q}_1$, but are independent of $\mathscr{X}$ and $\mathscr{Y}$) for which

$$\text{wid}_{\mathscr{M}}\,(\mathscr{X} \bullet \mathscr{Y}) \leq a_1\,\text{wid}_{\mathscr{M}}\,(\mathscr{X}, \mathscr{Y}) + a_2(\text{wid}\,(\boldsymbol{x}^{\text{B}}, \boldsymbol{y}^{\text{B}}))^2.$$

(Recall that the symbol "•" refers to the classical McCormick product described in Definition 22.) Define the following intermediate quantities:

$$n_1 := \underline{(\underline{y}^{\text{B}}\boldsymbol{x}^{\text{C}})} + \underline{(\underline{x}^{\text{B}}\boldsymbol{y}^{\text{C}})} - \underline{x}^{\text{B}}\underline{y}^{\text{B}}, \qquad n_2 := \underline{(\overline{y}^{\text{B}}\boldsymbol{x}^{\text{C}})} + \underline{(\overline{x}^{\text{B}}\boldsymbol{y}^{\text{C}})} - \overline{x}^{\text{B}}\overline{y}^{\text{B}},$$

$$n_3 := \overline{(\underline{y}^{\text{B}}\boldsymbol{x}^{\text{C}})} + \overline{(\overline{x}^{\text{B}}\boldsymbol{y}^{\text{C}})} - \overline{x}^{\text{B}}\underline{y}^{\text{B}}, \qquad n_4 := \overline{(\overline{y}^{\text{B}}\boldsymbol{x}^{\text{C}})} + \overline{(\underline{x}^{\text{B}}\boldsymbol{y}^{\text{C}})} - \underline{x}^{\text{B}}\overline{y}^{\text{B}}.$$

Using Lemma 7,

$$\begin{aligned}
\text{wid}_{\mathscr{M}}\,\mathscr{Z} &\leq \text{wid}\,\boldsymbol{z} \\
&= \lambda_i(n_3, n_4, p) - \nu_i(n_1, n_2, p) \\
&\leq \tfrac{1}{2}\left(\min\{n_3 + p, n_4\} + \min\{n_3, n_4 + p\}\right) - \tfrac{1}{2}\left(\max\{n_1 - p, n_2\} + \min\{n_1, n_2 - p\}\right) \\
&\leq \min\{n_3 + p, n_4 + p\} - \max\{n_1 - p, n_2 - p\} \\
&= \min\{n_3, n_4\} - \max\{n_1, n_2\} + 2p \\
&= \text{wid}_{\mathscr{M}}\,(\mathscr{X} \bullet \mathscr{Y}) + 2p.
\end{aligned}$$

Define the absolute value of any interval $\boldsymbol{a} \in \mathbb{IR}$ as $|\boldsymbol{a}| := \max\{|\underline{a}|, |\overline{a}|\} \geq 0$. Using [25, Equation 4.3],

$$\text{wid}\,(\boldsymbol{x}^{\text{B}}\boldsymbol{y}^{\text{B}}) \leq |\boldsymbol{x}^{\text{B}}|\,\text{wid}\,\boldsymbol{y}^{\text{B}} + |\boldsymbol{y}^{\text{B}}|\,\text{wid}\,\boldsymbol{x}^{\text{B}} \leq |\boldsymbol{q}_1|\,\text{wid}\,(\boldsymbol{x}^{\text{B}}, \boldsymbol{y}^{\text{B}}).$$

Thus,

$$p \leq a_p(|\boldsymbol{q}_1|\,\text{wid}\,(\boldsymbol{x}^{\text{B}}, \boldsymbol{y}^{\text{B}}))^2.$$

Combining the above results, Lemma 26, and (12),

$$\begin{aligned}
\text{wid}_{\mathscr{M}}\,(\mathscr{X}\mathscr{Y}) &= \text{wid}_{\mathscr{M}}\,(\mathscr{S}qu_i(\mathscr{Z})) \\
&\leq \text{wid}_{\mathscr{M}}\,\mathscr{Z} + 2p \\
&\leq \text{wid}_{\mathscr{M}}\,(\mathscr{X} \bullet \mathscr{Y}) + 4p \\
&\leq a_1\,\text{wid}_{\mathscr{M}}\,(\mathscr{X}, \mathscr{Y}) + a_2(\text{wid}\,(\boldsymbol{x}^{\text{B}}, \boldsymbol{y}^{\text{B}}))^2 + 4p \\
&\leq a_1\,\text{wid}_{\mathscr{M}}\,(\mathscr{X}, \mathscr{Y}) + (4a_p|\boldsymbol{q}_1|^2 + a_2)(\text{wid}\,(\boldsymbol{x}^{\text{B}}, \boldsymbol{y}^{\text{B}}))^2;
\end{aligned}$$

this yields the required result, since a_1, a_2, a_p, and $|\boldsymbol{q}_1|$ are each independent of $\mathscr{X}$ and $\mathscr{Y}$. □

Theorem 7 *Given some $i^* \in \{1,2\}$ and a factorable function $f : B \subset \mathbb{R}^n \to \mathbb{R}^m$ whose composed UIFs each satisfy Assumptions 1 and 2 with $i := i^*$, any natural $\mathscr{C}^{i^*}$ McCormick extension $\mathscr{F} : \mathbb{M}B_{\text{prop}} \to \mathbb{MR}^m$ of f is $(1,2)$-convergent. Any unconstrained $\mathscr{C}^{i^*}$ McCormick extension $\mathscr{F}_{\text{unc}} : \mathbb{M}B \to \mathbb{MR}^m$ of f is also $(1,2)$-convergent.*

Proof As discussed in [31, Section 3.9.7], the composition of $(1,2)$-convergent functions is itself $(1,2)$-convergent. The addition operation $+ : \mathbb{MR}^2_{\text{prop}} \to \mathbb{MR}_{\text{prop}}$ is $(1,2)$-convergent [31, Lemma 3.9.17], as is any UIF which satisfies Assumption 2 [31, Lemma 3.9.23]. Lemmata 26 and 27 show that the squashing operation and the multiplication operation described in Definition 13 are each $(1,2)$-convergent as well. Combining these results, $\mathscr{F}$ and $\mathscr{F}_{\text{unc}}$ are each $(1,2)$-convergent. □

References

1. Adjiman, C.S., Dallwig, S., Floudas, C.A., Neumaier, A.: A global optimization method, αBB, for general twice-differentiable constrained NLPs – I. Theoretical advances. Computers Chem. Engng **22**, 1137–1158 (1998)
2. Alefeld, G., Mayer, G.: Interval analysis: theory and applications. J. Comput. Appl. Math. **121**, 421–464 (2000)
3. Beckers, M., Mosenkis, V., Maier, M., Naumann, U.: Adjoint mode computation of subgradients for McCormick relaxations. Tech. Rep. AIB-2011-10, RWTH Aachen (2011)
4. Bertsekas, D.P.: Nondifferentiable optimization via approximation. In: M. Balinski, P. Wolfe (eds.) Mathematical Programming Study 3, pp. 1–25. North-Holland Publishing Company, Amsterdam (1975)
5. Bertsekas, D.P.: Nonlinear Programming, second edn. Athena Scientific, Belmont, MA (1999)
6. Bompadre, A., Mitsos, A.: Convergence rate of McCormick relaxations. J. Glob. Optim. **52**, 1–28 (2012)
7. Broyden, C.G., Dennis Jr., J.E., Moré, J.J.: On the local and superlinear convergence of quasi-Newton methods. J. Inst. Math. Appl. **12**, 223–245 (1973)
8. Chachuat, B.: MC++: A toolkit for bounding factorable functions, v1.0. `https://projects.coin-or.org/MCpp` (2014). Retrieved online on July 2, 2014
9. Du, K., Kearfott, R.B.: The cluster problem in multivariate global optimization. J. Glob. Optim. **5**, 253–265 (1994)
10. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems, vol. 2. Springer-Verlag New York, Inc., New York, NY (2003)
11. Gabriel, S.A., Moré, J.J.: Smoothing of mixed complementarity problems. Preprint MCS-P541-0995, Argonne National Laboratory (1995)
12. Galán, S., Feehery, W.F., Barton, P.I.: Parametric sensitivity functions for hybrid discrete/continuous systems. Appl. Numer. Math. **31**, 17–47 (1999)
13. Griewank, A., Rabier, P.J.: On the smoothness of convex envelopes. T. Am. Math. Soc. **322**, 691–709 (1990)
14. Griewank, A., Walther, A.: Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, 2nd edn. Other Titles in Applied Mathematics. SIAM, Philadelphia, PA (2008)
15. Hartman, P.: Ordinary Differential Equations, second edn. SIAM, Philadelphia, PA (2002)
16. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms I. A Series of Comprehensive Studies in Mathematics. Springer-Verlag, Berlin (1993)
17. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms II. A Series of Comprehensive Studies in Mathematics. Springer-Verlag, Berlin (1993)
18. Horst, R., Tuy, H.: Global Optimization: Deterministic Approaches, second edn. Springer-Verlag, Berlin (1993)
19. Kesavan, P., Allgor, R.J., Gatzke, E.P., Barton, P.I.: Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. Math. Program., Ser. A **100**, 517–535 (2004)
20. Kiwiel, K.C.: Methods of Descent for Nondifferentiable Optimization. Lecture Notes in Mathematics. Springer-Verlag, Berlin (1985)
21. Li, X., Tomasgard, A., Barton, P.I.: Nonconvex generalized Benders decomposition for stochastic separable mixed-integer nonlinear programs. J. Optim. Theory Appl. **151**, 425–454 (2011)
22. Liberti, L., Pantelides, C.C.: Convex envelopes of monomials of odd degree. J. Glob. Optim. **25**, 157–168 (2003)
23. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: Part I - Convex underestimating problems. Math. Program. **10**, 147–175 (1976)
24. Mitsos, A., Chachuat, B., Barton, P.I.: McCormick-based relaxations of algorithms. SIAM J. Optim. **20**, 573–601 (2009)

25. Moore, R.E.: Methods and Applications of Interval Analysis. SIAM, Philadelphia (1979)
26. Naumann, U.: The Art of Differentiating Computer Programs. SIAM, Philadelphia (2012)
27. Neumaier, A.: Interval Methods for Systems of Equations. Cambridge University Press, Cambridge (1990)
28. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Classics in Applied Mathematics. SIAM, Philadelphia (2000)
29. Qi, L., Sun, D.: Smoothing functions and smoothing Newton method for complementarity and variational inequality problems. J. Optim. Theory App. **113**, 121–147 (2002)
30. Rudin, W.: Principles of Mathematical Analysis, third edn. McGraw-Hill, Inc., New York, NY (1976)
31. Schaber, S.D.: Tools for dynamic model development. Ph.D. thesis, Massachusetts Institute of Technology (2014)
32. Scholz, D.: Theoretical rate of convergence for interval inclusion functions. J. Glob. Optim. **53**, 749–767 (2012)
33. Scott, J.K.: Reachability analysis and deterministic global optimization of differential-algebraic systems. Ph.D. thesis, Massachusetts Institute of Technology (2012)
34. Scott, J.K., Barton, P.I.: Convex and concave relaxations for the parametric solutions of semi-explicit index-one differential-algebraic equations. J. Optim. Theory App. **156**, 617–649 (2013)
35. Scott, J.K., Barton, P.I.: Improved relaxations for the parametric solutions of ODEs using differential inequalities. J. Global Optim. **57**, 143–176 (2013)
36. Scott, J.K., Chachuat, B., Barton, P.I.: Nonlinear convex and concave relaxations for the solutions of parametric ODEs. Optim. Control Appl. Meth. **34**, 145–163 (2013)
37. Scott, J.K., Stuber, M.D., Barton, P.I.: Generalized McCormick relaxations. J. Glob. Optim. **51**, 569–606 (2011)
38. Stuber, M.D., Scott, J.K., Barton, P.I.: Convex and concave relaxations of implicit functions. Optim. Method Softw. **In press** (2014)
39. Tsoukalas, A., Mitsos, A.: Multivariate McCormick relaxations. J. Glob. Optim. **59**, 633–662 (2014)
40. Wechsung, A.: Global optimization in reduced space. Ph.D. thesis, Massachusetts Institute of Technology (2014)
41. Wechsung, A., Schaber, S.D., Barton, P.I.: The cluster problem revisited. J. Glob. Optim. **58**, 429–438 (2014)
42. Wechsung, A., Scott, J.K., Watson, H.A.J., Barton, P.I.: Reverse propagation of McCormick relaxations. J. Glob. Optim. **In press** (2015)
43. Whitney, H.: Analytic extensions of differentiable functions defined in closed sets. Trans. Amer. Math. Soc. **36**, 63–89 (1934)