Sequence of polyhedral relaxations for nonlinear univariate functions

Kaarthik Sundar1 · Sujeevraja Sanjeevi2 · Harsha Nagarajan1

Received: 2 November 2020 / Revised: 23 February 2021 / Accepted: 23 February 2021 / Published online: 17 April 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
Given a nonlinear, univariate, bounded, and differentiable function \( f(x) \), this article develops a sequence of Mixed Integer Linear Programming (MILP) and Linear Programming (LP) relaxations that converge to the graph of \( f(x) \) and its convex hull, respectively. Theoretical convergence of the sequence of relaxations to the graph of the function and its convex hull is established. For nonlinear non-convex optimization problems, the relaxations presented in this article can be used to construct tight MILP and LP relaxations. These MILP and the LP relaxations can also be used with MILP-based and spatial branch-and-bound based global optimization algorithms, respectively.

Keywords Piecewise relaxation · Univariate function · Global optimization · Convex hull · Continuous relaxation

1 Introduction
Developing convex relaxations of non-convex functions is a core part of any global optimization algorithm for nonlinear, non-convex optimization problems (Tawarmalani and Sahinidis (2013)). Traditional global optimization algorithms rely on convex relaxations of non-convex functions to construct a relaxation of the optimization problem and obtain bounds on the optimal objective value with very little computational effort. These relaxations are then embedded into a spatial branch-and-bound
(sB&B) framework to obtain a global optimal solution (Smith and Pantelides (1997), Pardalos and Romeijn (2013), Floudas and Pardalos (2014)). The past decade has seen the advent of MILP-based global optimization algorithms (Wickson and Karimi (2008), Misener et al. (2011), Teles et al. (2013), Castillo Castillo et al. (2018), Nagarajan et al. (2016), (2019), D’Ambrosio et al. (2012)) that solve a sequence of MILPs or convex MINLPs based on successively tighter piecewise convex relaxations to compute the optimal solution. The main motivation of this approach is the meteoric improvement in the speed of off-the-shelf MILP solvers like Gurobi (2019). Both sB&B-based and MILP-based global optimization algorithms have one core sub-problem each. The sB&B-based algorithms rely on constructing and solving convex relaxations of non-convex structures and MILP-based algorithms rely on solution to piecewise convex relaxations to non-convex structures. Though commercial sB&B-based solvers like BARON (Sahinidis (1996)) are fairly mature, they still lack support for a wide variety of functions like trigonometric, hyperbolic etc. Similarly, given the nascency of MILP-based global optimization methods (Misener et al. (2011), Castro (2016), Nagarajan et al. (2019)), their support for non-convex structures is even lesser. In particular, Misener et al. (2011), Castro (2016), Nagarajan et al. (2019) have extensively looked at polynomial optimization problems and developed many piecewise convex relaxations for a single multilinear term. For other polynomial terms, these papers have relied on piecewise convex relaxations of the equation $y = x^2$ and the idea of factorable functions to build piecewise relaxations of other (higher-order) non-convex polynomial terms.

Before we present literature related to this work, we illustrate the importance of developing polyhedral and piecewise polyhedral relaxations for univariate functions in global optimization algorithms. To this end, we consider the 1730 NLP instances in MINLPLib. A total of 20 nonlinear functions (see Table 3 in https://www.gams.com/latest/docs/UG_Parameters.html for their definitions) occur in these instances. The number of instances that contain each of these functions are shown in Fig. 1. Among the 20 functions, all but ‘mul’, ‘div’, ‘power’, ‘centropy’, ‘mod’, and ‘min’ are unary nonlinear functions i.e. univariate functions. Given any factorable MINLP, i.e. each nonlinear function in the MINLP is a factorable function, the MINLP can be equivalently reformulated to the following problem via introduction of sufficient number of artificial variables:

$$\min : \quad c \cdot x + d \cdot y$$ \hspace{1cm} (1a)

subject to: $A_1 x + A_2 y \leq b,$ \hspace{1cm} (1b)

$y_i = f_i(x) \quad \forall f_i \in \text{univariate functions},$ \hspace{1cm} (1c)

$y_j = g_j(x) \quad \forall g_j \in \text{multilinear functions},$ \hspace{1cm} (1d)

$l \leq x \leq u,$ \hspace{1cm} (1e)
To solve the MINLP in Eq. (1) to global optimality using a sB&B-based method, a convex relaxation is required for each Eq. in (1c) and (1d). On the other hand, for a MILP-based method, a piecewise convex relaxation for each Eq. in (1c) and (1d), given partitions for a subset of variables in the decision variable vector $\mathbf{x}$, is required (see Misener et al. (2011), Castro (2016), Nagarajan et al. (2019)). In the literature, Castro (2016), Nagarajan et al. (2019), Sundar et al. (2021) have developed many piecewise convex relaxations for multilinear functions and $y = x^2$ i.e., ‘mul’, ‘sqr’, and ‘power’. This has enabled very efficient MILP-based global solution methods that are comparable to state-of-the-art spatial branch-and-bound solvers like BARON to certain classes of problems like polynomial optimization problems (see Nagarajan et al. (2019)). Extending these methods to other problems in the MINLPLib necessitates development of piecewise relaxations for the other non-convex functions shown in Fig. 1. All functions other than the ones mentioned above lack piecewise convex relaxations. On the other hand, developing convex relaxations for general univariate functions would aid sB&B-based solvers like BARON to support a wider family of nonlinear non-convex functions. This article focuses on nonlinear univariate functions that are bounded and differentiable in their domains, and develops both polyhedral relaxations and piecewise polyhedral relaxations for these functions that can directly be integrated into sB&B-based and MILP-based global optimization approaches, respectively. This class of functions is vast, and subsumes 14 of the 20 nonlinear functions in Fig. 1. Furthermore, ‘abs’ and ‘min’ which do not fall under the category of functions in this article, have disjunctive and convex reformulations that can directly be integrated into existing global methods. In summary, given a nonlinear, bounded, and differentiable univariate function $f: \mathbb{R} \rightarrow \mathbb{R}$ and a decision variable vector $\mathbf{x} \in \mathbb{R}^n$, Eq. (1f)

$$\mathbf{x} \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2} \times \{0, 1\}^{n_3}.$$
function \( y = f(x) \) with a bounded domain \([c, u]\), the contributions of this article are as follows: (i) we develop a sequence of MILP and LP relaxations that converge to the graph of \( f(x) \) and its convex hull, respectively. (ii) For a given partition of the domain, we construct an MILP and an LP relaxation that can be directly utilized in MILP-based and sB&B-based global optimization algorithms.

The idea of developing polyhedral and piecewise polyhedral relaxations for non-linear, univariate functions is not new and has been addressed in the literature (see Gounaris and Floudas (2008), Liberti and Pantelides (2003), Kallrath and Rebennack (2014), Rebennack (2016)). The difference between a polyhedral relaxation and a piecewise polyhedral relaxation is that for the latter, the domain of the function is partitioned into sub-domains, a polyhedral relaxation is constructed for each sub-domain and a disjunctive formulation to capture the relaxation for the whole domain is obtained. A polyhedral relaxation, on the other hand, constructs a polyhedron that contains the entire function in its domain. Hence, by definition, a piecewise polyhedral relaxation is an MILP. Gounaris and Floudas (2008) develop a series of tight convex \( \alpha \)-BB under-estimators for univariate functions in \( C^2 \) and prove that it converges to the convex hull on one side of the function with increasing partition density. In Liberti and Pantelides (2003), the authors focus on developing convex envelopes for monomial terms with odd-degree. Kallrath and Rebennack (2014) develop piecewise relaxations for univariate functions by solving multiple non-convex MINLPs to global optimality. Finally, in Rebennack (2016), the author develops piecewise over- and under-estimators for the function \( x^2 \), and these estimators are essentially the same as the polyhedral relaxations developed in this article. Our work differs from the aforementioned works in the sense that (i) it develops a single framework for all nonlinear, univariate, bounded, and differentiable functions (i.e., \( C^1 \)) which provides both MILP (piecewise polyhedral) and LP (polyhedral) relaxations for a fixed partition of the domain. (ii) It has theoretical convergence properties for both the LP and the MILP relaxations with increasing partition density, similar to Gounaris and Floudas (2008). The theoretical convergence of the LP relaxations to the convex hull is a novel result and, to the best of our knowledge, is the first result that takes a step in the direction of characterizing the convex hull of a nonlinear, non-convex, bounded, and differentiable univariate function. (iii) It is valid for a wider class of univariate, nonlinear functions unlike Liberti and Pantelides (2003), Rebennack (2016). (iv) It can easily be integrated into state-of-the-art global optimization approaches, be it sB&B-based or MILP-based.

The remainder of the article is organized as follows. Sec. 2 introduces the formal problem statement and notations. Sec. 3 introduces the main intuition behind the relaxations geometrically and presents the formulations. Sec. 4 presents the theoretical properties of the relaxations. Finally, Sec. 5 presents the computational results with conclusions in Section 6.

2 Problem statement

We are given a nonlinear, univariate function \( f : [c, u] \to \mathbb{R} \) that is bounded and differentiable. The objective of the article is to develop a sequence of MILP relaxations converging to the graph of \( f(x) \) i.e., \( \mathcal{X} = \{(x, y) \in [c, u] \times \mathbb{R} : y = f(x)\}, \)
with corresponding LP relaxations converging to the convex hull of \( \mathcal{X} \), \( \text{conv}(\mathcal{X}) \). Throughout the rest of the article, we always assume the function under investigation is nonlinear, univariate, bounded, and differentiable. We first introduce some definitions and notations that will be used throughout the rest of the article.

**Definition 1** Partition of a closed interval – Given a closed interval \([\ell, u] \subseteq \mathbb{R}\), a partition \( p \) of \([\ell, u] \) is an ordered sequence of real numbers \((x_0, x_1, \ldots, x_k)\) such that \( \ell = x_0 < x_1 < \ldots < x_k = u \). We denote the set of partition points of \( p \) by \( \mathcal{P}(p) \), i.e., \( \mathcal{P}(p) = \{x_0, \ldots, x_k\} \), and the set of all possible partitions of \([\ell, u]\) by \( \mathcal{A}[\ell, u] \). Given \( p \in \mathcal{A}[\ell, u] \), each closed interval \([x_i, x_{i+1}] \) is referred to as a sub-interval of \( p \), and we let \( \mathcal{S}(\cdot) \) be the function mapping \( p \) to its set of sub-intervals.

**Definition 2** Break point of a function – Given a nonlinear, bounded, and differentiable function \( f : [\ell, u] \rightarrow \mathbb{R} \), \( b \in [\ell, u] \) is referred to as a break point of the function, if at \( b \), the function \( f \) changes from being convex to concave or vice versa.

**Definition 3** Base partition of a function – Given a function \( f : [\ell, u] \rightarrow \mathbb{R} \), the partition \( p_0^f \in \mathcal{A}[\ell, u] \) is referred to as the base partition of \( f \) if it satisfies: (i) every break point of \( f \) in \([\ell, u]\) is in \( \mathcal{P}(p) \), (ii) for any sub-interval \([x_i, x_{i+1}] \) in \( \mathcal{S}(p_0^f) \), \( f'(x_i) \neq f'(x_{i+1}) \), and (iii) \( |\mathcal{P}(p_0^f)| \) is minimum.

The above definition makes \( \mathcal{P}(p_0^f) \) non-unique due to condition (ii). To make it unique, without loss of generality, we construct \( p_0^f \) with a partition that includes only the break points of \( f \) in its domain, and add a partition point at the mid-point of every sub-interval \([x_i, x_{i+1}] \), if \( f'(x_i) = f'(x_{i+1}) \). The inclusion of break points in the definition of \( p_0^f \) ensures that in each sub-interval of \( p_0^f \), the function \( f \) is either concave or convex.

**Definition 4** Refinement of a partition – Given two partitions \( p, q \in \mathcal{A}[\ell, u] \), \( q \) is defined to be the refinement of \( p \), if \( \mathcal{P}(p) \subset \mathcal{P}(q) \).

Any function that refines a partition is also referred to as a refinement scheme. In this article, we will present all the properties of the forthcoming polyhedral relaxations of \( f \) with respect to the interval-bisection refinement scheme defined below.

**Definition 5** Interval bisection – Given any partition \( p \in \mathcal{A}[\ell, u] \), a refinement scheme \( \mathcal{B} : \mathcal{A}[\ell, u] \rightarrow \mathcal{A}[\ell, u] \) is referred to as interval bisection if is satisfies the following properties: (i) \( \mathcal{P}(p) \subset \mathcal{P}(\mathcal{B}(p)) \), \( |\mathcal{P}(\mathcal{B}(p))| = |\mathcal{P}(p)| + 1 \) and (ii) the additional partition point in \( \mathcal{P}(\mathcal{B}(p)) \setminus \mathcal{P}(p) \) is a mid-point of some sub-interval in \( p \) i.e., \( \mathcal{R}(p) \).

**Definition 6** Admissible partitions for a function – Given \( f : [\ell, u] \rightarrow \mathbb{R} \) and a partition \( p \in \mathcal{A}[\ell, u] \), \( p \) is defined to be admissible to \( f \) if \( p = p_0^f \), or if \( p \) is obtained by
applying a sequence of interval bisections on $p_j^0$. The set of all admissible partitions of $f$ is denoted by $A_j[\ell', u]$.

Next, we present a systematic way to construct base MILP and LP relaxations for a function $f : [\ell', u] \to \mathbb{R}$ and an admissible partition $p \in A_j[\ell', u]$.

3 Geometry of polyhedral relaxations

An MILP relaxation for the function $f : [\ell', u] \to \mathbb{R}$ is constructed using an admissible partition $p \in A_j[\ell', u]$. The relaxation constructed using $p_j^0$ is referred to as the base MILP relaxation of $f$. The core idea of the relaxation is: if $f$ is convex or concave in any closed interval, construct a triangle whose one side is the secant of $f$ through endpoints of the interval, and the other two sides are the tangents of $f$ at the endpoints; the region bounded by this triangle is a trivial relaxation of $f$ in that closed interval. Given $p_j^0$, the base MILP relaxation is obtained as the disjunctive union of a chain of triangles, one for each sub-interval in $\mathcal{S}(p_j^0)$. This construction is guaranteed to result in a relaxation of $f$ since the definition 3 ensures that $f$ is either concave or convex in each sub-interval. It is also guaranteed to be a relaxation for any admissible partition $p \in A_j[\ell', u] \setminus \{p_j^0\}$ since $p$ has to necessarily be some refinement of $p_j^0$. Given a sub-interval $[x_i, x_{i+1}] \in \mathcal{S}(p_j^0)$, the equations of the two tangents and the secant represented as a piecewise linear function and linear equation, respectively, are as follows:

$$h(x_i) = f(x_i) + f'(x_i) \cdot (x - x_i) \quad (2a)$$

$$t_{i+1}(x) = \begin{cases} 
\max \{ h(x_i), h(x_{i+1}) \} & \text{if } f(x) \text{ is convex in } [x_i, x_{i+1}] \\
\min \{ h(x_i), h(x_{i+1}) \} & \text{if } f(x) \text{ is concave in } [x_i, x_{i+1}] 
\end{cases} \quad (2b)$$

$$s_{i+1}(x) = f(x_i) + \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \cdot (x - x_i) \quad (2c)$$

The Eq. (2a) is the equation of the tangent of the function $f$ at $x_i$. When $f$ is convex in the sub-interval $[x_i, x_{i+1}]$, $t_{i+1}(x)$ (resp. $s_{i+1}(x)$) is the lower (resp. upper) bounding function and when $f$ is concave in the sub-interval, $t_{i+1}(x)$ (resp. $s_{i+1}(x)$) becomes the upper (resp. lower) bounding function. The base relaxation obtained for the function $y = x^3$ in $[-1.5, 2]$ is shown in Fig. 2. The function is convex in $[0, 2]$ and concave in $[-1.5, 0]$. Hence, in the sub-intervals $[0, 2]$ and $[-1.5, 0]$ the sides of the triangle formed by the tangents form the lower and upper bounding functions, respectively. In the example shown in Fig. 2, vertices $v_0, v_1,$ and $v_2$ are points on $y = f(x)$ and obtained using the base partition $p_j^0$. The vertex $v_{i,i+1}$ for $i \in \{0, 1\}$ (in Fig. 2) is obtained as the intersection of tangents at $v_i$ and $v_{i+1}$. Extrapolating this notation, given $f$ and $p \in A_j[\ell', u]$, the vertices $v_0, v_1, \ldots, v_k$ $(k = |\mathcal{P}(p)| - 1)$ are points on the curve $y = f(x)$ at partition points of $p$ and the vertex $v_{i,i+1}$ for $i \in \{0, \ldots, k - 1\}$ is the intersection of the two tangents to the curve at $v_i$ and $v_{i+1}$. The MILP relaxation for
Sequence of polyhedral relaxations for nonlinear univariate…

$X$, denoted by $\mathcal{R}(f, p)$, is then given by the disjunctive union of a chain of $k$ triangles where the vertices of triangle $i$ are $\{v_{i-1}, v_i, v_{i-1,i}\}$. We now define the strength of the relaxation $\mathcal{R}(f, p)$, given $f$ and $p$ as follows:

**Definition 7** Strength of $\mathcal{R}(f, p)$ – Given $f : [\ell, u] \rightarrow \mathbb{R}$, $p \in A_f[\ell, u]$ and the corresponding MILP relaxation $\mathcal{R}(f, p)$, the strength of the relaxation denoted by $\epsilon$ is given by

$$
\epsilon(\mathcal{R}(f, p)) = \max_{i \in \{1, \ldots, k\}} \max_{x \in [x_{i-1}, x_i]} |t_i(x) - s_i(x)|
$$

(3)

where $k = |\mathcal{P}(p)| - 1$, and $t_i(x)$ and $s_i(x)$ are given by Eq. (2).

Informally, the relaxation strength $\epsilon(\mathcal{R}(f, p))$ is defined as the maximum vertical distance over the disjunctive union of the $k$ triangles formed by secants and tangents at each sub-interval defined by the admissible partition $p$. We remark that the strength of the relaxation is always non-negative; lower the value of $\epsilon(\mathcal{R}(f, p))$, the tighter the relaxation. Next, we formulate the disjunctive union of the chain of $k$ triangles as a MILP.

**A formulation for $\mathcal{R}(f, p)$**: Again, we are given $f : [\ell, u] \rightarrow \mathbb{R}$ and an admissible partition $p \in A_f[\ell, u]$. The number of triangles in $\mathcal{R}(f, p)$ is given by $k = |\mathcal{P}(p)| - 1$. Each triangle $i$ is associated with a binary variable $z_i$ and two non-negative continuous variables $\delta_1^i, \delta_2^i$. Given any vertex $v$, we let $v^x$ and $v^y$ represent the $x$- and $y$-coordinate of $v$. With this notation, we write the MILP formulation of the disjunctive union of triangles as:

$$
y = v_0^y + \sum_{i=1}^{k} \left\{ \delta_1^i(v_{i-1,i}^y - v_{i-1}^y) + \delta_2^i(v_i^y - v_{i-1}^y) \right\}
$$

(4a)
The formulation (4) is referred to as the “incremental formulation” (Yıldız and Vielma (2013)) since the portion of each triangle is filled by the variables $x$ and $y$.

The triangle $i + 1$ can be filled only if the triangle $i$ is already filled i.e. when $\delta_2^i = 1$.

While it is possible to develop other formulations to capture the disjunctive union of triangles with the same theoretical properties as the one in Eq. (4), we refrain from doing so as that is not the focus of this article (see Yıldız and Vielma 2013 for details). We also remark that the MILP formulation for $R(f, p)$ is in the variable space $(x, y, \delta, z)$ where $\delta$ and $z$ are vectors of additional continuous and binary variables in Eq. (4). We will use the notation $\text{Proj}_{x,y} R(f, p)$ to denote the projection of $R(f, p)$ to the $(x, y)$-variable space. It is not difficult to prove that $\text{Proj}_{x,y} R(f, p)$ is actually the disjunctive union of the $k$ triangles in the $(x, y)$ space.

### 4 Sequence of relaxations and theoretical properties

Before we present the sequence of polyhedral (MILP) relaxations for the function $f$, we invoke some results from literature that will be used to prove their theoretical properties.

**Proposition 1** (Lemma 2 in Thakur (1980)) Given the function $f : [\ell, u] \rightarrow \mathbb{R}$, a sub-interval $[a, b]$ of $[\ell, u]$ where $f$ is convex or concave, and the partition that consists of the points $a$ and $b$, i.e., $p = (a, b) \in A[a, b]$, the strength of the relaxation $R(f, p)$ is bounded above according to the following inequality:

$$\epsilon(R(f, p)) \leq \frac{(b-a)|f'(a) - f'(b)|}{4}$$

**Proof** See Thakur (1980).

### 4.1 Algorithm for generating sequence of relaxations

The sequence of MILP relaxations for the function $f : [\ell, u] \rightarrow \mathbb{R}$ is obtained using the following procedure. The base relaxation is given by $R(f, p_0)$ where $p_0 = p^0_f$. A
sequence of partitions \{p_0, p_1, p_2, \ldots\} is then created according to the following procedure: Given \( p_k \) and \( \mathcal{R}(f, p_k) \), the upper bound on the strength of the relaxation in every sub-interval \( i \in \mathcal{A}(p_k) \) is computed using Eq. (5); we let \( \epsilon_i \) denote this bound for sub-interval \( i \). The sub-interval \( i \) with the maximum \( \epsilon_i \) is then divided into two equal parts. This refinement of the sub-interval is an interval bisection and results in the partition \( p_{k+1} \). This procedure results in an infinite sequence of partitions \( p_0, p_1, p_2, \ldots \) such that \( p_0 = p_0 \subseteq \mathcal{P}(p_{k+1}) \) and \( |\mathcal{P}(p_{k+1})| = |\mathcal{P}(p_{k})| + 1 \) for every \( k \). The infinite sequence of MILP relaxations is then given by \( \mathcal{R}(f, p_k) \) for \( k \in \mathbb{Z}_+ \). We now state and prove the main theoretical properties of this sequence.

**Proposition 2** (Theorem 2.3 in Burkard et al. (1991)) Given \( f : [\ell, u] \to \mathbb{R} \), \( \delta > 0 \), and a sub-interval \([a, b]\) where \( f \) is convex or concave, there exists a finite \( N \in \mathbb{Z}_+ \) such that \( \epsilon(\mathcal{R}(f, (a, b))) < \delta \) for any \( k \geq N \); in particular

\[
N \leq \max \left( 2, \frac{3}{2} \sqrt{\frac{b-a}{2\delta} \cdot |f''(b) - f''(a)|} \right)
\]

**Proof** See Burkard et al. (1991).

**Theorem 1** \( \lim_{k \to \infty} \epsilon(\mathcal{R}(f, p_k)) = 0 \).

**Proof** Given \( \delta > 0 \), we need a finite \( M \in \mathbb{Z}_+ \) such that for any \( k \geq M \), \( \epsilon(\mathcal{R}(f, p_k)) < \delta \). For each sub-interval in \( i \in \mathcal{A}(p^0) \), let \( N_i \) be the finite number provided by proposition 2 for the given value of \( \delta \). Setting \( M = \sum_i N_i \) completes the proof.

**Corollary 1** \( \lim_{k \to \infty} \text{Proj}_{x,y} \mathcal{R}(f, p_k) = X \).

We now present some theoretical properties of the linear programming (LP) relaxation of (4). To do so, we let \( \mathcal{R}^{lp}(f, p) \) be the LP relaxation of \( \mathcal{R}(f, p) \), and \( \text{Proj}_{x,y} \mathcal{R}^{lp}(f, p) \) the projection of \( \mathcal{R}^{lp}(f, p) \) to the \((x, y)\) space. To prove some properties of this LP relaxation, we invoke the definition of a *locally ideal* MILP formulation, introduced by Padberg and Rijal (2012): an MILP formulation of a set \( Y \) is said to be locally ideal if the vertices of its corresponding LP relaxation satisfy all required integrality conditions. Additionally, in any locally ideal MILP formulation of the set \( Y \), the projection of its LP relaxation to the original space of variables \((x, y\) in this case) is exactly the convex hull of \( Y \) i.e., \( \text{conv}(Y) \) (Jeroslow and Lowe (1984)).

**Proposition 3** The incremental MILP formulation in Eq. (4) i.e., the MILP formulation for \( \mathcal{R}(f, p) \), is locally ideal.

**Proof** The matrix for the constraint system in Eq. (4), ignoring the constraints defining \( x \) and \( y \), is a network matrix and hence is totally unimodular. Therefore, all the
extreme points of the LP relaxation of Eq. (4) satisfy the integrality restrictions on the vector of \( z \) variables.

\[ \square \]

\textbf{Corollary 2} Given \( f : [\ell, u] \rightarrow \mathbb{R} \) and an admissible partition \( p \in \mathcal{A}_f[\ell, u] \), \( \text{Proj}_{x,y} \mathcal{R}^p(f, p) \) is the convex hull of \( k \) triangles, where \( k = |\mathcal{P}(p)| - 1 \).

We remark the only vertices of the triangle that can constitute the extreme points of \( \text{Proj}_{x,y} \mathcal{R}^p(f, p) \) are in the set \( \mathcal{V} = \{v_0, v_k\} \cup \{v_{0,1}, \ldots, v_{k-1,k}\} \) where \( k = |\mathcal{P}(p)| - 1 \). This is because the other vertices can be expressed as a convex combination of vertices in \( \mathcal{V} \). In fact, when \( f \) is convex, \( \mathcal{V} \) is exactly the set of extreme points of \( \text{Proj}_{x,y} \mathcal{R}^p(f, p) \). Hence, an equivalent linear programming formulation for \( \text{Proj}_{x,y} \mathcal{R}^p(f, p) \) is given by the following convex hull description of vertices in \( \mathcal{V} \):

\[
\text{Proj}_{x,y} \mathcal{R}^p(f, p) = \text{Proj}_{x,y} \left\{ (x, y, \lambda) \in [\ell, u] \times \mathbb{R} \times \Delta_{|\mathcal{V}|} : \left( \frac{x}{y} \right) = \sum_{i=1}^{|\mathcal{V}|} \lambda_i w_i \right\} \quad (7)
\]

where, \( w_1, w_2, \ldots, w_{|\mathcal{V}|} \) are vertices in \( \mathcal{V} \), and \( \Delta_{|\mathcal{V}|} \) is a \(|\mathcal{V}|\)-dimensional simplex.

\textbf{Corollary 3} Given \( f : [\ell, u] \rightarrow \mathbb{R} \) and an admissible partition \( p \in \mathcal{A}_f[\ell, u] \), \( \text{conv}(\mathcal{X}) \subseteq \text{Proj}_{x,y} \mathcal{R}^p(f, p) \) is the convex hull of \( k \) triangles, where \( k = |\mathcal{P}(p)| - 1 \).

\textbf{Proof} The proof follows from Corollary 2 and the fact that \( \mathcal{X} \subseteq \mathcal{R}(f, p) \).

\[ \square \]

\textbf{Proposition 4} Given \( f : [\ell, u] \rightarrow \mathbb{R} \) and any infinite sequence of partitions, \( \{p_0, p_1, \ldots\} \) generated by the algorithm in Sec. 4.1 i.e., using an interval bisection with the bisection occurring in the interval with maximum vertical distance at that iteration, \( \text{Proj}_{x,y} \mathcal{R}^p(f, p_{k+1}) \subseteq \text{Proj}_{x,y} \mathcal{R}^p(f, p_k) \) for any \( k \in \mathbb{Z}_{\geq 0} \).

\textbf{Proof} The proof follows from the observation that for any infinite sequence of partitions \( \{p_0, p_1, \ldots\} \) generated using the procedure in Sec. 4.1, \( \text{Proj}_{x,y} \mathcal{R}(f, p_{k+1}) \subseteq \text{Proj}_{x,y} \mathcal{R}(f, p_k) \). This is true because the partition \( p_0 \) is the base partition of the function \( f \) and for any sub-interval \( s \in \mathcal{S}(p_0) \), \( f \) is either convex or concave in that sub-interval. This, combined with the fact that the MILP formulation for \( \mathcal{R}(f, p_k) \) is locally ideal for any \( k \in \mathbb{Z}_{\geq 0} \), implies \( \text{Proj}_{x,y} \mathcal{R}^p(f, p_{k+1}) \subseteq \text{Proj}_{x,y} \mathcal{R}^p(f, p_k) \).

\[ \square \]

Corollary 3 along with Proposition 4 implies that as algorithm in Sec. 4.1 is applied to any \( f : [\ell, u] \rightarrow \mathbb{R} \), it generates an infinite sequence of partitions \( \{p_0, p_1, \ldots\} \) such that the projections of the LP relaxations of \( \mathcal{R}(f, p_k) \) to the \((x, y)\) variable space form a decreasing sequence of sets with every projection containing the \( \text{conv}(\mathcal{X}) \). We now present a main theorem that states that this decreasing sequence of sets formed by the projection of the LP relaxations i.e., \( \text{Proj}_{x,y} \mathcal{R}^p(f, p_k) \) indeed converges to \( \text{conv}(\mathcal{X}) \). To do so, we introduce the following definitions and some known results from the literature.
Definition 8  Given two closed sets $A$ and $B$, the Hausdorff distance between the sets $A$ and $B$, denoted by $d_H(A,B)$ is defined as

$$d_H(A,B) = \max \left\{ \max_{a \in A} \min_{b \in B} d(a,b), \max_{b \in B} \min_{a \in A} d(a,b) \right\} \quad \text{(8)}$$

where, $d(a,b)$ is the Euclidean distance between the points $a$ and $b$.

Proposition 5  If $A$ is a polytope and $B$ is a closed convex subset of $A$, then there exists a vertex $v \in A$ such that $d_H(A,B) = \min_{b \in B} d(v,b)$ i.e., the Hausdorff distance between $A$ and $B$ is achieved at a vertex of $A$.

Proof  First we observe that if $B \subset A$, then $d_H(A,B) = \max_{a \in A} \min_{b \in B} d(a,b)$. Note that when $B$ is a closed convex set, the function $g(a) = \min_{b \in B} d(a,b)$ is a convex function. Hence, the problem of computing $d_H(A,B)$ reduces to maximizing a convex function over the polytope $A$ and it is known that the maximum for this problem is achieved at a vertex of $A$ (see Boyd and Vandenberghe (2004)). \(\square\)

We now present the main theorem that proves the convergence of the sequence of LP relaxations of the MILPs to $\text{conv}(X)$.

Theorem 2  $\lim_{k \to \infty} \text{Proj}_{x,y} R^{lp}(f, p_k) = \text{conv}(X)$.

Proof  Proving the convergence of the monotonically decreasing sequence of sets $\text{Proj}_{x,y} R^{lp}(f, p_k)$ to $\text{conv}(X)$ is equivalent to proving the statement that as $k \to \infty$, $d_H(\text{Proj}_{x,y} R^{lp}(f, p_k), \text{conv}(X)) \to 0$ (see Salinetti and Wets 1979). For finite $k$, due to Proposition 5, the Hausdorff distance between $\text{Proj}_{x,y} R^{lp}(f, p_k)$ and $\text{conv}(X)$, if strictly greater than zero, is achieved at some vertex of $\text{Proj}_{x,y} R^{lp}(f, p_k)$ that is the intersection of two tangents to the function $f(x)$ and this vertex cannot be contained in $\text{conv}(X)$. Furthermore, Burkard et al. (1991) prove that the Hausdorff distance between the two sets is bounded above by relaxation strength $\varepsilon(\mathcal{R}(f, p_k))$ given in Eq. (3). This shows that the sequence of Hausdorff distances, being non-negative, is bounded from below by 0 and is bounded from above by a sequence that converges to 0 (by Theorem 1). Hence, the sequence of Hausdorff distances converges to zero as well, thereby establishing the set-theoretic convergence of $\lim_{k \to \infty} \text{Proj}_{x,y} R^{lp}(f, p_k)$ to the limit $\text{conv}(X)$. \(\square\)

5 Computational results

The overarching theme of results presented in this section is to demonstrate the use of our relaxations to solve general mixed-integer nonlinear programs. To that end, instances and univariate functions are hand-picked to show the utility of these relaxations in different contexts. Benchmarking the effectiveness of these relaxations on a wider class of test instances is beyond the scope of this article. All the computational
experiments were implemented using the Julia programming language (Bezanson et al. (2017)) using JuMP v0.21.0 (Dunning et al. (2017)) and run on a MacBook Pro with a 2.9 GHz Dual-Core Intel Core i5 processor and 16GB RAM. The relaxations presented in this article are also made publicly available as a Julia package (see https://github.com/sujeevraja/PolyhedralRelaxations.jl).

Strength and size of the MILP relaxations—The first set of results is aimed at demonstrating the strength of the MILP relaxation for different univariate functions. For each of the chosen univariate function, we generate MILP relaxations using the interval bisection scheme on the base partition with two stopping criteria (i) the strength of the MILP relaxation (see Eq. (3)) is less than a pre-specified tolerance, denoted by $\epsilon$, and (ii) the number of partitions excluding the base partition introduced by the refinement scheme does not exceed a pre-specified number, denoted by $b$. We report the number of partitions introduced in the relaxation and the strength of the relaxation obtained using the stopping criteria (i) and (ii), respectively. Table 1 presents these results for four nonlinear non-convex univariate functions. The functions $\sin x$, $x \cdot |x|$ and $S(x)$ were chosen because these univariate functions frequently occur in optimization problems concerning power networks (Lu et al. (2018)), natural gas networks (Sundar and Zlotnik (2018)) and logistic regression (Kleinbaum et al. (2002)), respectively. Note that in Table 1, the lower the value of $\epsilon$, the tighter the relaxation.

Visualization of the LP relaxations—Here, we show how the LP relaxations evolve with increasing number of partitions to illustrate the convergence of the sequence of LP relaxations to the convex hull of the graph of the univariate function. To that end, we choose three univariate functions $\sin x$, $x \cdot |x|$, and $x^4 - x^3$. Figs. 3, 4, 5 show the projection of the LP relaxation in the space of $(x, y)$ variables with increasing partition count for these functions. In all figures, $b$ is the number of partitions excluding the base partition of the univariate functions considered. The red curve denotes the actual function; the green and the blue lines are the over- and under-estimators, respectively.

Relaxations for separable Mixed-Integer Nonlinear Programs from MINLPLib—The results presented in Tables 2 and 3 aim to show the strength of the MILP and the LP relaxations on separable Mixed-Integer nonlinear programs (MINLPs) and nonlinear programs (NLPs) instances from MINLPLib in Bussieck et al. (2003). MINLPs/NLPs are said to be separable if each nonlinear function in the problem can be

---

| $f(x)$ | $[\ell, u]$ | $p^0_f$ | # partitions | Strength of relaxation |
|--------|-------------|--------|--------------|-----------------------|
|        |             |        | $\epsilon = \infty$ | $\epsilon = 0.1$ | $\epsilon = 0.01$ | $b = 0$ | $b = 50$ | $b = 100$ |
| $\sin x$ | $[0, 2\pi]$ | $[0, \pi, 2\pi]$ | 2 | 12 | 28 | 1.5707 | 0.0009 | 0.0009 |
| $x^3$ | $[-1, 1]$ | $[-1, 0, 1]$ | 2 | 6 | 26 | 0.7500 | 0.0014 | 0.0004 |
| $x \cdot |x|$ | $[-2, 2]$ | $[-2, 0, 2]$ | 2 | 16 | 32 | 2.0000 | 0.0078 | 0.0020 |
| $S(x)$ | $[-5, 5]$ | $[-5, 0, 5]$ | 2 | 6 | 14 | 0.3042 | 0.0009 | 4.32E-5 |
Fig. 3  LP relaxations of the convex hull of $y = \sin x$ with a domain of $[0, 2\pi]$.

Fig. 4  LP relaxations of the convex hull of $y = x \cdot |x|$ with a domain of $[-1, 1]$. 
decomposed into a sum of univariate functions. Four separable instances (i) \textit{trig} (ii) \textit{ex4\_1\_1}, and (iii) \textit{fo7} (iv) \textit{ramsey} are chosen to illustrate the effectiveness of relaxations presented in the article.

The \textit{trig}, \textit{ramsey} and \textit{fo7} instances are chosen to illustrate the relaxations’ effectiveness on a separable NLPs with trigonometric functions, log, \(x^a\) where \(a\) is fractional and MINLP respectively. The instance \textit{ex4\_1\_1} is chosen to compare the quality of relaxations on different reformulations of the problem. We present the original NLP of \textit{ex4\_1\_1} (Eq. (9)) and its two equivalent reformulations (Eq. (10) and Eq. (11)) below:

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|c|}
\hline
instance & g-opt & \multicolumn{3}{|c|}{LB (relative gap %)} & \multicolumn{3}{|c|}{Computation time (sec)} \\
\hline
 & & \(\epsilon = \infty\) & \(\epsilon = 0.1\) & \(\epsilon = 0.01\) & \(\epsilon = \infty\) & \(\epsilon = 0.1\) & \(\epsilon = 0.01\) \\
\hline
\textit{trig} & -3.7625 & -5.0749 (25.86) & -3.7943 (0.84) & -3.7694 (0.18) & 1.97 & 1.02 & 3.55 \\
\textit{ramsey} & -2.4874 & -3.0637 (18.80) & -3.0637 (18.80) & -2.5309 (1.7) & 0.0 & 0.0 & 0.01 \\
\textit{ex4\_1\_1} & -7.4873 & -15.8046 (52.63) & -7.5239 (0.48) & -7.4892 (0.02) & 0.07 & 1.14 & 3.49 \\
\textit{ex4\_1\_1} & -7.4873 & -27.24E4 (99.99) & -7.5633 (1.00) & -7.4941 (0.09) & 0.01 & 25.23 & 161.23 \\
\textit{fo7} & 20.7298 & 16.7051 (24.09) & 18.7674 (10.46) & 20.5481 (0.88) & 3.0 & 9.15 & 50.79 \\
\hline
\end{tabular}
\caption{MILP relaxation values on MINLPLib instances}
\end{table}
The reformulations differ in the way univariate functions are lifted by addition of auxiliary variables. When the MILP and the LP relaxations are applied to the two reformulations \textit{ex4} \_\textit{1} \_\textit{1} \textit{a} and \textit{ex4} \_\textit{1} \_\textit{1} \textit{b}, the strength of the respective relaxations differ in the sense that the reformulation \textit{ex4} \_\textit{1} \_\textit{1} \textit{a} always produces better solutions than the reformulation \textit{ex4} \_\textit{1} \_\textit{1} \textit{b}. On the other hand, some computational effort is required to obtain the base partition of the univariate function in Eq. (10b) unlike the five univariate functions in Eq. (11b)–(11c). The takeaway from this exercise is that it is better to find base partitions of a single univariate function that is expressed as a sum/difference of univariate terms, instead of breaking them up into multiple univariate functions and using partitions of the functions. In Tables 2 and 3, the relative gap % is the gap between the lower bound (LB) and the global optimal solution in the column “g-opt”. Results from Table 2 show that just the application of the vanilla MILP relaxation, without additional enhancements, constructed using the algorithm in Sec. 4.1 yields a solution within 1% of the global optimum for the chosen instances. The results in Table 3 indicate that quick LP-based lower bounds
can also be obtained by utilizing the formulation in Eq. (7) to relax the graph of the nonlinear univariate function.

Relaxations for erf(x) and Γ(x)—This set of results is aimed at demonstrating the efficacy of the proposed relaxations for NLPs with the gamma function, Γ(x), and the Gaussian error function, erf(x). Both these functions are only defined using integrals as follows:

\[ Γ(x) = \int_{0}^{\infty} t^{x-1} \exp(-t) \, dt, \quad x > 0, \]  
\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} \exp(-t^2) \, dt. \]  

(12) \hspace{2cm} (13)

Here, we demonstrate another facet of the proposed relaxations, namely the reliance only on oracles that provide values of the function and its derivative. In other words, we do not require closed-form expressions of a function or its derivative to formulate its relaxations. Both functions considered in this set of results lack closed form expressions. For the gamma function, even the derivative does not have a closed form expression. Nevertheless, there do exist oracles that can provide values of these functions and their derivatives up to an arbitrary tolerance. To demonstrate the effectiveness of the relaxations on NLPs with the above functions, we consider the following two optimization problems:

\[ \text{erf:} \quad \text{min} \quad \text{erf}(x + y) + \sin x \cdot \exp \left( -\frac{y}{2} \right) \]  
subject to: \[ x^2 + y^2 \geq 10, \quad 0 \leq x, y \leq 10. \]  

\[ \text{gamma:} \quad \text{min} \quad Γ(x) \]  
subject to: \[ 0.5 \leq x \leq 5. \]

(14a) \hspace{2cm} (14b) \hspace{2cm} (15)

The globally optimal solutions for the NLPs in Eq. (14) and (15) are 0.0 (see McCormick (1985)) and 0.8856 (Edwards Deming and Colcord (1935)), respectively. These optimization problems cannot be directly provided to state-of-the-art global optimization solvers like BARON or SCIP. In the NLP in Eq. (14), we note that the objective contains a bilinear function. We utilize the transformation \[ w_1 = \sin x + \exp(-y/2), \quad w_2 = \sin x - \exp(-y/2) \] to transform \( \sin x \cdot \exp(-y/2) \) to \( (w_1^2 - w_2^2)/4 \). This enables direct use of the formulation proposed in this paper to compute tight lower bounds. The strength of the MILP and the LP relaxations for these two NLPs are shown in Table 4. For this set of results, we do not report computation times since all runs were completed within a fraction of a second. In Table 4, MILP-LB and the LP-LB are values of the lower bound obtained by solving the piecewise polyhedral relaxations and the polyhedral relaxations, respectively, for the corresponding value of \( \epsilon \). As mentioned in the introduction, the relaxations for functions like the error function and the gamma function (shown to be very effective in Table 4) can directly be integrated into either sB&B-based or MILP-based global
optimization algorithms resulting in global solution techniques to MINLPs with a wider class of non-convex structures.

6 Conclusion

We present a systematic way to construct a sequence of MILP and LP relaxations for the graph of a univariate function in \( C^1 \left[ \ell, u \right] \) i.e., bounded, continuous, and differentiable with domain \( \left[ \ell, u \right] \). The sequence of MILP and LP relaxations is proved to converge to the graph of the function and to its convex hull when the domain of the function is partitioned with an interval bisection scheme. To the best of our knowledge, this is the first work in the literature that shows such convergence results for both MILP and the LP relaxations in a single framework. Future work would focus on studying the efficacy of these relaxations when put to use in both sB&B-based and MILP-based global optimization algorithms for general non-convex optimization problems.

Acknowledgements The work was funded LANL’s Directed Research and Development (LDRD) projects, “20170201ER: POD: A Polyhedral Outer-approximation, Dynamic-discretization optimization solver” and “20200603ECR: Distributed Algorithms for Large-Scale Ordinary Differential/Partial Differential Equation (ODE/PDE) Constrained Optimization Problems on Graphs”. This work was carried out under the U.S. DOE Contract No. DE-AC52-06NA25396.

References

Burkard RE, Hamacher HW, Günter R (1991) Sandwich approximation of univariate convex functions with an application to separable convex programming. Naval Res Logist (NRL) 38(6):911–924
Busseieck MR, Drud AS, Alexander M (2003) MINLPLib-a collection of test models for mixed-integer nonlinear programming. Inf J Comput 15(1):114–119
Castillo CPA, Castro PM, Vladimir M (2018) Global optimization of MIQCPs with dynamic piecewise relaxations. J Global Optim 71(4):691–716
Castro PM (2016) Normalized multiparametric disaggregation: an efficient relaxation for mixed-integer bilinear problems. J Global Optim 64(4):765–784
Claudia D, Jon L, Andreas W (2012) An algorithmic framework for MINLP with separable non-convexity Mixed integer nonlinear programming. Springer, New York, pp 315–347
Edwards DW, Cocord CG (1935) The minimum in the gamma function. Nature 135(3422):917–917
Floudas CA, Pardalos PM (2014) Recent advances in global optimization. Princeton University Press, Princeton
Gabriella S, Wets RJB (1979) On the convergence of sequences of convex sets in finite dimensions. SIAM Rev 21(1):18–33
Gounaris CE, Floudas CA (2008) Tight convex underestimators for mathematical\textsuperscript{2}-continuous problems: I univariate functions. J Global Optim 42(1):51–67
Gurobi Optimization. Gurobi 8 performance benchmarks, 2019. \url{https://www.gurobi.com/pdfs/benchmarks.pdf}
Harsha N, Mowen L, Emre Y, Russell B (2016) Tightening McCormick relaxations for nonlinear programs via dynamic multivariate partitioning. In International Conference on Principles and Practice of Constraint Programming, pages 369–387. Springer
Harsha N, Mowen L, Site W, Russell B, Kaarthik S (2019) An adaptive, multivariate partitioning algorithm for global optimization of nonconvex programs. J Global Optim 74(4):639–675
Iain D, Joey H, Miles L (2017) JuMP: a modeling language for mathematical optimization. SIAM Rev 59(2):295–320
Jeff B, Alan E, Stefan K, Shah VB (2017) Julia: a fresh approach to numerical computing. SIAM Rev 59(1):65–98
Jeroslow RG, Lowe JK (1984) Modelling with integer variables Mathematical programming at oberwolfach II. Springer, New York
Josef K, Steffen R (2014) Computing area-tight piecewise linear overestimators, underestimate and tubes for univariate functions Optim Sci Eng. Springer, New York, pp 273–292
Kaarthik S, Anatoly Z (2018) State and parameter estimation for natural gas pipeline networks using transient state data. IEEE Trans Control Syst Technol 27(5):2110–2124
Kaarthik S, Harsha N, Jeff L, Site W, Russell B (2021) Piecewise polyhedral formulations for a multilinear term. Op Res Lett 49(1):144–149
Kleinbaum DG, Dietz K, Gail M, Mitchell K, Mitchell K (2002) Logistic regression. Springer, New York
Leo L, Pantelides CC (2003) Convex envelopes of monomials of odd degree. J Global Optim 25(2):157–168
McCormick GP (1985) Global solutions to factorable nonlinear optimization problems using separable programming techniques
Misener R, Thompson JP, Floudas CA (2011) APOGEE: global optimization of standard, generalized, and extended pooling problems via linear and logarithmic partitioning schemes. Comput Chem Eng 35(5):876–892
Mohit T, Sahinidis NV (2013) Convexification and global optimization in continuous and mixed-integer nonlinear programming: theory, algorithms, software, and applications. Springer, New York
Mowen L, Harsha N, Russell B, Eksioglu SD, Mason SJ (2018) Tight piecewise convex relaxations for global optimization of optimal power flow. In Power Systems Computation Conference (PSCC), pages 1–7. IEEE
Padberg MW, Rijal MP (2012) Location, scheduling, design and integer programming, vol 3. Springer, New York
Pardalos PM, Edwin RH (2013) Handbook of global optimization. Springer, New York
Sahinidis NV (1996) Baron: a general purpose global optimization software package. J Global Optim 8(2):201–205
Steffen R (2016) Computing tight bounds via piecewise linear functions through the example of circle cutting problems. Math Methods Op Res 84(1):3–57
Stephen B, Lieven V (2004) Convex Optim. Cambridge University Press, Cambridge
Smith EMB, Pantelides CC (1997) Global optimisation of nonconvex MINLPs. Comput Chem Eng 21:S791–S796
Teles JP, Castro PM, Matos HA (2013) Univariate parameterization for global optimization of mixed-integer polynomial problems. Europ J Op Res 229(3):613–625
Thakur LS (1980) Error analysis for convex separable programs: bounds on optimal and dual optimal solutions. J Math Anal Appl 75(2):486–494
Wicaksono DS, Karimi IA (2008) Piecewise MILP under-and overestimators for global optimization of bilinear programs. AIChE J 54(4):991–1008
Yildiz S, Vielma JP (2013) Incremental and encoding formulations for mixed integer programming. Op Res Lett 41(6):654–658

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.