Optimization Letters (2022) 16:2221–2240
https://doi.org/10.1007/s11590-022-01852-2

ORIGINAL PAPER

Convex envelopes for ray-concave functions

Javiera Barrera[1] · **Eduardo Moreno**[1] · **Gonzalo Muñoz**[2]

Received: 7 May 2021 / Accepted: 12 January 2022 / Published online: 7 February 2022

Abstract

Convexification based on convex envelopes is ubiquitous in the non-linear optimization literature. Thanks to considerable efforts of the optimization community for decades, we are able to compute the convex envelopes of a considerable number of functions that appear in practice, and thus obtain tight and tractable approximations to challenging problems. We contribute to this line of work by considering a family of functions that, to the best of our knowledge, has not been considered before in the literature. We call this family *ray-concave* functions. We show sufficient conditions that allow us to easily compute closed-form expressions for the convex envelope of ray-concave functions over arbitrary polytopes. With these tools, we are able to provide new perspectives to previously known convex envelopes and derive a previously unknown convex envelope for a function that arises in probability contexts.

Keywords Convex envelopes · Nonlinear programming · Convex optimization

1 Introduction

Strong convex relaxations of complex optimization problems are a key aspect of the development of tractable computational techniques in the field. In this regard, a popular approach has been the study of *convex underestimators* of functions, that is, given an arbitrary function f, a convex function f' such that $f'(x) \leq f(x)\ \forall x \in P$, where P is a given convex set. Such a function can be used to *relax* a sub-level set $\{x \in P : f(x) \leq 0\}$ with the convex set $\{x \in P\ :\ f'(x) \leq 0\}$, and thus obtain a computationally

✉ Gonzalo Muñoz
gonzalo.munoz@uoh.cl

Javiera Barrera
javiera.barrera@uai.cl

Eduardo Moreno
eduardo.moreno@uai.cl

1 Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Santiago, Chile

2 Institute of Engineering Sciences, Universidad de O'Higgins, Rancagua, Chile

tractable approximation. The pointwise largest convex underestimator is known as the *convex envelope of f over P*, and the optimization community has allocated considerable efforts on finding such envelopes for various classes of functions f and sets P.

Definition 1 The convex envelope of a function f over a set P is given by

$$\text{conv } f(x) = \sup\{g(x) : g \text{ is convex and } g(x) \leq f(x) \, \forall x \in P\}$$

We also make references to the concave envelope, denoted conc f, which is analogously defined.

In this work, we consider the case where P is a polytope and study the convex envelope of a family of functions that are convex when restricted to the facets of P, and exhibit what we define as *ray-concavity*.

Definition 2 A function $f : P \to \mathbb{R}$ is *ray-concave* over P if, for every $x \in P$, the function f restricted to $\{\alpha x : \alpha \geq 0\} \cap P$ is concave.

We present sufficient conditions for deriving simple closed-form formulas of the convex envelopes of ray-concave functions over arbitrary polytopes in any dimension.

Our result is closely related to known results for general functions over polytopes. To the best of our knowledge, the vast majority of the work producing closed-form formulas of convex envelopes in arbitrary dimension either require a rectangular domain, or require f to be *edge-concave*, in which case the convex envelope is polyhedral[1]. With our result, through the concept of ray-concavity, we are able to explicitly construct convex envelopes which are not necessarily polyhedral, in any dimension, for a new family of functions that has not been explicitly exploited before in the literature.

Our result yields a previously unknown convex envelope of a function that appears in probability contexts.

Example 1 The function $f(x, y) = -\frac{xy}{x+y-xy}$ is ray-concave over any box $[0, u_x] \times [0, u_y]$ with $u_x, u_y \leq 1$.

This function is one of the the main motivations behind this work. Additionally, many functions for which convex envelope formulas are known exhibit ray-concavity (e.g., $f(x_1, x_2) = -x_1 x_2$ or $f(x_1, x_2) = x_1/x_2$ for $x_1, x_2 > 0$), and our result provide a new perspective on these expressions and alternative derivations.

2 Literature review

The literature of convex envelopes is vast. Probably the most well-known and used convex envelope is that of the bilinear function $f(x_1, x_2) = x_1 x_2$ over a rectangular region, for which the convex (and concave) envelope is obtained through the McCormick envelopes [1, 21].

[1] A function is *polyhedral* if its epigraph is a polyhedron.

To the best of our knowledge, the first method capable of constructing the convex envelope for a family of functions (as opposed to a particular function) is provided in [33]. Based on disjunctive programming, they show a general expression of the convex envelope for functions that are concave on one variable, convex on the rest, and defined over a rectangular region. Later on, in [7] the authors show how to compute the evaluation of conv f when f is an $(n-1)$-convex function (i.e., f is convex whenever one variable is fixed to any value) over a rectangular domain. The function evaluation requires the solution of a convex optimization problem. In [10, 11], the authors formulate the convex envelope of a lower semi-continuous function over a compact set as a convex optimization problem. They use this to compute, explicitly, the convex envelope for various functions that are the product of a convex function and a component-wise concave function, over a box. We remark that in all the aforementioned cases, the convex envelopes may be non-polyhedral, and that explicit calculations consider hyper-rectangular domains.

Considerable efforts have been put toward the case of polyhedral convex envelopes. In [30, 31], it is shown that *edge-concavity* of a function f (i.e., concavity over all edge directions of a polytope P) implies that the convex envelope of f over P is polyhedral. The construction of these convex envelopes is studied in [23]. In [25], necessary and sufficient conditions for the convex envelope to be polyhedral are also provided, and they are used to obtain the convex envelope of a multilinear function over the unit box (see also [26, 28]). In [22], the authors provide explicit expressions for the facets of the convex envelope of a trilinear monomial over a box. In [3], the authors design a cutting plane approach to generate, on-the-fly, the convex envelope of a bilinear function over a box. The strength of the convex underestimator of a bilinear function that is obtained from using a term-wise convex envelopes is analyzed in [20].

Other known results include the convex envelopes of odd-degree monomials over an interval [14] and the fractional function $f(x_1, x_2) = x_1/x_2$ over a rectangle [33, 34]. Recently, in [18] the author computed the convex envelope of cubic functions in two dimensions, over a rectangular region.

While a big portion of these works involve rectangular regions, there exists important work considering sets beyond boxes in two dimensions. In [29], the authors derive explicit formulas for the convex envelope of bilinear bivariate functions over a class of special polytopes called D-polytopes. The case of the fractional function x_1/x_2 over a trapezoid is studied in [12]. This was expanded in [4], where convex envelopes for bilinear and fractional bivariate functions over quadrilaterals are constructed. The convex envelope of a bilinear bivariate function over a triangle has been carefully studied in [2, 15, 29]. Such envelopes were tested computationally in [15] within a branching scheme for QCQPs with positive results. A closed-form expression for the convex envelope of x_1x_2 over a box intersected with a halfspace was obtained in [5]. In [19] it is shown how to evaluate the convex envelope, and obtain a supporting hyperplane, for bivariate functions over arbitrary polytopes. This approach involves solving a low-dimensional convex problem. This procedure was refined in [17], by shifting the calculations to the solution of a KKT system. These last techniques were extensively tested in [24] to improve general-purpose optimization routines. In [16], the author characterizes the convex envelope of various bivariate functions (including the bilinear and fractional functions) over arbitrary polytopes using a polyhedral sub-division of the polytopes. In some cases, the convex envelope in each element of the sub-division can be given explicitly.

To the best of our knowledge, there is no construction that can provide a closed-form formula for the convex envelope of Example 1, and almost no construction allowing the explicit computation of non-polyhedral convex envelopes over polytopes beyond boxes in dimension $n \geq 3$. The only exception that we are aware of is [32]. In this work, the authors derive explicit convex and concave envelopes of several functions on sub-sets of a hyper-rectangle, which are obtained through polyhedral subdivisions. In this case the authors can obtain, in closed form, the convex envelope of disjunctive functions of the form $xf(y)$, and the concave envelope of *concave-extendable* supermodular functions. This may produce non-polyhedral envelopes. We remark some similarities with their construction below, however, it is worth noting that the results in [32] cannot directly provide a formula for conv f for the function f in Example 1. On one hand, this function does not fit the disjunctive framework of [32], so we cannot apply their convex envelope construction. On the other hand, one could consider using their concave envelope results with $-f$, thus effectively constructing conv f. However, the function $-f$ is not concave-extendable from the vertices of the box $[0, u_x] \times [0, u_y]$ (this can be inferred from the expression of conc $(-f)$ we obtain in Sect. 4).

3 Convex envelopes for ray-concave functions

Overall, we consider a polytope $P \subset \mathbb{R}^n$ with non-empty interior.

Definition 3 For any $v \in \mathbb{R}^n \setminus \{0\}$ such that $\exists\, \alpha \geq 0, \alpha v \in P$ (i.e., the ray defined by v intersects the polytope) we define

$$v^+ = \alpha^+ v, \text{ where } \alpha^+ = \arg\max\{\alpha \,:\, \alpha \geq 0,\ \alpha v \in P\} \qquad \text{and}$$
$$v^- = \alpha^- v, \text{ where } \alpha^- = \arg\min\{\alpha \,:\, \alpha \geq 0,\ \alpha v \in P\}.$$

In simple words, v^+ and v^- are the intersections of the ray given by v with the boundary of P (see Fig. 2). Note that if $0 \in P$ then $v^- = 0$ for all $v \in P$.

We remark that $v^\pm$ are continuous functions of v. Below, we emphasize this functional aspect when taking derivatives.

Using this definition, a function $f : P \to \mathbb{R}$ is ray-concave iff f restricted to the segment $[v^-, v^+]$ is concave for all v where $v^\pm$ is well defined. Our main results provides an explicit characterization for the convex envelope of ray-concave functions that are convex on the facets of P.

Theorem 1 *Let $f : P \to \mathbb{R}$ be a continuously differentiable and ray-concave function over a polytope P, such that f is convex over the facets of P. Let $g : P \to \mathbb{R}$ be defined as*

$$g(v) = \alpha_v f(v^-) + (1 - \alpha_v) f(v^+), \tag{1}$$

where $\alpha_v \in [0, 1]$ *is such that* $v = \alpha_v v^- + (1 - \alpha_v) v^+$. *If* g *is positively homogeneous, then conv* $f = g$.

Remark 1 In Sect. 3.3 we provide more insights on the positively homogeneous requirement. For example, we show that whenever $0 \in P$, g is positively homogeneous iff $f(0) = 0$. The latter is not a restrictive requirement, as we can compute the convex envelope of $f - f(0)$ instead.

A linear interpolation of a similar type as (1) has been considered in multiple articles. The general result in [7], for example, shows that to evaluate conv f for an edge-convex function f over a box, it suffices to consider the lines passing through x where the function f is concave, similarly to our result. Each evaluation involves solving an optimization problem (see [7, Theorem 3.1]). Another example is given by [32], who construct envelopes explicitly using secants of a similar type. In [15], the author also uses such lines in his construction of convex envelopes of the bilinear function over triangles.

In our case, by considering ray-concavity, we only need to consider secants on the rays emanating from the origin in the envelope construction.

To prove Theorem 1, we first provide three lemmas about the convexity of the function g over different regions of the domain. We divide the polytope P into subregions using the rays that pass through the vertices of P.

Definition 4 Let $\mathcal{F}$ be the set of facets of P. If $0 \notin P$, for each pair of facets $F_i, F_j \in \mathcal{F}$ we define the region

$$B_{ij} = \{v \in P : v^- \in F_i, v^+ \in F_j\}.$$

We refer to the hyperplane containing the facet F_i as the *in-hyperplane of* B_{ij}, and to the hyperplane containing the facet F_j as the *out-hyperplane of* B_{ij}. Alternatively, if $0 \in P$, for each facet $F_j \in \mathcal{F}$ we define the region

$$B_{0j} = \{v \in P : v^+ \in F_j\}.$$

In this case we only define the *out-hyperplane of* B_{0j}. We denote by $\mathcal{B}$ the set of all full-dimensional regions B_{ij}.

In Fig. 1 we illustrate the regions we consider in $\mathcal{B}$, which clearly form a sub-division of P. Note that if $B_{ij} \neq \emptyset$ then $B_{ji} = \emptyset$.

Also note that every $B_{ij} \in \mathcal{B}$ is polyhedral: for example, in the case $0 \notin P$, it is not hard to see that

$$B_{ij} = \text{cone}(F_i) \cap \text{cone}(F_j) \cap P. \tag{2}$$

Polyhedrality follows since both F_i and F_j are polyhedra.

Remark 2 For a given region $B_{ij} \in \mathcal{B}$ we can provide an explicit formula for $v^{\pm}$. In fact, note that we can assume that the out-hyperplane of B_{ij} has the form $a^{+\mathsf{T}}\mathbf{x} = 1$. Since

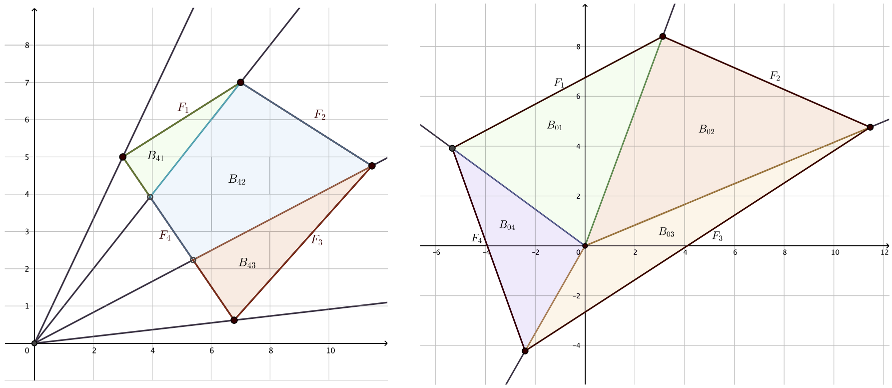

Fig. 1 Polyhedral sub-division of P into regions $\mathcal{B}$ according to intersection of rays with the boundary

v and v^+ lie on the same ray, we obtain $v^+ = \frac{1}{a^{+\mathsf{T}}v}v$ for any $v \in B_{ij}$. Similarly, for the case $0 \notin P$, we may assume that the in-hyperplane of B_{ij} has the form $a^{-\mathsf{T}}\mathbf{x} = 1$, and then $v^- = \frac{1}{a^{-\mathsf{T}}v}v$ for any $v \in B_{ij}$.

Moreover, since $v = \alpha_v v^- + (1-\alpha_v)v^+$, this implies that

$$\frac{\alpha_v}{a^{-\mathsf{T}}v} + \frac{1-\alpha_v}{a^{+\mathsf{T}}v} = 1 \tag{3}$$

3.1 Convexity and differentiability over a single region

To show convexity of g, we first prove that under the homogeneity assumption of Theorem 1, g is convex in each region $B \in \mathcal{B}$.

Lemma 1 *Let $B \in \mathcal{B}$ and let $g : B \subset \mathbb{R}^n \to \mathbb{R}$ as defined in (1). If g is positively homogeneous, then g is convex in B.*

Proof Let $u, w \in B$ and let $v = \lambda u + (1-\lambda)w$ for $\lambda \in [0, 1]$. By convexity of the region, $v \in B$ as well. To prove the convexity of g over B, we show that $g(v) \leq \lambda g(u) + (1-\lambda)g(w)$.

Recall that u^+ and w^+ belong to the same facet defining B, and u^- and w^- are either 0 (if $0 \in P$) or belong to the same facet defining B (if $0 \notin P$). Hence, there exist $\gamma, \varepsilon, \rho \in [0, 1]$ such that:

$$\begin{aligned} v &= \gamma v^- + (1-\gamma)v^+ \\ v^- &= \varepsilon u^- + (1-\varepsilon)w^- \\ v^+ &= \rho u^+ + (1-\rho)w^+ \end{aligned}$$

In Fig. 2 we illustrate these vectors.

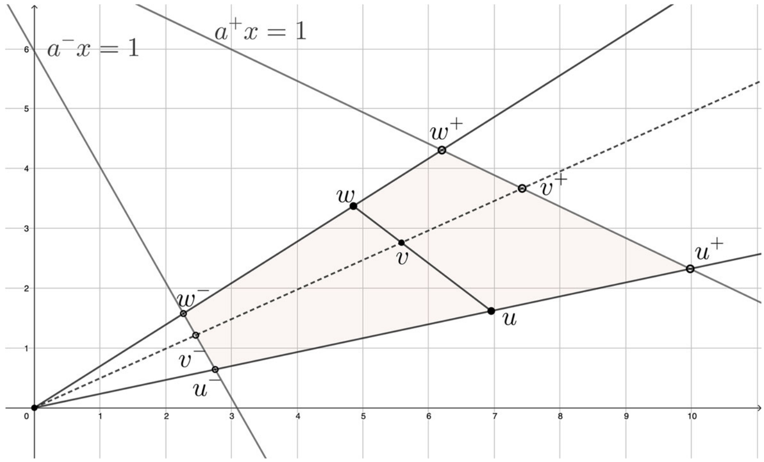

Fig. 2 Notation for Lemma 1

Since $g(v) = \gamma f(v^-) + (1-\gamma) f(v^+)$ and f is convex on the facets containing $\{u^+, v^+, w^+\}$ and $\{u^-, v^-, w^-\}$ (if $0 \notin P$), we know that

$$\begin{aligned} g(v) &= \gamma f(v^-) + (1-\gamma) f(v^+) && (4)\\ &\leq \gamma(\varepsilon f(u^-) + (1-\varepsilon) f(w^-)) + (1-\gamma)(\rho f(u^+) + (1-\rho) f(w^+)) && (5)\\ &= \gamma\varepsilon f(u^-) + (1-\gamma)\rho f(u^+) + \gamma(1-\varepsilon) f(w^-) + (1-\gamma)(1-\rho) f(w^+) && (6) \end{aligned}$$

Let ${a^+}^\intercal \mathbf{x} = 1$ be the out-hyperplane of B, i.e., the hyperplane that contains u^+, w^+, and v^+. By Remark 2, we know that

$$v^+ = \frac{1}{{a^+}^\intercal v} v = \frac{1}{{a^+}^\intercal v}(\lambda u + (1-\lambda) w) = \underbrace{\lambda \frac{{a^+}^\intercal u}{{a^+}^\intercal v}}_{\rho} u^+ + \underbrace{(1-\lambda)\frac{{a^+}^\intercal w}{{a^+}^\intercal v}}_{1-\rho} w^+$$

where we deduce $\rho = \lambda \frac{{a^+}^\intercal u}{{a^+}^\intercal v}$ because ${a^+}^\intercal v = \lambda {a^+}^\intercal u + (1-\lambda) {a^+}^\intercal w$. In a similar way, when $0 \notin P$ we can apply the same for v^- we obtain

$$\varepsilon = \lambda \frac{{a^-}^\intercal u}{{a^-}^\intercal v} \quad \text{and} \quad 1-\varepsilon = (1-\lambda)\frac{{a^-}^\intercal w}{{a^-}^\intercal v} \tag{7}$$

where ${a^-}^\intercal \mathbf{x} = 1$ is the in-hyperplane of B. If $0 \in P$, then $u^- = w^- = v^- = 0$ and thus ε can take any value in $[0, 1]$. To simplify the proof, we abuse notation and consider $\frac{{a^-}^\intercal u}{{a^-}^\intercal v} = \frac{{a^-}^\intercal w}{{a^-}^\intercal v} = 1$ for this case, so (7) still holds.

Substituting the values of ρ and ε into (6), we obtain

$$g(v) \leq \gamma\varepsilon f(u^-) + (1-\gamma)\rho f(u^+) + \gamma(1-\varepsilon) f(w^-) + (1-\gamma)(1-\rho) f(w^+)$$

$$
\begin{aligned}
&= \gamma\lambda \frac{a^{-\top}u}{a^{-\top}v} f(u^-) + (1-\gamma)\lambda \frac{a^{+\top}u}{a^{+\top}v} f(u^+) \\
&\quad + \gamma(1-\lambda)\frac{a^{-\top}w}{a^{-\top}v} f(w^-) + (1-\gamma)(1-\lambda)\frac{a^{+\top}w}{a^{+\top}v} f(w^+) \\
&= \lambda\left(\gamma\frac{a^{-\top}u}{a^{-\top}v} f(u^-) + (1-\gamma)\frac{a^{+\top}u}{a^{+\top}v} f(u^+)\right) \\
&\quad + (1-\lambda)\left(\gamma\frac{a^{-\top}w}{a^{-\top}v} f(w^-) + (1-\gamma)\frac{a^{+\top}w}{a^{+\top}v} f(w^+)\right). \qquad (8)
\end{aligned}
$$

What follows uses that g is positively homogeneous in order to rewrite the last inequality. To do so, note that

$$
\gamma\frac{a^{-\top}u}{a^{-\top}v}\cdot u^- + (1-\gamma)\frac{a^{+\top}u}{a^{+\top}v}\cdot u^+ = \left(\gamma\frac{1}{a^{-\top}v} + (1-\gamma)\frac{1}{a^{+\top}v}\right)u = u \qquad (9)
$$

because $v = \gamma v^- + (1-\gamma)v^+ = \left(\gamma\frac{1}{a^{-\top}v} + (1-\gamma)\frac{1}{a^{+\top}v}\right)v$. Let $\Omega = \gamma\frac{a^{-\top}u}{a^{-\top}v} + (1-\gamma)\frac{a^{+\top}u}{a^{+\top}v}$ —this is simply the sum of the weights in the leftmost linear combination of (9). By definition of g, and because we are assuming it to be positively homogeneous, we have that

$$
\begin{aligned}
g(u) &= g\left(\gamma\frac{a^{-\top}u}{a^{-\top}v}\cdot u^- + (1-\gamma)\frac{a^{+\top}u}{a^{+\top}v}\cdot u^+\right) && \text{(due to (9))} \\
&= \Omega\cdot g\left(\frac{\gamma\frac{a^{-\top}u}{a^{-\top}v}}{\Omega}\cdot u^- + \frac{(1-\gamma)\frac{a^{+\top}u}{a^{+\top}v}}{\Omega}\cdot u^+\right) && \text{(pos. homog.)} \\
&= \Omega\cdot\left(\frac{\gamma\frac{a^{-\top}u}{a^{-\top}v}}{\Omega}\cdot f(u^-) + \frac{(1-\gamma)\frac{a^{+\top}u}{a^{+\top}v}}{\Omega}\cdot f(u^+)\right) && \text{(def. of } g\text{)} \\
&= \gamma\frac{a^{-\top}u}{a^{-\top}v}\cdot f(u^-) + (1-\gamma)\frac{a^{+\top}u}{a^{+\top}v}\cdot f(u^+)
\end{aligned}
$$

A similar relation can be deduced for w obtaining

$$
\gamma\frac{a^{-\top}w}{a^{-\top}v}w^- + (1-\gamma)\frac{a^{+\top}w}{a^{+\top}v}w^+ = \left(\gamma\frac{1}{a^{-\top}v} + (1-\gamma)\frac{1}{a^{+\top}v}\right)w = w
$$

which implies

$$
\begin{aligned}
g(w) &= g\left(\gamma\frac{a^{-\top}w}{a^{-\top}v}w^- + (1-\gamma)\frac{a^{+\top}w}{a^{+\top}v}w^+\right) \\
&= \gamma\frac{a^{-\top}w}{a^{-\top}v}f(w^-) + (1-\gamma)\frac{a^{+\top}w}{a^{+\top}v}f(w^+).
\end{aligned}
$$

Using these expressions for $g(u)$ and $g(w)$ in (8) we obtain

$$g(v) = g(\lambda u + (1-\lambda)w) \leq \lambda g(u) + (1-\lambda)g(w).$$

This shows that g is convex in B. □

The previous lemma shows that g is convex in each region $B \in \mathcal{B}$. Before moving to convexity toward P, we show differentiability of g in each region and compute the corresponding gradient, which we rely on in the next section.

Lemma 2 *Let $B \in \mathcal{B}$ and $g : B \subset \mathbb{R}^n \to \mathbb{R}$ as defined in (1). Let ${a^{\pm}}^{\intercal} x = 1$ be the in-hyperplane and out-hyperplane of B. Then, g is a differentiable function in int(B). Moreover, the gradient is given by*

$$\nabla g(v) = \frac{\alpha_v}{{a^-}^{\intercal} v}\left(\delta^- a^- + \nabla f(v)|_{v=v^-}\right) + \frac{1-\alpha_v}{{a^+}^{\intercal} v}\left(\delta^+ a^+ + \nabla f(v)|_{v=v^+}\right) \quad (10)$$

where

$$\begin{aligned} \delta^- &= (\nabla f(v)|_{v=v^*} - \nabla f(v)|_{v=v^-})^{\intercal}\, v^- \leq 0 \\ \delta^+ &= (\nabla f(v)|_{v=v^*} - \nabla f(v)|_{v=v^+})^{\intercal}\, v^+ \geq 0 \end{aligned}$$

for a vector v^* contained on the segment $[v^-, v^+]$, and $\frac{1}{{a^-}^{\intercal} v}$:=0 in the case $0 \in P$.

Proof Consider $v \in \text{int}(B)$ arbitrary. In this proof, to aid the reader, we emphasize that $v^{\pm}$ and α_v are functions of v by referring to them as $v^{\pm}(v)$ and $\alpha_v(v)$, respectively. Note that these functions are differentiable in the interior of B.

Since $g(v) = \alpha_v(v) f(v^-(v)) + (1-\alpha_v(v)) f(v^+(v))$ and f is differentiable, g is also differentiable in the interior of B. The gradient of g is given by

$$\begin{aligned} \nabla g(v) &= \nabla\alpha_v(v) \cdot f(v^-(v)) + \alpha_v(v)\nabla f(v^-(v)) \\ &\quad + \nabla(1-\alpha_v(v)) \cdot f(v^+(v)) + (1-\alpha_v(v))\nabla f(v^+(v)) \\ &= \nabla\alpha_v(v) \cdot \left(f(v^-(v)) - f(v^+(v))\right) + \alpha_v(v)\nabla f(v^-(v)) \\ &\quad + (1-\alpha_v(v))\nabla f(v^+(v)) \end{aligned} \quad (11)$$

where

$$\nabla f(v^{\pm}(v)) = \mathcal{D}v^{\pm}(v)^{\intercal} \nabla f(v)|_{v=v^{\pm}} \quad (12)$$

and $\mathcal{D}v^{\pm}(v)$ is the Jacobian matrix of $v^{\pm}(v)$. Recall that we are assuming $v^{\pm}(v)$ intersects a facet of P contained in a hyperplane of equation ${a^{\pm}}^{\intercal}\mathbf{x} = 1$. Hence, by Remark 2 and defining $\frac{1}{{a^-}^{\intercal} v}$:=0 when $0 \in P$,

$$v^{\pm}(v) = \frac{1}{{a^{\pm}}^{\intercal} v} v, \quad \text{so} \quad \nabla\left(\frac{1}{{a^{\pm}}^{\intercal} v}\right) = \frac{-a^{\pm}}{({a^{\pm}}^{\intercal} v)^2} \quad (13)$$

and

$$\mathcal{D}v^{\pm}(v) = v\nabla\left(\frac{1}{a^{\pm\mathsf{T}}v}\right)^{\mathsf{T}} + \frac{1}{a^{\pm\mathsf{T}}v}I_n$$
$$= \frac{-va^{\pm\mathsf{T}}}{(a^{\pm\mathsf{T}}v)^2} + \frac{1}{a^{\pm\mathsf{T}}v}I_n = \frac{-1}{a^{\pm\mathsf{T}}v}\left(v^{\pm}(v)a^{\pm\mathsf{T}} - I_n\right).$$

Substituting into (12) we obtain

$$\nabla f(v^{\pm}(v)) = \left(\frac{-1}{a^{\pm\mathsf{T}}v}\left(v^{\pm}(v)a^{\pm\mathsf{T}} - I_n\right)\right)^{\mathsf{T}} \nabla f(v)|_{v=v^{\pm}}$$
$$= \frac{-1}{a^{\pm\mathsf{T}}v}\left(\left[\nabla f(v)|_{v=v^{\pm}}^{\mathsf{T}} v^{\pm}(v)\right]a^{\pm} - \nabla f(v)|_{v=v^{\pm}}\right)$$

Note that $\nabla v^{\pm}(v)^{\mathsf{T}}v = 0$ and $\nabla f(v^{\pm}(v))^{\mathsf{T}}v = 0$. This is expected because varying v over its ray does not change the position of $v^{\pm}(v)$ nor the value of $f(v^{\pm}(v))$. On the other hand, applying the gradient to (3) we obtain

$$\nabla\alpha_v(v) = \left(\frac{1}{a^{-\mathsf{T}}v} - \frac{1}{a^{+\mathsf{T}}v}\right)^{-1}\left(\alpha_v(v)\frac{a^-}{(a^{-\mathsf{T}}v)^2} + (1-\alpha_v(v))\frac{a^+}{(a^{+\mathsf{T}}v)^2}\right) \quad (14)$$

Finally, the mean value theorem ensures there exists v^* on the segment $[v^-(v), v^+(v)]$ such that

$$f(v^-(v)) - f(v^+(v)) = \nabla f(v)|_{v=v^*}^{\mathsf{T}}(v^-(v) - v^+(v)) \quad (15)$$
$$= \left(\frac{1}{a^{-\mathsf{T}}v} - \frac{1}{a^{+\mathsf{T}}v}\right)\nabla f(v)|_{v=v^*}^{\mathsf{T}} v \quad (16)$$

Grouping all the terms into (11), we obtain an explicit formula for the gradient of g at v as

$$\begin{aligned}
\nabla g(v) &= \nabla\alpha_v(v)\cdot\left(f(v^-(v)) - f(v^+(v))\right) + \alpha_v(v)\nabla f(v^-(v)) \\
&\quad + (1-\alpha_v(v))\nabla f(v^+(v)) \\
&= \left(\alpha_v(v)\frac{a^-}{(a^{-\mathsf{T}}v)^2} + (1-\alpha_v(v))\frac{a^+}{(a^{+\mathsf{T}}v)^2}\right)\nabla f(v)|_{v=v^*}^{\mathsf{T}} v \\
&\quad + \alpha_v(v)\left(\frac{-1}{a^{-\mathsf{T}}v}\left(\left[\nabla f(v)|_{v=v^-}^{\mathsf{T}} v^-(v)\right]a^- - \nabla f(v)|_{v=v^-}\right)\right) \\
&\quad + (1-\alpha_v(v))\left(\frac{-1}{a^{+\mathsf{T}}v}\left(\left[\nabla f(v)|_{v=v^+}^{\mathsf{T}} v^+(v)\right]a^+ - \nabla f(v)|_{v=v^+}\right)\right) \\
&= \frac{\alpha_v(v)}{a^{-\mathsf{T}}v}\left(\left[\left(\nabla f(v)|_{v=v^*}^{\mathsf{T}} v\right)\frac{1}{a^{-\mathsf{T}}v} - \left(\nabla f(v)|_{v=v^-}^{\mathsf{T}} v^-(v)\right)\right]a^- + \nabla f(v)|_{v=v^-}\right) \\
&\quad + \frac{1-\alpha_v(v)}{a^{+\mathsf{T}}v}\left(\left[\left(\nabla f(v)|_{v=v^*}^{\mathsf{T}} v\right)\frac{1}{a^{+\mathsf{T}}v} - \left(\nabla f(v)|_{v=v^+}^{\mathsf{T}} v^+(v)\right)\right]a^+ + \nabla f(v)|_{v=v^+}\right)
\end{aligned}$$

$$= \frac{\alpha_v(v)}{a^{-\mathsf{T}}v} \left(\underbrace{\left[(\nabla f(v)|_{v=v^*} - \nabla f(v)|_{v=v^-})^\mathsf{T} \, v^-(v) \right]}_{\delta^-} a^- + \nabla f(v)|_{v=v^-} \right)$$

$$+ \frac{1-\alpha_v(v)}{a^{+\mathsf{T}}v} \left(\underbrace{\left[(\nabla f(v)|_{v=v^*} - \nabla f(v)|_{v=v^+})^\mathsf{T} \, v^+(v) \right]}_{\delta^+} a^+ + \nabla f(v)|_{v=v^+} \right)$$

$$= \frac{\alpha_v(v)}{a^{-\mathsf{T}}v} \left(\delta^- a^- + \nabla f(v)|_{v=v^-} \right) + \frac{1-\alpha_v(v)}{a^{+\mathsf{T}}v} \left(\delta^+ a^+ + \nabla f(v)|_{v=v^+} \right)$$

Note that since f is ray-concave, it is concave on the segment $[v^-(v), v^+(v)]$, so for any v' on the segment, the directional derivatives of f on the direction v' satisfy

$$\nabla f(v)|_{v=v^-}^\mathsf{T} v' \geq \nabla f(v)|_{v=v^*}^\mathsf{T} v' \geq \nabla f(v)|_{v=v^+}^\mathsf{T} v'$$

because $\mathbf{0}$, v, $v^+(v)$ and $v^-(v)$ are collinear. Therefore,

$$(\nabla f(v)|_{v=v^*} - \nabla f(v)|_{v=v^-})^\mathsf{T} \, v' \leq 0$$
$$(\nabla f(v)|_{v=v^*} - \nabla f(v)|_{v=v^+})^\mathsf{T} \, v' \geq 0$$

for any $v' \in [v^-(v), v^+(v)]$, so $\delta^+ \geq 0$ and $\delta^- \leq 0$. □

As a side note, in the last lemma we only used differentiability of f to show the formula (10), therefore such formula is always valid for g defined as (1). Ray-concavity of f was only used to show the signs of $\delta^{\pm}$, and facet-convexity of f was not needed.

3.2 Convexity over the polytope

We know provide the last step which proves that g is convex in P.

Lemma 3 *Let $g : P \subset \mathbb{R}^n \to \mathbb{R}$ as defined in (1). If g is convex over each region $B \in \mathcal{B}$, then it is convex in P.*

Proof Our strategy to show convexity is to show *mid-point local convexity*, that is, for each $v \in P$, we show there is a neighborhood of v where g is mid-point convex. We remind the reader that mid-point convexity reads

$$g\left(\frac{1}{2}(v_1 + v_2)\right) \leq \frac{1}{2}\left(g(v_1) + g(v_2)\right) \qquad \forall v_1, v_2 \in P.$$

Mid-point convexity does not always imply convexity, but in this case it suffices as the function g is continuous. Therefore, establishing local mid-point convexity implies local convexity [8]. And since local convexity implies convexity (see e.g. [13]), we conclude that g is convex.

We now proceed to proving local mid-point convexity of g. Let us consider $v \in P$, $d \in \mathbb{R}^n$, and $\varepsilon > 0$ such that $v \pm \varepsilon d \in P$. We would like to show that

$$g(v) \leq \frac{1}{2}(g(v - \varepsilon d) + g(v + \varepsilon d)). \tag{17}$$

If $v \pm \varepsilon d \in \text{int}(B)$ the inequality follows from convexity of g within a region. Therefore, we may assume $v \in B_s \cap B_t$, $v - \varepsilon d \in B_s$ and $v + \varepsilon d \in B_t$ for some $B_s, B_t \in \mathcal{B}$.

Let $a_s^{+\mathsf{T}} x = 1$ be the out-hyperplane of B_s, and $a_s^{-\mathsf{T}} x = 1$ be its in-hyperplane. Similarly, we define $a_t^{\pm}$. Thus, $v^{\pm} = \frac{1}{a_s^{\pm\mathsf{T}} v} v = \frac{1}{a_t^{\pm\mathsf{T}} v} v$. Let ∇g_{B_s} and ∇g_{B_t} be the gradients of g in $\text{int}(B_s)$ and $\text{int}(B_t)$ respectively (see Lemma 2). Since these gradients are continuous, we can apply formula (10) to B_s and B_t. From here, we obtain

$$\nabla(g_{B_s}(v) - g_{B_t}(v)) = \frac{\alpha_v}{a_s^{-\mathsf{T}} v} \delta^- (a_s^- - a_t^-) + \frac{1-\alpha_v}{a_s^{+\mathsf{T}} v} \delta^+ (a_s^+ - a_t^+) \tag{18}$$

Now we focus on showing that $\nabla g_{B_s}(v)^{\mathsf{T}} d \leq \nabla g_{B_t}(v)^{\mathsf{T}} d$. Since $v + \varepsilon d \in B_t$,

$$(v + \varepsilon d)^{\pm} = \frac{v + \varepsilon d}{a_t^{\pm\mathsf{T}} (v + \varepsilon d)}.$$

We start exploring the facet contained in $a_t^{+\mathsf{T}} x = 1$. Recall that, by convexity of the polytope P, P is contained in the half space $\{x : a_i^{+\mathsf{T}} x \leq 1\}$ for all out-hyperplanes associated to a region $B_i \in \mathcal{B}$. Hence, $a_s^{+\mathsf{T}} (v + \varepsilon d)^+ \leq 1$ and $a_s^{+\mathsf{T}} (v + \varepsilon d) \leq a_t^{+\mathsf{T}} (v + \varepsilon d)$. Since $a_s^{+\mathsf{T}} v = a_t^{+\mathsf{T}} v$ we conclude that

$$(a_s^+ - a_t^+)^{\mathsf{T}} d \leq 0.$$

In a similar way, for the facet contained in $a_t^{-\mathsf{T}} x = 1$, by convexity of the polytope we get that $a_s^{-\mathsf{T}} (v + \varepsilon d)^- \geq 1$. So, $a_s^{-\mathsf{T}} (v + \varepsilon d) \geq a_t^{-\mathsf{T}} (v + \varepsilon d)$ and we conclude that

$$(a_s^- - a_t^-)^{\mathsf{T}} d \geq 0.$$

As $\delta^- \leq 0$ and $\delta^+ \geq 0$ (Lemma 2), we obtain that

$$\nabla(g_{B_s}(v) - g_{B_t}(v))^{\mathsf{T}} d = \frac{\alpha_v}{a_s^{-\mathsf{T}} v} \underbrace{\delta^- (a_s^- - a_t^-)^{\mathsf{T}} d}_{\leq 0} + \frac{1-\alpha_v}{a_s^{+\mathsf{T}} v} \underbrace{\delta^+ (a_s^+ - a_t^+)^{\mathsf{T}} d}_{\leq 0} \leq 0.$$

so we conclude that

$$\nabla g_{B_s}(v)^{\mathsf{T}} d \leq \nabla g_{B_t}(v)^{\mathsf{T}} d$$

Finally, we can use the first order characterization of convexity within each region and obtain

$$\begin{aligned} g(v + \varepsilon d) &= g_{B_t}(v + \varepsilon d) \geq g(v) + \varepsilon \nabla g_{B_t}(v)^{\mathsf{T}} d \geq g(v) + \varepsilon \nabla g_{B_s}(v)^{\mathsf{T}} d \\ g(v - \varepsilon d) &= g_{B_s}(v - \varepsilon d) \geq g(v) - \varepsilon \nabla g_{B_s}(v)^{\mathsf{T}} d \geq g(v) - \varepsilon \nabla g_{B_t}(v)^{\mathsf{T}} d \end{aligned}$$

These two inequalities imply (17). This completes the proof of local mid-point convexity of g which, as discussed at the beginning of this proof, implies convexity of g in P. □

Note that, similarly to Lemma 2, the latter proof does not explicitly rely on facet-convexity. The result mainly uses that g is convex on each region and that f is ray-concave (in order to use the signs of $\delta^{\pm}$ in the gradient formula).

Knowing that g defines a convex function over the domain, we can prove our main theorem, showing that it corresponds to the convex envelope of f over the polytope P.

Proof (Theorem 1) By previous lemma, we know that g is a convex function over the domain P. We show that g is an underestimator of f, that is, $g(v) \leq f(v)$ for all $v \in P$. For $v = 0$ it clearly holds. If $v \neq 0$, $v \in P$ implies that $\alpha_v \in [0, 1]$. Additionally, since f is concave over $[v^-, v^+]$ we know that

$$f(v) = f(\alpha_v v^- + (1 - \alpha_v)v^+) \geq \alpha_v f(v^-) + (1 - \alpha_v) f(v^+) = g(v).$$

Finally, we argue why g is the largest convex function that underestimates f. Let h be another convex function that underestimates f and let $v \in P$ such that $h(v) > g(v)$. Restricted to the segment $[v^-, v^+]$, the function h is also convex. But this is a contradiction, because f is concave on $[v^-, v^+]$, so the largest convex function underestimating f on this segment is the line interpolating $f(v^-)$ and $f(v^+)$, which is exactly g. □

3.3 On the positively homogeneity condition

In this section we present characterizations for when the function g constructed in (1) is positively homogeneous.

Lemma 4 *If $0 \in P$, then g is positively homogeneous if and only if $f(0) = 0$. In this case,*

$$g(v) = a^{+\mathsf{T}} v \cdot f(v^+),$$

where $a^{+\mathsf{T}} x = 1$ is the out-hyperplane of the region $B \ni v$ (see Remark 2).

Proof If $0 \in P$ then $v = \alpha_v \cdot 0 + (1 - \alpha_v)v^+ = \frac{1-\alpha_v}{a^{+\mathsf{T}} v} v$, so

$$g(v) = (1 - a^{+\mathsf{T}} v) f(0) + a^{+\mathsf{T}} v \cdot f(v^+)$$

If g is positively homogeneous, then $g(0) = 0 = f(0)$. To prove the other direction, if $f(0) = 0$ then $g(v) = a^{+\mathsf{T}} v \cdot f(v^+)$, which is homogeneous because for any $\lambda > 0$ such that $\lambda v \in P$, $(\lambda v)^+ = v^+$ so $g(\lambda v) = a^{+\mathsf{T}}(\lambda v) \cdot f(v^+) = \lambda g(v)$. □

As mentioned in Remark 1, the condition $f(0) = 0$ is not restrictive in the construction of convex envelopes when $0 \in P$. If $f(0) \neq 0$, it suffices to define $\hat{f} = f - f(0)$

and use our construction to derive conv $\hat{f}$. The desired convex envelope simply follows from noting that conv $f = \text{conv } \hat{f} + f(0)$. In a similar way, if $0 \notin P$ we can shift the domain by defining $\hat{f}(v) = f(v + v_0) - f(v_0)$ over $\hat{P} = \{v - v_0 : v \in P\}$ for any $v_0 \in P$. This translation preserves convexity over the facets and ensures the positive homogeneity of $\hat{g}$, but ray-concavity of $\hat{f}$ must be revised in order to apply Theorem 1. We illustrate the use of these transformations in Sect. 4.

Lemma 5 *If $0 \notin P$, then g is positively homogeneous iff, for every $v \in P$, $a^{-\mathsf{T}} v \cdot f(v^-) = a^{+\mathsf{T}} v \cdot f(v^+)$, where $a^{\pm\mathsf{T}} x = 1$ are the in-hyperplane and out-hyperplane of a region $B \ni v$ (see Remark 2). In this case,*

$$g(v) = a^{-\mathsf{T}} v \cdot f(v^-) = a^{+\mathsf{T}} v \cdot f(v^+).$$

Proof Since $v^{\pm} = \frac{1}{a^{\pm\mathsf{T}} v} v$, if g is homogeneous then

$$g(v) = g\left((a^{\pm\mathsf{T}} v) \cdot v^{\pm}\right) = a^{\pm\mathsf{T}} v \cdot g(v^{\pm}) = a^{\pm\mathsf{T}} v \cdot f(v^{\pm}).$$

For the other direction, if $a^{-\mathsf{T}} v \cdot f(v^-) = a^{+\mathsf{T}} v \cdot f(v^+)$, by (3) we obtain

$$\begin{aligned} g(v) &= \alpha_v f(v^-) + (1 - \alpha_v) f(v^+) \\ &= \alpha_v f(v^-) + \left(1 - \frac{\alpha_v}{a^{-\mathsf{T}} v}\right) (a^{+\mathsf{T}} v) f(v^+) \\ &= \alpha_v f(v^-) + \left(1 - \frac{\alpha_v}{a^{-\mathsf{T}} v}\right) (a^{-\mathsf{T}} v) f(v^-) \\ &= \alpha_v f(v^-) + (a^{-\mathsf{T}} v) f(v^-) - \alpha_v f(v^-) \\ &= (a^{-\mathsf{T}} v) f(v^-) = (a^{+\mathsf{T}} v) f(v^+) \end{aligned}$$

So, g is homogeneous because for any $\lambda > 0$ such that $\lambda v \in P$, $g(\lambda v) = a^{\pm\mathsf{T}}(\lambda v) \cdot f((\lambda v)^{\pm}) = \lambda (a^{\pm\mathsf{T}} v) \cdot f(v^{\pm}) = \lambda g(v)$. □

We note that our results have an unexpected consequence: when f is a homogeneous function, convexity of f over the facets of P implies convexity of f over P.

Corollary 1 *Let $f : P \to \mathbb{R}$ be continuously differentiable and convex (concave) over the facets of P. If f is positively homogeneous, then f is convex (concave) over P.*

Proof We show the proof for f convex on the facets; the concave case is almost identical. If f is positively homogeneous then in particular it is ray-linear. Hence $f(v) = (a^{+\mathsf{T}} v) \cdot f(v^+) = g(v)$. In addition, since f is convex on the facets of P, by Theorem 1 $g = \text{conv } f$, so f is convex over P. □

4 Examples of ray-concave functions and their envelopes

In this section, we provide the convex envelopes of various explicit functions. Some of these are new, and some have been provided in the literature before. In the latter case, our result provides new perspectives, and in some cases simpler derivations.

4.1 Example 1 revisited

Let us consider the function

$$f(x,y)=\frac{xy}{x+y-xy} \tag{19}$$

in a box $[0,u_x]\times[0,u_y]\subseteq[0,1]^2$. This function appears naturally in the context of network reliability optimization. In fact, if X, Y are independent Bernoulli random variables indicating the current state of two serial components, with reliabilities $p_X := \mathbb{P}(X=1)$ and $p_Y := \mathbb{P}(Y=1)$ then

$$f(p_X,p_Y)=\mathbb{P}\left(X\cdot Y=1|X+Y\geq 1\right).$$

corresponds to the resulting reliability of a *degree-2* reduction [27].

We compute the concave envelope of (19) via the convex envelope of $\hat{f}=-f$. The function $\hat{f}$ can be directly verified to be convex on the facets of $[0,u_x]\times[0,u_y]$. For instance

$$h(x)=\hat{f}(x,u_y)=-\frac{xu_y}{x+u_y-xu_y},$$

and a simple calculation shows

$$h''(x)=\frac{2(1-u_y)u_y^2}{(1-(1-u_y)(1-x))^3}\geq 0.$$

As for ray-concavity, we compute

$$\hat{f}(x,\lambda x)=-\frac{\lambda x^2}{x+\lambda x-\lambda x^2}\Rightarrow\frac{\partial^2\hat{f}}{\partial x^2}(x,\lambda x)=-\frac{2\lambda^2(1+\lambda)}{(1+\lambda(1-x))^3},$$

therefore $\frac{\partial^2 f}{\partial x^2}(x,\lambda x)\geq 0$ for $x\leq 1$ and $\lambda\geq 0$. By Theorem 1, the concave envelope of $f(x,y)$, denoted conc f, is given by

$$\begin{aligned}\text{conc}\, f(x,y)&=-\text{conv}\,\hat{f}(x,y)=-(a^{+\mathsf{T}}v)\hat{f}(v_x^+,v_y^+)\\&=\begin{cases}\frac{y}{u_y}\frac{x\frac{u_y}{y}\cdot u_y}{x\frac{u_y}{y}+u_y-x\frac{u_y}{y}\cdot u_y} & \text{if } y\geq\frac{u_y}{u_x}x\\ \frac{x}{u_x}\frac{u_x\cdot y\frac{u_x}{x}}{u_x+y\frac{u_x}{x}-u_x\cdot y\frac{u_x}{x}} & \text{if } y\leq\frac{u_y}{u_x}x\end{cases}\\&=\begin{cases}\frac{x\cdot y}{x+y-x\cdot u_y} & \text{if } y\geq\frac{u_y}{u_x}x\\ \frac{x\cdot y}{x+y-u_x\cdot y} & \text{if } y\leq\frac{u_y}{u_x}x\end{cases}\end{aligned}$$

Note that this procedure also computes, for free, the concave envelope of f on the non-rectangular polytopes $\{(x,y)\in[0,u_x]\times[0,u_y]\;:\;y\leq\frac{u_y}{u_x}x\}$ and $\{(x,y)\in[0,u_x]\times[0,u_y]\;:\;y\geq\frac{u_y}{u_x}x\}$.

Before moving on to the next examples, we would like to measure the advantage of using conc f, for f defined in (19), in comparison to other—perhaps easier to derive—concave overestimators. For simplicity, in what follows we use $u_x = u_y = 1$.

Using that f is factorable we can use the procedure by McCormick [21] to quickly compute a concave overestimator of f. One convenient way of rewriting f to apply such procedure is

$$f(x, y) = h_1(h_2(x, y)) - 1,$$

where $h_2(x, y) = \frac{xy}{x+y}$ and $h_1(z) = \frac{1}{1-z}$. We chose this factorization of f since both h_1, h_2 are well defined in the relevant domains: $h_2(x, y) \to 0$ when $(x, y) \to (0, 0)$, and the function h_1 is well-defined in $[0, 1/2]$, which is the range of h_2. This, unfortunately, does not hold in other direct factorizations of f.

The procedure by McCormick [21] produces the following concave overestimator

$$f(x, y) \le \operatorname{conc} h_1(\operatorname{mid}(\operatorname{conv} h_2(x, y), \operatorname{conc} h_2(x, y), z_{\max})) - 1$$

where $\operatorname{mid}(\cdot, \cdot, \cdot)$ is the midpoint of the three arguments and $z_{\max}$ is a maximizer of h_1. In our case, h_1 is a convex increasing function on $[0, 1/2]$, thus

$$\operatorname{conc} h_1(z) = z\frac{h_1(1/2) - h_1(0)}{1/2} + h_1(0) = 1 + 2z$$

and $z_{\max} = 1/2$. Additionally, h_2 is a concave function and $h_2(x, y) \le 1/2$, which implies

$$\operatorname{mid}(\operatorname{conv} h_2(x, y), \operatorname{conc} h_2(x, y), z_{\max}) = h_2(x, y).$$

The resulting concave overestimator of $f(x, y)$ is

$$f(x, y) \le \tilde{f}(x, y) := 2\frac{xy}{x + y}.$$

In Fig. 3 we plot functions f, $\tilde{f}$ and conc f, where the strong dominance of the latter with respect to $\tilde{f}$ can be appreciated.

We can also quantify the gap improvement of conc f with respect to $\tilde{f}$: we consider the *total gap* (see [9]) between $\tilde{f}$ and f defined as

$$\delta^{\tilde{f}} = \int_{[0,1]^2} (\tilde{f}(x, y) - f(x, y))dxdy.$$

Similarly, we define $\delta^{\operatorname{conc} f}$. A direct evaluation of these integrals (e.g. using Mathematica 12.3 [6]) shows that

$$\text{Gap improvement} := \frac{\delta^{\tilde{f}} - \delta^{\operatorname{conc} f}}{\delta^{\tilde{f}}} \approx 63.5\%$$

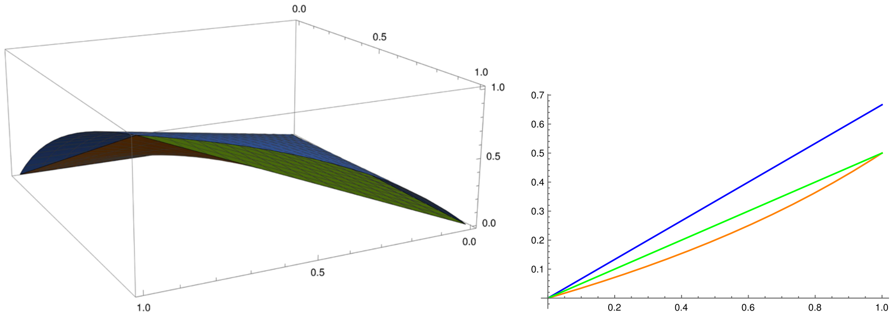

Fig. 3 Comparison between concave overestimator (blue) obtained using [21] and concave envelope (green) of function f (orange) defined in (19). Left: the plot of the three functions on the box $[0, 1] \times [0, 1]$. Right: functions restricted to the ray $\{(r/2, r) : r \in [0, 1]\}$ (color figure online)

which is substantial.

4.2 Additional examples

Example 2 Consider the function $f(x, y) = -xy$, whose convex envelope over $[l_x, u_x] \times [l_y, u_y]$ is well-known. In order to construct its convex envelope using Theorem 1, we first shift the domain by considering the function

$$\hat{f}(x, y) = f(x + l_x, y + l_y) + l_x \cdot l_y = -(x + l_x) \cdot (y + l_y) + l_x \cdot l_y$$

over the box $[0, u_x - l_x] \times [0, u_y - l_y]$. It is easy to verify that $\hat{f}$ is ray-concave and linear on the facet of any box $[0, u_x - l_x] \times [0, u_y - l_y]$. Theorem 1 implies that $\operatorname{conv} \hat{f}(v) = a^{+\mathsf{T}} v \cdot \hat{f}(v^+)$ and thus

$$\begin{aligned}\operatorname{conv} \hat{f}(x, y) &= \begin{cases} \frac{y}{u_y - l_y} \cdot \hat{f}\left(x \cdot \frac{u_y - l_y}{y}, y \cdot \frac{u_y - l_y}{y}\right) & \text{if } y \geq \frac{u_y - l_y}{u_x - l_x} x \\ \frac{x}{u_x - l_x} \cdot \hat{f}\left(x \cdot \frac{u_x - l_x}{x}, y \cdot \frac{u_x - l_x}{x}\right) & \text{if } y \leq \frac{u_y - l_y}{u_x - l_x} x \end{cases} \\ &= \begin{cases} -u_y x - l_x y & \text{if } y \geq \frac{u_y - l_y}{u_x - l_x} x \\ -l_y x - u_x y & \text{if } y \leq \frac{u_y - l_y}{u_x - l_x} x \end{cases}\end{aligned}$$

Since $\operatorname{conv} f(x, y) = \operatorname{conv} \hat{f}(x - l_x, y - l_y) - l_x \cdot l_y$, we obtain

$$\operatorname{conv} f(x, y) = \begin{cases} -u_y x - l_x y + l_x u_y & \text{if } y - l_y \geq \frac{u_y - l_y}{u_x - l_x}(x - l_x) \\ -l_y x - u_x y + l_y u_x & \text{if } y - l_y \leq \frac{u_y - l_y}{u_x - l_x}(x - l_x) \end{cases}$$

which corresponds to the McCormick envelopes for this function.

Note that we shift the domain in order to ensure the positive homogeneity condition due to $0 \in P$. This is not necessary for the particular case when $l_x = l_y = l$ and $u_x = u_y = u$. In this case, considering $\hat{f}(x, y) = -xy - lu$ we obtain

$$a^{-\mathsf{T}}v \cdot \hat{f}(v^-) = \begin{cases} \frac{x}{l}\left(-l \cdot y\frac{l}{x} - lu\right) = -ux - ly & \text{if } y \geq x \\ \frac{y}{l}\left(-x\frac{l}{y} \cdot l - lu\right) = -lx - uy & \text{if } y \leq x \end{cases}$$

$$a^{+\mathsf{T}}v \cdot \hat{f}(v^+) = \begin{cases} \frac{y}{u}\left(-x\frac{u}{y} \cdot u - lu\right) = -ux - ly & \text{if } y \geq x \\ \frac{x}{u}\left(-u \cdot y\frac{u}{x} - lu\right) = -lx - uy & \text{if } y \leq x \end{cases}$$

Hence, by Lemma 5,

$$\operatorname{conv} f(x, y) = \operatorname{conv} \hat{f}(x, y) + lu = \begin{cases} -ux - ly + lu & \text{if } y \geq x \\ -lx - uy + lu & \text{if } y \leq x \end{cases}$$

□

Example 3 Let us consider the following example from [19]. Let $f(x, y) = y/x$ and

$$P = \left\{(x, y) \in \mathbb{R}^2 : -x + 2y \leq 2, 1 \leq x \leq 2, 0 \leq y \leq 2\right\}.$$

Note that f is ray-linear and convex on the facets of P. However, if we try to use the construction in Theorem 1, the resulting function is $g(x, y) = y/x$ (the original function), which is not positively homogeneous and thus Theorem 1 does not apply. Moreover, g is not even convex.

In order to apply Theorem 1 in this example, we proceed as follows. We shift the domain by considering the function as $\hat{f}(x, y) = f(x + 1, y)$ and the polytope $\hat{P} = \{(x, y) \in \mathbb{R}^2 : (x + 1, y) \in P\}$. Note that $\hat{f}(x, y)$ is ray-concave because $\hat{f}(x, \lambda x) = \lambda \frac{x}{x+1}$ is concave for $x \geq 0$ and $\lambda \geq 0$. Convexity on the facets can be directly verified. Moreover, $0 \in \hat{P}$ and $\hat{f}(0) = 0$, therefore, Theorem 1 and Lemma 4 provide a construction guaranteed to yield conv $\hat{f}$. Since the outer facets of $\hat{P}$ are $x = 1$ and $-\frac{1}{3}x + \frac{2}{3}y = 1$, we obtain

$$\begin{aligned} \operatorname{conv} \hat{f}(x, y) &= \begin{cases} (-\frac{1}{3}x + \frac{2}{3}y)\hat{f}\left(\frac{x}{-\frac{1}{3}x+\frac{2}{3}y}, \frac{y}{-\frac{1}{3}x+\frac{2}{3}y}\right) & \text{if } y \geq 2x \\ x\hat{f}\left(\frac{x}{x}, \frac{y}{x}\right) & \text{if } y \leq 2x \end{cases} \\ &= \begin{cases} y\frac{-\frac{1}{3}x+\frac{2}{3}y}{x+\left(-\frac{1}{3}x+\frac{2}{3}y\right)} = y\frac{-x+2y}{2x+2y} & \text{if } y \geq 2x \\ \frac{1}{2}y & \text{if } y \leq 2x \end{cases} \end{aligned}$$

As conv $f(x, y) = \operatorname{conv} \hat{f}(x - 1, y)$ we obtain

$$\operatorname{conv} f(x, y) = \begin{cases} y\frac{1-x+2y}{2(x+y+1)} & \text{if } y \geq 2(x-1) \\ \frac{1}{2}y & \text{if } y \leq 2(x-1) \end{cases}$$

□

The following example shows how Corollary 1 can be used to prove the convexity of positively homogeneous functions.

Example 4 Let f be a 3-dimensional Cobb–Douglas function

$$f(x_1, x_2, x_3) = Ax_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$$

where $A, \alpha_1, \alpha_2, \alpha_3 > 0$, $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and $\mathbf{x} \in \mathbb{R}_+^3$.

It is known that the 2-dimensional Cobb–Douglas function is concave if $\alpha_i + \alpha_j < 1$, hence f is concave over the facets of the box $P = [l_1, u_1] \times [l_2, u_2] \times [l_3, u_3] \subset \mathbb{R}_+^3$.

Since $\sum_{i=1}^{3} \alpha_i = 1$, f is positively homogeneous, so by Corollary 1 we conclude that f is concave over P. □

Acknowledgements The research leading to these results received funding from grants ANID/CONICYT-Fondecyt Regular 1200809 (J.B., E.M.), MathAmsud 19-MATH-03 (J.B.) and ANID/CONICYT-Fondecyt Iniciación 11190515 (G.M.). We would also like to thank Felipe Serrano for helpful discussions and to the anonymous reviewer for their valuable feedback.

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

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Al-Khayyal, F.A., Falk, J.E.: Jointly constrained biconvex programming. Math. Oper. Res. **8**(2), 273–286 (1983). https://doi.org/10.1287/moor.8.2.273
2. Anstreicher, K.M., Burer, S.: Computable representations for convex hulls of low-dimensional quadratic forms. Math. Program. **124**(1), 33–43 (2010). https://doi.org/10.1007/s10107-010-0355-9
3. Bao, X., Sahinidis, N.V., Tawarmalani, M.: Multiterm polyhedral relaxations for nonconvex, quadratically constrained quadratic programs. Optim. Methods Softw. **24**(4–5), 485–504 (2009). https://doi.org/10.1080/10556780902883184
4. Benson, H.P.: On the construction of convex and concave envelope formulas for bilinear and fractional functions on quadrilaterals. Comput. Optim. Appl. **27**(1), 5–22 (2004). https://doi.org/10.1023/B:COAP.0000004976.52180.7f
5. Hijazi, H.: Perspective envelopes for bilinear functions. In: AIP Conference Proceedings, vol. 2070, p. 020017. AIP Publishing LLC (2019)
6. Inc., W.R.: Mathematica, Version 12.3.1. Champaign, IL (2021). https://www.wolfram.com/mathematica
7. Jach, M., Michaels, D., Weismantel, R.: The convex envelope of (n-1)-convex functions. SIAM J. Optim. **19**(3), 1451–1466 (2008). https://doi.org/10.1137/07069359X
8. Jensen, J.L.W.V.: Om konvekse funktioner og uligheder imellem middelvaerdier. Nyt tidsskrift for matematik **16**, 49–68 (1905)
9. Khajavirad, A., Michalek, J.J., Sahinidis, N.V.: Relaxations of factorable functions with convex-transformable intermediates. Math. Program. **144**(1), 107–140 (2014). https://doi.org/10.1007/s10107-012-0618-8
10. Khajavirad, A., Sahinidis, N.V.: Convex envelopes of products of convex and component-wise concave functions. J. Glob. Optim. **52**(3), 391–409 (2012). https://doi.org/10.1007/s10898-011-9747-5
11. Khajavirad, A., Sahinidis, N.V.: Convex envelopes generated from finitely many compact convex sets. Math. Program. **137**(1), 371–408 (2013). https://doi.org/10.1007/s10107-011-0496-5
12. Kuno, T.: A branch-and-bound algorithm for maximizing the sum of several linear ratios. J. Glob. Optim. **22**(1), 155–174 (2002). https://doi.org/10.1023/A:1013807129844

13. Li, Y.C., Yeh, C.C.: Some characterizations of convex functions. Comput. Math. Appl. **59**(1), 327–337 (2010). https://doi.org/10.1016/j.camwa.2009.05.020
14. Liberti, L., Pantelides, C.C.: Convex envelopes of monomials of odd degree. J. Glob. Optim. **25**(2), 157–168 (2003). https://doi.org/10.1023/A:1021924706467
15. Linderoth, J.: A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs. Math. Program. **103**(2), 251–282 (2005). https://doi.org/10.1007/s10107-005-0582-7
16. Locatelli, M.: Polyhedral subdivisions and functional forms for the convex envelopes of bilinear, fractional and other bivariate functions over general polytopes. J. Glob. Optim. **66**(4), 629–668 (2016). https://doi.org/10.1007/s10898-016-0418-4
17. Locatelli, M.: Convex envelopes of bivariate functions through the solution of KKT systems. J. Glob. Optim. **72**(2), 277–303 (2018). https://doi.org/10.1007/s10898-018-0626-1
18. Locatelli, M.: Convex envelope of bivariate cubic functions over rectangular regions. J. Glob. Optim. **76**(1), 1–24 (2020). https://doi.org/10.1007/s10898-019-00846-2
19. Locatelli, M., Schoen, F.: On convex envelopes for bivariate functions over polytopes. Math. Program. **144**(1), 65–91 (2014). https://doi.org/10.1007/s10107-012-0616-x
20. Luedtke, J., Namazifar, M., Linderoth, J.: Some results on the strength of relaxations of multilinear functions. Math. Program. **136**(2), 325–351 (2012). https://doi.org/10.1007/s10107-012-0606-z
21. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: part I—convex underestimating problems. Math. Program. **10**(1), 147–175 (1976). https://doi.org/10.1007/BF01580665
22. Meyer, C.A., Floudas, C.A.: Trilinear monomials with mixed sign domains: facets of the convex and concave envelopes. J. Glob. Optim. **29**(2), 125–155 (2004). https://doi.org/10.1023/B:JOGO.0000042112.72379.e6
23. Meyer, C.A., Floudas, C.A.: Convex envelopes for edge-concave functions. Math. Program. **103**(2), 207–224 (2005). https://doi.org/10.1007/s10107-005-0580-9
24. Muller, B., Serrano, F., Gleixner, A.: Using two-dimensional projections for stronger separation and propagation of bilinear terms. SIAM J. Optim. **30**(2), 1339–1365 (2020). https://doi.org/10.1137/19M1249825
25. Rikun, A.D.: A convex envelope formula for multilinear functions. J. Glob. Optim. **10**(4), 425–437 (1997). https://doi.org/10.1023/A:1008217604285
26. Ryoo, H.S., Sahinidis, N.V.: Analysis of bounds for multilinear functions. J. Glob. Optim. **19**(4), 403–424 (2001). https://doi.org/10.1023/A:1011295715398
27. Satyanarayana, A., Wood, R.K.: A linear-time algorithm for computing k-terminal reliability in series-parallel networks. SIAM J. Comput. **14**(4), 818–832 (1985). https://doi.org/10.1137/0214057
28. Sherali, H.D.: Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets. Acta Math. Vietnam **22**(1), 245–270 (1997)
29. Sherali, H.D., Alameddine, A.: An explicit characterization of the convex envelope of a bivariate bilinear function over special polytopes. Ann. Oper. Res. **25**(1), 197–209 (1990). https://doi.org/10.1007/BF02283695
30. Tardella, F.: On the existence of polyhedral convex envelopes. In: Floudas, C.A., Pardalos, P. (eds.) Frontiers in Global Optimization, pp. 563–573. Springer, Berlin (2004). https://doi.org/10.1007/978-1-4613-0251-3_30
31. Tardella, F.: Existence and sum decomposition of vertex polyhedral convex envelopes. Optim. Lett. **2**(3), 363–375 (2008). https://doi.org/10.1007/s11590-007-0065-2
32. Tawarmalani, M., Richard, J.P.P., Xiong, C.: Explicit convex and concave envelopes through polyhedral subdivisions. Math. Program. **138**(1), 531–577 (2013). https://doi.org/10.1007/s10107-012-0581-4
33. Tawarmalani, M., Sahinidis, N.V.: Semidefinite relaxations of fractional programs via novel convexification techniques. J. Glob. Optim. **20**(2), 133–154 (2001). https://doi.org/10.1023/A:1011233805045
34. Zamora, J.M., Grossmann, I.E.: A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms. J. Glob. Optim. **14**(3), 217–249 (1999). https://doi.org/10.1023/A:1008312714792