

Explicit convex and concave envelopes through polyhedral subdivisions*

Mohit Tawarmalani[†] Jean-Philippe P. Richard[‡] Chuanhui Xiong[§]

June 1, 2010

Abstract

In this paper, we derive explicit characterizations of convex and concave envelopes of several nonlinear functions over various subsets of a hyper-rectangle. These envelopes are obtained by identifying polyhedral subdivisions of the hyper-rectangle over which the envelopes can be constructed easily. In particular, we use these techniques to derive, in closed-form, the concave envelopes of concave-extendable supermodular functions and the convex envelopes of disjunctive convex functions.

1 Introduction and Motivation

A significant amount of research has been devoted to developing concave overestimators and convex underestimators of nonlinear functions $f(x)$ over the hypercube. One of the motivations for such research is that, whenever an optimization problem involves maximizing $f(x)$ (resp. minimizing $f(x)$) or contains an inequality $f(x) \geq r$ (resp. $f(x) \leq r$), replacing $f(x)$ by a concave overestimator (resp. convex underestimator) yields a convex relaxation of the problem. Such a relaxation can, for instance, be used in branch-and-bound algorithms for global optimization where convex relaxations must be constructed over successively refined partitions of the original variable space; see [37] for an exposition.

In order for branch-and-bound algorithms to produce globally optimal solutions, certain mild technical conditions are typically needed; see [16]. In particular, if one can guarantee that for a minimization problem the node with the lowest lower bound is chosen periodically, the volume of partition elements tends to zero, and the relaxations approach the original functions when the volume of the partition elements goes down to zero, branch-and-bound converges to a globally optimal solution. It is well-known, see for example [3], that the concave (resp. convex) envelope, i.e. the lowest (resp. highest) concave overestimator (resp. convex underestimator) of a function over a specified region, converges to this function as partition elements become smaller. As a result, deriving concave and convex envelopes of nonlinear functions over partition elements is a problem that is commonly encountered in the implementation of branch-and-bound algorithms for nonlinear programs. Further, since among all partitioning schemes in branch-and-bound algorithms, the rectangular partitioning scheme in which partition elements are hyper-rectangles is used most often, computing convex and concave envelopes of general functions $f(x)$ over a hyper-rectangle is a problem of crucial practical importance.

*The work was supported by NSF CMMI 0900065 and 0856605.

[†]Krannert School of Management, Purdue University. Email: mtawarma@purdue.edu. Corresponding author.

[‡]Department of Industrial and Systems Engineering, University of Florida.

[§]Krannert School of Management, Purdue University.

35 Since it is NP-Hard to maximize/minimize a multilinear function over the unit hypercube, see
36 [9], finding the concave/convex envelope of a generic function $f(x)$ is provably hard. Nevertheless,
37 for many practically useful functions, such as bilinear terms [1], various types of multilinear functions
38 [24, 30, 27, 5, 3], and the fractional term [36], concave envelopes have been derived in the literature.
39 Further, general theoretical frameworks for the construction of such envelopes [2, 10, 32, 7, 31,
40 38, 21, 23] have been proposed. It is noticeable however that, despite recent progress in the field,
41 there remain many practically useful functions for which concave envelopes are not known. As
42 an example, consider the function $d(x) = \frac{1}{a_0 + \sum_{i=1}^n a_i x_i}$ over the unit hypercube. This function
43 appears, for instance, in the formulation of the consistent biclustering problem [6]. If we assume
44 that $a_0 + \sum_{i=1}^n a_i x_i > 0$ whenever $0 \leq x \leq 1$, then f is well-defined over the relevant domain. A
45 standard procedure to relax $z = d(x)$ is to first introduce a new variable $y = a_0 + \sum_{i=1}^n a_i x_i$ and
46 then to relax $z = \frac{1}{y}$ by constructing the convex and concave envelopes of $\frac{1}{y}$. This leads to the
47 relaxation $z \geq \frac{1}{y}$ and $z \leq \frac{1}{y^L} + \frac{y^L - y}{y^U y^L}$. Here, y^L and y^U are computed respectively by minimizing and
48 maximizing $a_0 + \sum_{i=1}^n a_i x_i$ over the unit hypercube. Assuming $a_i \neq 0$ for $i = 1, \dots, n$ and $n > 1$,
49 this procedure yields a concave overestimator of $d(x)$ that is weaker than the concave envelope of
50 $d(x)$.

51 In this paper, we develop techniques for identifying the convex/concave envelopes of nonlinear
52 functions by investigating polyhedral subdivisions of the hyper-rectangle. Following this approach,
53 we provide streamlined and unified generalizations of a variety of results from the literature and
54 expose new convex/concave envelope characterizations and separation results for them. In Section 2,
55 we develop a general set of tools for the convexification of polyhedral functions providing a common
56 framework for the derivation of earlier results in [24, 30, 3]. In particular, we show that computing
57 the value of the concave envelope at a point is equivalent to solving a certain optimization problem.
58 Insights derived from this result allow us to describe polynomial separation procedures for a variety
59 of functions. For example, we show that the concave (resp. convex) envelopes of a maximum
60 (resp. minimum) of a collection of functions is polynomially separable if the concave (resp. convex)
61 envelopes of the individual functions are polynomially separable. The remainder of the paper studies
62 a variety of polyhedral subdivisions of the hyper-rectangle and gives insights regarding the classes
63 of functions for which they describe the convex/concave envelopes.

64 In Section 3, we show that by combining the results of [19, 42, 38] concave envelopes of super-
65 modular concave-extendable functions can be developed over a lattice family. This result gener-
66 alizes the explicit characterizations of convex/concave envelopes for specific functions described in
67 [30, 8, 5, 21, 26]. In addition, we show that this result has many, as yet unrealized, applications
68 in improving relaxations of factorable programs beyond the classical technique of [20] and its more
69 recent variants implemented in global optimization software [40, 18, 4]. To support this claim,
70 consider the function $d(x)$ described above. This function is of the form $f(x) = c(a_0 + \sum_{i=1}^n a_i x_i)$.
71 Our results allow the derivation of the concave (resp. convex) envelope of f over a hyper-rectangle
72 if $c(\cdot)$ is a convex (resp. concave) function. In factorable programming, products of variables are
73 replaced with new variables until a function of the form of $f(x)$ is obtained. Then, a variable, say
74 y , is introduced to replace $a_0 + \sum_{i=1}^n a_i x_i$ and $c(y)$ is overestimated using a linear function over
75 $[y^L, y^U]$ where the bounds y^L and y^U are derived from the bounds on x_i and the defining expression
76 for y . Assume $n > 1$, $c(\cdot)$ is strictly convex, and without loss of generality that $a_i > 0$ for all
77 i . Then, the factorable relaxation is clearly weaker than the aforementioned envelope because the
78 concave envelope matches the function value at $(x_1^U, \dots, x_{n-1}^U, x_n^L)$ whereas the factorable relaxation
79 overestimates the function value. This illustrates that exploiting the closed-form concave envelopes
80 we develop in this paper will help strengthen relaxations in commercial global optimization solvers.

81 In Section 4, we show that the orthogonal disjunctions theory [23] can be used to develop con-

82 vex envelopes of functions of the form $xg(y)$ over the unit hypercube when $g(\cdot)$ is a non-increasing
83 convex function. These relaxations are piecewise-conic and have a variety of applications in global
84 optimization. For example, we show that a variety of fractional, logarithmic, and polynomial func-
85 tions can be convexified using the approach. We also develop polyhedral subdivisions to convexify
86 a symmetric function of binary variables generalizing prior results in [30]. We then study situa-
87 tions where the envelope of a function obtained over the unit hypercube is similar/dissimilar to
88 its envelope over a subset of the hypercube. In particular, we describe two extreme situations. In
89 the first case, the envelope changes over the entire subregion and therefore an entirely new proof is
90 required. In the second case, the envelope remains the same over a portion of the feasible region
91 and, therefore, we leverage the proof of the envelope over the hyper-rectangle in our construction.
92 Throughout the section, we provide examples and sample illustrations of our results. We conclude
93 in Section 5 with comments on the applicability of the results developed in this paper and directions
94 of future research.

95 2 Preliminaries

96 In this section, we review and unify existing literature regarding the derivation of concave envelopes
97 over hyper-rectangles.

98 **Definition 2.1.** For a function $f : S \mapsto \mathbb{R}$, where S is a nonempty convex subset of \mathbb{R}^n , the function
99 $g(x) : S \rightarrow \mathbb{R}$ is the concave envelope of $f(x)$ over S if

- 100 1. $g(x)$ is concave over S
- 101 2. $g(x) \geq f(x)$ for all $x \in S$
- 102 3. If $h(x)$ is any concave function over $\text{conv}(S)$ that satisfies $h(x) \geq f(x)$ for all $x \in S$, then
103 $h(x) \geq g(x)$ for all $x \in S$.

104 We denote the concave envelope of f over a set S by $\text{conc}_S(f)$. If the region is clear from the
105 description, we sometimes will omit the subscript S .

106 In words, $\text{conc}_S(f)$ is the lowest concave overestimator of the function $f(x)$ over S . Similarly,
107 the convex envelope of a function is the highest convex underestimator of the function f over S . In
108 the remainder of the text, we will refer to the convex envelope as $\text{conv}_S(f)$.

109 We consider a continuous function $f(x) = f(x_1, x_2, \dots, x_n)$ over the hyper-rectangle $x_i^L \leq x_i \leq$
110 x_i^U . The conjugate of f is denoted as f^* . We assume without loss of generality (*wlog*) that $x_i^U > x_i^L$
111 for $i = 1, \dots, n$. Otherwise, the dimension of x can be reduced by fixing variables x_i with $x_i^U = x_i^L$.
112 We further assume that, for every i , $x_i^U = 1$ and $x_i^L = 0$, or else, the following linear transformation
113 can be used to transform x into x' :

$$114 \quad x' = T(x) = T(x_1, \dots, x_n) = \left(\frac{x_1 - x_1^L}{x_1^U - x_1^L}, \dots, \frac{x_n - x_n^L}{x_n^U - x_n^L} \right) \quad (1)$$

115 where $0 \leq x' \leq 1$. Transformation (1) will typically be without loss of generality for our study
116 although we mention that it might not preserve all useful properties of f . In the remainder of this
117 paper, we refer to the unit hypercube in \mathbb{R}^n as \mathcal{H}_n , i.e. $\mathcal{H}_n = [0, 1]^n$.

118 Concave envelopes can often be constructed by restricting the domain of the definition of f to
119 the extreme points of the hypercube. Definition 2.2, which is inspired by previous work on convex
120 extensions [37], formalizes this notion.

121 **Definition 2.2.** A function $f(x) : P \rightarrow \mathbb{R}$, where P is a polytope, is said to be concave-extendable
122 (resp. convex-extendable) from $X \subseteq P$ if the concave (resp. convex) envelope of $f(x)$ is only
123 determined by X , i.e., $\text{conc}(f)$ over P is also the concave envelope of \hat{f} over P , where \hat{f} is the
124 restriction of f to X that is defined as follows:

$$125 \quad \hat{f}(x) = \begin{cases} f(x) & x \in X \\ -\infty & \text{otherwise.} \end{cases}$$

126 □

127 It follows from Definition 2.2 that $\text{conv}(X) = P$. In particular, we will often encounter functions
128 that are concave-extendable or convex-extendable from the vertices of the unit hypercube, i.e.
129 $P = [0, 1]^n$ and $X = \text{vert}([0, 1]^n)$. Clearly, convex functions are always concave-extendable from
130 vertices. Examples of functions that are not convex but still concave-extendable from vertices
131 include multilinear functions [24] and, more generally, functions that are convex when restricted
132 to the space of each variable, i.e., the space created when all other variables are fixed to arbitrary
133 values within their domain. The concave envelope of any function that is concave-extendable from
134 vertices is polyhedral since it is completely determined by a finite number of points. A partial
135 converse is also known to be true: all continuously differentiable functions that have a polyhedral
136 concave envelope over the unit hypercube are concave-extendable from vertices; see Theorem 1.1 in
137 [24].

138 Concave envelopes of functions that are concave-extendable from the vertices of P are intimately
139 related to certain partitions of P . We describe these relations next.

140 **Definition 2.3** ([17]). Let $S \subseteq \mathbb{R}^n$. A set of n -dimensional polyhedra $S_1, \dots, S_m \subseteq S$ is a polyhedral
141 subdivision of S if $S = \bigcup_{i=1}^m S_i$ and $S_i \cap S_j$ is a (possibly empty) face of both S_i and S_j . □

142 In particular if each polyhedron in the subdivision is a simplex, then the polyhedral subdivision
143 is called a *triangulation*. In the optimization literature, triangulations are also known as *simplicial*
144 *covers*; see [5] for example. Observe that there is no requirement in Definition 2.3 that the extreme
145 points of S_i are also extreme points of S . However, in this paper, we will be most interested in
146 subdivisions where the extreme points of each polyhedron are also extreme points of S . We say
147 that these subdivisions *do not add vertices*.

148 Consider a finite collection of points $(v_1, \dots, v_m) \in \mathbb{R}^n$ such that $\text{aff}(\text{conv}(v_1, \dots, v_m)) = \mathbb{R}^n$.
149 Consider the corresponding matrix $V \in \mathbb{R}^{n \times m}$, whose j^{th} column V_j satisfies $V_j = v_j$. We denote
150 the submatrix of V that consists of columns in an index set J as $V(J)$. For simplicity of notation
151 and because it will be clear from the context, we also denote the set of points v_j corresponding
152 to the index set J as $V(J)$ and therefore we use $\text{conv}(V(J))$ to represent $\text{conv}\left(\bigcup_{j \in J} v_j\right)$. Let
153 $f(V) = (f(v_1), \dots, f(v_m))$ and let e denote the vector of all ones. Consider the following primal-
154 dual pair of linear programming problems:

$$155 \quad \begin{array}{ll} P(x) : & \min_{(a,b)} a^T x + b \\ & \text{s.t.} \quad a^T V + be \geq f(V) \\ & \quad a \in \mathbb{R}^n, b \in \mathbb{R} \end{array} \qquad \begin{array}{ll} D(x) : & \max_{\lambda} f(V)^T \lambda \\ & \text{s.t.} \quad V\lambda = x \\ & \quad e^T \lambda = 1 \\ & \quad \lambda \geq 0. \end{array}$$

156 The constraints of the primal problem $P(x)$ express that for the linear inequality $a^T x + b$ to be valid
157 for the concave envelope of f over $\text{conv}(V)$, its value at each of the points v_j must be larger than
158 $f(v_j)$. Given a point $x \in \mathbb{R}^n$, the dual problem searches to find, among all ways of describing x as a
159 convex combination of vectors v_j , one that yields the largest interpolated value. Let F denote the

feasible region of $P(x)$. Observe that F does not depend on x and that F is nonempty since b can be chosen arbitrarily large. Since $D(x)$ is feasible if $x \in \text{conv}(V)$ and since the feasible region of $D(x)$ is bounded, it follows from strong duality in linear programming that the optimal values of $P(x)$ and $D(x)$ are finite and equal for each $x \in \text{conv}(V)$. We denote this optimal value by $z(x)$. For a given $(a, b) \in F$, we let $J(a, b)$ denote the index set of constraints of F that are tight at (a, b) and let $R(a, b) = \text{conv}(V(J(a, b)))$. It follows from complementarity slackness conditions that if (a, b) is optimal for $P(x)$, then all optimal solutions λ to $D(x)$ belong to $R(a, b)$. In the following theorem, we record some relations between the above primal-dual pair and $\text{conc}(f)(x)$. Similar results have appeared in the literature. We will discuss these connections after the proof.

Theorem 2.4. *Consider a function $f : V \mapsto \mathbb{R}^n$ and let $\text{conc}(f)$ be its concave envelope over $\text{conv}(V)$. Also define $\mathcal{R} = \{R(a', b') \mid (a', b') \in \text{vert}(F)\}$. Then,*

1. $z(x) = \text{conc}(f)(x)$ for $x \in \text{conv}(V)$.
2. Let $(a^*, b^*) \in \text{vert}(F)$. Then, (a^*, b^*) is optimal for $P(x)$ if and only if $x \in R(a^*, b^*)$. Further, the extreme points of F are in one-to-one correspondence with the non-vertical facets of $\text{conc}(f)(x)$.
3. For each $(a', b') \in \text{vert}(F)$, $a'x + b' \geq f(x)$ defines a facet of $\text{conc}(f)$ over $R(a', b')$.
4. \mathcal{R} is a polyhedral subdivision of $\text{conv}(V)$. Further, $\text{conc}(f)$ can be computed by interpolating f affinely over each element of \mathcal{R} .

Proof. To prove (1), we consider $x' \in \text{conv}(V)$. Let λ' be any feasible solution of $D(x')$, then

$$\text{conc}(f)(x') = \text{conc}(f)(V\lambda') \geq \text{conc}(f)(V)^T \lambda' \geq f(V)^T \lambda' \quad (2)$$

where the equality follows from feasibility of λ' , the first inequality holds from concavity of $\text{conc}(f)$ and the second inequality is satisfied because $\text{conc}(f)(x) \geq f(x)$ for all $x \in \text{conv}(V)$. This implies that $\text{conc}(f)(x') \geq z(x')$ since λ' can be chosen to be an optimal solution of $D(x')$ in (2). Further, if (a', b') is feasible to F , then $a'^T x + b' \geq f(x)$ for all $x \in \{v_1, \dots, v_m\}$. Since affine functions are concave, we know that $a'x' + b' \geq \text{conc}(f)(x')$. This implies that $\text{conc}(f)(x') \leq z(x')$ since (a', b') can be chosen to be an optimal solution of $P(x')$. We conclude that $\text{conc}(f)(x') = z(x')$.

We now prove (2). Since $\text{aff}(\text{conv}(v_1, \dots, v_m)) = \mathbb{R}^n$ and $\text{rank}(V \mid e) = n + 1$, by Minkowski's representation theorem (see Theorem 4.8 in [22]), there exists an optimal solution (a^*, b^*) to $P(x)$ that is an extreme point of F . Consider any point $x'' \in R(a^*, b^*)$. Since x'' can be expressed as a convex combination of v_j , $j \in J(a^*, b^*)$, there exists a solution λ'' that is feasible to $D(x'')$ and that satisfies complementary slackness conditions with (a^*, b^*) . Therefore, (a^*, b^*) must be optimal to $P(x)$ for every $x \in R(a^*, b^*)$. Further, since (a^*, b^*) is an extreme point of F , at least $n + 1$ of the points in $V(J(a^*, b^*))$ are affinely independent. This implies that $a^*x + b^* \geq f(x)$ defines a facet of $\text{conc}(f)$. On the other hand, (a^*, b^*) cannot be optimal to $P(x'')$ if $x'' \notin R(a^*, b^*)$ since there does not exist a complementary dual feasible solution.

Consider a non-vertical facet G defined by $\tilde{a}x + \tilde{b} \leq f(x)$ and consider a point $(\tilde{x}, \tilde{a}\tilde{x} + \tilde{b})$ in the relative interior of this facet. First, note that (\tilde{a}, \tilde{b}) is feasible to F and $\tilde{a}\tilde{x} + \tilde{b} = \text{conc}(f)(\tilde{x}) = z(\tilde{x})$. Therefore, (\tilde{a}, \tilde{b}) is optimal to $P(\tilde{x})$. Since any underestimating inequality of $f(x)$ that is tight at $(\tilde{x}, \tilde{a}\tilde{x} + \tilde{b})$ is also tight everywhere on G and $\dim(G) = n$, it follows that the optimal solution for $P(\tilde{x})$ is unique. Since $P(\tilde{x})$ always has an extreme point solution, (\tilde{a}, \tilde{b}) must be an extreme point of F . Hence, there is a one-to-one correspondence between extreme points and facets of $\text{conc}(f)$.

We have shown that for each $x \in \text{conv}(V)$ there is an extreme point of F that optimizes $P(x)$ and the optimal value is $z(x)$. Therefore, \mathcal{R} is the subdivision of $\text{conv}(V)$ obtained by projecting

203 the hypograph of $z(x)$ to x -space. As proven above, the concave envelope is affine over each $R(a, b)$
 204 if $(a, b) \in \text{vert}(F)$ and $ax + b > \text{conc}(f)(x)$ if $x \notin R(a, b)$. Projecting the hypograph of a polyhedral
 205 function yields a (regular) polyhedral subdivision of the domain; see [17]. Further, for each extreme
 206 point, (a', b') , of F , $V(J(a', b')) \subseteq R(a', b')$ consists of at least $n + 1$ affinely independent points.
 207 Therefore, (a', b') can be recovered from $R(a', b')$ by solving the corresponding constraints of F . \square

208 Any polyhedral subdivision can be refined into a triangulation [17]. Therefore, by Theorem 2.4
 209 there exists a triangulation of the domain that is such that $\text{conc}(f)$ is affine over each simplex of the
 210 triangulation and $\text{conc}(f)(x) = f(x)$ at all extreme points x of the simplices of the triangulation.
 211 Theorem 2.4 can be partially extended to general nonlinear functions by expanding the set of
 212 constraints to include an inequality for each feasible point (or, more precisely, each point in the
 213 generating set); see [37] for details. The main idea is that since $b \geq f(x) - a^T x$ for all x , it follows
 214 that the objective is minimized when $b = (-f)^*(-a^T)$; see [25]. Then, $\inf\{a^T x + b\} = \inf\{a^T x +$
 215 $(-f)^*(-a^T)\} = -\sup\{-a^T x - (-f)^*(-a^T)\} = -(-f)^{**}(x)$. If the underlying set is compact and
 216 $f(x)$ is upper-semicontinuous, $f(x)$ is bounded from above. Therefore, $-(-f)^{**}(x) = \text{conc } f(x)$ by
 217 Theorem 1.3.5 in [14]. The advantage of restricting the result to finite point sets is that F has
 218 finitely many constraints, and, as a result, one can identify the facets of the concave envelope as
 219 well as the simplices of the corresponding triangulation by studying the basic feasible solutions of F .
 220 When Theorem 2.4 is applied to functions that are concave-extendable from vertices of a hypercube,
 221 the number of constraints defining F is exponentially large, since a constraint is created for each
 222 extreme point of the hypercube. As a result, identifying the basic feasible solutions of F can be
 223 computationally difficult. In this paper, we identify situations where these basic feasible solutions
 224 can be identified explicitly. We now relate Theorem 2.4 to existing results in the literature.

225 Concave-extendability has been used in [3] to develop an algorithmic approach for the derivation
 226 of concave envelopes. In particular, the authors designed a column-generation algorithm to find a
 227 facet of the concave envelope of a function that is concave-extendable from vertices by separating
 228 the envelope from a pre-specified point. They also proved, using a slightly different proof technique,
 229 the following result that establishes the correspondence between the facets of the concave envelope
 230 and the basic solutions of $P(x)$.

231 **Corollary 2.5** (Theorem 2.4 in [3]). *$z = a^{*T}x + b^*$ defines a non-vertical facet of the concave*
 232 *envelope of the multilinear function $f(x)$ over $P = \prod_{i=1}^n [l_i, u_i]$ if and only if (a^*, b^*) is a basic*
 233 *feasible solution of the following linear programming problem:*

$$\begin{aligned}
 & \min_{(a,b)} && a^T x + b \\
 & \text{s.t.} && a^T v^j + b \geq f(v^j) \quad \forall v^j \in \text{vert}(P) \\
 & && a \in \mathbb{R}^n, b \in \mathbb{R}.
 \end{aligned} \tag{3}$$

235 *Proof.* Multilinear functions are concave-extendable from vertices of hypercubes; see [24]. Letting
 236 $V = \text{vert}(P)$, the result follows directly from Theorem 2.4. \square

237 **Corollary 2.6** (Lemma 1.1 in [24]). *Let $f(x)$ be a continuously differentiable function on an n -*
 238 *dimensional convex polytope P . Assume $\text{conc}(f)(x)$ over P is a polyhedral function. Let $h(x) =$*
 239 *$ax + b$ be an affine function and assume that there exist $v^i, i = 1, \dots, n+1$, $n+1$ affinely independent*
 240 *vertices of P , such that $h(v^i) = f(v^i), i = 1, \dots, n+1$ and $h(x) \geq f(x)$ for all $x \in \text{vert}(P)$. Then,*
 241 *$h(x)$ is an element of $\text{conc}(f)$ and, in particular, $h(x)$ defines the concave envelope of $f(x)$ over*
 242 *$\text{conv}(v^1, \dots, v^{n+1})$.*

243 *Proof.* For a continuously differentiable function, $\text{conc}(f)$ is polyhedral if and only if f is concave-
 244 extendable from vertices; see Theorem 1.1 in [24]. Note that $ax + b \geq f(x)$ for all $x \in \text{vert}(P)$ and
 245 $a^T v^i + b = f(v^i)$ for $n + 1$ affinely independent vertices establish that (a, b) is an extreme point of
 246 F . Since $\text{conv}(v^1, \dots, v^{n+1}) \subseteq R(a, b)$, the result follows from Theorem 2.4. \square

247 We now exploit Theorem 2.4 to study functions constructed by affine extensions over triangula-
 248 tions. Formally, let $\mathcal{S} = (S_1, \dots, S_m)$ be a triangulation of $\text{conv}(V)$ that does not add new vertices,
 249 where S_i is a simplex for each i and J_i denotes the index set of vertices of S_i . We construct the
 250 function $f^{\mathcal{S}} : S \mapsto \mathbb{R}$ by interpolating the function f affinely over each simplex S_i . More precisely,
 251 given a point $x \in S$, there exists an i such that $x \in S_i$. Since S_i is a simplex, there exists a unique
 252 λ that is feasible to $D(x)$ and is such that $\lambda_j = 0$ for all $j \notin J_i$. Then, we define $f^{\mathcal{S}}(x) = f(V)^T \lambda$.
 253 Note that this definition is consistent because if $x \in S_i \cap S_{i'}$, then x belongs to a common face of
 254 S_i and $S_{i'}$, and $\lambda_j = 0$ for all $j \notin J_i \cap J_{i'}$.

255 **Corollary 2.7.** *Consider a function $f : V \mapsto \mathbb{R}$, and let \mathcal{S} be a triangulation of $\text{conv}(V)$ that does
 256 not add vertices. Then, $f^{\mathcal{S}}$ is the concave envelope of f over $\text{conv}(V)$ if and only if $f^{\mathcal{S}}$ is concave.*

257 *Proof.* Clearly, $f^{\mathcal{S}}$ is a concave envelope of f only if it is concave. Now, we show the converse.
 258 By construction, $f^{\mathcal{S}}(x)$ is the objective value of a feasible solution in $D(x)$. Then, it follows from
 259 Theorem 2.4 that for any $x \in \text{conv}(V)$, $f^{\mathcal{S}}(x) \leq \text{conc}(f)(x)$. Further, $f^{\mathcal{S}}(x) = f(x)$ whenever $x \in V$
 260 and so $f^{\mathcal{S}}(x) \geq f(x)$. Since $f^{\mathcal{S}}$ is concave, $f^{\mathcal{S}}(x) \geq \text{conc}(f)(x)$. Therefore, for any $x \in \text{conv}(V)$,
 261 $f^{\mathcal{S}}(x) = \text{conc}(f)(x)$. \square

262 The ideas in Corollary 2.7 can be extended to more general settings using the notion of barycen-
 263 tric coordinates or inclusion certificates; see [34]. Theorem 2.4 was proven with a finite point set
 264 and can be used to construct concave envelopes of functions restricted to this set. If the optimal
 265 value function of $P(x)$ turns out to be the concave envelope of the unrestricted f over $\text{conv}(V)$,
 266 then it follows that f must be concave-extendable from V . This observation is formalized below.

267 **Corollary 2.8.** *Consider a function $f : \text{conv}(V) \mapsto \mathbb{R}$. Then, there exists a triangulation \mathcal{S} using
 268 only the vertices in V such that $f^{\mathcal{S}}$ is the concave envelope of f over $\text{conv}(V)$ if and only if f is
 269 concave-extendable from V .*

270 *Proof.* If f is concave-extendable from V , then the result follows directly from Theorem 2.4 and
 271 the fact that any polyhedral subdivision can be refined into a triangulation. For the converse,
 272 let \mathcal{S} be a triangulation for which $f^{\mathcal{S}}$ is the concave envelope of f over $\text{conv}(V)$. It follows that,
 273 $f^{\mathcal{S}}(x) \leq z(x) \leq \text{conc}(f)(x) = f^{\mathcal{S}}(x)$, where the first inequality is satisfied because $f^{\mathcal{S}}(x)$ corresponds
 274 to a feasible solution for $D(x)$, the second inequality follows from Theorem 2.4 where it is shown
 275 that $z(x)$ is the concave envelope of f restricted to V , and the last equality holds because of our
 276 assumption. Therefore, the equality holds throughout. Then, $z(x) = \text{conc}(f)(x)$ which in turn
 277 implies by Theorem 2.4 that f is concave-extendable from V . \square

278 Consider the problem $M(r, s) = \max\{f(x) - r^T x - s \mid x \in V\}$. The ability to construct the
 279 concave envelope of $f(x)$ is closely related to the ability to solve $M(r, s)$.

280 **Corollary 2.9.** *If $M(r, s)$ can be solved in polynomial time, then $P(x)$ can also be solved in poly-
 281 nomial time. Further, if there is a polynomial-time separation algorithm for $\text{conv}(V)$, a polynomial-
 282 time algorithm to find an optimal solution for $D(x)$, and a polynomial-time algorithm to solve $P(x)$,
 283 then $M(r, s)$ can be solved in polynomial time.*

284 *Proof.* We first show the first statement of the corollary. Assume there exist a polynomial-time
 285 algorithm to solve $M(r, s)$. We show that a polynomial-time separation algorithm can be constructed
 286 for $P(x)$. For any solution (a, b) , we solve $M(a, b)$. If the optimal value $M(a, b)$ is nonpositive, then
 287 $f(x) \leq a^T v + b$ for all $v \in V$ and therefore $(a, b) \in F$. Otherwise, the optimal solution of $M(a, b)$
 288 gives a hyperplane separating (a, b) from F . Therefore, the optimization oracle for $M(r, s)$ yields a
 289 separation oracle for $P(x)$. Then, the result follows from Theorem 6.4.9 in [12].

290 We now prove the second statement of the corollary. Define $M'(r, s)$ as $\max\{\text{conc}(f)(x) - r^t x - s \mid$
 291 $x \in \text{conv}(V)\}$, where $\text{conc}(f)(x)$ is the concave envelope of $f(x)$ over $\text{conv}(V)$. We show that the
 292 optimal value of $M(r, s)$ is the same as that of $M'(r, s)$. Clearly, the optimal value of $M(r, s)$ is
 293 no larger than that of $M'(r, s)$. For the converse, consider the optimal solution x' to the $M'(r, s)$.
 294 Let λ' be the optimal solution to $D(x')$. Then, $(\text{conc}(f)(x') - r^t x' - s)e^t \geq f(V)^T - r^t V - se^t$,
 295 where $e \in \mathbb{R}^m$ is a vector of all ones. Since $(\text{conc}(f)(x') - r^t x' - s)e^t \lambda' = \text{conc}(f)(x') - r^t x' - s =$
 296 $(f(V)^T - r^t V - se^t)\lambda'$, it follows that $\text{conc}(f)(x') - r^t x' - s = f(v) - r^t v - s$ for any v in the support
 297 of λ' . Therefore, given the optimal solution to $M'(r, s)$, λ' can be computed in polynomial time and,
 298 as a result, a solution to $M(r, s)$ can be computed. Now, we solve $M'(r, s)$ by reformulating it as
 299 $M''(r, s)$ which is defined as $\max\{t \mid \text{conc}(f)(x) - r^t x - s - t \geq 0, x \in \text{conv}(V)\}$. Using Theorem 6.4.9
 300 in [12], it suffices to construct a strong separation oracle for $M''(r, s)$. Given (\bar{t}, \bar{x}) , if $\bar{x} \notin \text{conv}(V)$
 301 we can use the separation algorithm for $\text{conv}(V)$. Otherwise, solve $P(\bar{x})$ and let (\bar{a}, \bar{b}) be its optimal
 302 solution. Then, define $a' = \bar{a} - r$ and $b' = \bar{b} - s - \bar{t}$. It follows that $a'^t x + b' \geq \text{conc}(f)(x) - r^t x - s - \bar{t}$
 303 for all $x \in \text{conv}(V)$ and $a'^t \bar{x} + b' = \text{conc}(f)(\bar{x}) - r^t \bar{x} - s - \bar{t}$. Therefore, $a'^t \bar{x} + b' \geq 0$ if and only
 304 if (\bar{t}, \bar{x}) is feasible. Otherwise, if $a'^t \bar{x} + b' < 0$, we find a separating hyperplane $a'^t x + b' \geq 0$ that
 305 separates the feasible region of $M''(r, s)$ from (\bar{t}, \bar{x}) . \square

306 Although the proof that an algorithm to solve $M(r, s)$ can be used to solve $P(x)$ uses the
 307 ellipsoid algorithm, it is possible develop a Dantzig-Wolfe decomposition algorithm (albeit without
 308 polynomial time complexity) for the solution of $D(x)$ using the algorithm for $M(r, s)$; see Bao
 309 et al. [3] for details. The proof technique used to show that $M(r, s)$ can be solved using algorithms
 310 for separation of $\text{conv}(V)$ and optimization routines for $D(x)$ and $P(x)$ is similar to that used in
 311 [12] for showing that submodular function minimization is polynomially solvable. Corollary 2.9
 312 is also related to Theorem 1 in [33] in that the author discusses the equivalence of the concave
 313 envelopes of two functions f and f' if the optimization problems $\max\{f(x) - r^t x - s \mid x \in V\}$ and
 314 $\max\{f'(x) - r^t x - s \mid x \in V\}$ have the same optimal value.

315 The formulation of the concave envelope as in Theorem 2.4 enables one to compute the concave
 316 envelope for functions defined as a maximum of other functions. Consider $f_i : V \mapsto \mathbb{R}$, $i \in 1, \dots, k$.
 317 We denote $P(x)$, $D(x)$, and F associated with f_i as $P(f_i, x)$, $D(f_i, x)$, and $F(f_i)$ respectively.

318 **Corollary 2.10.** *Consider a collection of functions $f_i : V \mapsto \mathbb{R}$, $i \in 1, \dots, k$. If there exists a*
 319 *polynomial-time algorithm to solve $P(f_i, x)$ for each i and $x \in \text{conv}(V)$, and a polynomial-time*
 320 *strong separation algorithm for $\text{conv}(V)$, then there exists a polynomial-time algorithm to optimize*
 321 *a linear function over $F(\max\{f_1, \dots, f_k\})$, and hence to solve $P(\max\{f_1, \dots, f_k\}, x)$.*

322 *Proof.* Consider the optimization problem $P'(f_i, x, r)$ defined as $\min\{a^T x + br \mid (a, b) \in F(f_i)\}$.
 323 Denote its optimal value by $z(f_i, x, r)$. We first construct a strong optimization oracle for $P'(f_i, x, r)$
 324 [12], *i.e.*, an oracle that provides an optimal solution if one exists, otherwise it returns a recession
 325 direction in which the objective function decreases. Since $F(f_i) \neq \emptyset$, the recession cone of $F(f_i)$,
 326 denoted as $0^+(F(f_i))$, is given by $\{(a, b) \mid av + b \geq 0 \text{ for all } v \in V\}$.

327 Since $z(f_i, x, r)$ is positively homogeneous in (x, r) , by scaling if necessary, we may assume that
 328 r is 1, -1 , or 0. If $x \in \text{conv}(V)$ and $r = 1$, the oracle is assumed to be available. If $x \notin \text{conv}(V)$
 329 and $r = 1$, then using the separation routine for $\text{conv}(V)$ we can find in polynomial time a ρ

330 such that $\rho^T x < c$ and $\rho^T v \geq c$ for all $v \in V$. Then, $(\rho^T, -c) \in 0^+(F(f_i))$ and is the desired
331 recession direction. Now, we assume that $r = 0$. If $x = 0$ then the optimal solution of $P(f_i, 0)$
332 is optimal to $P'(f_i, x, r)$. Otherwise, there exists an x_k such that $x_k \neq 0$. If $x_k < 0$, use the
333 strong separation oracle of $\text{conv}(V)$ to compute $x_k^L = \min\{x'_k \mid x' \in \text{conv}(V)\}$; see Theorem
334 6.4.9 in [12]. Then, $v_k - x_k^L \geq 0$, for all $v \in V$ and therefore $(e_k^T, -x_k^L) \in 0^+(F)$ is the desired
335 recession direction, where e_k is the k^{th} principal vector. On the other hand, if $x_k > 0$, then compute
336 $x_k^U = \max\{x'_k \mid x' \in \text{conv}(V)\}$ and, as before, $(-e_k^T, x_k^U)$ is the desired recession direction. Now,
337 assume that $r = -1$. Then, $(0, 1) \in 0^+(F(f_i))$ is the desired recession direction.

338 Since $F(\max\{f_1, \dots, f_k\}) = \bigcap_{i=1}^k F(f_i)$, the strong optimization oracles can be used to opti-
339 mize a linear function over $F(\max\{f_1, \dots, f_k\})$ and hence to solve $P(\max\{f_1, \dots, f_k\}, x)$ using the
340 ellipsoid algorithm; see Corollary 14.1d in [28]. \square

341 In most applications, the underlying polyhedron $\text{conv}(V)$ will typically be simple and so the
342 corresponding separation algorithm will be trivial. We will describe, in the forthcoming sections,
343 various types of functions for which concave envelopes can be obtained in polynomial time. It follows
344 from Corollary 2.10 that the concave envelope of the maximum of any subset of these functions can
345 also be computed in polynomial time.

346 The above algorithm is polynomial-time only if k is treated as part of the input. Otherwise,
347 as we will describe later, the convex envelope over $[0, 1]^n$ of a function that is submodular when
348 restricted to $\{0, 1\}^n$ can be expressed as a maximum of exponentially many linear functions. Since
349 $\text{conv}(f) \leq f$, it follows easily that $\text{conc}(\text{conv}(f)) \leq \text{conc}(f)$. Further, since each point in $\{0, 1\}^n$
350 belongs to $\text{vert}([0, 1]^n)$, it follows that $\text{conv}(f) = f$ at each $v \in V$. Therefore, $\text{conc}(\text{conv}(f)) \geq$
351 $\text{conc}(f)$. Combining, $\text{conc}(\text{conv}(f)) = \text{conc}(f)$. If k was not part of input, Corollary 2.10 would
352 imply that $P(x)$ can be solved in polynomial time for a submodular function, giving a polynomial-
353 time separation routine for maximizing a submodular function. This, in turn, is not possible unless
354 $P = NP$.

355 Corollary 2.10 can also be proven using disjunctive programming if an explicit polynomial-sized
356 characterization of the facets of f_i is available for each i . The main idea would be to express
357 the hypograph of $\max\{f_1, \dots, f_k\}$ as the convex hull of the union of hypographs for each f_i in a
358 lifted space; see Theorem 16.5 in [25]. This would provide an explicit polynomial-sized polyhedral
359 representation of the concave envelope in a higher-dimensional space.

360 **3 Supermodular function that is concave-extendable from vertices**

361 In this section, we use a result of Lovász [19] to derive the triangulation associated with the concave
362 envelope of supermodular functions. This allows us to construct closed-form expressions for the
363 concave envelopes of supermodular functions over the hypercube assuming that these functions are
364 concave-extendable from vertices. We then demonstrate the utility of this construction in two ways.
365 First, we provide a direct and unified derivation of many recent results in the literature (each of
366 which was initially proven using a different technique) as a consequence of this simple construction.
367 Second, we show that it can be used to improve the relaxations currently used in existing factorable
368 programming solvers; see [39, 18, 4]. In particular, factorable programming techniques [20] typically
369 use variable substitution to relax a function expressed as a composition of a convex function with a
370 linear function during the construction of relaxations. We will show, among many other examples,
371 that the techniques described in this section apply to this structure.

372 It follows from our discussion in Section 2 that the facets of the concave envelope of any function
373 that is concave-extendable from the vertices of a polytope P can be obtained through the solution of
374 a linear program, $P(x)$, which has a constraint for every vertex of P . As a result, the linear program

375 typically has an exponential number of constraints, limiting the applicability of the technique.
 376 However, if the function under study is well-structured, we show that it is sometimes possible
 377 to deduce the triangulation associated with its concave envelope by explicitly characterizing the
 378 solution of the linear program. Supermodularity is one such function structure that permits an
 379 a-priori derivation of the corresponding triangulation.

380 **Definition 3.1** ([42]). *A function $f(x) : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be supermodular if $f(x' \vee x'') +$
 381 $f(x' \wedge x'') \geq f(x') + f(x'')$ for all $x', x'' \in S$, where $x' \vee x''$ denotes the component-wise maximum
 382 and $x' \wedge x''$ denotes the component-wise minimum of x' and x'' . \square*

383 An important special case of the above definition is encountered when $S = \{0, 1\}^n$. In this
 384 case, any element x of S is of the form $x = \sum_{i \in K} e_i$ where e_i is the i^{th} unit vector in \mathbb{R}^n and
 385 $K \subseteq \{1, \dots, n\}$. Then, f can also be viewed as a set function in the following way. We define
 386 $f' : 2^N \rightarrow \mathbb{R}$ as $f'(K) = f(\sum_{j \in K} e_j)$. Then, $f(x)$ is supermodular if and only if $f'(A \cap B) + f'(A \cup$
 387 $B) \geq f'(A) + f'(B)$.

388 Given a function $f : \{0, 1\}^n \mapsto \mathbb{R}$ that is supermodular, it follows from Theorem 2.4 that there
 389 is a triangulation of the hypercube that yields the concave envelope of f . We show in Theorem 3.3
 390 that this triangulation is in fact Kuhn's triangulation. A triangulation $\mathcal{K} = \{\Delta_1, \dots, \Delta_n!\}$ is said
 391 to be *Kuhn's triangulation* of the hypercube, $[0, 1]^n$, if the simplices of \mathcal{K} are in a one-to-one
 392 correspondence with the permutations of $\{1, \dots, n\}$ as discussed next. Given a permutation, π of
 393 $\{1, \dots, n\}$, the $n + 1$ vertices of the corresponding simplex Δ_π are $\{(0, \dots, 0) + \sum_{j=1}^k e_{\pi(j)} \mid k =$
 394 $0, \dots, n\}$; see [17]. Observe that the origin is a vertex of each of the simplices composing Kuhn's
 395 triangulation.

396 We define the *Lovász extension* [19] of a function $f(x)$ as $f^{\mathcal{K}}(x)$. Given any $x \in [0, 1]^n$, we can
 397 find a permutation π of $\{1, \dots, n\}$ such that $x_{\pi(1)} \geq x_{\pi(2)} \geq \dots \geq x_{\pi(n)}$ by sorting the components
 398 of x . It is clear that x belongs to Δ_π since it can be expressed as the following convex combination
 399 of its extreme points: $x = (1 - x_{\pi(1)})0 + \sum_{j=1}^{n-1} (x_{\pi(j)} - x_{\pi(j+1)}) \left(\sum_{r=1}^j e_{\pi(r)} \right) + x_{\pi(n)} \left(\sum_{r=1}^n e_{\pi(r)} \right)$. It
 400 follows that

$$\begin{aligned}
 401 \quad f^{\mathcal{K}}(x) &= (1 - x_{\pi(1)})f(0) + \sum_{j=1}^{n-1} (x_{\pi(j)} - x_{\pi(j+1)}) f\left(\sum_{r=1}^j e_{\pi(r)}\right) + x_{\pi(n)} f\left(\sum_{r=1}^n e_{\pi(r)}\right) \\
 402 \quad &= \sum_{i=1}^n \left(f\left(\sum_{j=1}^i e_{\pi(j)}\right) - f\left(\sum_{j=1}^{i-1} e_{\pi(j)}\right) \right) x_{\pi(i)} + f(0) \tag{4}
 \end{aligned}$$

403 for all $x \in \Delta_\pi$.

404 We next present a result that is crucial in developing the concave envelope of a supermodular
 405 function that is concave-extendable from the vertices of the unit hypercube. Because it plays an
 406 important role in the subsequent development, we provide here a self-contained proof using the
 407 techniques of Section 2. We note however that this lemma was first stated, although not explicitly
 408 proven, in Lovász [19].

409 **Lemma 3.2** (Proposition 4.1 in [19]). *$f^{\mathcal{K}}$ is concave if and only if f restricted to $\{0, 1\}^n$ is super-*
 410 *modular.*

411 *Proof.* Given $S \subseteq \{1, \dots, n\}$, let $\chi(S)$ be the indicator vector of S . Consider two arbitrary subsets,
 412 X and Y , of $\{1, \dots, n\}$. Then, if $f^{\mathcal{K}}$ is concave, the following argument shows that f restricted to

413 $\{0, 1\}^n$ is supermodular:

$$\begin{aligned}
& \frac{1}{2}f(\chi(X)) + \frac{1}{2}f(\chi(Y)) = \frac{1}{2}f^{\mathcal{K}}(\chi(X)) + \frac{1}{2}f^{\mathcal{K}}(\chi(Y)) \leq f^{\mathcal{K}}\left(\frac{1}{2}(\chi(X) + \chi(Y))\right) \\
& = f^{\mathcal{K}}\left(\frac{1}{2}\chi(X \cup Y) + \frac{1}{2}\chi(X \cap Y)\right) = \frac{1}{2}f(\chi(X \cup Y)) + \frac{1}{2}f(\chi(X \cap Y)).
\end{aligned}$$

415 Here, the first inequality follows from concavity of $f^{\mathcal{K}}(x)$, the second equality is satisfied since
416 $\chi(X) + \chi(Y) = \chi(X \cup Y) + \chi(X \cap Y)$, and the last equality holds because $f^{\mathcal{K}}$ is affine over the line
417 segment $[\chi(X \cap Y), \chi(X \cup Y)]$ since this line segment is completely contained in at least one of the
418 simplices Δ_π .

419 Now, we argue that if f restricted to $\{0, 1\}^n$ is supermodular then $f^{\mathcal{K}}(x)$ is concave. To this
420 end, we will show that $f^{\mathcal{K}}(x) = z(x)$, where $z(x)$ is the optimal value of $P(x)$. Since $z(x)$ is the
421 minimum of affine functions of x , one for each $(a, b) \in F$, it will follow that $f^{\mathcal{K}}(x)$ is concave.
422 Consider $x' \in [0, 1]^n$ and assume without loss of generality, by reordering the components of x'
423 if necessary that $x'_1 \geq \dots \geq x'_n$. Since the multipliers $(1 - x'_1), (x'_1 - x'_2), \dots, x'_n$ yield a feasible
424 solution to $D(x')$, it follows from weak duality that $f^{\mathcal{K}}(x') \leq z(x')$.

425 To show that $z(x') \leq f^{\mathcal{K}}(x')$, we show that $a'_i = f\left(\sum_{r=1}^i e_r\right) - f\left(\sum_{r=1}^{i-1} e_r\right)$ and $b' = f(0)$ solves
426 $P(x')$ and has objective value $f^{\mathcal{K}}(x')$. To this end, we show first that $(a', b') \in F$, i.e., $a'^T v + b' \geq$
427 $f(v)$ for all $v \in \{0, 1\}^n$ by induction on $\|v\|_1$. The base case is clear since $v = 0$ is the only vector
428 with $\|v\|_1 = 0$ and since $b' = f(0)$. For the inductive step, consider $v \in \{0, 1\}^n$ and assume that
429 the result holds for all $w \in \{0, 1\}^n$ with $\|w\|_1 < \|v\|_1$. Define k to be the largest index for which
430 $v_k = 1$. Then,

$$\begin{aligned}
431 \quad a'^T v + b' &= a'^T(v - e_k) + b' + a'^T e_k \geq f(v - e_k) + f\left(\sum_{r=1}^k e_r\right) - f\left(\sum_{r=1}^{k-1} e_r\right) \\
432 \quad &= f\left(v \wedge \sum_{r=1}^{k-1} e_r\right) + f\left(v \vee \sum_{r=1}^{k-1} e_r\right) - f\left(\sum_{r=1}^{k-1} e_r\right) \geq f(v),
\end{aligned}$$

433 where the first inequality follows from the inductive hypothesis and the definition of a'_k , the second
434 equality follows from the definition of k , and the second inequality holds because of the supermod-
435 ularity of f . By construction, see also (4), $a'^T x' + b' = f^{\mathcal{K}}(x')$ and therefore $z(x') \leq f^{\mathcal{K}}(x')$. \square

436 It seems that Lemma 3.2 was originally motivated by Edmonds' greedy algorithm for optimizing
437 linear function over extended polymatroids [11]. Although, in the proof of Lemma 3.2, we replaced
438 this optimization problem with $P(x)$, the proof still makes use of Edmonds' algorithm implicitly.
439 We discuss the connections next. First, note that F reduces to an extended polymatroid when b is
440 restricted to be zero and $V = \{0, 1\}^n$. In general, if b is assumed to be zero in $P(x)$, then the optimal
441 value function $z(x)$ of $P(x)$ yields the tightest positively homogeneous concave overestimator of f
442 instead of its concave envelope; see, for example, Proposition 2 in [23]. If $f(x)$ is supermodular, the
443 concave envelope is positively homogeneous as long as $f(0) = 0$, an assumption that can be made
444 without loss of generality by translating f if necessary. For more general functions, however, the
445 concave envelope may not be positively homogeneous over the domain and assuming $b = 0$ would
446 be restrictive in those cases. If $f(x)$ is supermodular, in the light of Theorem 2.4, the above proof
447 shows that $f^{\mathcal{K}}(x) = \text{conc}(f)(x)$. This fact can be derived from Lemma 3.2 using Corollary 2.7.

448 **Theorem 3.3.** *Consider a function $f : [0, 1]^n \mapsto \mathbb{R}^n$. The concave envelope of f over $[0, 1]^n$ is*
449 *given by $f^{\mathcal{K}}(x)$ if and only if f is supermodular when restricted to $\{0, 1\}^n$ and concave-extendable*
450 *from the vertices of the unit hypercube.*

451 *Proof.* If f is concave-extendable from the vertices of the unit hypercube and supermodular when
 452 restricted to $\{0, 1\}^n$ then it follows from Lemma 3.2 and Corollary 2.7 that $f^{\mathcal{K}}(x)$ is the concave
 453 envelope of $f(x)$. On the other hand, if $f^{\mathcal{K}}(x)$ is the concave envelope of $f(x)$, then it follows
 454 from Lemma 3.2 and Corollary 2.8 that f restricted to $\{0, 1\}^n$ is supermodular and f is concave-
 455 extendable from $\{0, 1\}^n$. \square

456 Theorem 3.3 establishes that the concave envelope of a function that is concave-extendable from
 457 the vertices of the unit hypercube and that is supermodular when restricted to $\{0, 1\}^n$ is its Lovász
 458 extension. It follows from the proof of Lemma 3.2 that each of the linear functions (4) is valid for
 459 $\text{conc}_{[0,1]^n} f(x)$ and therefore

$$460 \quad \text{conc}_{[0,1]^n} f(x) = \min_{\pi \in \Pi} \sum_{i=1}^n \left(f \left(\sum_{j=1}^i e_{\pi(j)} \right) - f \left(\sum_{j=1}^{i-1} e_{\pi(j)} \right) \right) x_{\pi(i)} + f(0) \quad (5)$$

461 where Π is the set of permutations of $\{1, \dots, n\}$. By encoding the permutations differently, we can
 462 also establish that

$$463 \quad \text{conc}_{[0,1]^n} f(x) = \min_{\pi \in \Pi} \sum_{i=1}^n \left(f \left(\sum_{j|\pi(j) \leq \pi(i)} e_j \right) - f \left(\sum_{j|\pi(j) < \pi(i)} e_j \right) \right) x_i + f(0) \quad (6)$$

464 an expression that is sometimes easier to use.

465 Next, we show that supermodularity can also help to obtain the concave envelope of certain
 466 functions over sets other than the unit hypercube (or more generally a hyper-rectangle). To this
 467 end, consider a directed graph $G = (V, E)$ where $V = \{1, \dots, n\}$ and let I_0 and I_1 be non-intersecting
 468 subsets of $\{1, \dots, n\}$. Consider the sets $C = \bigcap_{(i,j) \in E} \{x \mid x_i \geq x_j\}$, $C_0 = \bigcap_{i \in I_0} \{x \mid x_i = 0\}$, and
 469 $C_1 = \bigcap_{i \in I_1} \{x \mid x_i = 1\}$. Define

$$470 \quad S = [0, 1]^n \cap C \cap C_0 \cap C_1.$$

471 The matrix associated with the constraints in C is composed of the node-edge incidence matrix of a
 472 directed graph appended with identity matrices. Therefore, it is totally unimodular. It follows that,
 473 whenever S is nonempty, its vertices are binary. Further, Kuhn's triangulation gives a polyhedral
 474 subdivision of S . This can be seen by considering a point $x \in S$. Sort the coordinates of x in a
 475 non-decreasing order extending the pre-order defined by G . If σ is the corresponding permutation
 476 of $\{1, \dots, n\}$, then x clearly belongs to the associated simplex of Kuhn's triangulation, i.e. $x \in \Delta_\sigma$.
 477 Let T be the face of Δ_σ such that $x \in \text{ri}(T)$. Let $v \in \text{vert}(T)$. Then, it can be verified that
 478 $v \in \{0, 1\}^n \cap S$. Further, note that if x and y belong to S , then so do $x \vee y$ and $x \wedge y$. Thus, the set
 479 S is the convex hull of the incidence vectors of a lattice family, where a lattice family is a family of
 480 sets \mathcal{C} such that if $A, B \in \mathcal{C}$, then $A \cap B$ and $A \cup B$ also belong to \mathcal{C} . By a slight modification of
 481 Proposition 10.3.3 in Grötschel et al. [12], it can be shown that the incidence vectors of a finitely-
 482 sized lattice family can be expressed as the vertices of S by appropriately defining C , C_0 , and C_1 . A
 483 function f is said to be supermodular for a lattice family \mathcal{C} or the corresponding incidence vectors,
 484 $\text{vert}(S)$, if $f(A \cap B) + f(A \cup B) \geq f(A) + f(B)$ for all $A, B \in \mathcal{C}$.

485 **Corollary 3.4.** *Let $f : S \mapsto \mathbb{R}^n$ be supermodular when restricted to $\text{vert}(S)$ and concave-extendable*
 486 *from the vertices of S . Then, for any $x \in S$, $f^{\mathcal{K}}(x)$ is well-defined and forms the concave envelope*
 487 *of f over S .*

488 *Proof.* Because of the form of S and the Corollary's assumption, f restricted to $\text{vert}(S)$ can be
 489 extended to $\bar{f} : \{0, 1\}^n \mapsto \mathbb{R}$ in such a way that \bar{f} is supermodular when restricted to $\{0, 1\}^n$; see

490 Theorem 49.2 in [29]. Let $x' \in S$. Then, $x' \in \text{ri}(T)$ where T is a face of Δ_σ and σ is an ordering
491 of coordinates of x' consistent with the pre-ordering of coordinates defining S and such that the
492 coordinates of x' are sorted in non-decreasing order. Since the vertices of T belong to S , it follows
493 that $f^{\mathcal{K}}(x')$ is well-defined and $\bar{f}^{\mathcal{K}}(x') = f^{\mathcal{K}}(x')$. Let $h(x)$ be the concave envelope of $f(x)$ over S .
494 By Theorem 3.3, $\bar{f}^{\mathcal{K}}(x)$ is the concave envelope of \bar{f} over $[0, 1]^n$. Therefore, by concave-extendability
495 of f from $\text{vert}(S)$, it follows that $f^{\mathcal{K}}(x') = \bar{f}^{\mathcal{K}}(x') \geq h(x')$. However $f^{\mathcal{K}}(x')$ is also a feasible solution
496 to $D(x')$ for $V = \text{vert}(S)$. Therefore, $f^{\mathcal{K}}(x') \leq h(x')$. In other words, $f^{\mathcal{K}}(x') = h(x')$. \square

497 As was exploited in the proof of Corollary 3.4, an extension of f restricted to $\text{vert}(S)$, say \bar{f} , can
498 be constructed that is supermodular when restricted to $\{0, 1\}^n$. Instead, if f itself can be extended
499 to $[0, 1]^n$ such that the resulting function is not only supermodular when restricted to $\{0, 1\}^n$ but
500 is also concave-extendable from $\{0, 1\}^n$, then the concave-extendability of f from $\text{vert}(S)$ follows.
501 This is because $\bar{f}^{\mathcal{K}}(x) = \text{conc}_{[0,1]^n} \bar{f}(x) \geq \text{conc}_S f(x) \geq f^{\mathcal{K}}(x)$, where the first equality follows from
502 Theorem 3.3, the first inequality since $S \subseteq [0, 1]^n$, and the second inequality since $f^{\mathcal{K}}(x)$ is a feasible
503 solution to $D(x)$. But, as argued above, $f^{\mathcal{K}}(x) = \bar{f}^{\mathcal{K}}(x)$. Therefore, the equality holds throughout
504 and, as a result, f is concave-extendable from $\text{vert}(S)$.

505 **Remark 3.5.** Consider a polyhedral subdivision of $\text{conv}(V)$, namely $\bigcup_{i \in I} S_i$, which defines the con-
506 cave envelope of $f(x) : V \mapsto \mathbb{R}^n$. Let $V' \subseteq V$ and S'_i be a polytope that is a subset of S_i and whose
507 vertices belong to V' . Then, $\text{conc}_{S'_i}(f) \leq \text{conc}_{\text{conv}(V')}(f)$. Note that $\text{conc}_{S'_i}(f) = \text{conc}_{S_i}(f) =$
508 $\text{conc}_{\text{conv}(V)}(f)$ where the first equality follows by affinity of $\text{conc}_{S_i}(f)$ and the second from the
509 structure of the polyhedral subdivision. It follows that $\text{conc}_{S'_i}(f) = \text{conc}_{\text{conv}(V')}(f)$. Therefore if
510 $V' = \bigcup_{i \in I} S'_i$, then the concave envelope of f over V' is obtained by restricting the concave envelope
511 of f over V to V' . This observation was the key to the proof of Corollary 3.4. We will encounter
512 various other applications of this observation in the remainder of the paper. \square

513 It can be shown that Theorem 3.3 and Corollary 3.4 generalize many results that have been
514 developed for specific functions. To demonstrate the applicability of Theorem 3.3, we will now derive
515 a variety of results from the literature as a consequence. Theorem 3.3 asserts that, for a given f , the
516 concave envelope of f over the unit hypercube is $f^{\mathcal{K}}(x)$ if and only if f is supermodular and concave-
517 extendable from vertices. Proofs in the literature typically demonstrate that $f^{\mathcal{K}}(x)$ is the concave
518 envelope directly. However, the latter properties are often much easier to prove as we illustrate
519 below. In these discussions, the following result is useful in establishing the supermodularity of
520 nonlinear functions.

521 **Lemma 3.6** (Lemma 2.6.4 in [42]). Consider a lattice X and let $K = \{1, \dots, k\}$. Let $f_i(x)$, $i \in K$,
522 be increasing supermodular (resp. submodular) functions on X , and Z_i , $i \in K$, be convex subsets
523 of \mathbb{R} . Assume $Z_i \supseteq \{f_i(x) \mid x \in X\}$. Let $g(z_1, \dots, z_k, x)$ be supermodular in (z_1, \dots, z_k, x) on $Z_1 \times$
524 $\dots \times Z_k \times X$. If for all $i \in K$, $\bar{z}_{i'} \in Z_{i'}$ for $i' \in K \setminus \{i\}$, and $\bar{x} \in X$, $g(\bar{z}_1, \dots, \bar{z}_{i-1}, z_i, \bar{z}_{i+1}, \dots, \bar{z}_k, \bar{x})$
525 is increasing (decreasing) and convex in z_i on Z_i , then $g(f_1(x), \dots, f_k(x), x)$ is supermodular on
526 X . \square

527 By choosing $g(z_1, \dots, z_k, x)$ appropriately as $z_1 z_2 \dots z_k$ or $-z_1 z_2 \dots z_k$, it follows easily that a
528 product of nonnegative, increasing (decreasing) supermodular functions is also nonnegative increas-
529 ing (decreasing) and supermodular; see Corollary 2.6.3 in [42]. Also, it follows trivially that a conic
530 combination of supermodular functions is supermodular.

531 We now use Theorem 3.3 and Corollary 3.4 to derive the concave envelope of some multilinear
532 functions over certain polytopes and apply this general result to derive various results of the liter-
533 ature. More precisely, we define $G \subseteq \mathbb{R}^{\sum_{i=1}^n d_i}$, where each $y \in G$ is expressed as (y_1, \dots, y_n) , and

534 $y_i = (y_{i1}, \dots, y_{id_i}) \in \mathbb{R}^{d_i}$, as:

$$535 \quad G = \left\{ y \in \mathbb{R}^{\sum_{i=1}^n d_i} \mid \sum_{r=1}^{d_i} y_{ir} \leq 1 \forall i; y_{ir} \geq 0 \forall (i, r) \right\},$$

536 i.e. G is a set of points in $\mathbb{R}^{\sum_{i=1}^n d_i}$ that satisfy n non-overlapping generalized upper bound con-
537 straints. Note that since we can choose $d_i = 1$ for all i , G also include hypercubes. For each i , let
538 $D_i = \{1, \dots, d_i\}$ and T_i be a chain (by inclusion) of subsets of D_i where $\emptyset = T_{i0} \subset \dots \subset T_{id_i} = D_i$.
539 Without loss of generality, by relabeling the variables if necessary, we assume that $T_{ir} = \{1, \dots, r\}$.
540 Consider the multiset M where each i in $\{1, \dots, n\}$ has d_i copies. Let Π denote the set of dis-
541 tinct arrangements of M . Then, each $\pi \in \Pi$ is a permutation of $\{1, \dots, \sum_{i=1}^n d_i\}$, where we may
542 additionally assume that, for each $i \in \{1, \dots, d_i\}$, $\pi_{i1} \geq \dots \geq \pi_{id_i}$. For $r \in \{1, \dots, d_i\}$, we let
543 $e(i, r) \in \mathbb{R}^{\sum_{i=1}^n d_i}$ represent the r^{th} principal vector in the i^{th} subspace. Further, let $e(i, d_i + 1)$ be
544 the zero vector in $\mathbb{R}^{\sum_{i=1}^n d_i}$. For a given π , i and i' in $\{1, \dots, n\}$, and $r \in \{1, \dots, d_i\}$, if there exists
545 an index $j \in \{1, \dots, d_{i'}\}$ such that $\pi_{i'j} \leq \pi_{ir}$, we define $w_\pi^{ir}(i') = \min\{j \mid \pi_{i'j} \leq \pi_{ir}\}$, otherwise we
546 set $w_\pi^{ir}(i') = d_{i'} + 1$.

547 Next we introduce an example we will use to illustrate the above notation and the result of
548 Corollary 3.8.

549 **Example 3.7.** Consider the function

$$550 \quad \hat{f}(y_{11}, y_{12}, y_{21}, y_{22}) = 2(1 + y_{11})(2 + y_{21} + y_{22}) + 3(y_{11} + y_{12})y_{21}$$

551 over the polytope

$$552 \quad \hat{G} = \{y \in \mathbb{R}_+^4 \mid y_{11} + y_{12} \leq 1, y_{21} + y_{22} \leq 1\}.$$

553 For the set above, the arrangements $(\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22})$ in Π are $(2, 1, 4, 3)$, $(3, 1, 4, 2)$, $(3, 2, 4, 1)$,
554 $(4, 1, 3, 2)$, $(4, 2, 3, 1)$ and $(4, 3, 2, 1)$. In particular, for $\pi = (4, 1, 3, 2)$, we have $w_\pi^{11}(1) = 1$, $w_\pi^{12}(1) =$
555 2 , $w_\pi^{21}(1) = 2$, $w_\pi^{22}(1) = 2$ and $w_\pi^{11}(2) = 1$, $w_\pi^{12}(2) = 3$, $w_\pi^{21}(2) = 1$, $w_\pi^{22}(2) = 2$. \square

556 **Corollary 3.8.** Consider the function $f(y) = \sum_{k \in K} a_k \prod_{i=1}^n (b_{ik} + \sum_{j \in T_{ir_{ik}}} y_{ij})$ over G , where for
557 each k , $r_{ik} \in D_i \cup \{0\}$, $a_k \geq 0$, and $b_{ik} \geq 0$. Then, the concave envelope of $f(y)$ over G is given by:

$$558 \quad \min_{\pi \in \Pi} \sum_{i=1}^n \sum_{j=1}^{d_i} y_{ij} \left[\sum_{p=j}^{d_i} \left(f \left(\sum_{i'=1}^n e(i', w_\pi^{ip}(i')) \right) - f \left(\sum_{i'=1}^n e(i', w_\pi^{ip}(i')) - e(i, p) + e(i, p+1) \right) \right) \right] + f(0). \quad (7)$$

559 In particular, consider $f'(y) = \sum_{k \in K} a_k \prod_{i \in I_k} \sum_{j \in T_{ir_{ik}}} y_{ij}$, where, for each k , $I_k \subseteq \{1, \dots, n\}$.
560 Then, the concave envelope of f' over G is:

$$561 \quad \sum_{k \in K} a_k \min_{i \in I_k} \left(\sum_{j \in T_{ir_{ik}}} y_{ij} \right). \quad (8)$$

562 *Proof.* Consider the invertible linear transformation of G obtained by defining $Y_{ir} = \sum_{j=1}^r y_{ij}$ for
563 $r = 1, \dots, d_i$ and by setting Y_{i0} to zero for notational convenience. The linear transformation G' of
564 G has the form:

$$565 \quad G' = \left\{ Y \in \mathbb{R}^{\sum_{i=1}^n d_i} \mid 0 \leq Y_{i1} \leq \dots \leq Y_{id_i} \leq 1 \forall i \right\}.$$

566 It is easy to verify that \bar{f} defined over G' is computed as $\bar{f}(Y) = \sum_{k \in K} a_k \prod_{i=1}^n (b_{ik} + Y_{ir_k})$ satisfies
567 $\bar{f}(Y) = f(y)$. Clearly \bar{f} is supermodular since it is a conic combination of multilinear terms (see

568 Lemma 3.6 and the following discussion) and concave-extendable over $0 \leq Y \leq 1$ (see Theorem 2.1
569 in [24]). It follows from Corollary 3.4 that the concave envelope of \bar{f} over G' is obtained as $\bar{f}^K(Y)$.
570 Therefore, for any permutation π in Π , we obtain a corresponding facet of the concave envelope in
571 the space of Y variables using the expression (6). In particular, for $i \in 1, \dots, n$ and $j \in \{1, \dots, d_i\}$,
572 the coefficient α_{ij} of variable Y_{ij} is given by

$$\begin{aligned}
573 \quad \alpha_{ij} &= \bar{f} \left(\sum_{(i',j') | \pi_{i',j'} \leq \pi_{ij}} e(i',j') \right) - \bar{f} \left(\sum_{(i',j') | \pi_{i',j'} < \pi_{ij}} e(i',j') \right) \\
574 &= \bar{f} \left(\sum_{i'=1}^n \sum_{j' | \pi_{i',j'} \leq \pi_{ij}} e(i',j') \right) - \bar{f} \left(\sum_{i'=1}^n \sum_{j' | \pi_{i',j'} < \pi_{ij}} e(i',j') - e(i,j) \right) \\
575 &= \bar{f} \left(\sum_{i'=1}^n \sum_{j'=w_{\pi}^{ij}(i')}^{d_i} e(i',j') \right) - \bar{f} \left(\sum_{i'=1}^n \sum_{j'=w_{\pi}^{ij}(i')}^{d_i} e(i',j') - e(i,j) \right) \\
576 &= f \left(\sum_{i'=1}^n e(i', w_{\pi}^{ij}(i')) \right) - f \left(\sum_{i'=1}^n e(i', w_{\pi}^{ij}(i')) - e(i,j) + e(i,j+1) \right). \\
577
\end{aligned}$$

578 It then remains to convert this expression back to the space of y variables. For $i \in 1, \dots, n$ and
579 $j \in \{1, \dots, d_i\}$, the coefficient that y_{ij} receives is $\sum_{p=j}^{d_i} \alpha_{ip}$ showing (7).

580 Now, consider f' and its term $f'_k = a_k \prod_{i \in I_k} \left(\sum_{j \in T_{ir_k}} y_{ij} \right)$. Then, $f'_k \left(\sum_{i'=1}^n e(i', w_{\pi}^{ip}(i')) \right) = a_k$
581 if $\pi_{i'r_{i'k}} \leq \pi_{i,p}$ for all $i' \in I_k$ and 0 otherwise. Similarly,

$$582 \quad f_k \left(\sum_{i'=1}^n e(i', w_{\pi}^{ip}(i')) - e(i,p) + e(i,p-1) \right) = \begin{cases} 0 & \pi_{i'r_{i'k}} > \pi_{i,p} \text{ for some } i' \in I_k \setminus i \text{ or } p \geq r_{ik} \\ a_k & \text{otherwise.} \end{cases}$$

583 Simplifying (7), the result follows. \square

584 Note that (7) gives the concave envelope of any function that is supermodular in Y_{ir} for $i =$
585 $1, \dots, n$ and $r = 1, \dots, d_i$ over G' , which is a lattice family, and concave-extendable from the vertices
586 of G' .

587 **Example 3.9.** Consider the function \hat{f} of Example 3.7. Applying the result of Corollary 3.8, we
588 obtain for $\pi = (4, 1, 3, 2)$ that

$$\begin{aligned}
589 \quad \alpha_{11}^{\pi} &= f(e(1,1) + e(2,1)) + f(e(1,2) + e(2,3)) - f(e(1,2) + e(2,1)) - f(e(2,3) + e(1,3)) = 6 \\
590 \quad \alpha_{12}^{\pi} &= f(e(1,2) + e(2,3)) - f(e(2,3) + e(1,3)) = 0 \\
591 \quad \alpha_{21}^{\pi} &= f(e(1,2) + e(2,1)) + f(e(1,2) + e(2,2)) - f(e(1,2) + e(2,2)) - f(e(1,2) + e(2,3)) = 5 \\
592 \quad \alpha_{22}^{\pi} &= f(e(1,2) + e(2,2)) - f(e(1,2) + e(2,3)) = 2.
\end{aligned}$$

593 It follows that $6y_{11} + 5y_{2,1} + 2y_{22} + 4$ defines a facet of the concave envelope of \hat{f} over \hat{G} . \square

594 Next, we discuss several results in the literature that are a special case of Corollary 3.8. Let
595 $D = \{1, \dots, \sum_{i=1}^n d_i\}$. For $d \in D$, let $i(d) = \min\{i \mid \sum_{i'=1}^i d_i \geq d\}$ and $j(d) = d - \sum_{i'=1}^{i(d)-1} d_{i'}$. For
596 an element d of D , the pair $(i(d), j(d))$ yields the index of the variable of G that would be in d^{th}
597 position if the variables were ordered as $y_{1,1}, \dots, y_{1,d_1}, \dots, y_{n1}, \dots, y_{nd_n}$.

598 **Corollary 3.10** (Theorem 4 and Theorem 6 in [30]). Consider the function $\phi^m(y) : \text{vert}(G) \mapsto \mathbb{R}$
599 defined as $\sum_{J \subseteq D, |J|=m} [\prod_{d \in J} y_{i(d), j(d)}]$, where $m \leq n$. The concave envelope of $\phi^m(y)$ over G is
600 given by:

$$601 \quad \min \left\{ \sum_{k=m}^n \binom{k-1}{m-1} \sum_{j=1}^{d_{i_k}} y_{i_k j} \mid \{i_m, \dots, i_n\} \subseteq \{1, \dots, n\} \right\}.$$

602 If $d_i = 1$ for all i , then $\text{conc}_{\text{vert}(G)} \phi^m(y)$ is also the concave envelope of $\phi^m(y) : G \mapsto \mathbb{R}$ over G .

603 *Proof.* Let $N = \{1, \dots, n\}$. We may restrict the summation in $\phi^m(y)$ to those subsets J of D that
604 are such that, for any d and d' in J , $i(d) \neq i(d')$. This is because if a certain subset J does not
605 satisfy this condition, then $\prod_{d \in J} y_{i(d), j(d)}$ equals zero for every $y \in \text{vert}(G)$. If $d_i = 1$ for all i , this
606 condition holds trivially.

607 Therefore, we may rewrite

$$608 \quad \phi^m(y) = \sum_{U=\{i_1, \dots, i_m\} \subseteq N} \sum_{j_1=1}^{d_{i_1}} \sum_{j_2=1}^{d_{i_2}} \cdots \sum_{j_m=1}^{d_{i_m}} y_{i_1, j_1} y_{i_2, j_2} \cdots y_{i_m, j_m} = \sum_{U \subseteq N, |U|=m} \left[\prod_{i \in U} \sum_{j=1}^{d_i} y_{i, j} \right].$$

609 The concave envelope of $\phi^m(y)$ is of the form (8) derived in Corollary 3.8:

$$610 \quad \sum_{U \subseteq N, |U|=m} \min_{i \in U} \left(\sum_{j=1}^{d_i} y_{i, j} \right) = \sum_{U \subseteq N, |U|=m} \min_{i \in U} (S_i)$$

611 where $S_i = \sum_{j=1}^{d_i} y_{i, j}$ and $S = (S_1, \dots, S_n)$. Let $\{\pi_1, \dots, \pi_n\}$ be the permutation of $\{1, \dots, n\}$ that
612 sorts S_i in increasing order, i.e. $S_{\pi_1} \leq S_{\pi_2} \leq \dots \leq S_{\pi_n}$. Since S_{π_p} is the p^{th} smallest among all S s,
613 it will be minimum in every subset U that does not contain $\{\pi_1, \pi_2, \dots, \pi_{p-1}\}$. Observe that there
614 are $\binom{n-p}{m-1}$ such sets when $1 \leq p \leq n - m + 1$ and 0 otherwise. It follows that the concave envelope
615 is given by

$$616 \quad \min_{\pi \in \Pi} \sum_{p=1}^{n-m+1} \binom{n-p}{m-1} S_{\pi_p} = \min_{\pi \in \Pi} \sum_{k=m}^n \binom{k-1}{m-1} S_{\pi_{n-k+1}},$$

617 where Π is the set of permutations of $\{1, \dots, n\}$. The expression in the Corollary follows by noticing
618 that the underestimating affine function does not depend on the permutation but only on the subset
619 $\{\pi_1, \dots, \pi_{n-m+1}\}$. \square

620 Note that it is necessary to restrict $\phi^m(x)$ to the extreme points of G when d_i is not equal to
621 1 for some i . For example, consider xy over $\{(x, y) \in \mathbb{R}^2 \mid x + y \leq 1, x, y \geq 0\}$. The function in
622 Corollary 3.10 can be reduced to this case by setting $n = 1$, $d_1 = 2$, and $m = 2$. It can be argued
623 that the concave envelope is $\frac{xy}{x+y}$ if $x + y > 0$ and 0 if $(x, y) = (0, 0)$. This function is non-polyhedral
624 and not concave-extendable from vertices.

625 **Corollary 3.11** ([21]). Let $N = \{1, \dots, n\}$ and $\Gamma = 2^N$. The concave envelope of $\phi(x) =$
626 $\sum_{T \subseteq \Gamma} a_T \prod_{i \in T} x_i$ where $a_T \geq 0$ for all $T \subseteq \Gamma$ over the unit hypercube is given by:

$$627 \quad \sum_{T \subseteq \Gamma} a_T \min\{x_i : i \in T\}.$$

628 *Proof.* Follows directly from Corollary 3.8 by setting $d_i = 1$ for all i . \square

629 **Corollary 3.12** (Theorem 1 in [26]). *Consider the set:*

$$630 \quad X = \left\{ (x, t) \in \mathbb{R}^{n+1} \mid t \leq \sum_{1 \leq i < j \leq n} q_{ij} x_i x_j, x \in \{0, 1\}^n \right\}$$

631 where $q_{ij} \geq 0$ for $i, j = 1, \dots, n$ and $q_{ij} = q_{ji}$. Then,

$$632 \quad \text{conv}(X) = \left\{ (x, t) \in \mathbb{R}^{n+1} \mid t \leq \sum_{i=2}^n \sum_{j=1}^{i-1} q_{\pi(j)\pi(i)} x_{\pi(i)}, x \in [0, 1]^n \forall \pi \in \Pi \right\}$$

633 where Π is the set of permutations of $\{1, \dots, n\}$.

634 *Proof.* Follows directly from Corollary 3.11 by allowing only quadratic terms. \square

635 Observe that the result of Corollary 3.12 can be trivially extended to allow terms of the form
 636 $q_{ii}x_i^2$ where $q_{ii} > 0$ since the function is still concave-extendable and therefore $q_{ii}x_i^2$ can be replaced
 637 with $q_{ii}x_i$ before the envelope is constructed. The supermodularity of the resulting function follows
 638 directly.

639 **Corollary 3.13** (Theorem 1 in [5]). *The concave envelope of $m(x) = \prod_{i=1}^n x_i$ over $\prod_{i=1}^n [L_i, U_i]$,
 640 where $L_i \geq 0$ for all i , is given by:*

$$641 \quad \min_{\pi \in \Pi} \sum_{i=1}^n \left(\left(\prod_{\pi(j) < \pi(i)} U_j \right) \left(\prod_{\pi(j) > \pi(i)} L_j \right) (x_i - L_i) \right) \quad (9)$$

642 where Π is the set of permutations of $\{1, \dots, n\}$.

643 *Proof.* Clearly, $m(x)$ is supermodular and concave-extendable from $\prod_{i=1}^n \{L_i, U_i\}$. Let $m'(x') =$
 644 $m(x)$ where $x' = T(x)$; see (1). This transformation does not alter supermodularity or concave-
 645 extendability. Therefore, it follows that the concave envelope can be constructed as in Theorem 3.3.
 646 Then, following (6), the concave envelope of m' over $[0, 1]^n$ is given by

$$647 \quad \min_{\pi \in \Pi} \sum_{i=1}^n \left(\left(\prod_{\pi(j) \leq \pi(i)} U_j \right) \left(\prod_{\pi(j) > \pi(i)} L_j \right) - \left(\prod_{\pi(j) < \pi(i)} U_j \right) \left(\prod_{\pi(j) \geq \pi(i)} L_j \right) \right) x'_i.$$

648 Factoring out $\left(\prod_{\pi(j) < \pi(i)} U_j \right) \left(\prod_{\pi(j) > \pi(i)} L_j \right)$ and substituting $x'_i = \frac{x_i - L_i}{U_i - L_i}$, we obtain (9). \square

649 Linear transformations can often be used to make functions supermodular. For example, Corol-
 650 lary 3.8 uses a transformation that maps G to S and uses supermodularity of the corresponding
 651 transformed function. Another useful transformation, which we refer to as *switching*, involves trans-
 652 forming a variable from x to $1 - x$. For a given $x \in \mathbb{R}^n$ and $T \subseteq \{1, \dots, n\}$, we denote by $x(T)$ the
 653 vector in \mathbb{R}^n obtained as $x(T)_i = 1 - x_i$ if $i \in T$ and $x(T)_i = x_i$ otherwise. Further, for a function
 654 $f : \{0, 1\}^n \mapsto \mathbb{R}$ we define $f(T) : \{0, 1\}^n \mapsto \mathbb{R}$ such that $f(T)(x) = f(x(T))$. It is easy to verify that
 655 $\text{conc}(f)(x) = \text{conc}(f(T))(x(T))$. Let $\mathcal{S} = \bigcup_{i \in I} P_i$ be a polyhedral subdivision of $[0, 1]^n$, where each
 656 P_i is a polyhedron. Then for each i , define $P_i(T) = \{x \mid x(T) \in P_i\}$. and let $\mathcal{S}(T) = \bigcup_{i \in I} P_i(T)$ be
 657 the corresponding polyhedral subdivision of $[0, 1]^n$.

658 As we discussed in Section 1, functions of the type $f(a_0 + \sum_{i=1}^n a_i x_i)$ appear commonly as an
 659 intermediate step in the construction of relaxations of factorable programs. Typically, the weakening

660 step of substituting $a_0 + \sum_{i=1}^n a_i x_i$ with a new variable y is performed before the actual relaxation
661 is obtained. In the following corollary, we show that such a step is unnecessary by deriving the
662 concave envelope of $f(a_0 + \sum_{i=1}^n a_i x_i)$ over the unit hypercube. We show later in Example 3.23 that
663 the relaxation obtained by using Corollary 3.14 indeed has the potential to improve the relaxations
664 used in factorable programming.

665 **Corollary 3.14.** *Let $g(x) = f(L(x)) : [0, 1]^n \mapsto \mathbb{R}$ where f is convex and $L(x) = a_0 + \sum_{i=1}^n a_i x_i$.
666 Let $T = \{i \mid a_i < 0\}$. Then, $g(T)(x)$ is concave-extendable from $\{0, 1\}^n$ and supermodular. The
667 concave envelope of $g(x)$ is determined by $\mathcal{K}(T)$.*

668 *Proof.* The convexity of g and, hence, of $g(T)$ follows from the assumptions in the corollary. There-
669 fore, $g(T)$ is concave-extendable from $\{0, 1\}^n$. First assume that $T = \emptyset$. Let $x', x'' \in [0, 1]^n$ and as-
670 sume without loss of generality that $L(x') \leq L(x'')$. Then, $L(x' \wedge x'') \leq L(x') \leq L(x'') \leq L(x' \vee x'')$.
671 Further, $L(x') + L(x'') = L(x' \wedge x'') + L(x' \vee x'')$ since $L(\cdot)$ is affine. Using Hardy-Littlewood-
672 Polyá/Karamata's inequality, we obtain that $f(L(x')) + f(L(x'')) \leq f(L(x' \wedge x'')) + f(L(x' \vee x''))$
673 since the sequence $(L(x' \wedge x''), L(x' \vee x''))$ is majorized by $(L(x'), L(x''))$ and f is convex; see
674 Section 3.17 in [13]. The result then follows from Theorem 3.3. Now, assume that $T \neq \emptyset$.
675 Applying the corollary to $g(T)$, we conclude that the concave envelope of $g(T)$ is defined by \mathcal{K} .
676 Since $\text{conc}(g)(x) = \text{conc}(g(T))(x(T))$, we conclude that $\text{conc}(g)(x)$ is described by the triangulation
677 $\mathcal{K}(T)$. \square

678 The following result is a direct consequence of Theorem 3.3 that is well suited for applications
679 involving disjunctions.

680 **Corollary 3.15.** *Consider a function $f(y, x) = f(y, x_1, \dots, x_n) : \{0, 1\}^{n+1} \mapsto \mathbb{R}$ and define $f_0(x) :=$
681 $f(0, x)$ and $f_1(x) := f(1, x)$. Then, $f(y, x)$ is supermodular if and only if f_0 and f_1 are supermodular,
682 and $f_1(x) - f_0(x)$ is a non-decreasing function of x . Assume f_0 and f_1 are supermodular and
683 $f_1(x) - f_0(x)$ is monotone. Then, the concave envelope of f over $[0, 1]^{n+1}$ is described by $\mathcal{K}(T)$
684 where $T = \emptyset$ if $f_1(x) - f_0(x)$ is non-decreasing and $T = \{1\}$ if $f_1(x) - f_0(x)$ is non-increasing.*

685 *Proof.* For the direct implication, note that f_0 and f_1 have to be supermodular if f is supermodular.
686 Further, for any $x' \geq x$, $f(1, x) + f(0, x') \leq f(1, x') + f(0, x)$ as f is supermodular and $x \vee x' = x$
687 and $x \wedge x' = x'$. This shows that $f_1(x) - f_0(x)$ is non-decreasing. For the reverse implication,
688 consider two arbitrary points (y', x') and (y'', x'') in $\{0, 1\}^n$. If $y' = y''$, then $f(y', x') + f(y'', x'') \leq$
689 $f((y', x') \wedge (y'', x'')) + f((y', x') \vee (y'', x''))$ by supermodularity of f_0 and f_1 . Without loss of generality,
690 we assume $y' = 0$ and $y'' = 1$. Then,

$$\begin{aligned}
691 \quad f(y', x') + f(y'', x'') &= f_0(x') + f_0(x'') + f_1(x'') - f_0(x'') \\
692 &\leq f_0(x' \wedge x'') + f_0(x' \vee x'') + f_1(x' \vee x'') - f_0(x' \vee x'') \\
693 &= f_0(x' \wedge x'') + f_1(x' \vee x'') \\
694 &= f((y', x') \wedge (y'', x'')) + f((y', x') \vee (y'', x'')),
\end{aligned}$$

695 where the first inequality holds because f_0 is supermodular and because $f_1(x) - f_0(x)$ is non-
696 decreasing and the last equality holds because $y' \wedge y'' = 0$ and $y' \vee y'' = 1$. The rest of the result
697 follows from Theorem 3.3 after switching y if $f_1(x) - f_0(x)$ is non-increasing. \square

698 In the statement of Corollary 3.15, we emphasize that the polyhedral subdivision $\mathcal{K}(\{1\})$ is
699 obtained from Kuhn's triangulation by switching the first variable of the function f , i.e. it is
700 obtained by switching the variable y and not the variable x_1 .

701 Corollary 3.15 also applies to certain nonlinear functions that do not intrinsically exhibit a
702 disjunctive structure. Consider $f(y, x) = f_0(x) + y(f_1(x) - f_0(x))$. When x is fixed, the function

703 is linear in y . Therefore, it suffices to restrict $y \in \{0, 1\}$. Then, Corollary 3.15 yields the concave
704 envelope of $f(y, x)$ when $f_0(\cdot)$ and $f_1(\cdot)$ are supermodular and concave-extendable from vertices and
705 $f_1(\cdot) - f_0(\cdot)$ is non-decreasing. In fact, the proof of Corollary 3.15 can be easily generalized to show
706 that $f(y, x) = f_0(x) + y(f_1(x) - f_0(x))$ is supermodular over $[0, 1]^{n+1}$. Assume $0 \leq y' \leq y'' \leq 1$.
707 Then,

$$\begin{aligned}
708 \quad f(y', x') + f(y'', x'') &= f(y', x') + f(y', x'') + f(y'', x'') - f(y', x'') \\
709 &\leq f(y', x' \vee x'') + f(y', x' \wedge x'') + (y'' - y')(f_1(x'') - f_0(x'')) \\
710 &\leq f(y', x' \vee x'') + f(y', x' \wedge x'') + (y'' - y')(f_1(x' \vee x'') - f_0(x' \vee x'')) \\
711 &= f(y', x' \vee x'') + f(y', x' \wedge x'') + f(y'', x' \vee x'') - f(y', x' \vee x'') \\
712 &= f(y'', x' \vee x'') + f(y', x' \wedge x''),
\end{aligned}$$

713 where the first inequality follows from the supermodularity of $f_0(x)$ and $f_1(x)$ and the second
714 inequality follows since $y'' \geq y'$, $f_1(x) - f_0(x)$ is non-decreasing, and $x' \vee x'' \geq x''$. Since the
715 concave-extendability of $f(y^L, x)$ and $f(y^U, x)$ follows from [35], it follows that we can develop the
716 concave envelope of $f(y, x)$ over $[y^L, y^U] \times [0, 1]^n$ using Theorem 3.3 for $0 \leq y^L \leq y^U \leq 1$.

717 In Corollary 3.16, we particularize the result of Corollary 3.15 to situations where $f(y, x) =$
718 $yg(x)$, for example $\frac{y}{1 + \sum_{i=1}^n x_i}$ and $y \log(1 + \sum_{i=1}^n x_i)$. The result also applies to $\frac{y}{y + \sum_{i=1}^n x_i}$ and
719 $y \log(y + \sum_{i=1}^n x_i)$ if one restricts the region to $y + \sum_{i=1}^n x_i \geq 1$. This is a natural restriction
720 when the variables y and x_i are binary; see [6] for applications in consistent biclustering problems.
721 The supermodularity of these functions for a fixed y follows from Corollary 3.14 and, therefore,
722 Corollary 3.15 applies.

723 **Corollary 3.16.** *Consider a function $f(y, x) = f(y, x_1, \dots, x_n) : \{0, 1\}^{n+1} \mapsto \mathbb{R}$, where $f(0, x) = 0$
724 and $f(1, x) = f_1(x)$. Assume $f_1(x)$ is non-increasing and supermodular. Then, $\text{conc}_{[0,1]^{n+1}}(f)$ is
725 described by $\mathcal{K}(\{1\})$. Let $W = \{(y, x) \in [0, 1]^{n+1} \mid y + \sum_{i=1}^n x_i \geq 1\}$. Then, for any $(y, x) \in W$,
726 $\text{conc}_W(f)(y, x) = \text{conc}_{[0,1]^{n+1}}(f)(y, x)$.*

727 *Proof.* It follows from Corollary 3.15 that $\text{conc}_{[0,1]^{n+1}}(f)(y, x)$ is described by $K(\{1\})$. Since $W \subseteq$
728 $[0, 1]^{n+1}$, $\text{conc}_{[0,1]^{n+1}}(f)(y, x) \geq \text{conc}_W(f)(y, x)$. Observe that $\text{conc}_{[0,1]^{n+1}}(f)(y, x)$ is linear for $x \in$
729 $Y = \{(y, x) \mid 0 \leq x_1, \dots, x_n \leq 1 - y \leq 1\}$. However, Y is obtained as a union of simplices in $\mathcal{K}(\{1\})$.
730 In particular, if K_π is the simplex associated with permutation π (after replacing y with $1 - \bar{y}$),
731 then $Y = \bigcup_{\pi \in \Pi'} K_\pi$, where Π' is the set of permutations of $\{1, \dots, n + 1\}$ that are restricted to
732 have 1 as the first element. Let $W' = \text{cl}([0, 1]^{n+1} \setminus Y)$. Since $\text{vert}(W) = \text{vert}(W')$ and W is convex,
733 it follows that $W = \text{conv}(W')$. Let $W'' = Y \cap \{(y, x) \mid y + \sum_{i=1}^n x_i \geq 1\}$. It is easy to see that
734 W'' is the convex hull of $\{(0, x) \in [0, 1]^{n+1} \mid \sum_{i=1}^n x_i \geq 1\}$ and $(1, 0)$. Therefore, W'' has binary
735 extreme points. It can now be easily verified that, for any $(y, x) \in W$, $\text{conc}_{[0,1]^{n+1}}(f)(y, x)$ is a
736 feasible solution to $D(y, x)$. Therefore, $\text{conc}_{[0,1]^{n+1}}(f)(y, x) \leq \text{conc}_W(f)(y, x)$. It follows that, for
737 any $(y, x) \in W$, $\text{conc}_W(f)(y, x) = \text{conc}_{[0,1]^{n+1}}(f)(y, x)$. \square

738 Corollary 3.16 can also be derived as a consequence of Theorem 3.3 applied to $f_1(x)$ along with
739 Theorem 4.1, which will be proven later and describes the concave envelope of $yg(x)$ under more
740 general conditions.

741 **Example 3.17.** *Let $g(z)$ be a convex non-increasing function and $f(y, x) = yg(\sum_{i=1}^n x_i)$. Assume
742 $x \in \{0, 1\}^n$. Then, $g(\sum_{i=1}^n x_i)$ is supermodular by Corollary 3.14. By definition, it is concave-
743 extendable from the vertices. The concave envelope is therefore given by Corollary 3.16. In par-
744 ticular, if Π is the set of permutations of $\{1, \dots, n\}$ then $\bigcup_{\pi \in \Pi, 0 \leq m \leq n} S(\pi, m)$ gives the polyhedral
745 division of $\{0, 1\}^n$ that defines the concave envelope of $f(y, x)$ where $S(\pi, m) = \{(y, x) \mid x \in$
746 $K_\pi, x_{\pi(m)} \geq 1 - y \geq x_{\pi(m+1)}\}$. Here, we assume $x_{\pi(0)} = 1$ and $x_{\pi(n+1)} = 0$. Further the concave*

747 envelope of $f(y, x)$ can be computed as $\min_{\pi \in \Pi, 0 \leq m \leq n} h^{S(\pi, m)}(y, x)$ where $h^{S(\pi, m)}(y, x)$ is the facet
 748 of $\text{conc}_{[0,1]^{n+1}}$ that is tight over $S(\pi, m)$ and is given by:

$$749 \quad h^{S(\pi, m)}(y, x) = g(0) + \sum_{i=1}^m (g(i) - g(i-1))x_{\pi(i)} - g(m)(1-y).$$

750 The restriction of the concave envelope to $W = \{(y, x) \in [0, 1]^{n+1} \mid y + \sum_{i=1}^n x_i \geq 1\}$ gives the
 751 concave envelope over W . In particular, consider $f(y, x) = \frac{y}{y + \sum_{i=1}^n x_i}$ where $(y, x) \in W \cap \{0, 1\}^{n+1}$.
 752 Then, the concave envelope of $f(y, x)$ over W is given by:

$$753 \quad \min_{\pi \in \Pi, 0 \leq m \leq n} \left(1 - \sum_{i=1}^m \frac{1}{i(i+1)} x_{\pi(i)} - \frac{1}{m+1} (1-y) \right). \quad (10)$$

754 This fractional function appears in the formulation of consistent bichustering problems [6]. The
 755 standard factorable relaxation introduces $z = \frac{1}{y + \sum_{i=1}^n x_i}$ and $w = yz$. Let $u(x, y) = y + \sum_{i=1}^n x_i$.
 756 Then, $z = \frac{1}{u(x, y)}$ is relaxed over $u(x, y) \in [1, n+1]$ as $z \leq \frac{n+2}{n+1} - \frac{u(x, y)}{n+1}$. Finally, $w \leq \min\{y, \frac{1}{n+1}y +$
 757 $z - \frac{1}{n+1}\}$ which, equivalently, yields $w \leq \min\{y, \frac{1}{n+1}y - \frac{1}{n+1}u(x, y) + 1\}$. The same relaxation is
 758 obtained if the concave envelope of $\frac{y}{u(x, y)}$ is constructed directly over $[0, 1] \times [1, n+1]$; see [36].
 759 Clearly, the concave envelope developed in (10) is tight when $y = 1$ and $x_i = 1$ for all $i \in I$,
 760 where $\emptyset \subsetneq I \subsetneq N$ (evaluates to $\frac{1}{1+|I|}$) whereas the factorable relaxation is not tight at these points
 761 (evaluates to $\frac{n+1-|I|}{n+1}$). It can also be directly verified that the concave envelope is tighter relative to
 762 the factorable relaxation at these points by observing that $(n - |I|)|I| > 0$ for $1 \leq |I| \leq n - 1$.

763 Corollary 3.16 exemplifies a situation where restricting attention to $y + \sum_{i=1}^n x_i$ does not result
 764 in a substantial change in the triangulation. This may appear surprising when one considers the
 765 origin is a vertex of every simplex in Kuhn's triangulation. However, a more careful observation
 766 reveals that the removing the origin does not have a significant impact in Corollary 3.16 because
 767 the triangulation is given after switching y , i.e., it is $K(\{1\})$ and not K .

768 When the concave envelope is determined by Kuhn's triangulation, the envelope will typically
 769 change drastically if the origin is removed from the underlying region. We next describe a situa-
 770 tion that illustrates this phenomenon. Corollary 3.14 shows that if $f(\cdot)$ is a convex function then
 771 $f(\sum_{i=1}^n x_i)$ is supermodular and concave-extendable from vertices and, therefore, its concave en-
 772 velope is defined by Kuhn's triangulation. In various situations, it will be useful to construct the
 773 concave envelope over $\sum_{i=1}^n x_i \geq 1$, a situation where the origin is no longer an extreme point of the
 774 underlying polytope. Next, we study this situation by considering the slightly more general case
 775 where we seek to determine the concave envelope of $f(\sum_{i=1}^n x_i)$ assuming that $f(\cdot)$ that is convex
 776 over $[1, n]$ but $\frac{(n-1)}{n}f(0) + \frac{1}{n}f(n) < f(1)$, i.e., f is nonconvex because its value at 0 is below what
 777 is required for convexity.

778 We first introduce a polyhedral subdivision of $[0, 1]^n$ that we will prove in Theorem 3.18 yields
 779 the concave envelope of f . For $k = 0, \dots, n$ we define Π^k to be the set of permutations of exactly
 780 k elements of $\{1, \dots, n\}$. In other words, π belongs to Π^k if $\pi : \{1, \dots, k\} \rightarrow \{1, \dots, n\}$ and
 781 $\pi(i) \neq \pi(j)$ for $i \neq j$. For such a permutation, we set $|\pi| = k$ and use the notation $i \notin \pi$ to signify
 782 that $i \notin \{\pi(1), \pi(2), \dots, \pi(k)\}$. We also use the notation $\tilde{\Pi} = \bigcup_{k=0}^{\max(n-2, 0)} \Pi^k$. For $\pi \in \tilde{\Pi}$, we define

$$783 \quad S_\pi = \left\{ x \in \mathbb{R}^n \left| \begin{array}{l} 0 \leq x_{\pi(1)} \leq \dots \leq x_{\pi(|\pi|)} \leq 1 \\ \sum_{i \notin \pi} x_i \geq 1 + (n - |\pi| - 1)x_{\pi(|\pi|)} \\ \sum_{i \notin \pi} x_i \leq 1 + (n - |\pi| - 1)x_j, \forall j \notin \pi \end{array} \right. \right\},$$

784 where $x_{\pi(0)}$ is assumed to be 0. Let $\Delta = \{x \in [0, 1]^n \mid \sum_{i=1}^n x_i \leq 1\}$. Next, we define $\mathcal{K}^{-0} =$
785 $\{\Delta, \bigcup_{\pi \in \tilde{\Pi}} S_\pi\}$. We will prove in Theorem 3.18 that \mathcal{K}^{-0} is a polyhedral subdivision of $[0, 1]^n$.
786 Here, we argue the weaker result that \mathcal{K}^{-0} covers $[0, 1]^n$ by constructing, for each $x \in [0, 1]^n \setminus \Delta$, a
787 permutation $\bar{\pi} \in \tilde{\Pi}$ for which $x \in S_{\bar{\pi}}$. For an arbitrary $x \in [0, 1]^n \setminus \Delta$, we first sort the components
788 of x in increasing order, thereby obtaining a permutation π of $\{1, \dots, n\}$ for which $0 \leq x_{\pi(1)} \leq$
789 $\dots \leq x_{\pi(n)} \leq 1$. For $j = 0, \dots, n-1$, define $C(j) = \sum_{i=j+1}^n (x_{\pi(i)} - x_{\pi(j)}) - (1 - x_{\pi(j)})$. Clearly,
790 $C(j)$ is decreasing in j . Further, since $x \in [0, 1]^n \setminus \Delta$, it follows that $C(0) > 0$ and $C(n-1) \leq 0$.
791 Define now $\bar{j} = \max\{j \mid C(j) > 0\}$. It is easy to see that $x \in S_{\bar{\pi}}$ where $\bar{\pi}$ is the permutation of
792 $\{1, \dots, \bar{j}\}$ where $\bar{\pi}(t) = \pi(t)$ for $t = 1, \dots, \bar{j}$.

793 It can be verified that, for all $\pi \in \tilde{\Pi}$, S_π is a simplex with $\text{vol}(S_\pi) = \frac{n-1-|\pi|}{n!}$. Further, the
794 vertices of S_π are e_i for all $i \notin \pi$, $\sum_{i \notin \pi} e_i + \sum_{j=|\pi|+1-r}^{|\pi|} e_{\pi(j)}$ for $r = 0, \dots, |\pi|$. Given a function f ,
795 we define

$$796 \quad h_\Delta(x) = (f(1) - f(0)) \sum_{i=1}^n x_i + f(0),$$

797 to be the interpolation of f over the vertices of Δ and, for each $\pi \in \tilde{\Pi}$,

$$798 \quad h_\pi(x) = \sum_{i=1}^{|\pi|} (f(n-i+1) - f(n-i)) x_{\pi(i)} + \frac{f(n-|\pi|) - f(1)}{n-|\pi|-1} \sum_{i \notin \pi} x_i + \frac{(n-|\pi|)f(1) - f(n-|\pi|)}{n-|\pi|-1}.$$

799 to be the interpolation of f over the vertices of S_π .

800 **Theorem 3.18.** *Let $g(x) = f(\sum_{i=1}^n x_i)$ where $f(z)$ is a convex function over $z \in [1, n]$. Assume*
801 *that g is concave-extendable from $\{0, 1\}^n$ and that $(n-1)f(0) \leq nf(1) - f(n)$. Then, $\text{conc}_{[0,1]^n}(f)$*
802 *is described by the polyhedral subdivision \mathcal{K}^{-0} and*

$$803 \quad \text{conc}_{[0,1]^n} f(x) = \min \left\{ h_\Delta(x), \min_{\pi \in \tilde{\Pi}} h_\pi(x) \right\}.$$

804 *Proof.* Consider the following sets

$$805 \quad W_1 = \left\{ x \left| \begin{array}{l} x_{\pi(1)} = \dots = x_{\pi(|\pi|)} = 0 \\ \sum_{i \notin \pi} x_i \geq 1 \\ \sum_{i \notin \pi} x_i \leq 1 + (n - |\pi| - 1) \min_{i \notin \pi} x_i \end{array} \right. \right\} \text{ and } W_2 = \left\{ x \left| \begin{array}{l} 0 \leq x_{\pi(1)} \leq \dots \leq x_{\pi(|\pi|-1)} \leq 1 \\ x_{\pi(|\pi|)} = 1 \\ x_i = 1 \ \forall i \notin \pi \end{array} \right. \right\}.$$

806 Then, by introducing variables $\bar{x}_i = 1 - x_i$ for $i \notin \pi$, W_1 and W_2 become orthogonal sets. It is easy
807 to verify by using Theorem 1 in [41] that $S_\pi = \text{conv}(W_1 \cup W_2)$. Further, h_π is tight at all the extreme
808 points of W_1 and W_2 . Therefore, if we prove that $h_\pi(x) \geq f(x)$, it will follow from Theorem 2.4
809 that h_π defines the concave envelope of $f(x)$ over S_π . First, we verify that $f(0) \leq h_\pi(0)$. Since f
810 is convex, $\frac{(n-|\pi|)f(1) - f(n-|\pi|)}{n-|\pi|-1}$ is increasing in $|\pi|$. Therefore, the minimum value is attained when
811 $|\pi| = 0$. However, by assumption $(n-1)f(0) \leq nf(1) - f(n)$, therefore, $f(0) \leq h_\pi(0)$. Without
812 loss of generality, we may assume that $\pi = (1, \dots, |\pi|)$. Then, by convexity of f , it follows that

$$813 \quad \frac{f(n-|\pi|) - f(1)}{n-|\pi|-1} \leq f(n-|\pi|+1) - f(n-|\pi|) \leq \dots \leq f(n) - f(n-1).$$

814 Therefore, $h_\pi(x)$ may be rewritten as: $c_0 + \sum_{i=1}^n c_i \sum_{j \geq i} x_j$ where $c_i \geq 0$ for all $i \in [1, n]$. In
815 particular, it is easy to verify that $\min\{h_\pi(x) \mid \sum_{i=1}^n x_i = y\} = r(y) = c_0 + \sum_{i=1}^n c_i (i - n + y)^+$,

816 where $r(y) = f(y)$ for $y \in \{1, n - |\pi|, \dots, n\}$. Since, $r(y)$ is linear between consecutive integer values,
817 it follows that $r(y) \geq f(y)$. In other words, $h_\pi(x) \geq f(\sum_{i=1}^n x_i)$. If $f(\cdot)$ is a strictly convex function
818 for $i \in [1, n]$ and $(n - 1)f(0) < nf(1) - f(n)$ then it is easy to verify that this inequality is strict
819 when $x \notin \text{vert}(W_1) \cup \text{vert}(W_2)$. Therefore, it follows that $\Delta \cup \bigcup_{\pi \in \Pi} S_\pi$ is a polyhedral subdivision
820 of $[0, 1]^n$ that defines $\text{conc}_{[0,1]^n} f$. \square

821 **Example 3.19.** Consider the function $f : \{0, 1\}^5 \rightarrow \mathbb{R}$ where $f(x) = 3 - \log_2\left(\sum_{i=1}^5 x_i\right)$ when
822 $x \neq 0$ and $f(x) = 0$ when $x = 0$. Clearly, this function satisfies the assumptions of Theorem 3.18.
823 We now derive two facets of $\text{conc}_{[0,1]^5}(f)$. For $\pi^a \in \Pi^0$, we have

$$824 \quad S_{\pi^a} = \{x \in \mathbb{R}^5 \mid x_1 + x_2 + x_3 + x_4 + x_5 \geq 1, x_1 + x_2 + x_3 + x_4 + x_5 \leq 1 + 4x_j, \forall j = 1, \dots, 5\}.$$

825 The corresponding facet of $\text{conc}_{[0,1]^5}(f)$ is given by

$$826 \quad h_{\pi^a}(x) = \frac{f(5) - f(1)}{4} \sum_{i=1}^5 x_i + \frac{5f(1) - f(5)}{4} = -\frac{\log_2(5)}{4}(x_1 + x_2 + x_3 + x_4 + x_5) + \frac{\log_2(5)}{4} + 3$$

827 For $\pi^b \in \Pi^2$ with $\pi^b(1) = 1, \pi^b(2) = 2$ we have

$$828 \quad S_\pi = \{x \in \mathbb{R}^5 \mid 0 \leq x_1 \leq x_2, x_3 + x_4 + x_5 \geq 1 + 2x_2, x_3 + x_4 + x_5 \leq 1 + 2x_j, \forall j = 3, \dots, 5\}.$$

829 The corresponding facet of $\text{conc}_{[0,1]^5}(f)$ is given by

$$830 \quad h_{\pi^b}(x) = (f(5) - f(4))x_1 + (f(4) - f(3))x_2 + \frac{f(3) - f(1)}{2} \sum_{i=3}^5 x_i + \frac{3f(1) - f(3)}{2}$$

$$831 \quad = -(\log_2(5) - 2)x_1 - (2 - \log_2(3))x_2 - \frac{\log_2(3)}{2}(x_3 + x_4 + x_5) + \frac{\log_2(3)}{2} + 3. \quad \square$$

832 **Example 3.20.** Let $g(x) = \frac{1}{\sum_{i=1}^n x_i}$ where $x_i \in \{0, 1\}$ and $\sum_{i=1}^n x_i \geq 1$. We define $g(0) = 0$.
833 Since $S_\pi \subseteq W \subseteq [0, 1]^n$, it follows that $\text{conc}_{S_\pi} g(x) \leq \text{conc}_W g(x) \leq \text{conc}_{[0,1]^n} g(x)$. For each
834 $x \in W$, there exists π such that $x \in S_\pi$ and, by Theorem 3.18, $\text{conc}_{S_\pi} g(x) = \text{conc}_{[0,1]^n} g(x)$; see also
835 Remark 3.5. Therefore, $\max_{\pi \in \Pi} \text{conc}_{S_\pi} g(x) = \text{conc}_W g(x)$. Incidentally, the same concave envelope
836 is also obtained if $x_i \in [0, 1]$ since $g(x)$ is a convex function and, therefore, concave-extendable from
837 the vertices. \square

838 Although it is in general NP-Hard to identify supermodular functions [9], some special classes
839 of functions can be easily identified to be supermodular. It is well-known, for instance, that the
840 function

$$841 \quad \sum_{J \subseteq N} a_J \prod_{j \in J} x_j + \sum_{I \subseteq N} b_I \prod_{i \in I} (1 - x_i) \quad (11)$$

842 is supermodular if a_J, b_I are nonnegative for all $I, J \subseteq N$; see also Lemma 3.6 and the following
843 discussion. A multilinear function is called *unimodular* if by switching variables x_i in some subset
844 K of N , it can be recast into the form (11). It is shown in [9] that unimodular functions can
845 be recognized by solving a linear programming problem. This linear program yields a polynomial
846 time recognition technique for unimodular functions. Combined with Theorem 3.3, this allows
847 construction of concave envelopes of many multilinear functions. In certain cases, it is easy to
848 recognize that the function is unimodular. The following result illustrates one such example.

849 **Corollary 3.21** (Theorem 15 in [8]). Consider $f(x, y) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} x_i y_j$ where $x \in [0, 1]^n$ and
850 $y \in [0, 1]^m$. Then $\text{conc}_{[0,1]^{n+m}}(f)(x, y) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} \min\{x_i, y_j\}$ and $\text{conv}_{[0,1]^{n+m}}(f)(x, y) =$
851 $\sum_{i=1}^n \sum_{j=1}^m a_{ij} (x_i + y_j - 1)^+$.

852 *Proof.* The concave envelope follows directly from Corollary 3.11. Now, switch the y variables to
853 write $f(x, \bar{y}) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} x_i (1 - \bar{y}_j)$. Since $f(x, \bar{y})$ is submodular (negative of a supermodular
854 function), the convex envelope follows directly from Corollary 3.11. \square

855 **Example 3.22.** Let $f(x) = \sum_{i=1}^k a_i \prod_{j \in J_i} f_{ij}(x_j)$ where $a_i \geq 0$, each f_{ij} is nonnegative, convex and
856 for each i , either $f_{ij}(x_j)$ is increasing or decreasing for all $j \in J_i$. The convexity of $f_{ij}(\cdot)$ implies
857 that $f(x)$ is concave-extendable from the vertices of the hypercube. Since the product of nonnegative
858 increasing (decreasing) univariate functions is supermodular, the concave envelope of $f(x)$ follows
859 from Theorem 3.3. As a concrete example, we may set $f_{ij}(x_j) = x_j^{q_{ij}}$ where $q_{ij} \geq 1$ for all j or
860 $q_{ij} < 0$ for all j . Observe that this example extends the class of functions treated in (11) and in
861 Corollary 3.11. \square

862 **Example 3.23.** Let $f(x) = \sum_{i=1}^k g_i \left(a_i + \sum_{j=1}^n a_{ij} x_j \right)$ where for each j either $a_{ij} \geq 0$ or $a_{ij} \leq 0$
863 for all i , and, for each i , g_i is a convex function. It follows from Corollary 3.14 that the concave
864 envelope of $f(x)$ is given by $K(T)$ where $T = \{j \mid a_{ij} \leq 0 \forall i\}$. As an example, let $c_i \geq 0$ for all i and
865 set $g_i(\cdot) = -c_i \log(\cdot)$. In particular, consider `hs62` from `globallib` which was originally formulated
866 in [15].

$$867 \quad \min \quad -32.174 \left(255 \log \left(\frac{0.03 + x + y + z}{0.03 + 0.09x + y + z} \right) + 280 \log \left(\frac{0.03 + y + z}{0.03 + 0.07y + z} \right) \right. \\
868 \quad \quad \quad \left. + 290 \log \left(\frac{0.03 + z}{0.03 + 0.13z} \right) \right) \\
869 \quad \text{s.t.} \quad x + y + z = 1 \\
870 \quad \quad \quad x, y, z \geq 0.$$

871 If we solve the factorable relaxation, we obtain a lower bound of -83126.9 . Instead, constructing
872 the concave envelope of

$$873 \quad f(x, y, z) = 255 \log \left(\frac{1}{0.03 + 0.09x + y + z} \right) + 280 \log \left(\frac{1}{0.03 + 0.07y + z} \right) \\
\quad \quad \quad + 290 \log \left(\frac{1}{0.03 + 0.13z} \right) \quad (12)$$

874 using Corollary 3.14 gives a lower bound of -52944.9 . Observe that the above technique does not
875 give the concave envelope of (12) over the feasible region. Instead, if one further realizes that the
876 triangle $\{(x, y) \mid x + y + z = 1, x, y, z \geq 0\}$ can be transformed to a lattice family (in a manner
877 similar to Corollary 3.8) by introducing $u = x$, $v = x + y$ and $w = x + y + z = 1$, then (12) can be
878 written as:

$$879 \quad 255 \log \left(\frac{1}{0.12 - 0.91u} \right) + 280 \log \left(\frac{1}{1.03 - 0.07u - 0.93v} \right) + 290 \log \left(\frac{1}{0.16 - 0.13v} \right). \quad (13)$$

880 The feasible region in the (u, v) space is given by $\{(u, v) \mid 0 \leq u \leq v \leq 1\}$. Since the coefficients of
881 u and v are nonpositive, we introduce $\bar{u} = 1 - u$ and $\bar{v} = 1 - v$. Notice that a lattice family remains
882 a lattice family if all the sets are complemented. Then, the concave envelope of (13) and hence (12)
883 over the feasible region can be developed using Theorem 3.3 as:

$$884 \quad f(x, y, z) \geq 535 \log(103) - 490 \log(2) - 1650 \log(5) + (825 \log(3) - 535 \log(103) - 650 \log(2))x \\
885 \quad \quad \quad + (280 \log(5) - 280 \log(103) - 880 \log(2) + 290 \log(3))y.$$

886 The concave envelope could have also been developed simply by realizing that (12) is convex and the
887 feasible region is a triangle. However, we chose to develop it in the above way to demonstrate the
888 techniques developed in this section. With the concave envelope introduced into the formulation, the
889 lower bound improves to -42429.2 . The global minimum has an objective value of -26272.5 . It is
890 interesting to observe that the proposed relaxation leads to a 53% improvement without recognizing
891 the lattice family and 71.5% improvement after recognizing the lattice family when compared to the
892 standard factorable relaxation. \square

893 4 Convex envelopes of disjunctive functions

894 As shown in Sections 2 and 3, if the envelope of a nonlinear function is polyhedral, it can be described
895 using polyhedral subdivisions. However, it may not be apparent that polyhedral subdivisions also
896 play an important role in characterizing non-polyhedral envelopes of certain functions. In this
897 section, we provide an example by considering a function of the form $xf(y)$ where $f(\cdot)$ is convex
898 and non-increasing. Such a function is typically not convex, even in the simple case where $f(y) =$
899 $-y$. However, since $xf(y)$ is convex for any fixed x , the convex envelope can be formed over
900 the hypercube using disjunctive programming. This structure appears commonly in factorable
901 programming. However, it is not typically exploited since the convex envelope can only be described
902 in a lifted space. In Theorem 4.1, we show that the convex envelope can be written in the original
903 space without introducing additional variables when $f(y)$ is non-increasing and the lower bound on
904 x is 0. In this description, we use the recession function $f0^+(y)$ of f where $f0^+(y) = \sup\{f(x +$
905 $y) - f(x) \mid x \in \text{dom } f\}$; see Section 8 in [25].

906 **Theorem 4.1.** Consider a function $g(x, y) = xf(y)$ where $(x, y) \in [0, 1] \times [0, 1]^n$. Let $f(y)$ be
907 a convex non-increasing function and (x', y') be a point in the domain. Let $y'' = (y''_i)_{i=1}^n$, where
908 $y''_i = \min(y'_i, x')$. Then,

$$909 \quad \text{conv}(g)(x', y') = h(x', y') = \begin{cases} x' f\left(\frac{y''}{x'}\right) & \text{if } x' > 0 \\ f0^+(y'') & \text{if } x' = 0 \\ \infty & \text{otherwise.} \end{cases} \quad (14)$$

910 *Proof.* Since $xf(y)$ is linear in x for any fixed value of $y \in [0, 1]^n$, it suffices to consider $x \in \{0, 1\}$
911 when building the convex envelope of this function over $[0, 1]^{n+1}$. For a given subset J of N define
912 $W_0(J) = \{(0, y) \in [0, 1]^{n+1} \mid y_i = 0, \forall i \in J\}$ and $W_1(J) = \{(1, y) \in [0, 1]^{n+1} \mid y_i = 1, i \notin J\}$. First,
913 we construct the convex envelope of $g(x, y)$ over $W' = \text{conv}(W_0(J) \cup W_1(J))$. This convex envelope
914 is obtained by convexifying the two disjunctions

$$915 \quad \begin{array}{l|l} z = 0 & z \geq xf\left(\frac{y}{x}\right) \\ x = 0 & x = 1 \\ y_J = 0 & 0 \leq y_J \leq 1 \\ 0 \leq y_{N \setminus J} \leq 1 & y_{N \setminus J} = 1. \end{array}$$

916 Observe that the above two sets are orthogonal and $h(x', y')$ is a closed positively homogeneous
917 function (see Theorem 8.2 in [25]). Therefore, by Theorem 1 in [41], it follows that the convex
918 envelope (highest convex underestimator that is lower-semicontinuous) of $g(x, y)$ over $W' = \{(x, y) \mid$
919 $0 \leq y_i \leq x \leq y_j \leq 1 \forall i \in J, j \in N \setminus J\}$ has the form of (14). For $y \geq 0$,

$$920 \quad f0^+(y) = \lim_{\lambda \uparrow \infty} \frac{f(\lambda y) - f(0)}{\lambda} \leq 0$$

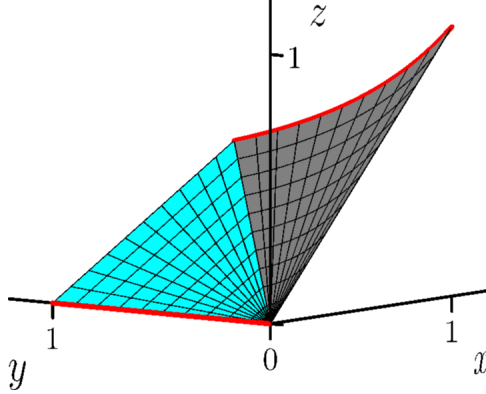


Figure 1: Convex Envelope of $\frac{x}{1+y}$ over $[0, 1]^2$

921 where the equality follows by definition (see Corollary 8.5.2 in [25]) and the inequality because
 922 f is non-increasing and $\lambda y \geq 0$. Since the convex envelope is independent of $y_{N \setminus J}$ and $g(x, y)$
 923 is non-increasing in y , it follows that $\text{conv}_{W'}(g)(x, y) \leq g(x, y)$ for all $(x, y) \in \{0, 1\} \times [0, 1]^n$.
 924 Since $\text{conv}_{W'}(g)$ is convex, $\text{conv}_{W'}(g)(x, y) \leq \text{conv}_{[0, 1]^{n+1}}(g)(x, y)$. However, $W' \subseteq [0, 1]^{n+1}$.
 925 Therefore, $\text{conv}_{W'}(g)(x, y) \geq \text{conv}_{[0, 1]^{n+1}}(g)(x, y)$. Combining these results, we conclude that
 926 $\text{conv}_{W'}(g)(x, y) = \text{conv}_{[0, 1]^{n+1}}(g)(x, y)$. \square

927 We next provide some geometrical insights into the proof of Theorem 4.1, discuss settings in
 928 which it can be generalized, and describe some applications.

929 The convex envelope of $xf(y)$ developed in Theorem 4.1 has an interesting structure. It is
 930 expressed as the maximum of a finite set of positively homogeneous functions. Each function
 931 attains the maximum over one of the polytopes in the subdivision $\bigcup_{J \subseteq N} S_J$ of $[0, 1]^{n+1}$, where
 932 $S_J = \{(x, y) \mid 0 \leq y_j \leq x \forall j \in J, x \leq y_j \leq 1 \forall j \in N \setminus J\}$. We illustrate this feature on the following
 933 example.

934 **Example 4.2.** Consider the function $g : [0, 1]^2 \mapsto \mathbb{R}$ defined as $g(x, y) = \frac{x}{1+y}$. The convex envelope
 935 of g can be obtained by convexifying its restrictions to $x = 0$ and $x = 1$, restrictions that are depicted
 936 as red thick lines in Figure 1. The proof of Theorem 4.1 argues that the convex envelope of g can
 937 be obtained by first constructing the convex envelope of g over $S_\emptyset = \{(x, y) \mid 0 \leq x \leq y \leq 1\}$, which
 938 is depicted in cyan, and gluing it to the convex envelope of g over $S_{\{1\}} = \{(x, y) \mid 0 \leq y \leq x \leq 1\}$,
 939 which is depicted in gray. More precisely, applying the formulas described in Theorem 4.1 yields
 940 that $\text{conv}_{[0, 1]}(g)(x, y) = \frac{x^2}{x + \min\{x, y\}}$ if $x > 0$ and $\text{conv}_{[0, 1]}(g)(x, y) = 0$ if $x = 0$. \square

941 Note that the convex envelope derived in Example 4.2 was obtained earlier in [36] in a more gen-
 942 eral setting using disjunctive programming. We used this example solely to illustrate the polyhedral
 943 subdivision that is at the core of the proof.

944 We next describe settings for which Theorem 4.1 can be adapted and/or generalized. First
 945 observe that, if $f(y)$ is non-decreasing, the convex envelope of $xf(y)$ over the unit hypercube can
 946 still be derived using Theorem 4.1 by replacing y_i with $1 - \bar{y}_i$. Second, note that if $y' > y''$ and

947 $f(\cdot)$ is non-increasing, then $xf\left(\frac{\min(y',x)}{x}\right) \leq xf\left(\frac{\min(y'',x)}{x}\right)$. Therefore, Theorem 4.1 can be applied
 948 sequentially to convexify functions such as $f(y) \prod_{i=1}^m x_i$. Further, the result of Theorem 4.1 also
 949 applies to more general functions $g(x, y)$ that are such that (i) $g(0, y) = 0$, (ii) $\text{conv}_{[0,1]^{n+1}} g(1, y)$ is
 950 known explicitly and non-increasing, (iii) $g(x, y')$ is concave as a function of x for a fixed y is fixed
 951 at y' . Next we demonstrate applications of Theorem 4.1 in such contexts.

952 **Corollary 4.3.** Let $g : [0, 1]^{n+1} \mapsto \mathbb{R}$ be defined as $g(x, y) = \frac{x}{ax + \sum_{i=1}^n b_i y_i + c}$ where $a \in \mathbb{R}$, $b \in \mathbb{R}^n$,
 953 and $c \in \mathbb{R}$. Define $N = \{1, \dots, n\}$, $N^+ = \{i \in N \mid b_i \geq 0\}$, and $N^- = N \setminus N^+$. Assume that
 954 $c + \sum_{i \in N^-} b_i > 0$ and $a \geq 0$. Then,

$$955 \quad \text{conv}_{[0,1]^{n+1}} g(x, y) = \begin{cases} \frac{x^2}{(a+c)x + \sum_{i \in N^+} b_i \min\{x, y_i\} + \sum_{i \in N^-} b_i (x + y_i - 1)^+} & \text{if } x > 0 \\ 0 & \text{if } x = 0. \end{cases}$$

956 *Proof.* Note that $\min\{ax + \sum_{i=1}^n b_i y_i + c \mid x \in [0, 1], y \in [0, 1]^n\} = c + \sum_{i \in N^-} b_i > 0$. Therefore, the
 957 function $g(x, y)$ is well-defined over $[0, 1]^{n+1}$. Further, observe that

$$958 \quad \frac{\partial^2 g(x, y)}{\partial x^2} = -\frac{2a(c + \sum_{i=1}^n b_i y_i)}{(ax + \sum_{i=1}^n b_i y_i + c)^3} \leq 0.$$

959 The inequality follows since $a \geq 0$, $c + \sum_{i=1}^n b_i y_i > 0$ and $ax + \sum_{i=1}^n b_i y_i + c > 0$. Therefore, $g(x, \bar{y})$
 960 is concave in x for any fixed \bar{y} . The result then follows from Theorem 4.1 after complementing the
 961 variables y_i for $i \in N^-$. \square

962 An argument similar to Corollary 4.3 yields the concave envelope of $g(x, y) = x \log(ax +$
 963 $\sum_{i=1}^n b_i y_i + c)$. In this case, using the proof technique on $-g(x, y)$ we obtain

$$964 \quad \text{conc}_{[0,1]^{n+1}} g(x, y) = \begin{cases} -x \log(x) + \\ x \log((a+c)x + \sum_{i \in N^+} b_i \min\{x, y_i\} + \sum_{i \in N^-} b_i (x + y_i - 1)^+) & \text{if } x > 0 \\ 0 & \text{if } x = 0. \end{cases}$$

965 Observe that the concave envelope of $\frac{x}{ax + \sum_{i=1}^n b_i y_i + c}$ and the convex envelope of $x \log(ax + \sum_{i=1}^n b_i y_i +$
 966 $c)$ can also be obtained by using Corollary 3.16. Next, we show that Theorem 4.1 yields convex
 967 envelopes of many polynomial functions over the unit hypercube.

968 **Corollary 4.4.** Consider a function $g(x, y) = x \left(c + \sum_{i=1}^n \sum_{j=1}^k a_{ij} y_i^{p_{ij}}\right)$ where $a_{ij} \in \mathbb{R}_+$ and
 969 $p_{ij} - 1 \in \mathbb{R}_+$. Then the concave envelope of $g(x, y)$ over $[0, 1]^{n+1}$ is given by:

$$970 \quad \text{conv}(g)_{[0,1]^{n+1}}(x, y) = \begin{cases} cx + \sum_{i=1}^n \sum_{j=1}^k a_{ij} x^{1-p_{ij}} \max[x + y_i - 1, 0]^{p_{ij}} & \text{if } x > 0 \\ 0 & \text{if } x = 0. \end{cases}$$

971 The concave envelope of $g(x, y)$ over $[0, 1]^{n+1}$ is given by:

$$972 \quad \text{conc}(g)_{[0,1]^{n+1}}(x, y) = cx + \sum_{i=1}^n \min[y_i, x] \sum_{j=1}^k a_{ij}.$$

973 *Proof.* The convex envelope is obtained using Theorem 4.1 after complementing the variables y_i .
 974 For the concave envelope, note that $g(x, y)$ is supermodular and concave-extendable from vertices.
 975 Therefore, the result follows from Theorem 3.3. \square

976 Theorem 4.1 easily yields polyhedral subdivisions defining the convex envelope of $xf(\cdot)$ if $f(\cdot)$
977 has a polyhedral convex envelope. We consider a special case of $f(y)$ where y_i are binary valued to
978 expose the techniques involved. First, we will consider certain symmetric convex functions of binary
979 variables and develop their convex envelopes. These functions appear by themselves in nonlinear
980 integer programming and we discuss some of these applications. Then, we develop convex envelopes
981 of $xf(y)$, where $f(y)$ is such a symmetric function and y are binary. Subsequently, we will discuss
982 applications of this disjunctive form and consider alterations to the polyhedral subdivision when
983 the underlying region is restricted to a subset of the hypercube.

984 In order to develop the convex envelope of the symmetric function, we will need an exclusion
985 property that helps in identifying the convex envelopes of convex functions restricted to nonconvex
986 sets. Although, we will not need the full power of Proposition 4.5 in our subsequent development,
987 we include it here for other potential applications.

988 **Proposition 4.5.** *Consider a closed set X and an upper-semicontinuous (lower-semicontinuous)
989 concave (convex) function $f : \text{conv}(X) \mapsto \mathbb{R}$. Let $f|_X$ be the restriction of f to X . There exists a
990 $V \subseteq X$, where $\text{conv}(V) \setminus V \cap X = \emptyset$, and $|V| = \dim(V) + 1$ such that the optimal solution $z(x)$ of
991 $D(x)$ ($D'(x)$) equals $\text{conc}(f|_X)(x)$. Here $D'(x)$ is the same as $D(x)$ except that the maximization
992 is replaced with minimization.*

993 *Proof.* We denote the problem $D(x)$ with vertex set V as $D_V(x)$ and the corresponding optimal
994 value as $z_V(x)$. The existence of a V' such that $z_{V'}(x) = \text{conc}(f|_X)(x)$ and $|V'| = n + 1$ follows by
995 Carathéodory's theorem. Let V be such that $\text{conv}(V)$ is the minimum volume simplex in $\text{conv}(V')$
996 that satisfies this property. There exists a minimum since each point is chosen from a compact
997 feasible region $\text{conv}(V') \setminus X$, the multipliers are chosen from a compact set, and $V^T \lambda$ and volume
998 are continuous functions, and $f(V)^T \lambda$ is upper-semicontinuous. If this volume is zero, first note
999 that we can drop one point from V since any extreme solution of $D_V(x)$ will have a support
1000 at no more than $\dim(V) + 1$ points. We now reiterate to find the minimum volume simplex,
1001 where volume is now computed in $\text{aff}(V)$. Therefore, we may assume that there does not exist
1002 V'' such that $\text{conv}(V'') \subsetneq \text{conv}(V)$ and $z_{V''}(x) = \text{conc}(f)(x)$. Assume now, by contradiction, that
1003 $x' \in \text{conv}(V) \setminus V \cap X$. Let λ be the optimal solution of $D_V(x)$. By minimality of volume, it follows
1004 that $\lambda_i > 0$ for all i . Let λ' be a feasible solution of $D(x')$ and $r = \min_i \{\frac{\lambda_i}{\lambda'_i} \mid \lambda'_i > 0\}$. Further, let
1005 i' be the index that achieves this minimum. Clearly, $0 < r$. Then,

$$1006 \quad \text{conc}(f)(x) = f(V)^T \lambda = f(V)^T (\lambda - r\lambda') + rf(V)^T (\lambda') \leq f(V)^T (\lambda - r\lambda') + rf(x') \leq \text{conc}(f)(x),$$

1007 where the first inequality follows from concavity of f and the second inequality since $x' \in X$,
1008 $\lambda - r\lambda' \geq 0$, and $e^T (\lambda - r\lambda') + r = 1$. Therefore, equality holds throughout. This yields a
1009 contradiction since $V'' = V \setminus \{v_{i'}\} \cup x'$ is such that $\text{conv}(V'') \subsetneq V$ and $z_{V''}(x)$ equals $\text{conc}(f)(x)$. \square

1010 In Theorem 4.6 we consider a symmetric function of binary variables, $f(\|x\|_1)$, where f is a
1011 convex function, and show that its convex envelope is easy to characterize.

1012 **Theorem 4.6.** *Consider a function $g(x) : [0, 1]^n \mapsto \mathbb{R}$, that is convex-extendable from vertices.
1013 Then, the polyhedral subdivision $[0, 1]^n = \bigcup_{i=1}^n P_i$, where $P_i = \{x \mid i - 1 \leq \sum_{j=1}^n x_j \leq i, 0 \leq x \leq 1\}$
1014 describes the convex envelope of $g(x)$ if and only if its restriction to $\{0, 1\}^n$ can be written as
1015 $f(\sum_{i=1}^n x_i) + \sum_{i=1}^n a_i x_i$ for some convex function f . The corresponding convex envelope is:*

$$1016 \quad \max_{i \in \{1, \dots, n\}} (f(i) - f(i-1)) \sum_{j=1}^n x_j + if(i-1) - (i-1)f(i) + \sum_{j=1}^n a_j x_j. \quad (15)$$

1017 *Proof.* (\Leftarrow) Since $g(x)$ is convex-extendable from $\{0, 1\}^n$ it suffices to restrict $g(x)$ to $\{0, 1\}^n$ and
 1018 therefore we may assume that $g(x) = f(\sum_{i=1}^n x_i) + \sum_{i=1}^n a_i x_i$ for some convex function f . Consider
 1019 the set $W_i = \{x \in \mathbb{R}^n \mid \sum_{j=1}^n x_j = i\}$. The function $g(x)$ is linear over W_i . Since, each extreme
 1020 point of W_i is also an extreme point of $[0, 1]^n$, the convex envelope is tight at each such point.
 1021 Therefore, the convex envelope is also tight over each W_i . In other words, the convex envelope is
 1022 the convex envelope of $g(x)$ restricted to $\bigcup_{i=0}^n W_i$. It follows from Proposition 4.5 that the convex
 1023 envelope is then described by $\bigcup_{i=1}^n P_i$.

1024 (\Rightarrow) For the direct implication, consider any function $g(x)$ whose convex envelope is described
 1025 by $\bigcup_{i=1}^n P_i$. Therefore, the function is convex-extendable from $\{0, 1\}^n$ and the restriction of $g(x)$ to
 1026 $\{0, 1\}^n$ must be linear over each P_i . Let $l^i(x) = a_0^i + \sum_{j=1}^n a_j^i x_j$ equal $g(x)$ at the extreme points
 1027 of P_i . Note that P_1 is a simplex. Therefore, $l^1(x)$ is uniquely defined by the extreme points of P_1 .
 1028 Then, since $l^i(x)$ and $l^{i+1}(x)$ match at the extreme points of W_i , it follows that they also match
 1029 everywhere on $\text{aff}(W_i)$. In other words, $l^{i+1}(x) - l^i(x) = \alpha^{i+1}(\sum_{j=1}^n x_j - i)$ for $i = 1, \dots, n-1$.
 1030 Further, by convexity of the envelope, $\alpha^{i+1} \geq 0$, otherwise $l^i(x)$ overestimates the function at
 1031 the extreme points of W_{i+1} . In other words, $g(x) = a_0^1 + \sum_{i=1}^n a_i^1 x_i + \sum_{i=2}^n \alpha^i (\sum_{j=1}^n x_j - i)^+$ at
 1032 each point in $\{0, 1\}^n$, where $(\cdot)^+$ denotes $\max\{0, \cdot\}$. Since the second term is a convex function of
 1033 $\sum_{j=1}^n x_j$, the result follows. \square

1034 In fact, we have shown the following result.

1035 **Corollary 4.7.** *Consider a function $g(x) : P \mapsto \mathbb{R}$, that is convex-extendable from vertices of P ,
 1036 where $P \subseteq [0, 1]^n$ is a polytope. Assume that for each $i \in \{1, \dots, n-1\}$, $W_i = \{x \in P \mid \sum_{j=1}^n x_j = i\}$
 1037 is integral. Then, the polyhedral subdivision $P = \bigcup_{i=1}^n P_i$, where $P_i = \{x \in P \mid i-1 \leq \sum_{j=1}^n x_j \leq i\}$
 1038 describes the convex envelope of $g(x)$ if its restriction to $\text{vert}(P)$ can be written as $f(\sum_{i=1}^n x_i) +$
 1039 $\sum_{i=1}^n a_i x_i$ for some convex function f . The convex envelope is given by (15).*

1040 *Proof.* Note that W_0 and W_n are either empty or integral by definition. The remaining proof is just
 1041 as that of Theorem 4.6. \square

1042 We next give applications of Theorem 4.6 and Corollary 4.7 in the derivation of convex envelopes
 1043 of various functions. In the following result, we use the same notation as that used in Corollary 3.10.

1044 **Corollary 4.8** (Theorem 3 and 5 in Sherali [30]). *Consider the function $\phi^m(y) : \text{vert}(G) \mapsto \mathbb{R}$
 1045 defined as $\sum_{J \subseteq D, |J|=m} [\prod_{d \in J} y_{i(d), j(d)}]$, where $m \leq n$. Then, the convex envelope of $\phi^m(y)$ over G
 1046 is given by*

$$1047 \phi_C^m(x) = \max \left\{ 0, \binom{k}{m-1} \sum_{i=1}^n x_j - (m-1) \binom{k+1}{m} \mid k = m-1, \dots, n-1 \right\}. \quad (16)$$

1048 *If $d_i = 1$ for all i , then $\phi_C^m(x)$ is also the convex envelope of $\phi^m(y) : G \mapsto \mathbb{R}$ over G .*

1049 *Proof.* As in the proof of Corollary 3.10, we may restrict attention to J such that if d and d'
 1050 belong to J , then $i(d) \neq i(d')$. Note that $\{x \mid \sum_{j=1}^n \sum_{r=1}^{d_j} y_{jr} = i, \sum_{r=1}^{d_j} y_{jr} \leq 1 \forall j\}$ is an integral
 1051 polytope since the corresponding matrix is totally unimodular (see for example, Corollary 2.8 in
 1052 [22]). Note that $\phi^m(x)$ is supermodular and expressible as $\binom{\sum_{i=1}^n x_i}{m}$ where $\binom{u}{m}$ is defined as zero
 1053 if $u < m$. The convexity of ϕ^m as a function of $\sum_{j=1}^n x_j$ then follows from Proposition 5.1 in
 1054 [19] which states that a function of the form $g(|X|)$ is supermodular, where $|X|$ is the cardinality
 1055 of a set X if and only if g is convex. The convexity of ϕ^m can also be verified by directly since
 1056 $\binom{i}{m} - \binom{i-1}{m} = \binom{i-1}{m-1}$ which is a non-decreasing function of i . The convex envelope then follows from
 1057 Corollary 4.7. Then, substituting $f(i) = \binom{i}{m}$ in (15), we obtain (16). The last statement follows
 1058 just as in Corollary 3.10. \square

1059 **Example 4.9.** Consider the function $f(x) = \frac{1}{\sum_{i=1}^n x_i}$ where $x_i \in \{0, 1\}$, and $P = \{x \in [0, 1]^n \mid$
1060 $\sum_{i=1}^n x_i \geq 1\}$. The standard factorable programming relaxation uses the function itself as the
1061 convex underestimator. The function, $f(x)$, appears in the formulation of the consistent biclustering
1062 problem [6], where the authors relax $f(x)$ over P by cross-multiplying with the denominator and then
1063 relaxing $x_i f(x)$ over $[0, 1] \times [1, \frac{1}{n}]$. Since this relaxation is valid even when $x_i \in [0, 1]$ and since it
1064 is polyhedral, it is weaker than the factorable relaxation discussed above. Further, note that $f(x)$
1065 is convex and $W_i = \{x \in P \mid \sum_{j=1}^n x_j = i\}$ are clearly integral. Therefore, Corollary 4.7 applies
1066 and provides a description of the convex envelope of $f(x)$ over P . Observe that the factorable
1067 programming relaxation, which is non-polyhedral, is weaker than the polyhedral relaxation obtained
1068 from Corollary 4.7 when $\sum_{i=1}^n x_i \notin \mathbb{Z}$. It may be noted that the concave envelope of $f(x)$ was
1069 previously described in Example 3.20.

1070 As mentioned before, Theorem 4.1 also provides a constructive derivation of the polyhedral
1071 subdivision describing the convex envelope of $xf(y)$ when $f(y)$ has a polyhedral envelope. We next
1072 illustrate the constructions involved for the case where the function $f(y)$ is of the form $f(\|y\|_1)$,
1073 where $y \in \{0, 1\}^n$.

1074 **Corollary 4.10.** Consider $g(x, y) = xf(\sum_{i=1}^n y_i)$. Let f be a non-increasing convex function and
1075 $y \in \{0, 1\}^n$. For $I \subseteq N$ and $0 < l \leq |I|$, let

$$1076 \quad S(I, l) = \left\{ (x, y) \mid 0 \leq y_i \leq x \leq y_j \leq 1, \forall i \in I, j \in N \setminus I, (l-1)x \leq \sum_{i \in I} y_i \leq lx \right\}.$$

1077 Then, the polyhedral subdivision $\bigcup_{\substack{I \subseteq N \\ 0 < l \leq |I|}} S(I, l)$ defines the convex envelope of $g(x, y)$. In particular,
1078 the convex envelope of $g(x, y)$ over $S(I, l)$ is given by:

$$1079 \quad \left(f(l + |I^c|) - f(l - 1 + |I^c|) \right) \sum_{i \in I} y_i + \left(lf(l - 1 + |I^c|) - (l - 1)f(l + |I^c|) \right) x \quad (17)$$

1080 where $I^c = N \setminus I$.

1081 *Proof.* First note that when $x = 1$, the function $f(y)$ satisfies the conditions of Theorem 4.6.
1082 Therefore, the polyhedral subdivision is given by $\bigcup_{i=1}^n W'_i$, where $W'_i = \{y \in \mathbb{R}^n \mid i - 1 \leq \sum_{i=1}^n y_i \leq$
1083 $i\}$. In particular, over W'_i

$$1084 \quad \text{conv}_{[0, 1]^n}(f)(y) = h(y) := \left(f(i) - f(i - 1) \right) \sum_{j=1}^n y_j + \left(if(i - 1) - (i - 1)f(i) \right) \quad (18)$$

1085 Clearly, $\text{conv}_{[0, 1]^{n+1}}(xf(y)) = \text{conv}_{[0, 1]^{n+1}}(xh(y))$. Now, the situation fits the setting of Theorem 4.1.
1086 Therefore, the convex envelope over $S(I, l)$ is given by $xh\left(\frac{y'}{x}\right)$, where $y'_i = \min(y_i, x)$. By definition
1087 of $S(I, l)$, $y'_i = y_i$ for $i \in I$ and $y'_i = x$ for $i \in N \setminus I$. Expanding using (18) one obtains (17). It follows
1088 by choosing $f(x, y)$ to be a strictly convex and decreasing function (such as $\frac{1}{1+y_1+\dots+y_n}$) that the
1089 convex envelope of $g(x, y)$ is only tight at the binary points that belong to $\text{vert}(S(I, l))$. Therefore,
1090 $\bigcup_{\substack{I \subseteq N \\ 0 < l \leq |I|}} S(I, l)$ gives a polyhedral subdivision of $[0, 1]^{n+1}$. \square

1091 In Section 3, we discussed a situation where removing the origin from the underlying polytope
1092 changed the associated polyhedral subdivision completely. As we mentioned, this was because
1093 each simplex in the triangulation contained the origin as a vertex. For the function addressed in
1094 Corollary 4.10, it can be easily verified that the origin is still a vertex of each polyhedron in the

1095 subdivision. However, in this case the structure of the convex envelope is not completely altered
 1096 when the origin is removed from the underlying region. An intuitive reason for this is that the
 1097 polytopes that form the subdivision described in Corollary 4.10 are not simplices. Therefore, even
 1098 if the origin is removed from a polytope, it may still have sufficient points to describe the convex
 1099 envelope over a subregion. Theorem 4.11 exemplifies this phenomenon. We discuss an application
 1100 of this result in Example 4.12.

1101 **Theorem 4.11.** Consider $g(x, y) = xf(\sum_{i=1}^n y_i)$, where $f(z) : \mathbb{R} \mapsto \mathbb{R}$ is a convex non-increasing
 1102 function. Assume that $(x, y) \in \{0, 1\}^{n+1}$ and $(x, y) \neq (0, 0)$. Let $W = \{(x, y) \in [0, 1]^{n+1} \mid x +$
 1103 $\sum_{i=1}^n y_i \geq 1\}$. Then, the polyhedral subdivision $\mathcal{S} = \bigcup_{i=0}^{n-1} S(i) \cup \bigcup_{\substack{I \subseteq N \\ 0 \leq k \leq |I|-1}} T(I, k)$ describes the
 1104 convex envelope of $g(x, y)$ over W where

$$1105 \quad S(i) = \left\{ (x, y) \left| \begin{array}{l} 0 \leq y \leq 1 \\ 0 \leq x \leq 1 \\ 1 + (i-1)x \leq \sum_{j=1}^n y_j \leq 1 + ix \\ \sum_{j \in C} y_j \leq 1 + (|C| - 1)x \forall C \subseteq N \end{array} \right. \right\}$$

1106 and

$$1107 \quad T(I, k) = \left\{ (x, y) \left| \begin{array}{l} 0 \leq y_i \leq x \forall i \in I \\ x \leq y_j \leq 1 \forall j \in I^c \\ kx \leq \sum_{j \in I} y_j \leq (k+1)x \\ \sum_{j \in I^c} y_j \geq 1 + (|I^c| - 1)x \end{array} \right. \right\}.$$

1108 In particular,

$$1109 \quad \text{conv}_W(g(x, y)) = \max \left\{ \max_{0 \leq i \leq n-2} h^{S(i)}(x, y), \max_{\substack{I \subseteq N \\ 0 \leq k \leq |I|-1}} h^{T(I, k)}(x) \right\},$$

1110 where $h^{S(i)}(x, y) = (if(i) - (i-1)f(i+1))x - (f(i+1) - f(i))(1 - \sum_{j=1}^n y_j)$ and $h^{T(I, k)}(x, y) =$
 1111 $(f(|I^c| + k + 1) - f(|I^c| + k)) \sum_{j \in I} y_j + ((k+1)f(|I^c| + k) - kf(k+1 + |I^c|))x$.

1112 *Proof.* We first show that \mathcal{S} covers the unit hypercube. Consider $(x', y') \in W$. There are two cases.
 1113 First assume that $\sum_{j \in C} y'_j \leq 1 + (|C| - 1)x'$ for all $C \subseteq N$. Since this inequality holds for $C = N$,
 1114 we have that $\sum_{j=1}^n y'_j \leq 1 + (n-1)x'$. Further, since $(x', y') \in W$, we have that $\sum_{j=1}^n y'_j \geq 1 - x'$.
 1115 It follows that $(x', y') \in S(i)$ for some $i \in \{0, \dots, n-1\}$. Second, assume that there exists $J \in C$
 1116 such that $\sum_{j \in J^c} y'_j > 1 + (|J^c| - 1)x'$. Define $I = J \setminus \{j \in J \mid y'_j \geq x'\} \cup \{j \in J^c \mid y'_j < x'\}$. It is
 1117 easily verified that $y'_j \leq x'$ for $j \in I$, $y'_j \geq x'$ for $j \in I^c$, and that $\sum_{j \in I^c} y'_j > 1 + (|I^c| - 1)x'$.
 1118 Further, by construction of I , we have that $\sum_{j \in I} y'_j \leq |I|x'$. It follows that $(x', y') \in T(I, k)$ where
 1119 $k \in \{0, \dots, |I| - 1\}$.

1120 Next, we show that $S(i)$ has 0-1 extreme points. In fact, we will show that $S(i) = \text{conv}(W_1 \cup W_2)$
 1121 where $W_1 = \{(0, y) \mid 0 \leq y \leq 1, \sum_{j=1}^n y_j = 1\}$ and $W_2 = \{(1, y) \mid i \leq \sum_{i=1}^n y_i \leq i+1\}$. To this end,
 1122 we will show that, independent of the choice of objective coefficients b and c , the following linear

1123 program

$$\begin{aligned}
1124 \quad & P(S) \quad \min \quad bx + cy \\
1125 \quad & \text{s.t.} \quad 0 \leq y_j \leq 1 \quad j = 1, \dots, n \quad (\alpha_j) \\
1126 \quad & \quad \quad 0 \leq x \leq 1 \quad (\beta) \\
1127 \quad & \quad \quad 1 + (i-1)x \leq \sum_{j=1}^n y_j \leq 1 + ix \quad (\delta) \\
1128 \quad & \quad \quad \sum_{j \in C} y_j \leq 1 + (|C| - 1)x \quad \forall C \subseteq N \quad (\gamma_C) \\
1129 \quad &
\end{aligned}$$

1130 has an integer optimal solution. In the linear program $P(S)$, α , β , δ , and γ are the dual variables
1131 corresponding to the constraints. Each of the variables α_j , β , and δ corresponds to two constraints.
1132 Among these, the appropriate constraint depends on the sign of the associated dual variable.

1133 We assume without loss of generality that $c_1 \leq \dots \leq c_n$. Let $N(t) = \{1, \dots, t\}$. There are two
1134 cases. Assume first that $c_{i+1} \geq 0$. Define $\delta = \max\{0, c_i\}$, $\gamma_{N(t)} = c_t - c_{t+1}$ for $t = 1, \dots, i-1$,
1135 $\gamma_{N(i)} = \min\{0, c_i\}$, $\alpha_j = c_j - c_i$ for $j > i$. Let all other α and γ dual variables be set to 0. Adding
1136 the resulting (weighted) inequalities, we obtain

$$1137 \quad \sum_{j=1}^n c_j y_j - x \sum_{j=2}^i c_j \geq c_1. \quad (19)$$

1138 Let $\beta = b + \sum_{j=2}^i c_j$. If $\beta > 0$, adding the corresponding (weighted) constraint to (19) shows that
1139 $bx + cy \geq c_1$ for all feasible solutions of $P(S)$. Therefore, the integer solution $x = 0$, $y_1 = 1$, and
1140 $y_j = 0$ for $j > 1$, whose objective value is c_1 , is optimal for $P(S)$. If $\beta \leq 0$, we proceed similarly
1141 to show that $bx + \sum_{j=1}^n c_j y_j \geq b + \sum_{j=1}^i c_j$ for all feasible solutions of $P(S)$. Therefore, the integer
1142 solution $x = 1$, $y_j = 1$ for $j \leq i$, and $y_j = 0$ for $j > i$ is optimal for $P(S)$.

1143 Now, assume that $c_{i+1} < 0$. Define $\delta = c_{i+1}$, $\gamma_{N(t)} = c_t - c_{t+1}$ for $t = 1, \dots, i$, and $\alpha_j = c_j - c_{i+1}$
1144 for $j > i+1$. Let the remaining α and γ dual variables be set to zero. Adding the resulting
1145 (weighted) inequalities, we obtain that $\sum_{j=1}^n c_j y_j - x \sum_{j=2}^{i+1} c_j \geq c_1$. Let $\beta = b + \sum_{j=2}^{i+1} c_j$. If $\beta > 0$,
1146 we conclude that $bx + \sum_{j=1}^n c_j y_j \geq c_1$ and so the integer solution $x = 0$, $y_1 = 1$, and $y_j = 0$ for
1147 $j > 1$ is optimal for $P(S)$. If $\beta \leq 0$, we obtain similarly that $bx + \sum_{j=1}^n c_j y_j \geq b + \sum_{j=1}^{i+1} c_j$ and
1148 so the integer solution $x = 1$, $y_j = 1$ for $j \leq i+1$, and $y_j = 0$ for $j > i+1$ is optimal for $P(S)$.
1149 Hence, $S(i) = \text{conv}(W_1 \cup W_2)$. It follows in a manner similar to Theorem 4.1 and Corollary 4.10 by
1150 applying Theorem 1 of [41] that the extreme points of $T(I, k)$ are binary.

1151 Clearly, $h^{S(i)}(x, y) \leq 0$ if $x = 0$ and $\sum_{j=1}^n y_i \geq 1$ with equality when $\sum_{j=1}^n y_i = 1$. Also,
1152 $h^{S(i)}(x, y) = f(i) + (i-r)(f(i) - f(i+1))$ if $x = 1$ and $\sum_{j=1}^n y_j = r$. Then, it follows by convexity
1153 of f that $h^{S(i)}(x) \leq f(r)$ with equality if $r \in \{i, i+1\}$. Therefore, by Theorem 2.4, $h^{S(i)}(x, y) \leq$
1154 $\text{conv}_W g(x, y)$, with equality over $S(i)$.

1155 From Corollary 4.10, it follows that $\text{conv}_{[0,1]^{n+1}} g(x, y)$ over $T(I, k)$ is given by $h^{T(I, k)}(x, y)$.
1156 Therefore, $h^{T(I, k)}(x, y) \leq g(x, y)$. Further, $\text{vert}(T(I, k)) \subseteq \text{vert}(S(I, k+1))$, where $S(I, k+1)$ is
1157 defined as in Corollary 4.10. Therefore, $h^{T(I, k)}(x, y) = g(x, y)$ for $(x, y) \in \text{vert}(T(I, k))$. It follows
1158 then from Theorem 2.4 that $h^{T(I, k)}(x, y) \leq \text{conv}_W g(x, y)$ with equality over $T(I, k)$.

1159 Choosing $f(\cdot)$ to be a strictly convex and decreasing function, it can be verified that $h^{S(i)}(x, y)$
1160 is not tight at any binary point that is not an extreme point of $S(i)$. Similarly, as in Corollary 4.10,
1161 $h^{T(I, k)}(x, y)$ is not tight at any binary point that is not an extreme point of $T(I, k)$. Therefore,
1162 $\bigcup_{i=0}^{n-1} S(i) \cup \bigcup_{\substack{I \subseteq N \\ 0 \leq k \leq |I|-1}} T(I, k)$ is a polyhedral subdivision of W . \square

1163 **Example 4.12.** Consider $g(x, y) = \frac{x}{x + \sum_{i=1}^n y_i}$, where $(x, y) \in \{0, 1\}^{n+1}$ and $x + \sum_{i=1}^n y_i \geq 1$. This
 1164 function appears along with the specified constraint in the consistent biclustering problem [6]. The
 1165 convex envelope for $g(x, y)$ over W is described by the polyhedral division of Theorem 4.11. In
 1166 particular,

$$1167 \quad h^{S(i)}(x, y) = \frac{1}{(i+1)(i+2)} \left[(2i+1)x - \sum_{j=1}^n y_j + 1 \right]$$

1168 and

$$1169 \quad h^{T(I,k)}(x, y) = \frac{1}{(|I^c| + k + 2)(|I^c| + k + 1)} \left[(|I^c| + 2k + 2)x - \sum_{j \in I} y_j \right].$$

1170 Because for all feasible solutions $\frac{1}{x + \sum_{i=1}^n y_i} \in \left[\frac{1}{n+1}, 1 \right]$, the factorable relaxation of $g(x, y)$ takes the
 1171 form $\max \left\{ \frac{1}{n+1}x, x + u(x, y) - 1 \right\}$ where $u(x, y)$ is a convex underestimator of $\frac{1}{x + \sum_{i=1}^n y_i}$ over the
 1172 feasible region. If this convex underestimator is obtained without using the fact that variables are
 1173 binary, as is typical in global optimization software, $u(x, y)$ would be chosen equal to $\frac{1}{x + \sum_{i=1}^n y_i}$ and
 1174 the resulting factorable relaxation would therefore be non-polyhedral. Such relaxation can be verified
 1175 to be weaker than the relaxations that can be obtained from Corollary 4.10 and Theorem 4.11. To
 1176 illustrate the difference, consider the special case $g(x, y) = \frac{x}{x+y}$. At the point $(1, 0.5)$, the factorable
 1177 relaxation obtained without using integrality of the variables evaluates to $\frac{2}{3}$ while the relaxation of
 1178 Corollary 4.10 obtained by defining $g(x, y) = 0$ when $x = 0$ evaluates to $\frac{3}{4}$, a value that can be
 1179 computed after selecting $I = \{1\}$ and $l = 1$. Further, at the point $(0.5, 0.5)$, the factorable relaxation
 1180 obtained without using integrality evaluates to $\frac{1}{4}$. The relaxation using Corollary 4.10 also evaluates
 1181 to $\frac{1}{4}$. However, the relaxation of Theorem 4.11 (in particular, $h^{S(0)}(x, y)$) evaluates to $\frac{1}{2}$ at this
 1182 point. This example illustrates that, for this type of functions, Theorem 4.11 produces a relaxation
 1183 that is tighter over W than the relaxation obtained using Corollary 4.10. This relaxation is in turn
 1184 tighter than the traditional factorable relaxation. \square

1185 5 Conclusion

1186 We studied the problem of developing convex and concave envelopes of nonlinear functions over
 1187 subsets of a hyper-rectangle. In particular, we showed that the optimal value of a primal-dual pair
 1188 of linear optimization problems yields the concave envelope when it has a polyhedral structure. We
 1189 then showed that existence of polynomial-time separation algorithms for the concave envelopes of
 1190 a set of functions imply polynomial-time separability for the concave envelope of the maximum of
 1191 these functions.

1192 Next, we showed that a result of Lovász [19] allows construction of concave envelopes of super-
 1193 modular functions over a hyper-rectangle if the function is concave-extendable from the vertices of
 1194 the hyper-rectangle. We generalized this construction to consider supermodular functions over a
 1195 lattice family and demonstrated that this result yields simple derivations and extensions of results
 1196 in the literature [30, 8, 5, 21, 26]. As a particular application, we constructed the concave envelope
 1197 of the composition of a univariate convex function with a linear function, a structure commonly
 1198 encountered when deriving convex relaxations of factorable programs.

1199 We then showed that the convex envelope of certain functions that have a disjunctive property
 1200 can be developed by convexifying their restrictions over carefully selected orthogonal disjunctions.
 1201 As a consequence of this result, we developed convex envelopes for a variety of fractional and

1202 polynomial expressions over the unit hypercube. We then considered a convex function restricted to
 1203 a nonconvex set. We derived an exclusion property that limits the subsets that need to be considered
 1204 while evaluating the convex envelope outside the nonconvex set. We used this property to identify
 1205 the polyhedral subdivision that characterizes the convex envelope of a symmetric function of binary
 1206 variables that depends only on the cardinality of the set of binary variables that assume a value
 1207 of one. This result generalizes some earlier results discovered in [30] and has other applications as
 1208 well; see [6]. Then, we used these symmetric functions to define disjunctive functions, for which we
 1209 combined our previous results to derive their convex envelopes. This construction demonstrated that
 1210 polyhedral subdivisions are naturally obtained by using our convexification scheme for disjunctive
 1211 functions. Finally, we discussed applications of these disjunctive functions in relaxing the consistent
 1212 biclustering problem described in [6].

1213 The derivation of concave envelopes for nonconcave functions f yields ways to obtain convex
 1214 relaxations for constraints of the form $f(x) \geq r$. Investigating the computational advantages that
 1215 these new relaxations offer over those currently used in software implementations is an important
 1216 direction of future research. On the theoretical side, investigating whether stronger relaxations of
 1217 $f(x) \geq r$ can be obtained in closed-form is also an interesting avenue for future work.

1218 References

- 1219 [1] F. A. Al-Khayyal and J. E. Falk. Jointly constrained biconvex programming. *Mathematics of*
 1220 *Operations Research*, 8:273–286, 1983.
- 1221 [2] E. Balas and J. B. Mazzola. Nonlinear 0–1 programming: I. Linearization techniques. *Math-*
 1222 *ematical Programming*, 30:1–21, 1984.
- 1223 [3] X. Bao, N. V. Sahinidis, and M. Tawarmalani. Multiterm polyhedral relaxations for nonconvex,
 1224 quadratically constrained quadratic programs. *Optimization Methods and Software*, 24:485–504,
 1225 2009.
- 1226 [4] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Waechter. Branching and bounds tightening
 1227 techniques for non-convex MINLP. *Optimization Methods and Software*, 24:597–634, 2009.
- 1228 [5] H. P. Benson. Concave envelopes of monomial functions over rectangles. *Naval Research*
 1229 *Logistics*, 51:467–476, 2004.
- 1230 [6] S. Busygin, O. A. Prokopyev, and P. M. Pardalos. Feature selection for consistent biclustering
 1231 via fractional 0-1 programming. *Journal of Combinatorial Optimization*, 10:7–21, 2005.
- 1232 [7] S. Ceria and J. Soares. Convex programming for disjunctive convex optimization. *Mathematical*
 1233 *Programming*, 86:595–614, 1999.
- 1234 [8] D. Coppersmith, O. Günlük, J. Lee, and J. Leung. A polytope for a product of real linear
 1235 functions in 0/1 variables. *IBM Research Report*, 2003.
- 1236 [9] Y. Crama. Recognition problems for special classes of polynomials in 0-1 variables. *Mathemat-*
 1237 *ical Programming*, 44:139–155, 1989.
- 1238 [10] Y. Crama. Concave extensions for nonlinear 0–1 maximization problems. *Mathematical Pro-*
 1239 *gramming*, 61:53–60, 1993.
- 1240 [11] J. Edmonds. Submodular functions, matroids, and certain polyhedra. In *Combinatorial Struc-*
 1241 *tures and Their Applications*, pages 69–87. Gordan and Breach, 1970.

- 1242 [12] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*. Princeton Mathematical Series. Springer-Verlag, Berlin, New York, 1988.
1243
- 1244 [13] G. Hardy, J. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, 1988.
- 1245 [14] J-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis*. Springer, 2001.
- 1246 [15] W. Hock and K. Schittkowski. *Test examples for nonlinear programming codes*. Springer Verlag, 1981.
1247
- 1248 [16] R. Horst and H. Tuy. *Global optimization: Deterministic approaches*. Springer Verlag, Berlin, Third edition, 1996.
1249
- 1250 [17] C. W. Lee. Subdivisions and triangulations of polytopes. In J. E. Goodman and J. O'Rourke, editors, *Handbook of discrete and computational geometry*, chapter 14. CRC Press, 1997.
1251
- 1252 [18] LINDO Systems Inc. LINGO 11.0 optimization modeling software for linear, nonlinear, and integer programming. Available at <http://www.lindo.com>, 2008.
1253
- 1254 [19] L. Lovász. Submodular functions and convexity. In M. Grötschel and B. Korte, editors, *Mathematical Programming: The State of the Art*, pages 235–257. Springer, 1982.
1255
- 1256 [20] G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. *Mathematical Programming*, 10:147–175, 1976.
1257
- 1258 [21] C. A. Meyer and C. A. Floudas. Convex envelopes for edge-concave functions. *Mathematical Programming*, 103:207–224, 2005.
1259
- 1260 [22] G. L. Nemhauser and L. A. Wolsey. *Integer and combinatorial optimization*. Wiley-interscience series in discrete mathematics and optimization. John Wiley and Sons, 1988.
1261
- 1262 [23] J.-P. P. Richard and M. Tawarmalani. Lifted inequalities: A framework for generating strong cuts for nonlinear programs. *Mathematical Programming*, 121:61–104, 2010.
1263
- 1264 [24] A. D. Rikun. A convex envelope formula for multilinear functions. *Journal of Global Optimization*, 10:425–437, 1997.
1265
- 1266 [25] R. T. Rockafellar. *Convex analysis*. Princeton Mathematical Series. Princeton University Press, 1970.
1267
- 1268 [26] C.-D. Rodrigues, D. Quadri, P. Michelon, and S. Gueye. A t-linearization scheme to exactly solve 0-1 quadratic knapsack problems. In *Proceedings of the European Workshop on Mixed Integer Programming*, pages 251–260. CIRM, Marseille, France, 2010.
1269
1270
- 1271 [27] H. S. Ryoo and N. V. Sahinidis. Analysis of bounds of multilinear functions. *Journal of Global Optimization*, pages 403–424, 2001.
1272
- 1273 [28] A. Schrijver. *Theory of linear and integer programming*. Chichester, New York, Wiley, 1986.
- 1274 [29] A. Schrijver. *Combinatorial optimization, polyhedra and efficiency*. Springer-Verlag, Berlin, Heidelberg, 2003.
1275
- 1276 [30] H. D. Sherali. Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets. *Acta Mathematica Vietnamica*, 22:245–270, 1997.
1277

- 1278 [31] H. D. Sherali and H. Wang. Global optimization of nonconvex factorable programming prob-
1279 lems. *Mathematical Programming*, 89:459–478, 2001.
- 1280 [32] R. Stubbs and S. Mehrotra. A branch-and-cut method for 0–1 mixed convex programming.
1281 *Mathematical Programming*, 86:515–532, 1999.
- 1282 [33] F. Tardella. Existence and sum decomposition of vertex polyhedral convex envelopes. *Opti-
1283 mization Letters*, 2:363–375, 2008.
- 1284 [34] M. Tawarmalani. Polyhedral basis and disjunctive programming. Working paper, 2005.
- 1285 [35] M. Tawarmalani. Simultaneous convexification of functions. Working paper, 2010.
- 1286 [36] M. Tawarmalani and N. V. Sahinidis. Semidefinite relaxations of fractional programs via novel
1287 convexification techniques. *Journal of Global Optimization*, 20:137–158, 2001.
- 1288 [37] M. Tawarmalani and N. V. Sahinidis. *Convexification and global optimization in continu-
1289 ous and mixed-integer nonlinear programming*. Kluwer Academic Publishers, Dordrecht, The
1290 Netherlands, 2002.
- 1291 [38] M. Tawarmalani and N. V. Sahinidis. Convex extensions and envelopes of lower semi-continuous
1292 functions. *Mathematical Programming*, 93:247–263, 2002.
- 1293 [39] M. Tawarmalani and N. V. Sahinidis. Global optimization of mixed-integer nonlinear programs:
1294 A theoretical and computational study. *Mathematical Programming*, 99:563–591, 2004.
- 1295 [40] M. Tawarmalani and N. V. Sahinidis. A polyhedral branch-and-cut approach to global opti-
1296 mization. *Mathematical Programming*, 103:225–249, 2005.
- 1297 [41] M. Tawarmalani, J.-P. P. Richard, and K. Chung. Strong valid inequalities for orthogonal
1298 disjunctions and bilinear covering sets. *Mathematical Programming* (to appear), 2010.
- 1299 [42] D. M. Topkis. *Supermodularity and complementarity*. Princeton University Press, Princeton,
1300 NJ, 1998.