NEW RELAXATIONS FOR COMPOSITE FUNCTIONS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Taotao He

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Mohit Tawarmalani, Chair

Krannert School of Management

Dr. Yanjun Li

Krannert School of Management

Dr. Andrew(Lu) Liu

Krannert School of Management

Dr. J. George Shanthikumar

Krannert School of Management

Approved by:

Dr. Yanjun Li

Head of the School Graduate Program

To my family

ACKNOWLEDGMENTS

First and foremost, I would like to thank my adviser, Professor Mohit Tawarmalani. I would never have started my academic journey if I were not lucky enough to meet you and have you as an adviser. Thank you for giving me the opportunity to see how cute ideas are developed. Thank you for working with me to perfect results and papers. Thank you for making me feel confident in situations where I never imagined that I could feel in that way. Thank you for investing so much time, especially summer time, in talking with me. Thank you for letting me grow at my own speed.

I would also like to thank my thesis committee Professor Yanjun Li, Professor Andrew Liu and Professor George Shanthikumar for their feedback.

It was great 6 years in West Lafayette with my officemates and soccer teammates at Purdue, Tao, Didun, Xiangyu, Guangyuan, Zixin, Chipeng, Pengcheng, Weiran, Kai, Shou, Bohao, Yucheng, Yifan, and all PCFC members.

Last but not least, I am lucky enough to have the most supportive family. My mother, father, brother, mother-in-law, and father-in-law built for me a home without worries so that I could focus on my studies. More importantly, Yihui, thank you for making West Lafayette feel like a home.

TABLE OF CONTENTS

Page

LIST OF FIGURES

ABSTRACT

He, Taotao Ph.D., Purdue University, August 2019. New Relaxations for Composite Functions. Major Professor: Mohit Tawarmalani.

Mixed-integer nonlinear programs are typically solved using branch-and-bound algorithms. A key determinant of the success of such methods is their ability to construct tight and tractable relaxations. The predominant relaxation strategy used by most state-of-the-art solvers is the factorable programming technique. This technique recursively traverses the expression tree for each nonlinear function and relaxes each operator over a bounding box that covers the ranges for all the operands. While it is versatile, and allows finer control over the number of introduced variables, the factorable programming technique often leads to weak relaxations because it ignores operand structure while constructing the relaxation for the operator.

In this thesis, we introduce new relaxations, called composite relaxations, for composite functions by convexifying the outer-function over a polytope, which models an ordering structure of outer-approximators of inner functions. We devise a fast combinatorial algorithm to separate the hypograph of concave-extendable supermodular outer-functions over the polytope, although the separation problem is NP-Hard in general. As a consequence, we obtain large classes of inequalities that tighten prevalent factorable programming relaxations. The limiting composite relaxation obtained with infinitely many outer-approximators for each inner-function is shown to be related to the solution of an optimal transport problem. Moreover, composite relaxations can be seamlessly embedded into a discretization scheme to relax nonlinear programs with mixed-integer linear programs. Combined with linearization, composite relaxations provide a framework for deriving cutting planes used in relaxation hierarchies and more.

1. INTRODUCTION

A wide variety of problems involving managerial decisions, engineering design, and government policy in diverse domains such as energy, healthcare, and economics can be naturally formulated as mathematical optimization problems. Among these, a prominent and heavily researched class of problems is the mixed-integer nonlinear programs (MINLPs), a class of optimization problems that involve nonlinearities and discrete choices. Nonlinearities arise from models that capture the underlying science, human behavior, or the prevalent uncertainty. Discrete variables often model combinatorial choices, piecewise functions, or indivisibilities.

MINLPs are typically solved using branch-and-bound (B&B); an algorithm that, in the branching step, refines a partition of the variable domains and then, in the bounding step, chooses one partition element to construct a relaxation for the problem. A key determinant of the success of such method is their ability to construct tight and tractable relaxations. The predominant relaxation strategy used by most state-of-the-art solvers [1–4] is the factorable programming (FP) technique. This technique recursively traverses the expression tree for each nonlinear function and relaxes each operator over a bounding box that covers the ranges for all the operands. More specifically, FP treats each function as a recursive sum and/or product of univariate functions. The technique then relaxes bilinear terms over variable bounds using McCormick envelopes [5] and relaxes each univariate function using its function-specific structure over the range of the independent variable.

The following three features of the factorable relaxation scheme have resulted in its widespread adoption. First, as the partition size is refined, the relaxation converges asymptotically to the original function, a property needed for its successful use in a convergent B&B algorithms [6]. Second, the scheme imposes few restrictions on the types of functions that can be relaxed, besides boundedness, making it suitable for automatically relaxing large classes of MINLPs. Third, the number of variables introduced in the relaxation is in direct correspondence with the nonlinearities in the problem, and, as a result, the size of relaxations is not much larger than the original MINLP formulation.

Nevertheless, the primary deficiency of factorable programming is that it often produces weak relaxations [7, 8]. Significant research has been devoted to argument, rather than supplant, the basic factorable programming by exploiting *special* structure. For various types of multilinear functions [9–13], the fractional terms [14], and other useful functions [15–18], envelopes have been derived. Valid inequalities and cutting planes for multilinear functions have been proposed in [19–24]. Tighter convex relaxations for structured sets have been derived in [25–31].

Two *general-purpose* (structure-free) techniques are available to improve the FP relaxations. The first strategy is to construct relaxation hierarchies [32–34], doing so increases the relaxation size considerably because hierarchies introduce additional product operators and require variables to linearize the products. The second strategy is to discretize the box domain of each operator into a collection of sub-boxes and to outer-approximate the operator over each sub-box [35–37]. Then, the resulting relaxation is modeled as a union of polyhedra, and, thus, can be represented as a mixed integer linear program (MIP).

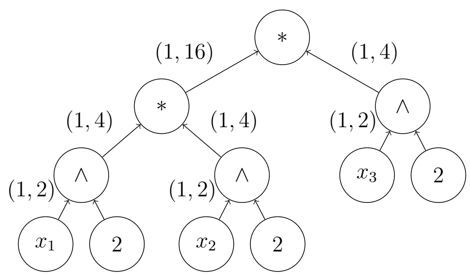
All above mentioned general-purpose techniques ignore operand structure while constructing relaxations for an operator. In this thesis, we provide a general-purpose relaxation framework that extracts operand structure to improve factorable programming relaxations, relaxation hierarchies, and MIP relaxations.

Example 1.0.1 *Consider a monomial* $x_1^2x_2^2x_3^2$ *over the box* $[1,2]^3$*, and consider an expression tree depicted in Figure 1.1a, whose edges are labeled with lower and upper bounds of tail nodes. Using information that the function* x_i^2 *is bounded from below (resp. above) by* 1 *(resp.* 4*) over* $[1,2]$*, factorable programming yields a convex underestimator for the root node's left child, that is,* $x_1^2x_2^2 \geq \max\{x_1^2+x_2^2-1, 4x_1^2+4x_2^2-16\}$*.*

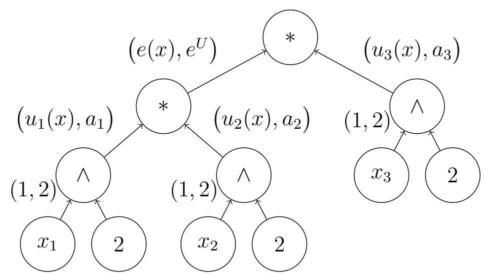
Then, ignoring descendant structure of the root node but bounds of its children, factorable programming constructs the following convex underestimator for the root node:

$$x_1^2x_2^2x_3^2 \geq \max\left\{\begin{array}{l}\max\{x_1^2+x_2^2-1, 4x_1^2+4x_2^2-16\}+x_3^2-1\\ 4\max\{x_1^2+x_2^2-1, 4x_1^2+4x_2^2-16\}+16x_3^2-64\end{array}\right\}.$$

However, exploiting a certain ordering relation between the right child of the root node and its underestimators, e.g., $x_1^2+x_2^2-1 \leq x_1^2x_2^2$ *and* $x_1^2+x_2^2-1 \leq 7$ *for all* $x \in [1,2]^3$*, this thesis produces a new convex underestimator* $16x_1^2+16x_2^2+7x_3^2-64$*, which is not majorized by the factorable one.*



(a) Bound propagation in expression tree



(b) Estimator propagation in expression tree

More generally, consider an expression tree depicted in Figure 1.1b, whose edges are labeled with underestimators of tail nodes and their upper bounds over the box domain $[1,2]^2$*, where for* $i \in \{1,2,3\}$*,* $u_i(x)$ *denotes* $(1, 2x_i-1, x_i^2)$ *and* a_i *denotes* $(1,3,4)$*. Exploiting a certain ordering relation in pairs* $\big(u_1(x), a_1\big)$ *and* $\big(u_2(x), a_2\big)$*, our relaxation techniques produce the following underestimators for the left child of the root node:*

$$x_1^2x_2^2 \geq \max\left\{\begin{array}{l} e_0(x) := 1\\ e_1(x) := u_{12}(x)+u_{22}(x)-1\\ e_2(x) := 2u_{11}(x)+u_{12}(x)+2u_{21}(x)+u_{22}(x)-9\\ e_3(x) := 3u_{11}(x)+u_{12}(x)+3u_{22}(x)-12\\ e_4(x) := 3u_{12}(x)+3u_{21}(x)+u_{22}(x)-12\\ e_5(x) := u_{11}(x)+3u_{12}(x)+u_{21}(x)+3u_{22}(x)-15\\ e_6(x) := \max\big\{e_0,\ldots,e_5, 4u_{12}(x)+4u_{22}(x)-16\big\}\end{array}\right\}.$$

Notice that, for $i \in \{0, \ldots, 6\}$ and for $x \in [1,2]^3$, we obtain that $e_i(x) \leq e_i^U$, where $e^U := (1, 7, 11, 13, 13, 15, 16)$. Using an ordering structure in pairs $\big(e(x), e^U\big)$ and $\big(u_3(x), a_3\big)$, this thesis generates fruitful convex underestimators for the root node:

$$x_1^2 x_2^2 x_3^3 \geq \max \left\{ \begin{array}{l} e_6(x) + u_{32}(x) - 1 \\ 2e_1(x) + e_6(x) + 6u_{31}(x) + u_{32}(x) - 21 \\ 3e_1(x) + e_6(x) + 7u_{32}(x) - 28 \\ e_2(x) + e_6(x) + 10u_{31}(x) + u_{32}(x) - 33 \\ \max\{e_3(x), e_4(x)\} + e_6(x) + 12u_{31}(x) + u_{32}(x) - 39 \\ e_1(x) + 2e_2(x) + e_6(x) + 4u_{31}(x) + 7u_{32}(x) - 40 \\ 3e_2(x) + e_6(x) + 11u_{32}(x) - 44 \\ 2e_5(x) + e_6(x) + 14u_{31}(x) + u_{32}(x) - 45 \\ e_1(x) + 2\max\{e_3(x), e_4(x)\} + e_6(x) + 6u_{31}(x) + 7u_{32}(x) - 46 \\ 3e_6(x) + 15u_{31}(x) + u_{32}(x) - 48 \\ e_2(x) + 2\max\{e_3(x), e_4(x)\} + e_6(x) + 2u_{31}(x) + 11u_{32}(x) - 50 \\ e_1(x) + 2e_5(x) + e_6(x) + 8u_{31}(x) + 7u_{32}(x) - 52 \\ 3\max\{e_3(x), e_4(x)\} + e_6(x) + 13u_{32}(x) - 52 \\ e_1(x) + 3e_6(x) + 9u_{31}(x) + 7u_{32}(x) - 55 \\ e_2(x) + e_5(x) + e_6(x) + 4u_{31}(x) + 11u_{32}(x) - 56 \\ \max\{e_3(x), e_4(x)\} + 2e_5(x) + e_6(x) + 2u_{31}(x) + 13u_{32}(x) - 58 \\ e_2(x) + 3e_6(x) + 5u_{31}(x) + 11u_{32}(x) - 59 \\ 3e_5(x) + e_6(x) + 15u_{32}(x) - 60 \\ \max\{e_3(x), e_4(x)\} + 3e_6(x) + 3u_{31}(x) + 13u_{32}(x) - 61 \\ e_5(x) + 3e_6(x) + u_{31}(x) + 15u_{32}(x) - 63 \\ 4e_6(x) + 16u_{32}(x) - 64 \end{array} \right\}. \tag{1.1}$$

*Next, consider the degree-*6 *RLT relaxation for the graph of* $x_1^2x_2^2x_3^2$ *over* $[1,2]^3$*, which is obtained by linearizing monomials in the following polynomial constraints with new variables:*

$$(x_1-1)^{\alpha_1}\cdots(x_3-1)^{\alpha_3}(2-x_1)^{\beta_1}\cdots(2-x_3)^{\beta_3}\geq 0 \quad \textit{for} \ \sum_{i=1}^{3}(\alpha_i+\beta_i)\leq 6.$$

*It can be verified that the degree-*6 *RLT does not implies some inequalities in (1.1), e.g.* $x_1^2x_2^2x_3^2 \geq 2e_1(x)+e_6(x)+6u_{31}(x)+u_{32}(x)-21$*.*

Now, suppose we discretize the range $[1,16]$ *of the root node's left child* $x_1^2x_2^2$ *along points* $\{1,13,16\}$*, and let*

$$\begin{aligned} G_1 &:= \bigl\{(x,\mu) \bigm| \mu \geq fx_3^2,\ f=x_1^2x_2^2,\ 1\leq f\leq 13,\ x\in[1,2]^3\bigr\},\\ G_2 &:= \bigl\{(x,\mu) \bigm| \mu \geq fx_3^2,\ f=x_1^2x_2^2,\ 13\leq f\leq 16,\ x\in[1,2]^3\bigr\}. \end{aligned}$$

The standard strategy constructs relaxation R_i *for* G_i *using factorable programming, where*

$$R_1 := \left\{ (x,\mu) \;\middle|\; \begin{array}{l} \mu \geq \max\bigl\{f+u_{32}(x)-1, 4f+13u_{32}(x)-52\bigr\},\ x\in[1,2]^3 \\ u_6(x)\leq f,\ 1\leq f\leq 13 \end{array} \right\},$$

$$R_2 := \left\{ (x,\mu) \;\middle|\; \begin{array}{l} \mu \geq \max\bigl\{f+13u_{32}(x)-13, 4f+16u_{32}(x)-64\bigr\},\ x\in[1,2]^3 \\ u_6(x)\leq f,\ 13\leq f\leq 16 \end{array} \right\}.$$

As discussed above, our approach, exploiting structure regarding underestimators, yields a relaxation for G_i *that is tighter than* R_i*.*

1.1 Outline and contributions

In Chapter 2, we introduce a new relaxation framework for MINLPs. These new relaxations, called *composite relaxations*, are tighter than factorable programming relaxations while using the same set of auxiliary variables. Consider the composite function $\phi\circ(f_1,\ldots,f_d)$ where, we refer to $\phi:\mathbb{R}^d\mapsto\mathbb{R}$ as the outer-function and $(f_1,\ldots,f_d)$ as the inner-functions. Our relaxation procedure, first, encodes the inner-functions $f(\cdot)$ into a polytope P in a generic fashion, and, then, relaxes the graph of the outer-function $\phi(\cdot)$ over this polytope P into a convex set. We derive the inequalities explicitly for the case where the outer-function is a bilinear term and there is only one non-trivial underestimator for each inner-function. Instead of the four inequalities that describe the McCormick envelopes, we obtain 12 non-redundant inequalities for the bilinear term. This result shows factorable programming relaxations can be easily tightened without introducing new variables. The main theoretical result in this chapter is the polynomial-time equivalence of separations between the convex hull of the outer-function $\phi(\cdot)$ over P and that over a structured subset Q of P. When d, the number of inner-functions, is constant and the convex hull of the graph of $\phi(\cdot)$ is determined by its value at extreme points of Q, we derive an LP formulation for the separation problem over Q, which is polynomially-sized in terms of the number of estimators due to a favorable vertex structure of Q. In contrast, a direct LP formulation for the separation problem over P is exponentially-sized in terms of the number of estimators.

The sole example of composite relaxations in Chapter 2 for which explicit inequalities are available concerns the product of two bounded functions each furnished with an underestimator. In Chapter 3, we describe ways in which we extend these results beyond the above example setting. First, we treat a larger class of outer-functions, in particular, those that are supermodular and concave-extendable over Q. Second, we allow arbitrarily many estimators for each inner-function. Though numerous, since these inequalities are generated using a fast combinatorial separation algorithm, they can be derived iteratively, with little computational overhead, to cut off infeasible regions from MINLP relaxations. Third, we extend our algorithm to allow simultaneous separation of a vector of composite functions, each with an outer-function that is supermodular and concave-extendable over Q. Fourth, we consider infinitely many estimators for each inner-function, assuming additionally that the outer-function is convex when all but one of its arguments are fixed. We show that, in this case, the composite relaxation arises as the solution of an optimal transport problem [38].

Composite relaxations not only improve factorable programming relaxations but also lay a foundation for improving relaxation hierarchies and MIP relaxations. In Chapter 4, we show that composite relaxations are not implied by the reformulation-linearization technique (RLT) [39]. Furthermore, combined with linearization, composite relaxations provide a framework for deriving cutting planes used in relaxation hierarchies and more. Last, we argue that the composite relaxations are particularly well-suited for constructing MIP relaxations via a discretization scheme. Our proposed MIP relaxations are obtained by reinterpreting the incremental formulation [35] and combining it with the composite relaxations, and are provable to be tighter than the standard ones in [36, 37]. For composite functions with discrete domains and univariate inner-functions, we obtain ideal logarithmic MIP formulations for their graphs. In contrast, the state-of-the-art formulation from [40] is not ideal. By exploiting our results for discrete domains, we show that, for certain compositions of univariate functions, we can construct a sequence of polyhedral relaxations that converge, in the limit, to the concave envelope.

2. A NEW FRAMEWORK TO RELAX COMPOSITE FUNCTIONS IN NONLINEAR PROGRAMS

In this chapter, we consider the composite function $\phi \circ (f_1, \ldots, f_d)$ where, we refer to $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ as the outer function and $(f_1, \ldots, f_d)$ as the inner functions. We outer-approximate the graph of f_i with n estimators and derive bounds for estimators. Then, we relax $\phi(\cdot)$ over a polytope, P, that models the ordering relationship between the functions, their estimators, and the bounds. Although P's structure is quite complex, it has as its subset a product of simplices, Q, which captures all the interesting structure of the convex hull of $\phi(\cdot)$ over P. We give a fast combinatorial algorithm, with complexity $O(dn)$ and a separation oracle call for the convex hull of $\phi(\cdot)$ over Q, that solves the separation problem for the graph of $\phi(\cdot)$ over P. Moreover, when d is constant, and the convex hull of the graph of $\phi(\cdot)$ is determined by its value at extreme points of Q, we derive an LP formulation for the separation problem, which is polynomially-sized in terms of the number of estimators.

We derive the inequalities explicitly for the case of a bilinear term, where $d = 2$ and there is only one non-trivial underestimator. Instead of the four inequalities that describe the McCormick envelopes, we obtain 12 non-redundant inequalities for the bilinear term. This result shows factorable programming relaxations can be easily tightened without introducing new variables. We also show that access to a separation oracle that generates facet-defining inequalities for the graph of $\phi(\cdot)$ over Q and its faces, which are affinely isomorphic to Q and obtained by ignoring some of the estimators, can be used to generate facet-defining inequalities that separate a point from the convex hull of graph of $\phi(\cdot)$ over P. Finally, we show that these results extend to simultaneous convexification of composite functions.

2.1 Improving factorable relaxations using outer-approximations

Factorable programming (FP) expresses each function as a recursive sum and product of constituent functions, where the key step involves relaxing a product of two functions. Consider a function $h : X \mapsto \mathbb{R}$, where X is a convex subset of $\mathbb{R}^m$, expressible as $h(x) = f_1(x)f_2(x)$ and consider its epigraph defined as $\text{epi}(h) := \big\{(x,\mu) \bigm| \mu \geq f_1(x)f_2(x), x \in X\big\}$. FP assumes that, for $i = 1, 2$, there are a convex function $c_{f_i}(x)$ and a concave function $C_{f_i}(x)$ and constants f_i^L and f_i^U so that, for $x \in X$, $f_i^L \leq f_i(x) \leq f_i^U$ and $c_{f_i}(x) \leq f_i(x) \leq C_{f_i}(x)$. Assume, without loss of generality, that $f_i^L \leq c_{f_i}(x)$ and $C_{f_i}(x) \leq f_i^U$. Then, FP relaxes the epigraph of $h(x)$, by introducing variables f_1 and f_2, as follows:

$$\left\{ (x, f, \mu) \;\middle|\; \begin{array}{l} \mu \geq \max\big\{f_1^L f_2 + f_1 f_2^L - f_1^L f_2^L,\ f_1^U f_2 + f_1 f_2^U - f_1^U f_2^U\big\} \\ c_{f_i}(x) \leq f_i \leq C_{f_i}(x),\ x \in X \end{array} \right\}.$$

The relaxation technique currently used in most global optimization solvers augments the above relaxation with results for specially structured problems; see [2–4, 41].

In this section, we improve FP relaxation of $f_1(x)f_2(x)$, when $f_i(x) \in [f_i^L, f_i^U]$ for all $x \in X$. We begin by assuming that $f_i^L \geq 0$. If the functions can be negative, $f_i(x) - f_i^L$ can be used instead. Information on $f_i(x)$ will be captured using underestimators $v_{ij}(x)$ and any upper bounds a_{ij} available for them.

Example 2.1.1 *Consider the function* $x_1^2x_2^2$ *over* $[0,2] \times [0,2]$ *and, for* $j \in J$, *let* $p(j) \in [0,2]$ *for* $j \in J$. *The epigraph of* x_i^2 *satisfies the tangent inequality* $x_i^2 \geq v_{ij} := p(j)^2 + 2p(j)\big(x_j - p(j)\big)$. *The inequality* $x_1^2x_2^2 \geq v_{1j_1}(x_1)v_{2j_2}(x_2)$ *does not hold, in general. For example,* $x_1^2 \geq 2x_1 - 1$ *and* $x_2^2 \geq 2x_2 - 1$ *but* $x_1^2x_2^2 \geq (2x_1-1)(2x_2-1)$ *is violated at* $(0,0)$. *Nevertheless, if* $u_{ij_i}(x) = \max\{0, v_{ij_i}(x)\}$. *then* $x_1^2x_2^2 \geq u_{1j_1}(x)u_{2j_2}(x)$. *From the definition of* $u_{ij_i}(x)$, $0 \leq u_{ij_i}(x) \leq p(j_i)\big(4 - p(j_i)\big) \leq 4$. *Let* $a_{i,j_i} = p(j_i)\big(4 - p(j_i)\big)$. *Then, McCormick inequalities can be used to underestimate* $u_{1j_1}(x)u_{2j_2}(x)$ *over* $\prod_{i=1,2}[0, a_{ij_i}]$ *to obtain* $x_1^2x_2^2 \geq u_{1j_1}(x)u_{2j_2}(x) \geq \max\big\{0, a_{1j_1}u_{2j_2}(x) + a_{2j_2}u_{1j_1}(x) - a_{1j_1}a_{2j_2}\big\} \geq a_{1j_1}v_{2j_2}(x) + a_{2j_2}v_{1j_1}(x) - a_{1j_1}a_{2j_2}$. *For* $p(j_1) = p(j_2) = 1$, *this reduces to* $x_1^2x_2^2 \geq 6x_1 + 6x_2 - 15$. *It can be easily verified that* $x_1 = 1.5$, $x_2 = 1.5$ *and* $\mu = 2$

satisfies the factorable relaxation, $\{(x, y, \mu) \mid 0 \leq \mu,\ 4x^2 + 4y^2 - 16 \leq \mu,\ 0 \leq x \leq 2,\ 0 \leq y \leq 2\}$*, but not* $\mu \geq 6x_1 + 6x_2 - 15$*.*

Example 2.1.1 used a *two-step procedure*, where we first relaxed x_i^2 using its lower bound and a tangent inequality. Then, in the second step, we relaxed the product of underestimators using McCormick envelopes. We concluded that the resulting inequality is not implied by using McCormick envelopes on the original product $x_1^2 x_2^2$, the *one-step procedure* typically used in factorable relaxations. This two-step procedure improves the quality of relaxation when the outer-function is monotonic, which is guaranteed in Example 2.1.1 by non-negativity of inner functions. Now, we consider a generalization that does not require the monotonicity assumption. We will consider a special case first, which already improves the factorable programming scheme. More specifically, we consider the bilinear term, $f_1 f_2$, in the presence of non-trivial underestimator $u_i(x)$ for $f_i(x)$, for $i = 1, 2$.

Theorem 2.1.1 *Let* $f_1^L \leq a_1 \leq f_2^U$ *and* $f_2^L \leq a_2 \leq f_2^U$*. Then, consider the set:*

$$P = \left\{(u, f) \;\middle|\; f_1^L \leq u_1 \leq \min\{f_1, a_1\}, f_1 \leq f_1^U, f_2^L \leq u_2 \leq \min\{f_2, a_2\}, f_2 \leq f_2^U\right\}.$$

The following linear inequalities are valid for the epigraph of $f_1 f_2$ *over* P*:*

$$f_1 f_2 \geq \max \left\{ \begin{aligned} & e_1 := f_1 f_2^U + f_2 f_1^U - f_1^U f_2^U \\ & e_2 := (f_2^U - a_2)u_1 + (f_1^U - a_1)u_2 + a_2 f_1 + a_1 f_2 + a_1 a_2 - a_1 f_2^U - f_1^U a_2 \\ & e_3 := (f_2^U - f_2^L)u_1 + f_2^L f_1 + a_1 f_2 - a_1 f_2^U \\ & e_4 := (f_1^U - f_1^L)u_2 + a_2 f_1 + f_1^L f_2 - f_1^U a_2 \\ & e_5 := (a_2 - f_2^L)u_1 + (a_1 - f_1^L)u_2 + f_2^L f_1 + f_1^L f_2 - a_1 a_2 \\ & e_6 := f_1^L f_2 + f_1 f_2^L - f_1^L f_2^L \end{aligned} \right\}.$$

Proof We show that e_3, e_4, and e_5 are valid underestimators for $f_1 f_2$ over P using the procedure in Example 2.1.1. That e_3 underestimates $f_1 f_2$ follows from:

$$
\begin{aligned}
f_1 f_2 &= (f_1 - f_1^L)(f_2 - f_2^L) + f_1^L f_2 + f_2^L f_1 - f_1^L f_2^L \\
&\geq (u_1 - f_1^L)(f_2 - f_2^L) + f_1^L f_2 + f_2^L f_1 - f_1^L f_2^L \\
&\geq (u_1 - f_1^L)(f_2^U - f_2^L) + (a_1 - f_1^L)(f_2 - f_2^L) - (a_1 - f_1^L)(f_2^U - f_2^L) \\
&\qquad\quad + f_2^L f_1 + f_1^L f_2 - f_1^L f_2^L \\
&= e_3,
\end{aligned}
$$

where the first equality shifts f_1 and f_2 so that we relax a product of non-negative functions, the first inequality underestimates the product $(f_1 - f_1^L)(f_2 - f_2^L)$, and the second inequality uses the McCormick relaxation for $(u_1 - f_1^L)(f_2 - f_2^L)$. The derivation for e_4 is symmetric, with the role of f_1 and f_2 is interchanged. Similarly, to show that e_5 is an underestimator:

$$
\begin{aligned}
f_1 f_2 &= (f_1 - f_1^L)(f_2 - f_2^L) + f_1^L f_2 + f_2^L f_1 - f_1^L f_2^L \\
&\geq (u_1 - f_1^L)(u_2 - f_2^L) + f_1^L f_2 + f_2^L f_1 - f_1^L f_2^L \\
&\geq (u_1 - f_1^L)(a_2 - f_2^L) + (a_1 - f_1^L)(u_2 - f_2^L) - (a_1 - f_1^L)(a_2 - f_2^L) \\
&\qquad\quad + f_2^L f_1 + f_1^L f_2 - f_1^L f_2^L \\
&= e_5.
\end{aligned}
$$

Observe that e_1 and e_6 are derived by using the functions f_1 and f_2 themselves as their underestimators.

To show the second inequality is valid, define $s_2 = \max\big\{u_2, f_2^L + \frac{a_2 - f_2^L}{f_2^U - f_2^L}(f_2 - f_2^L)\big\}$ and note that $a_2 - s_2 + f_2 - f_2^U = \frac{f^2 - f_2^U}{f_2^U - f_2^L}(f_2^U - a_2) \leq 0$. Then, $f_1 f_2 = e_2 + (s_2 - u_2)(f_1^U - a_1) + (u_1 - a_1)(a_2 - s_2 + f_2 - f_2^U) + (f_1 - u_1)(f_2 - s_2) + (f_1 - f_1^U)(s_2 - a_2) \geq e_2$, completing the proof. ∎

If $a_i \in \{f_i^L, f_i^U\}$ for $i \in \{1, 2\}$, the factorable relaxation is the convex hull of the epigraph of $f_1 f_2$ over P. Thus, to improve the factorable relaxation using Theorem 2.1.1, $a_i < f_i^U$ for at least some i. We show in [42] that the inequalities in

Theorem 2.1.1 describe the convex envelope of $f_1 f_2$ over P. Observe that the inequalities in Theorem 2.1.1 share many properties of the FP relaxations. First, they apply to all factorable programming problems. Second, the coefficients of u_i in the inequalities are non-negative. Therefore, u_i can be substituted with any convex underestimator of their defining relation, $u_i(x)$, to yield convex inequalities. Third, if there are n_i estimators of f_1, Theorem 2.1.1 yields $2n_1n_2 + n_1 + n_2 + 2$ inequalities underestimating $f_1 f_2$ instead of the two inequalities typically used in FP.

The underestimator e_2 is not obtained using the two-step procedure in Example 2.1.1. In the next example, we demonstrate that underestimator e_2 is not dominated by other inequalities.

Example 2.1.2 *Consider again the monomial $x_1^2 x_2^2$ over $[0,2]^2$. Then, let $u_i = \max\{0, 2x_i - 1\}$. Then, Theorem 2.1.1 yields the following relaxation, after substitutions:*

$$\left\{ (x,y,\mu) \;\middle|\; \begin{array}{l} \mu \geq e_1 = 4x_1^2 + 4x_2^2 - 16,\ \mu \geq e_2 = 2x_1 + 2x_2 + 3x_1^2 + 3x_2^2 - 17 \\ \mu \geq e_3 = 8x_1 + 3x_2^2 - 16,\ \mu \geq e_4 = 3x_1^2 + 8x_2 - 16 \\ \mu \geq e_5 = 6x_1 + 6x_2 - 15,\ \mu \geq e_6 = 0 \\ x \in [0,2]^2 \end{array} \right\}.$$

Observe that the inequalities $\mu \geq e_3$, $\mu \geq e_4$, and $\mu \geq e_5$ can be obtained by first underestimating x_1^2 and x_2^2 and then using the McCormick inequalities. However, the inequality $\mu \geq e_2$ is not obtained using the two-step procedure and is not redundant. For example, at $(x_1, x_2) = (1.6, 1.6)$, the highest underestimator is e_2 which equals 4.76, *while the remaining underestimators are below* 4.5.

We remark that we use x^2y^2 over $[0,2]^2$ only to illustrate our techniques and that for x^2y^2 over $[0,2]^2$, an explicit formulation of the convex hull is available [43]. Existence of e_2 shows that the epigraph of $f_1 f_2$ over P satisfies inequalities besides those obtained using the two-step procedure. In the rest of this chapter, we will study a significant generalization of P introduced above by including one or more underestimating and/or overestimating inequalities.

2.2 A relaxation framework for composite functions

In this section, we introduce the generalized version of the setup in Theorem 2.1.1, which will be subject of study for most of the remaining chapter. In the following, we shall denote the convex hull of set S by $\mathrm{conv}(S)$, the convex (resp. concave) envelope of $f(x)$ over S by $\mathrm{conv}_S(f)$ (resp. $\mathrm{conc}_S(f)$), the projection of a set S to the space of x variables by $\mathrm{proj}_x(S)$, the extreme points of S by $\mathrm{vert}(S)$, the dimension of the affine hull of S by $\dim(S)$, and the relative interior of S by $\mathrm{ri}(S)$.

Our setup will generalize that of Theorem 2.1.1 in the following way. First, we replace the bilinear term with an arbitrary function, ϕ, and consider relaxations of $\phi \circ f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$, where $f : X \mapsto \mathbb{R}^d$ is a vector of *bounded* functions over X. We shall write $f(x) := \big(f_1(x), \ldots, f_d(x)\big)$ and refer to f as inner functions while ϕ will be referred to as the outer function. Second, we will be derive convex relaxations for the graph of $\phi \circ f$ (instead of just the epigraph). Formally, we will relax the set: $\mathrm{gr}(\phi \circ f) = \big\{(x, \phi) \bigm| \phi = \phi(f(x)),\ x \in X\big\}$. We describe the generalization of polytope P in Section 2.2.1.

2.2.1 Polyhedral abstraction of outer-approximation

Let $\big(n(1), \ldots, n(d)\big) \in \mathbb{Z}^d$. We consider a vector of bounded functions $u : \mathbb{R}^m \mapsto \mathbb{R}^{\sum_{i=1}^d \left(n(i)+1\right)}$ defined as $u(x) = (u_1(x), \ldots, u_d(x))$, where $u_i(x) : \mathbb{R}^m \mapsto \mathbb{R}^{n(i)+1}$, and a vector $a = (a_1, \ldots, a_d) \in \mathbb{R}^{\sum_{i=1}^d \left(n(i)+1\right)}$, where $a_i \in \mathbb{R}^{n(i)+1}$. For all i and for every $x \in X$, assume that $f_i(x) \in [a_{i0}, a_{in_i}]$ and the pair $\big(u(x), a\big)$ satisfies the following inequalities:

$$
\begin{aligned}
& a_i^L \leq a_{i0} \leq \cdots a_{in(i)} \leq a_i^U, \\
& \text{for each } j \in A_i : \quad a_i^L \leq u_{ij}(x) \leq \min\{f_i(x), a_{ij}\}, \\
& \text{for each } j \in B_i : \quad \max\big\{f_i(x), a_{ij}\big\} \leq u_{ij}(x) \leq a_i^U, \\
& u_{i0}(x) = a_{i0}, \quad u_{in(i)}(x) = f_i(x),
\end{aligned}
\tag{2.1}
$$

where the pair (A_i, B_i) is a partition of $\{0, \ldots, n(i)\}$ so that $\{0, n(i)\} \subseteq A_i$. The first requirement that the elements of a_i are ordered in a non-decreasing order and contained within $[a_i^L, a_i^U]$ is only for notational convenience. The second (resp. third) requirement states that for $j \in A_i$ (resp $j \in B_i$), $u_{ij}(x)$ is an underestimator (resp. overestimator) for $f_i(x)$, which is bounded from above (resp. below) by a_{ij}, a value no larger (resp. smaller) than $a_{in(i)}$ (resp. a_{i0}), the upper bound (resp. lower bound) for $f_i(x)$. Finally, the fourth requirement states that $u_{i0}(x)$ is a constant function matching the lower bound on $f_i(x)$ and $u_{in(i)}$ is the function $f_i(x)$ itself.

Notice that we do not explicitly specify different lower bounds for the underestimators and upper bounds for the overestimators. This is because they are not important in constructing the relaxations and do not change the quality of the relaxation. We will show in Proposition 2.2.1 that, without loss of generality, a_i^L and a_i^U can be set to be the lower bound a_{i0} for $f_i(x)$ and upper bound $a_{in(i)}$ for $f_i(x)$ respectively. We remark that, for $j \in A_i$ (resp. $j \in B_i$), a_{ij} need not be the tightest upper (resp. lower) bound of $u_{ij}(x)$ over X. Tighter bounds for estimators and/or inner-functions can be obtained using the techniques of discussed in [44, 45]. Last, although the number of estimators for each function can be different, we will assume without loss of generality and for notational simplicity that $n(1) = \cdots = n(d)$.

Now, we describe our generalization of P formally. The polytope P is denoted, in general, as $P(a, a^L, a^U, B) := \prod_{i=1}^d P_i(a_i, a_i^L, a_i^U, B_i)$, where $B = \prod_{i=1}^d B_i$, $a^L = (a_1^L, \ldots, a_d^L)$ and $a^U = (a_1^U, \ldots, a_d^U)$

$$P_i(a_i, a_i^L, a_i^U, B_i) := \left\{ u_i \;\middle|\; \begin{array}{l} \text{for each } j \in A_i : u_{ij} \le u_{in} \text{ and } a_i^L \le u_{ij} \le a_{ij} \\ \text{for each } j \in B_i : u_{in} \le u_{ij} \text{ and } a_{ij} \le u_{ij} \le a_i^U \\ u_{i0} = a_{i0},\ a_{i0} \le u_{in} \le a_{in} \end{array} \right\}. \tag{2.2}$$

We will typically not write the arguments of P_i and P since they will be apparent from the context. We will refer to the polytope P_i as the *abstraction* of outer-approximators of function f_i and P as the *abstraction* of outer-approximators for vector of functions f. Essentially, the polytope P_i is obtained by introducing a variable u_{ij} for the estimator $u_{ij}(x)$ and replacing the ordering relationships between the functions with those between the introduced variables.

We remark that our construction treats estimators of inner functions abstractly, and exploits various bounding relationships while being oblivious of the precise dependence on x. Thus, our methods apply to general nonlinear programming problems while providing flexibility, which can be tailored to exploit specific problem structure, *e.g.* properties of the outer function $\phi(\cdot)$. The rest of the chapter will focus on studying the graph, Φ^P, of $\phi(\cdot)$ over the polytope P, and its convex hull, where:

$$\Phi^P = \big\{(u, \phi) \bigm| \phi = \phi(u_{1n}, \ldots, u_{dn}),\ u \in P\big\}.$$

2.2.2 An overview of the main polynomial equivalence result

Since P generalizes the standard hypercube $[0, 1]^d$, by considering the special case where ϕ is a bilinear function, it follows that there does not exist a polynomial time algorithm to solve the separation problem for Φ^P, unless P = NP. But, Theorem 2.1.1 shows that special instances of this problem are solvable, sometimes in closed-form, which improve relaxations for nonlinear programs. We now describe the main structural insights, we derive for Φ^P, in this chapter. We will show that convexifying Φ^P is polynomially equivalent to convexifying a simpler object Φ^Q, the graph of ϕ over a subset of P, which is termed Q here onwards and described formally in Section 2.3. We remark that this equivalence is not just a mapping of hard instances of Φ^P to hard instances of Φ^Q. Clearly, existence of such a mapping follows from NP-Hardness of these problems. Rather, this equivalence will provide an algorithm to convexify Φ^P using an oracle to convexify a specific related instance of Φ^Q. In [42], we will show that there is a large family of tractable instances for Φ^Q, which will then, by our results here, yield a tractable separation algorithm for the corresponding family of Φ^P.

To show this polynomial equivalence, we devise a separation oracle for $\text{conv}(\Phi^P)$ by augmenting the separation oracle for $\text{conv}(\Phi^Q)$ with a fast-polynomial-time combinatorial algorithm. The key ingredient of the combinatorial algorithm is a lifting procedure that lifts Φ^P into a higher dimensional space, a problem, we show, is equivalent to solving d two-dimensional convexification problems. A direct consequence of our results will be that the problem of convexifying Φ^P, when its extreme points project to the extreme points of P, is polynomially solvable in n, when the dimension d is fixed. This result is interesting for practice because compositions often involve only a few functions or can be recursively decomposed as such while each inner function has many outer-approximators. Instead, a direct formulation of the convex hull of Φ^P would be exponential because, even for a fixed d, the number of extreme points of P is exponential in n.

The resulting algorithm has many interesting features besides tractability. Consider the family of polytopes of the form Q that result when subsets of the outer-approximators of $f_i(x)$ are considered. Assume, for the purpose of illustration, that ϕ is a multilinear function, which would imply that Φ^P and Φ^Q are polyhedral sets [9]. If for each Q in this family, we have access to a separation oracle that separates from Φ^Q by generating a facet-defining inequality, the combinatorial algorithm can be used to devise an algorithm that separates points from Φ^P by facet-defining inequalities. Moreover, the polynomial equivalence carries over to the problem of simultaneously convexifying a collection of functions and, so does the property of generating facet-defining inequalities.

2.2.3 Projecting out introduced estimator variables

The convex hull of Φ^P is the intersection of the epigraph of the convex envelope and hypograph of concave envelope of $\phi(u_{1n}, \dots, u_{dn})$ over P, that is,

$$\operatorname{conv}(\Phi^P) = \Big\{(u, \phi) \;\Big|\; \operatorname{conv}_P(\phi)(u) \le \phi \le \operatorname{conc}_P(\phi)(u),\; u \in P\Big\}. \tag{2.3}$$

Observe that although the function ϕ depends only on $(u_{1n}, \dots, u_{dn})$, the functions $\operatorname{conv}_P(\phi)$ and $\operatorname{conc}_P(\phi)$ depend on all the variables u. In the following lemma, we consider the concave envelope $\operatorname{conc}_P(\phi)(u)$ and establish certain monotonicity properties for it. A similar property, albeit reversed in direction, can be obtained for $\operatorname{conv}_P(\phi)(u)$ since $\operatorname{conv}_P(\phi)(u) = -\operatorname{conc}_P(-\phi)(u)$.

Lemma 2.2.1 *If the concave envelope* $\text{conc}_P(\phi)(u)$ *is closed then it is non-increasing in* u_{ij} *for all* i *and* $j \in A_i \setminus \{0, n\}$ *and non-decreasing in* u_{ij} *for all* i *and* $j \in B_i$.

Proof We will only prove that $\text{conc}_P(\phi)(u)$ is non-increasing in u_{ij} for all i and $j \in A_i \setminus \{0, n\}$ since a similar argument shows that it is non-decreasing in u_{ij} for all i and $j \in B_i$. Let $\phi \leq \langle \alpha, u \rangle + b$ be a valid inequality of $\text{conc}_P(\phi)(u)$. Let $J_i := \{j' \in A_i \setminus \{0, n\} \mid \alpha_{ij'} > 0\}$. By considering $(\tilde{\alpha}, b')$, where $\tilde{\alpha}_{ij} = 0$ for all i and $j \in J_i$, $\tilde{\alpha}_{ij} = \alpha_{ij}$ otherwise, and $b' = b + \sum_{i=1}^{d} \sum_{j \in J_i} \alpha_{ij} a_i^L$, it is easy to construct a valid inequality $\phi \leq \langle \tilde{\alpha}, u \rangle + b'$ of $\text{conc}_P(\phi)(u)$ such that $\tilde{\alpha}_{ij} \leq 0$ for all i and $j \in A_i \setminus \{0, n\}$, and $\langle \tilde{\alpha}, u \rangle + b' \leq \langle \alpha, u \rangle + b$ for every $u \in P$.

Now assume that $\text{conc}_P(\phi)(u)$ is closed. For any function $\xi(\alpha)$, we will denote its convex conjugate by $\xi^*(\alpha)$. We prove that $\text{conc}_P(\phi)(u) = \psi(u)$, where $\psi(u) := \inf_\alpha \big\{ \langle \alpha, u \rangle + (-\phi)_P^*(-\alpha) \;\big|\; \alpha_{ij} \leq 0 \; \forall i \in \{1, \ldots, d\} \; \forall j \in A_i \setminus \{0, n\} \big\}$. This will show what we seek to prove since, by definition, $\psi(u)$ is non-increasing in u_{ij} for all i and $j \in A_i \setminus \{0, n\}$, being the infimum over α of linear functions $\langle \alpha, u \rangle + (-\phi)_P^*(-\alpha)$, all of which satisfy this property. Since $\text{conc}_P(\phi)(u)$ is assumed to be closed, by Theorem 1.3.5 in [46], we have

$$\text{conc}_P(\phi)(u) = \inf_\alpha \big\{ \langle \alpha, u \rangle + (-\phi)_P^*(-\alpha) \big\}. \tag{2.4}$$

It follows readily that $\psi(u) \geq \text{conc}_P(\phi)(u)$ because $\psi(u) \geq \inf_\alpha \big\{ \langle \alpha, u \rangle + (-\phi)_P^*(-\alpha) \big\}$. To show $\text{conc}_P(\phi)(u) \geq \psi(u)$, we consider a point $\bar{u} \in P$. By (2.4), there exists a sequence $\bar{\alpha}^k$ so that the inequality $\text{conc}_P(\phi)(u) \leq \langle \bar{\alpha}^k, u \rangle + (-\phi)_P^*(-\bar{\alpha}^k)$ is valid for all k and $\lim_{k \to \infty} \langle \bar{\alpha}^k, \bar{u} \rangle + (-\phi)_P^*(-\bar{\alpha}^k) = \text{conv}_P(\phi)(\bar{u})$. If $\bar{\alpha}_{ij}^k > 0$ for some i and $j \in A_i \setminus \{0, n\}$, we have shown that there exists a valid inequality $\phi \leq \langle \tilde{\alpha}^k, u \rangle + b'$ of $\text{conc}_P(\phi)(u)$ such that $\langle \tilde{\alpha}^k, u \rangle + b' \leq \langle \bar{\alpha}^k, u \rangle + (-\phi)_P^*(-\bar{\alpha}^k)$ for all $u \in P$ and $\tilde{\alpha}_{ij}^k \leq 0$ for all i and $j \in A_i \setminus \{0, n\}$. Therefore, we have

$$\begin{aligned} \text{conc}_P(\phi)(\bar{u}) &= \lim_{k \to \infty} \langle \bar{\alpha}^k, \bar{u} \rangle + (-\phi)_P^*(-\bar{\alpha}^k) \geq \lim_{k \to \infty} \langle \tilde{\alpha}^k, \bar{u} \rangle + b' \\ &\geq \lim_{k \to \infty} \langle \tilde{\alpha}^k, \bar{u} \rangle + (-\phi)_P^*(-\tilde{\alpha}^k) \geq \psi(\bar{u}), \end{aligned}$$

where first equality and first inequality are established above, second inequality follows because the validity of $\phi \leq \langle \tilde{\alpha}^k, u \rangle + b'$ for $u \in P$ implies $(-\phi)_P^*(-\tilde{\alpha}^k) =$

$\sup_{u\in P}\{\phi(u) - \langle\tilde{\alpha}^k, u\rangle\} \le b'$, and the last inequality holds since $\tilde{\alpha}^k \le 0$ for all i and $j \in A_i \setminus \{0, n\}$ implies $\tilde{\alpha}$ is feasible in the optimization problem defining $\psi(\bar{u})$. Thus, the proof is complete. ■

We now use monotonicity of $\mathrm{conc}_P(\phi)(u)$ and $\mathrm{conv}_P(\phi)(u)$ to construct relaxations for $\mathrm{gr}(\phi\circ f)$ in the space of $(x, \phi, u_{\cdot n})$ variables, where $u_{\cdot n} = (u_{1n}, \dots, u_{dn})$, by substituting the convex underestimators and concave overestimators with their defining relationships in the cuts valid for $\mathrm{conv}(\Phi^P)$.

Theorem 2.2.1 *Let $\phi\circ f$ be a composite function, where $\phi: \mathbb{R}^d \mapsto \mathbb{R}$ is a continuous function and $f: \mathbb{R}^m \mapsto \mathbb{R}^d$ is a vector of functions which are bounded over $X \subseteq \mathbb{R}^m$. Given a pair $\big(a, u(x)\big)$ satisfying (2.1), we have $\mathrm{gr}(\phi\circ f) \subseteq \mathrm{proj}_{(x,\phi)}(R)$, where*

$$R := \Big\{(x, u_{\cdot n}, \phi) \;\Big|\; \big(\tilde{u}_1(x, u_{1n}), \dots, \tilde{u}_d(x, u_{dn}), \phi\big) \in \mathrm{conv}(\Phi^P),\; u_{\cdot n} = f(x),\; x \in X\Big\},$$

and $\tilde{u}_i(x, u_{in}) = \big(u_{i0}(x), \dots, u_{i(n-1)}(x), u_{in}\big)$. The relaxation is convex if $u_{ij}(x)$ is convex for $j \in A_i \setminus \{n\}$ and concave for $j \in B_i$ and $\{u_{\cdot n} = f(x),\; x \in X\}$ is outer-approximated by a convex set.

Proof Define $\tilde{H} := \big\{(x, u, \phi) \;\big|\; \mathrm{conv}_P(\phi)(u) \le \phi \le \mathrm{conc}_P(\phi)(u), u_{\cdot n} = f(x), (x, u) \in U\big\}$, where $U := \big\{(x, u) \;\big|\; u_{i0} = a_{i0},\; u_{ij}(x) \le u_{ij}\; \forall j \in A_{ij} \setminus \{n\},\; u_{ij} \le u_{ij}(x)\; \forall j \in B_{ij}, x \in X\big\}$. We first show that $R = \mathrm{proj}_{(x,u_{\cdot n},\phi)}(\tilde{H})$. Clearly, $R \subseteq \mathrm{proj}_{(x,u_{\cdot n},\phi)}(\tilde{H})$ because, given $(x, u_{\cdot n}, \phi) \in R$, we may define $u_i := \tilde{u}_i(x, u_{in})$ for all i and obtain $(x, u, \phi) \in \tilde{H}$. We show $\mathrm{proj}_{(x,u_{\cdot n},\phi)}(\tilde{H}) \subseteq R$ by considering a point $(x, u, \phi) \in \tilde{H}$ and proving that $(x, u_{\cdot n}, \phi) \in R$. Clearly, $u_{\cdot n} = f(x)$ and $x \in X$. Since $\phi(u_{1n}, \dots, u_{dn})$ is continuous over P, we have

$$\phi \le \mathrm{conc}_P(\phi)(u_1, \dots, u_d) \le \mathrm{conc}_P(\phi)\big(\tilde{u}_1(x, u_{1n}), \dots, \tilde{u}_d(x, u_{dn})\big),$$

where the second inequality holds because of the monotonicity of $\mathrm{conc}_P(\phi)(u)$ shown in Lemma 2.2.1, and $u_{ij}(x) \le u_{ij}$ for $j \in A_{ij} \setminus \{n\}$, $u_{ij} \le u_{ij}(x)$ for $j \in B_{ij}$, and because the last component of $\tilde{u}_i(x, u_{in})$ equals u_{in}. Similarly, we obtain that

$$\mathrm{conv}_P\big(\tilde{u}_1(x, u_{1n}), \dots, \tilde{u}_d(x, u_{dn})\big) \le \phi.$$

Then, by (2.3) we conclude that $\big(\tilde{u}_1(x, u_{1n}), \ldots, \tilde{u}_d(x, u_{dn}), \phi\big) \in \operatorname{conv}(\Phi^P)$. Hence, $(x, u_{\cdot n}, \phi) \in R$.

Next, we show that $\operatorname{gr}(\phi \circ f) \subseteq \operatorname{proj}_{(x,\phi)}(\tilde{H})$, which implies that $\operatorname{gr}(\phi \circ f) \subseteq \operatorname{proj}_{(x,\phi)}(R)$. Let $(x, \phi) \in \operatorname{gr}(\phi \circ f)$ and define $u = u(x)$, where, in particular, $u_{\cdot n} = f(x)$. It follows readily that $u \in P$ because, by construction, we have $\{u \mid u = u(x), x \in X\} \subseteq P$. Moreover, since $(x, \phi) \in \operatorname{gr}(\phi \circ f)$, we have $\phi = \phi\big(f(x)\big) = \phi(u_{\cdot n})$, which implies that $(u, \phi) \in \Phi^P$ and, thus, $(u, \phi) \in \operatorname{conv}(\Phi^P)$. In other words, $\operatorname{conv}_P(\phi)(u) \leq \phi \leq \operatorname{conc}_P(\phi)(u)$. Therefore, $(x, u, \phi) \in \tilde{H}$ and, thus, $\operatorname{gr}(\phi \circ f) \subseteq \operatorname{proj}_{(x,\phi)} \tilde{H}$.

Last, assume that $\{u_{\cdot n} = f(x),\ x \in X\}$ is outer-approximated by a convex set, and $u_{ij}(x)$ is either a convex underestimator or concave overestimator for all i and $j < n$. Then, $\tilde{H}$ is convex, and therefore, R is convex being a projection of $\tilde{H}$. ■

2.2.4 Abstraction simplification

We first simplify $P(a, a^L, a^U, B)$ to show that it suffices to consider the case where $a_i^L = a_{i0}$ and $a_i^U = a_{in}$ to treat the general case.

Proposition 2.2.1 *Define u' so that $u'_{ij} = \min\big\{\max\{u_{ij}, a_{i0}\}, a_{in}\big\}$. Then, we have*

$$\operatorname{conc}_P(\phi)(u') = \operatorname{conc}_P(\phi)(u).$$

Consider u'' such that $a_{i0} \leq u''_{ij} \leq a_{in}$. Then, for any $a^L, a^U, \bar{a}^L, \bar{a}^U$ such that $a_i^L, \bar{a}_i^L \leq a_{i0}$ and $a_i^U, \bar{a}_i^U \geq a_{in}$ we have

$$\operatorname{conc}_{P(a,a^L,a^U,B)}(\phi)(u'') = \operatorname{conc}_{P(a,\bar{a}^L,\bar{a}^U,B)}(\phi)(u'').$$

Define $\tilde{a}_i^L = a_{i0}$ and $\tilde{a}_i^U = a_{in}$. Then, for any $u \in P(a, a^L, a^U, B)$ and $\phi \in \mathbb{R}$, $(u, \phi) \in \operatorname{conv}\big(\Phi^{P(a,a^L,a^U,B)}\big)$ if and only if $(u', \phi) \in \operatorname{conv}\big(\Phi^{P(a,\tilde{a}^L,\tilde{a}^U,B)}\big)$.

Proof For any $j \in B_i$, $u_{ij} \geq u_{in} \geq a_{i0}$. Therefore, $u'_{ij} \leq u_{ij}$. Similarly, for $j \in A_i$, $\max\{u_{ij}, a_{i0}\} \leq u_{in} \leq a_{in}$. Therefore, $u'_{ij} \geq u_{ij}$. Let $P = P(a, a^L, a^U, B)$. It follows from Lemma 2.2.1 that $\operatorname{conc}_P(\phi)(u') \leq \operatorname{conc}_P(\phi)(u)$. Now, we show that

$\text{conc}_P(\phi)(u') \geq \text{conc}_P(\phi)(u)$. Let $J_i(u) = \{j \mid u_{ij} < a_{i0}\}$ and $K_i(u) = \{j \mid u_{ij} > a_{in}\}$. We perform induction on $\sum_{i=1}^d \bigl(|J_i(u)| + |K_i(u)|\bigr)$. The base case is trivial because $|J_i(u)| + |K_i(u)| = 0$ implies $u' = u$ and the inequality is trivially satisfied. Let i' be such that there exists a $j' \in J_{i'}(u) \cup K_{i'}(u)$. Since $a_{i'0} \leq u_{i'n} \leq a_{i'n}$, it follows that $j' \neq n$. We assume $j' \in J_{i'}(u)$ as a similar argument applies when $j' \in K_{i'}(u)$. Since $u \in P$, by the definition of $\text{conc}_P(\phi)$, there exist convex multipliers λ_k and $(u^k, \phi^k) \in \Phi^P$ such that $\bigl(u, \text{conc}_P(\phi)(u)\bigr) = \sum_k \lambda^k (u^k, \phi^k)$. Define $\bar{u}^k_{i'j'} = a_{i'0}$ and $\bar{u}^k_{ij} = u^k_{ij}$ otherwise. Since $j' \neq n$, it can be verified easily that $(\bar{u}^k, \phi^k) \in \Phi^P$. Then, define $\bigl(\tilde{u}, \text{conc}_P(\phi)(u)\bigr) = \sum_k \lambda^k(\bar{u}^k, \phi^k)$ and observe that the representation shows that $\text{conc}_P(\phi)(\tilde{u}) \geq \text{conc}_P(\phi)(u)$, $\tilde{u}_{i'j'} = a_{i'0}$, and $\tilde{u}_{ij} = u_{ij}$ otherwise. However, since $j' \notin J_{i'}(\tilde{u})$, it follows that $\sum_{i=1}^d \bigl(|J_i(\tilde{u})| + |K_i(\tilde{u})|\bigr) = \sum_{i=1}^d \bigl(|J_i(u)| + |K_i(u)|\bigr) - 1$. Therefore, it follows that $\text{conc}_P(\phi)(u) \leq \text{conc}_P(\phi)(\tilde{u}) \leq \text{conc}_P(\phi)(u')$, where the last inequality is by the induction hypothesis, proving the second statement in the result.

Let $P^1 := P(a, a^L, a^U, B)$ and $P^2 = P(a, \bar{a}^L, \bar{a}^U, B)$. We show that $\text{conc}_{P^1}(\phi)(u'') \leq \text{conc}_{P^2}(\phi)(u'')$ since the result follows by interchanging the roles of P^1 and P^2. Let $\phi'' = \text{conc}_{P^1}(\phi)(u'')$. Then, there exist $(u^l, \phi^l) \in \Phi^{P^1}$ and convex multipliers γ_l so that $(u'', \phi'') = \sum_l \gamma_l (u^l, \phi^l)$. Let $\tilde{u}^l_{ij} = \min\bigl\{\max\{u^l_{ij}, a_{i0}\}, a_{in}\bigr\}$. Then, it follows that $(\tilde{u}^l, \phi^l) \in \Phi^{P_2}$. Therefore, $(\bar{u}, \phi'') = \sum_l \gamma_l(\tilde{u}^l, \phi^l) \in \text{conv}(\Phi^{P_2})$. It follows that $\phi'' \leq \text{conc}_{P^2}(\phi)(\bar{u})$. Moreover, $\bar{u}_{ij} \geq u''_{ij}$ for $j \in A_i \backslash \{0, n\}$ and $\bar{u}_{ij} \leq u''_{ij}$ for $j \in B_i$. Then, by Lemma 2.2.1, it follows that $\text{conc}_{P^1}(\phi)(u'') = \phi'' \leq \text{conc}_{P^2}(\phi)(\bar{u}) \leq \text{conc}_{P^2}(\phi)(u'')$ proving the fourth statement of the result.

Now, we show the last statement of the result. Let $P' = P(a, a^L, a^U, B)$ and $P'' = P(a, \tilde{a}^L, \tilde{a}^U, B)$. By the second statement of the result, it follows that $\text{conc}_{P'}(\phi)(u) = \text{conc}_{P'}(\phi)(u')$ and, because $u' \in P''$, it follows from the fourth statement of the result that $\text{conc}_{P'}(\phi)(u') = \text{conc}_{P''}(\phi)(u')$. Therefore, $\text{conc}_{P'}(\phi)(u) = \text{conc}_{P''}(\phi)(u')$. By considering $-\phi$, a similar argument shows that $\text{conv}_{P'}(\phi)(u) = \text{conv}_{P''}(\phi)(u')$. It follows then that $\text{conv}_{P'}(\phi)(u) \leq \phi \leq \text{conc}_{P'}(\phi)(u)$ if and only if $\text{conv}_{P''}(\phi)(u') \leq \phi \leq \text{conc}_{P''}(\phi)(u')$. ■

By Proposition 2.2.1, it suffices to construct $\mathrm{conv}\big(\Phi^{P(a,\tilde{a}^L,\tilde{a}^U,B)}\big)$ to characterize $\mathrm{conv}\big(\Phi^{P(a,a^L,a^U,B)}\big)$. The primary role played by a_i^L and a_i^U is to ensure that the estimating functions are bounded. Therefore, without loss of generality, we will assume in the foregoing discussion, unless specified otherwise, that $a_i^L = a_{i0}$ and $a_i^U = a_{in}$. In other words, we will use Proposition 2.2.1 to simplify P. Proposition 2.2.1 has another subtle value. It turns out that one of the main results in this paper can be viewed as a sharpening of Proposition 2.2.1. This sharper version will show that $(u, \phi) \in \mathrm{conv}(\Phi^P)$ if and only if a certain point $(s, \phi) \in \mathrm{conv}(\Phi^P)$, where s is larger than the u' constructed in the statement of Proposition 2.2.1. The transformation from u to s is significantly more involved and is, arguably, one of the cornerstones of the development later. The more general result will have much further reaching consequences since it will also enable a significant simplification of P.

We now further simplify the structure of polytope P. First, we show that we may assume that the index set B_i defining P_i in (2.2) is empty. To achieve this simplification, we will transform overestimators into underestimators (and vice-versa) and thus show that $B_i = \emptyset$ can be assumed without loss of generality. Consider an affine transformation $T : \mathbb{R}^{d\times(n+1)} \mapsto \mathbb{R}^{d\times(n+1)}$ defined as follows:

$$T(u)_{ij} = u_{ij} \quad \text{for } j \in A_i \qquad \text{and} \qquad T(u)_{ij} = a_{ij} - u_{ij} + u_{in} \quad j \in B_i. \tag{2.5}$$

Recall that A_i and B_i are the index sets for underestimators and overestimators of f_i respectively. Let $\big(u(x), a\big)$ be a pair satisfying (2.1). It follows that $\big(T(u(x))\big)_{ij}$ is an underestimator of $f_i(x)$ bounded from above by a_{ij} because $a_{ij} \le u_{ij}(x)$ for all i and $j \in B_i$. Clearly, the transformation T_i is invertible. More specifically, given a vector $t \in \mathbb{R}^{d\times(n+1)}$, we have $T^{-1}(t)_{ij} = t_{ij}$ for $j \in A_i$ and $T^{-1}(t)_{ij} = a_{ij} - t_{ij} + t_{in}$ for $j \in B_i$, where we have used $n \in A_i$ to substitute t_{in} for u_{in}. Since the transformation T is such that $T(u)_{ij}$ depends only on u_{ij} and u_{in}, we may write $T(u) := \big(T_1(u_1), \dots, T_d(u_d)\big)$, where $T_i : \mathbb{R}^{n+1} \mapsto \mathbb{R}^{n+1}$. Similarly, we write $T^{-1}(t) := \big(T_1^{-1}(t_1), \dots, T^{-1}(t_d)\big)$. Then,

$$\hat{P}_i := T(P_i) = \left\{ t_i \in \mathbb{R}^{n+1} \,\middle|\, \begin{array}{l} \text{for each } j \in A_i : t_{ij} \leq t_{in} \text{ and } a_i^L \leq t_{ij} \leq a_{ij} \\ \text{for each } j \in B_i : t_{ij} \leq t_{in} \text{ and } t_{in} + a_{ij} - a_i^U \leq t_{ij} \leq a_{ij} \\ t_{i0} = a_{i0},\ a_{i0} \leq t_{in} \leq a_{in} \end{array} \right\}.$$

We further relax $\hat{P}_i$ to $\bar{P}_i$ as follows:

$$\bar{P}_i = \left\{ t_i \in \mathbb{R}^{n+1} \,\middle|\, \begin{array}{l} t_{ij} \leq t_{in} \text{ and } a'_{iL} \leq t_{ij} \leq a_{ij} \\ t_{i0} = a_{i0},\ a_{i0} \leq t_{in} \leq a_{in} \end{array} \right\},$$

where $a'_{iL} = \min\{a_i^L, a_{i0} + a_{ij} - a_i^U\}$. Let $\hat{P} = \prod_{i=1}^d \hat{P}_i$ and $\bar{P} = \prod_{i=1}^d \bar{P}_i$.

Proposition 2.2.2 *Let* $G := \big\{(t, \phi) \,\big|\, t \in \hat{P}, (t, \phi) \in \operatorname{conv}(\Phi^{\bar{P}})\big\}$. *Then,* $\operatorname{conv}(\Phi^{\hat{P}}) = G$. *Moreover,* $\operatorname{conv}(\Phi^P) = \big\{(u, \phi) \,\big|\, u \in P, \big(T(u), \phi\big) \in \operatorname{conv}(\Phi^{\bar{P}})\big\}$.

Proof Since $\Phi^{\hat{P}} \subseteq G$ and G is convex, it follows that $\operatorname{conv}(\Phi^{\hat{P}}) \subseteq G$. Now, we show that $\operatorname{conv}(\Phi^{\hat{P}}) \supseteq G$. Let $(t, \phi) \in G$. It follows that $t \in \hat{P}$ and there exist convex multipliers λ_k and $(t^k, \phi^k) \in \Phi^{\bar{P}}$ such that $(t, \phi) = \sum_k \lambda_k (t^k, \phi^k)$. Let $\bar{t}_{ij}^k = \max\big\{t_{ij}^k, a_i^L, t_{in}^k + a_{ij} - a_i^U\big\}$. We show that $(\bar{t}_{ij}^k)_{j=1}^{n+1} \in \hat{P}_i$. This is easily verified since $\bar{t}_{in}^k = t_{in}^k$, $\bar{t}_{i0}^k = t_{i0}^k$, and $\max\big\{t_{ij}^k, a_i^L, t_{in}^k + a_{ij} - a_i^U\big\} \leq \min\big\{t_{in}^k, a_{ij}\big\}$. Since $t_{in}^k = \bar{t}_{in}^k$, it follows that $(\bar{t}^k, \phi^k) \in \operatorname{conv}(\Phi^{\hat{P}})$. Then, $(\bar{t}, \phi) := \sum_k \lambda_k (\bar{t}^k, \phi^k) \in \operatorname{conv}(\Phi^{\hat{P}})$. However, $\bar{t} \in \hat{P}$ and $\bar{t} \geq t$. Therefore, $\operatorname{conv}(\Phi^{\hat{P}}) \supseteq G$ because $t \in \hat{P}$ and

$$\operatorname{conv}_{\hat{P}}(\phi)(t) \leq \operatorname{conv}_{\hat{P}}(\phi)(\bar{t}) \leq \phi \leq \operatorname{conc}_{\hat{P}}(\phi)(\bar{t}) \leq \operatorname{conc}_{\hat{P}}(\phi)(t).$$

The second and third inequality follow from $(\bar{t}, \phi) \in \operatorname{conv}(\Phi^{\hat{P}})$. We will now show the first inequality, and the argument for the last inequality is similar. To see that $\operatorname{conv}_{\hat{P}}(\phi)(t) \leq \operatorname{conv}_{\hat{P}}(\phi)(\bar{t})$, let $\bar{u} = T^{-1}(\bar{t})$ and $u = T^{-1}(t)$ and observe that

$$\operatorname{conv}_{\hat{P}}(\phi)(t) = \operatorname{conv}_P(\phi)(u) \leq \operatorname{conv}_P(\phi)(\bar{u}) = \operatorname{conv}_{\hat{P}}(\phi)(\bar{t}),$$

where the first and last equality is because convex envelopes do not change when the domain and argument undergo the same invertible affine transformation and the inequality follows from Lemma 2.2.1 for $-\phi$ because, for $j \in A_i \backslash \{0, n\}$ (resp. $j \in B_i$), $u_{ij} \leq \bar{u}_{ij}$ (resp. $u_{ij} \geq \bar{u}_{ij}$), $u_{i0} = \bar{u}_{i0}$, and $u_{in} = \bar{u}_{in}$. The last statement in the result follows since T is an affine transformation and $T(u) \in \hat{P}$ if and only if $u \in P$. ∎

Since $\bar{P}$ is a special case of P with $B_i = \emptyset$ for all i, Proposition 2.2.2 shows that this special case is sufficient to treat the general case. Combined with Proposition 2.2.1, it suffices to consider $\bar{P}$ with $a_i^L = a_{i0}$. We may further assume that $a_{10} < a_{i1} < \cdots < a_{in}$, *i.e.*, we may assume that a_{ij}s are strictly increasing in j since we may replace with u_{ij} the maximum of all underestimators that share the same bounds. Unless specified otherwise, we will, from here onwards, let $P(a)$ denote $\prod_{i=1}^d P_i(a_i)$ for $a_{i0} < \cdots < a_{in}$, where

$$P_i(a_i) = \left\{ u_i \in \mathbb{R}^{n+1} \;\middle|\; \begin{array}{l} u_{ij} \leq u_{in} \text{ and } a_{i0} \leq u_{ij} \leq a_{ij} \\ u_{i0} = a_{i0},\ a_{i0} \leq u_{in} \leq a_{in} \end{array} \right\}. \tag{2.6}$$

We showed in Example 2.1.1 that redundant underestimators of inner functions can help improve the relaxation of the composite function. However, when the underestimators are obtained using a convex combination derivation, we will show that they do not improve the quality of the relaxation. For each pair $\big(u(x), a\big)$ satisfying (2.1), we denote by $\mathcal{R}_{\mathrm{CR}}\big(u(x), a\big)$, the relaxation of composite function $\phi \circ f$ obtained as in Theorem 2.2.1. Now, consider a vector of underestimators $u'(x)$ and their upper bounds a' obtained by taking convex combinations of $u(x)$ and a. More precisely, let Λ_i be a nonnegative matrix in $\mathbb{R}^{(n+1)\times(n'+1)}$, where rows indexes are in $\{0, \ldots, n\}$ and columns indexes are in $\{0, \ldots, n'\}$ such that

$$\Lambda_i \geq 0,\ e^{n+1}\Lambda_i = e^{n'+1},\ \Lambda_i(\cdot, 0) = (1, 0, \ldots, 0)^\top,\ \Lambda_i(\cdot, n') = (0, \ldots, 0, 1)^\top, \tag{2.7}$$

where e^k is the all-ones row vector in $\mathbb{R}^k$ and $\Lambda_i(\cdot, k)$ is the k^{th} column of the matrix Λ_i. We let $u_i(x)$ (resp. $u_i'(x)$) denote a row vector of $n+1$ (resp. $n'+1$) functions $\big(u_{i0}(x), \ldots, u_{in}(x)\big)$ (resp. $\big(u'_{i0}(x), \ldots, u'_{in'}(x)\big)$). Then, we define $u'(x) : \mathbb{R}^d \mapsto \mathbb{R}^{d\times(n'+1)}$ such that the i^{th} row of $u'(x)$ is $u_i'(x) := u_i(x)\Lambda_i$ and $a' \in \mathbb{R}^{d\times(n'+1)}$ such that the i^{th} row of a' is $a_i' := a_i\Lambda_i$. First two conditions in (2.7) imply that, for all i and $k \in \{0, \ldots, n'\}$, $u'_{ik}(x)$ and the corresponding bound a'_{ik} are obtained by taking a convex combination of $u_{i0}(x), \ldots, u_{in}(x)$ and their bounds $a_{i0}, \ldots, a_{in}$ respectively. The third (resp. fourth) requirement in (2.7) ensure that $u'_{i0} = u_{i0}$ and $a'_{i0} = a_{i0}$ (resp. $u'_{in'} = u_{in}$ and $a'_{in'} = a_{in}$). Therefore, the new pair $\big(u'(x), a'\big)$ satisfies the requirements in (2.1). Thus, we can derive a valid relaxation $\mathcal{R}_{\mathrm{CR}}(u'(x), a')$ for the composite function $\phi \circ f$ using $\big(u'(x), a'\big)$ as in Theorem 2.2.1. In the following proposition, we show that $\mathcal{R}_{\mathrm{CR}}\big(u(x), a\big) \subseteq \mathcal{R}_{\mathrm{CR}}\big(u'(x), a'\big)$.

Proposition 2.2.3 *Let $\big(u(x), a\big)$ be a pair which satisfies conditions in (2.1) and Λ_i be a matrix defined in (2.7). Define $u'(x) \in \mathbb{R}^d \mapsto \mathbb{R}^{d\times(n'+1)}$ such that $u'_{ij}(x) = u_i(x)\Lambda_i(\cdot, j)$ and $a' \in \mathbb{R}^{d\times(n'+1)}$ such that $a'_{ij} = a_i\Lambda_i(\cdot, j)$. Then, $\mathcal{R}_{CR}\big(u(x), a\big) \subseteq \mathcal{R}_{CR}\big(u'(x), a'\big)$.*

Proof Let $(x, u_{\cdot n}, \phi) \in \mathcal{R}_{\mathrm{CR}}\big(u(x), a\big)$. Define $u = u(x)$ and thus $(u, \phi) \in \operatorname{conv}\big(\Phi^{P(a)}\big)$. Let $u' = u'(x)$. By the definition of $u'(x)$, $u'_i = u_i(x)\Lambda_i$. To show that $\mathcal{R}_{\mathrm{CR}}\big(u(x), a\big) \subseteq \mathcal{R}_{\mathrm{CR}}\big(u'(x), a'\big)$, we only need to show that $(u', \phi) \in \operatorname{conv}\big(\Phi^{P(a')}\big)$. Let $\mathcal{A}$ be the affine transform used to obtain (u', ϕ) from (u, ϕ). Then, the result follows because:

$$(u', \phi) = \mathcal{A}(u, \phi) \in \mathcal{A}\operatorname{conv}\big(\Phi^{P(a)}\big) = \operatorname{conv}\big(\mathcal{A}\Phi^{P(a)}\big) \subseteq \operatorname{conv}\big(\Phi^{P(a')}\big),$$

where the first equality is by definition, the second equality is because $\mathcal{A}$, being an affine transform, commutes with convexification. All that remains to show is the final containment. This follows if we show that, for each $(\bar{u}, \bar{\phi}) \in \Phi^{P(a)}$, the point $(\bar{u}', \bar{\phi}) := \mathcal{A}(\bar{u}, \bar{\phi}) \in \Phi^{P(a')}$, where we recall that $\mathcal{A}$ is such that $\bar{u}'_i = \bar{u}_i\Lambda_i$. Observe that $a_{i0}e^{n+1}\Lambda_i \leq \bar{u}_i\Lambda_i \leq a_i\Lambda_i$ shows that $a_{i0}e^{n'+1} \leq \bar{u}'_i \leq a'_i$ and $\bar{u}_i\Lambda_i \leq \bar{u}_{in+1}e^{n+1}\Lambda_i$ shows that $\bar{u}'_i \leq \bar{u}_{in+1}e^{n'+1} = \bar{u}'_{in+1}e^{n'+1}$. Finally, since $\bar{u}_{in} = \bar{u}'_{in}$, it follows that $\bar{\phi} = \phi(\bar{u}_{1n}, \ldots, \bar{u}_{dn}) = \phi(\bar{u}'_{1n}, \ldots, \bar{u}'_{dn})$. Therefore, $(\bar{u}', \bar{\phi}) \in \Phi^{P(a')}$ and the result follows. ∎

Example 2.2.1 *As in Example 2.1.2, consider the monomial $x_1^2x_2^2$ over $[0, 2]^2$ and the underestimator $u_i = \max\{0, 2x_i - 1\}$, which is bounded from above by 3. Now, let*

$w_i(x_i) := \frac{2}{3} \times u_i(x_i) + \frac{1}{3} \times 0 = \max\{\frac{4}{3}x_i - \frac{2}{3}, 0\}$. *Clearly,* $w_i(x_i)$ *is an underestimator of* x_i^2 *bounded from above by* 2. *Let*

$$a' = \begin{bmatrix} 0 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \end{bmatrix} \qquad \textit{and} \qquad u'(x) = \begin{bmatrix} 0 & w_1(x) & u_1(x) & x_1^2 \\ 0 & w_2(x) & u_2(x) & x_2^2 \end{bmatrix}.$$

In this example, we qualify the underestimating functions of Theorem 2.1.1 with the underestimators used to derive them. For example, $e_2(w_1, f_1, u_2, f_2)$ *denotes the underestimator* e_2 *obtained therein using the underestimator* $w_1(x)$ *instead of* $u_1(x)$. *One of non-vertical facet-defining inequalities for the convex envelope of the bilinear term* f_1f_2 *over the polytope* $P(a')$ *is given by:*

$$\phi \geq w_1 + 3f_1 + 2u_2 + 2f_2 - 14, \tag{2.8}$$

which is obtained from Theorem 2.1.1 as $f_1f_2 \geq e_2(w_1, f_1, u_2, f_2)$. *By substituting* w_1, f_1, u_2, *and* f_2 *with* $\frac{4}{3}x_1 + \frac{2}{3}$, x_1^2, $2x_2 - 1$, *and* x_2^2 *respectively, we obtain:*

$$\phi \geq \frac{4}{3}x_1 + 4x_2 + 3x_1^2 + 2x_2^2 - 16 - \frac{2}{3}.$$

Even though this inequality is not directly obtained from Theorem 2.1.1 when $u_1(x)$ *is the underestimator for* $f_1(x)$, *it is redundant and is obtained as* $\phi \geq \frac{2}{3}e_2(u_1, f_1, u_2, f_2) + \frac{1}{3}e_4(u_1, f_1, u_2, f_2)$, *after substitution, derived as in Example 2.1.2. A tighter cut can be obtained from (2.8) by substituting* w_1 *with* $(4-2\sqrt{2})x_1 + (2-\sqrt{2})^2$, *which is the subgradient of* x_1^2 *at* $2-\sqrt{2}$ *and is bounded from above by* 2. *However, this underestimator is not a convex combinations of* 0, $u_1(x_1)$, *and* $f_1(x_1)$.

2.3 Polynomial time equivalence of separations

To efficiently utilize $\text{conv}(\Phi^P)$ for constructing relaxations, we must solve the *separation problem* of $\text{conv}(\Phi^P)$, that is, given a vector $(\bar{u}, \bar{\phi})$ we need to determine if $(\bar{u}, \bar{\phi}) \in \text{conv}(\Phi^P)$ and, if not, find a hyperplane that separates $(\bar{u}, \bar{\phi})$ from $\text{conv}(\Phi^P)$. In this section, we will prove that the separation problem of $\text{conv}(\Phi^P)$ can be solved in polynomial time, given a polynomial time separation oracle for $\text{conv}(\Phi^Q)$, where

$Q := \prod_{i=1}^{d} Q_i$ for a certain subset Q_i of P_i, which will be formally defined in (4.16), and

$$\Phi^Q := \big\{(s_1, \ldots, s_d, \phi) \bigm| \phi = \phi(s_{1n}, \ldots, s_{dn}),\ (s_1, \ldots, s_d) \in Q_1 \times \cdots \times Q_d\big\}.$$

This result admits many applications because the structure of the vertex set $\text{vert}(Q)$ is much simpler than that of $\text{vert}(P)$. In particular, each Q_i is a simplex and, thus, Q is a cartesian product of simplices. As a result, there are conditions under which we are able to devise a polynomial-time separation algorithm of $\text{conv}(\Phi^Q)$, while a similar algorithm for $\text{conv}(\Phi^P)$ is not directly apparent.

2.3.1 An extended formulation for polyhedral abstraction

In this subsection, we introduce a subset Q_i of P_i, characterize its vertices and facet-defining inequalities, and relate the two sets, Φ^P and Φ^Q, as projections of a common extended formulation.

Let a_i be a strictly increasing vector in $\mathbb{R}^{n+1}$ and let $V_i := (v_{i0}, \ldots, v_{in})$ be a symmetric matrix in $\mathbb{R}^{(n+1)\times(n+1)}$ whose j^{th} column is:

$$v_{ij} = (a_{i0}, \ldots, a_{ij-1}, a_{ij}, \ldots, a_{ij}) \qquad j = 0, \ldots, n. \tag{2.9}$$

By a slight abuse of notation, we will also treat V_i as a set of points $\{v_{i0}, \ldots, v_{in}\}$ so that $\text{conv}(V_i)$ denotes the convex hull of columns of V_i, *i.e.*, $\text{conv}(V_i) := \text{conv}(v_{i0}, \ldots, v_{in})$. Define $Q_i := \text{conv}(V_i)$. Since $V_i \subseteq P_i$, it follows that $Q_i \subseteq P_i$.

Now, we show that Q_i can be seen as an invertible affine transform of Δ_i defined as:

$$\Delta_i := \big\{z_i \in \mathbb{R}^{n+1} \bigm| 0 \leq z_{in} \leq \cdots \leq z_{i1} \leq z_{i0} = 1\big\}. \tag{2.10}$$

Denote by ζ_{ij} the vector $\sum_{j'=0}^{j} e_{j'}$ for all $j = 0, \ldots, n$, where e_j is the j^{th} principal vector in $\mathbb{R}^{n+1}$ with $e_0 = (1, 0, \ldots, 0)^\top$. It can be verified that $\text{vert}(\Delta_i) = \{\zeta_{ij}\}_{j=0}^{n}$. Moreover, it follows readily that points $\zeta_{i0}, \ldots, \zeta_{in}$ are affinely independent since the matrix $\zeta_i \in \mathbb{R}^{(n+1)\times(n+1)} := (\zeta_{i0}, \ldots, \zeta_{in})$ is invertible, being the upper triangular matrix of all ones. Thus, $\dim(\Delta_i) = n$. The affine transformation that maps Q_i to Δ_i is $Z_i : \mathbb{R}^{n+1} \mapsto \mathbb{R}^{n+1}$ and is defined as follows:

$$z_{i0} = 1 \qquad \text{and} \qquad z_{ij} = \frac{s_{ij} - s_{ij-1}}{a_{ij} - a_{ij-1}} \quad \text{for } j = 1, \ldots, n. \tag{2.11}$$

To verify that $\Delta_i = Z_i(Q_i)$ observe that Z_i maps v_{ij} to ζ_{ij}. Then, Z_i^{-1} is given by:

$$s_{ij} = a_{i0}z_{i0} + \sum_{k=1}^{j}(a_{ik} - a_{ik-1})z_{ik} \quad \text{for } j = 0, \ldots, n, \tag{2.12}$$

and $Q_i = Z_i^{-1}(\Delta_i)$. We now characterize the extreme points and facet-defining inequalities of Q_i.

Lemma 2.3.1 *Let $a_i := (a_{i0}, \ldots, a_{in})$ be a strictly increasing vector in $\mathbb{R}^{n+1}$ and let $Q_i := \mathrm{conv}(V_i)$. Then, Q_i is a simplex so that $\mathrm{vert}(Q_i) = V_i$, and*

$$Q_i = \left\{ (s_{ij})_{j=0}^{n} \;\middle|\; s_{i0} = a_{i0},\ 0 \le \frac{s_{in} - s_{in-1}}{a_{in} - a_{in-1}} \le \cdots \le \frac{s_{i1} - s_{i0}}{a_{i1} - a_{i0}} \le 1 \right\}, \tag{2.13}$$

where all the inequalities are facet-defining. For j, j', and j'', satisfying $0 \le j < j' < j'' \le n$, each point in Q_i satisfies $0 \le \frac{s_{ij''} - s_{ij'}}{a_{ij''} - a_{ij'}} \le \frac{s_{ij'} - s_{ij}}{a_{ij'} - a_{ij}} \le 1$.

Proof The inequality description follows from the preceding discussion since Q_i and Δ_i are related by an invertible affine transform, which maps v_{ij} to ζ_{ij} and the constraints in (2.13) transform those in (2.10). Now, we show the last statement in the result. For any $0 \le j < j' \le 1$, it follows that $\frac{s_{ij'} - s_{ij}}{a_{ij'} - a_{ij}} = \sum_{k=j+1}^{j'} \frac{s_{ik} - s_{ik-1}}{a_{ik} - a_{ik-1}} \frac{a_{ik} - a_{ik-1}}{a_{ij'} - a_{ij}}$, which is a convex combination of $\frac{s_{ik} - s_{ik-1}}{a_{ik} - a_{ik-1}}$ for $k = j+1, \ldots, j'$. Therefore, it follows from (2.13) that $0 \le \frac{s_{ij''} - s_{ij'}}{a_{ij''} - a_{ij'}} \le \frac{s_{ij'} - s_{ij}}{a_{ij'} - a_{ij}} \le 1$. ■

Next, we lift simplex Q_i into the space of (u_i, s_i) variables by imposing the ordering constraints $a_{i0}e \le u_i \le s_i$, $u_{i0} = s_{i0}$, and $u_{in} = s_{in}$, which yields a polytope

$$PQ_i = \left\{ (u_i, s_i) \;\middle|\; u_{i0} = s_{i0} = a_{i0},\ u_{in} = s_{in},\ a_{i0}e \le u_i \le s_i,\ s_i \in Q_i \right\}, \tag{2.14}$$

where e is the all-ones vector in $\mathbb{R}^{n+1}$. Let $\Phi^{PQ} := \left\{ (u, s, \phi) \;\middle|\; \phi = \phi(s_{1n}, \ldots, s_{dn}),\ (u, s) \in PQ \right\}$. We now show that Φ^{P} and Φ^{Q} are projections of Φ^{PQ}.

Lemma 2.3.2 *Let a_i be a strictly increasing vector in $\mathbb{R}^{n+1}$. Then,* $\text{proj}_{u_i}(PQ_i) = P_i$ *and* $\text{proj}_{s_i}(PQ_i) = Q_i$. *Moreover,* $\Phi^P = \text{proj}_{(u,\phi)}(\Phi^{PQ})$ *and* $\Phi^Q = \text{proj}_{(s,\phi)}(\Phi^{PQ})$.

Proof We first argue that $\text{proj}_{s_i}(PQ_i) = Q_i$. If $(u_i, s_i) \in PQ_i$ then $s_i \in Q_i$, and therefore $\text{proj}_{s_i}(PQ_i) \subseteq Q_i$. To show $Q_i \subseteq \text{proj}_{s_i}(PQ_i)$, we consider a point $s_i \in Q_i$ and observe $(s_i, s_i) \in PQ_i$. Second, we argue that $\text{proj}_{u_i}(PQ_i) = P_i$. To show $P_i \subseteq \text{proj}_{u_i}(PQ_i)$, let $u_i \in P_i$ and define $s_{ij} = \min\{a_{ij}, u_{in}\}$ for all j. We will show $(u_i, s_i) \in PQ_i$. It follows readily that, for $j \in \{0, n\}$, $u_{ij} = s_{ij}$ and $u_i \leq s_i$. In addition, observe that there exists a j' such that $a_{ij'-1} \leq u_{in} \leq a_{ij'}$. By direct computation, $s_i = \lambda v_{ij'-1} + (1-\lambda) v_{ij'}$, where $\lambda = \frac{a_{ij'} - u_{in}}{a_{ij'} - a_{ij'-1}}$. In other words, $s_i \in Q_i$. To prove $\text{proj}_{u_i}(PQ_i) \subseteq P_i$, we consider a point $(u_i, s_i) \in PQ_i$ and show $u_i \in P_i$. Clearly, $u_{i0} = a_{i0}$. Also, for $j = 1, \ldots, n$, $a_{i0} \leq u_{ij} \leq s_{ij} \leq \min\{u_{in}, a_{ij}\}$, where the first two inequalities hold by the definition of PQ_i and the last inequality holds because $u_{in} = s_{in}$ and the inequality, $s_{ij} \leq \min\{a_{ij, s_{in}}\}$, is satisfied by all extreme points of Q_i.

The last two statements follow similarly because Φ^P, Φ^Q, Φ^{PQ} are obtained from P, Q, and PQ respectively, by adding a coordinate ϕ which depends only on coordinates shared by P, Q, and PQ, namely $(u_{1n}, \ldots, u_{dn}) = (s_{1n}, \ldots, s_{dn})$. ■

Lemma 2.3.2 implies that $\text{conv}(\Phi^P) = \text{conv}\big(\text{proj}_{(u,\phi)}(\Phi^{PQ})\big) = \text{proj}_{(u,\phi)} \text{conv}\big((\Phi^{PQ})\big)$, where the second equality holds because $\text{conv}(AS) = A\,\text{conv}(S)$ for any affine mapping A and a set S. Similarly, $\text{conv}(\Phi^Q) = \text{proj}_{(u,\phi)} \text{conv}\big(\Phi^{PQ}\big)$.

Now, we characterize $\text{conv}(\Phi^{PQ})$ by using $\text{conv}(\Phi^Q)$ as follows:

$$\text{conv}(\Phi^{PQ}) = \Big\{(u, s, \phi) \;\Big|\; (s, \phi) \in \text{conv}(\Phi^Q),\; (u, s) \in PQ\Big\}. \tag{2.15}$$

We will establish the equality in a more general setting, so we can also apply the result when we consider the convex hull of a vector of functions over PQ in Section 2.4. Let $\mathcal{X} := \{(x, \mu) \in \mathbb{R}^{m+k} \mid \mu = f(x), x \in X\}$ be the graph of a vector of functions $f : \mathbb{R}^m \mapsto \mathbb{R}^k$ over a non-empty subset X of $\mathbb{R}^m$. Let $l : \mathbb{R}^m \mapsto \mathbb{R}^n$ and $h : \mathbb{R}^m \mapsto \mathbb{R}^n$ be vectors of affine functions. We consider a set $\mathcal{Z}$ defined as:

$$\mathcal{Z} = \Big\{(x, y, \mu) \in \mathbb{R}^{m+n+k} \;\Big|\; l(x) \leq y \leq h(x),\; (x, \mu) \in \mathcal{X}\Big\}.$$

Clearly, Φ^{PQ} is of the form of $\mathcal{Z}$, where $\mathcal{X} = \Phi^Q$, $h(s) = s$, $l_{in}(s) = s_{in}$, and $l_{ij}(s) = a_{i0}$ if $j < n$. Interpreting $\mathcal{Z}$ as such, the following result implies the equality in (2.15).

Lemma 2.3.3 *Assume that $l(x) \leq h(x)$ for all $x \in X$. Then, we have* $\operatorname{conv}(\mathcal{Z}) = \bigl\{(x, y, \mu) \bigm| (x, \mu) \in \operatorname{conv}(\mathcal{X}),\ l(x) \leq y \leq h(x)\bigr\}$.

Proof Let $R := \bigl\{(x, y, \mu) \bigm| (x, \mu) \in \operatorname{conv}(\mathcal{X}),\ l(x) \leq y \leq h(x)\bigr\}$. The inclusion $\operatorname{conv}(\mathcal{Z}) \subseteq R$ follows from $\mathcal{Z} \subseteq R$ since R is convex. We now show that $R \subseteq \operatorname{conv}(\mathcal{Z})$. Let $(x', y', \mu') \in R$. Clearly, (x', y', μ') lies in the hypercube, $H := \bigl\{(x, y, \mu) \bigm| (x, \mu) = (x', \mu'),\ l(x') \leq y \leq h(x')\bigr\}$. So, it suffices to show that the vertex set of H belongs to $\operatorname{conv}(\mathcal{Z})$. We only show that vertices of the form $\bigl(x', l(x'), \mu'\bigr)$ lie in $\operatorname{conv}(\mathcal{Z})$ since a similar argument applies to the remaining vertices.

Let $\mathcal{Z}' := \bigl\{(x, y, \mu) \bigm| (x, \mu) \in \mathcal{X},\ y = l(x)\bigr\}$ and express it as the image of $\mathcal{X}$ under the affine transformation $A : (x, \mu) \to \bigl(x, l(x), \mu\bigr)$. Therefore, $\operatorname{conv}(\mathcal{Z}') = \operatorname{conv}\bigl(A(\mathcal{X})\bigr) = A\bigl(\operatorname{conv}(\mathcal{X})\bigr)$ and, so, $\bigl(x', l(x'), \mu'\bigr) \in A\bigl(\operatorname{conv}(\mathcal{X})\bigr) = \operatorname{conv}(\mathcal{Z}') \subseteq \operatorname{conv}(\mathcal{Z})$, where the last containment follows because $\mathcal{Z}' \subseteq \mathcal{Z}$ is implied by our assumption that $l(x) \leq h(x)$ for all $x \in X$. ■

2.3.2 A combinatorial algorithm for polynomial equivalence

Before establishing polynomial equivalence of the separation problems associated with $\operatorname{conv}(\Phi^P)$ and $\operatorname{conv}(\Phi^Q)$, we discuss an application of this result. Since $\operatorname{vert}(Q) = \prod_{i=1}^{d} \operatorname{vert}(Q_i)$ and $|\operatorname{vert}(Q_i)| = n + 1$, it follows that $|\operatorname{vert}(Q)| = (n + 1)^d$. In other words, if d is a constant, the number of vertices of Q is polynomial in n. More generally, assuming r and d are constants, the number of r dimensional faces of Q is polynomial in n, being upper-bounded by $(n + 1)^d \binom{nd}{r}$. Since convex/concave-envelopes of many functions depend only on the function value at the vertices or, more generally, low-dimensional faces of the domain over which the function is convex, we can construct polynomial-sized formulations for separation of $\operatorname{conv}(\Phi^Q)$ and, therefore, of $\operatorname{conv}(\Phi^P)$. For simplicity, we only discuss the implication of the equivalence of separating $\text{conc}_Q(\phi)$ and $\text{conc}_P(\phi)$ since a similar discussion directly applies to $\text{conv}_Q(\phi)$ and $\text{conv}_P(\phi)$ by considering $-\phi$ instead.

Definition 2.3.1 ([16]) *A function $g : D \mapsto \mathbb{R}$, where D is a polytope, is concave-extendable (resp. convex-extendable) from $X \subseteq D$ if the concave (resp. convex) envelope of $g(x)$ is determined by X, that is, the concave envelope of g and $g|_X$ over D are identical, where $g|_X$ is the restriction of g to X:*

$$g|_X = \begin{cases} g(x) & x \in X \\ -\infty & \textit{otherwise.} \end{cases}$$

If ϕ is concave-extendable from vertices of Q then, using Theorem 2.4 in [16], we can separate $(\bar{s}, \bar{\phi}) \in \mathbb{R}^{d\times(n+1)} \times \mathbb{R}$ from the hypograph of $\text{conc}_Q(\phi)(s)$ by solving the following linear program:

$$\begin{aligned} \min_{(\alpha,b)} \quad & \langle \alpha, \bar{s} \rangle + b \\ \text{s.t.} \quad & \langle \alpha, v \rangle + b \geq \phi\big(v_{1n}, \ldots, v_{dn}\big) \ \forall v \in \text{vert}(Q) \\ & \alpha \in \mathbb{R}^{d\times(n+1)}, \ b \in \mathbb{R}, \end{aligned} \tag{2.16}$$

where an extreme point solution (α^*, b^*) yields a facet-defining inequality of $\text{conc}_Q(\phi)(s)$ tight at $\bar{s}$, *i.e.* $\langle \alpha^*, \bar{s} \rangle + b^* = \text{conc}_Q(\phi)(\bar{s})$. We retain b although it can be absorbed in α_{i0} if $a_{i0} \neq 0$. Since $|\text{vert}(Q)| = (n+1)^d$, it follows that the size of the LP (2.16) is polynomial in n for a fixed d. As an example of the usefulness of this construction, observe that multilinear functions are convex and concave extendable from the vertices of Q (see [9]). Then, as Theorem 2.3.1 shows, the above LP gives a tractable approach to separate $\text{conv}(\Phi^Q)$, when ϕ is multilinear and d is fixed. Techniques in [47] can be used to identify whether a function is concave-extendable (convex-extendable) from vertices of Q.

Theorem 2.3.1 *Assume that $\phi(s_{1n}, \ldots, s_{dn})$ is concave-extendable from* $\text{vert}(Q)$ *and d is a fixed constant. For any given $\bar{s} \in \mathbb{R}^{d\times(n+1)}$, there exists a polynomial time procedure to generate a facet-defining inequality of* $\text{conc}_Q(\phi)(s)$ *that is tight at $\bar{s}$.*

Proof Given $\bar{s} \in Q$, the LP (2.16) can be solved in polynomial time by using an interior point algorithm. Moreover, by Lemma 6.5.1 in [48], an optimal extreme point solution of linear program (2.16) can be found in polynomial time. Then, the result follows from Theorem 2.4 in [16]. ∎

Remark 2.3.1 *Although the separation problem of* $\operatorname{conc}_Q(\phi)(u)$ *can be directly formulated as a LP of polynomial size, the similar LP formulation for* $\operatorname{conc}_P(\phi)(u)$ *using the construction of (2.16) is exponentially-sized in* n *because the* $|\operatorname{vert}(P_i)|$ *is exponential in* n. *To see this, for* $i \in \{1, \dots, d\}$, *consider the face of* P_i *defined as* $F_i := P_i \cap \{u_i \mid u_{in} = a_{in}\}$. *Since* F_i *coincides with the hypercube* $\{u_i \mid u_{in} = a_{in},\ a_{i0} \le u_{ij} \le a_{ij}\ j = 1, \dots, n-1\}$ *and* $a_{i0} < a_{ij}$ *for* $j = 1, \dots, n-1$, *it follows that* $|\operatorname{vert}(F_i)| = 2^{n-1}$. *Because* $\operatorname{vert}(F_i) \subseteq \operatorname{vert}(P_i)$, $|\operatorname{vert}(P_i)| \ge 2^{n-1}$. *Since* $\operatorname{vert}(P) = \prod_{i=1}^{d} |\operatorname{vert}(P_i)|$, $|\operatorname{vert}(P)| \ge 2^{d(n-1)}$.

Remark 2.3.2 *A convex program, similar to the above LP, can be written to treat the separation problem of* $\operatorname{conc}_Q(\phi)(s)$ *for more general cases. For example, consider the case when* $\operatorname{conc}_Q(\phi)$ *is determined by its value over polynomially many faces of* Q *(for example, faces of dimension* r *or less, for some constant* r*) and* ϕ *is concave over those faces. To treat this case, we replace the constraint in (2.16) with* $b \ge \sup_{x \in F}\{\phi(x) - \langle \alpha, x \rangle\} = (-\phi)^*_F(-\alpha)$, *for each face* F *of* Q *which is required in the computation of* $\operatorname{conc}_Q(\phi)$. *Here,* $(-\phi)^*_F$ *denotes the Fenchel conjugate of* $-\phi$ *with its domain restricted to* F.

We now show that separating $\operatorname{conv}(\Phi^P)$ and $\operatorname{conv}(\Phi^Q)$ are polynomially equivalent. To this end, we devise a combinatorial algorithm that solves the separation problem of $\operatorname{conv}(\Phi^P)$ in $\mathcal{O}(dn)$ and a separation oracle call for $\operatorname{conv}(\Phi^Q)$. We also derive many structural insights relating $\operatorname{conv}(\Phi^P)$ and $\operatorname{conv}(\Phi^Q)$ that are useful in applications and illustrate their use to bilinear term in Section 2.3.3.

We start by presenting a brief preview of our construction. Given a point $(\bar{u}, \bar{\phi}) \notin \operatorname{conv}(\Phi^P)$, where $\bar{u} := (\bar{u}_1, \dots, \bar{u}_d) \in P$, we first devise a lifting procedure, Algorithm 1, to lift $\bar{u}_i$ to a particular point $(\bar{u}_i, \bar{s}_i) \in PQ_i$ for all $i = 1, \dots, d$. For this particular pair $(\bar{u}_i, \bar{s}_i)$, we show in Proposition 2.3.2 that $\bar{s}_{ij}$ can be expressed as a convex combination of $\bar{u}_{i0}, \ldots, \bar{u}_{in}$ for all $i = 1, \ldots, d$ and $j = 0, \ldots, n$. Second, we show that $(\bar{s}, \bar{\phi}) \notin \operatorname{conv}(\Phi^Q)$. Third, we augment the separation oracle for $\operatorname{conc}_Q(\phi)(s)$ to generate a cut which satisfies a certain sign condition on coefficients and cuts off $(\bar{s}, \bar{\phi})$. Last, given a cut for $\operatorname{conc}_Q(\phi)(s)$ satisfying these sign conditions, we derive a valid cut for $\operatorname{conc}_P(\phi)(u)$. The lifting procedure is a cornerstone in the proof architecture. Before presenting the lifting procedure formally, we illustrate, in the next example, the main idea behind the procedure.

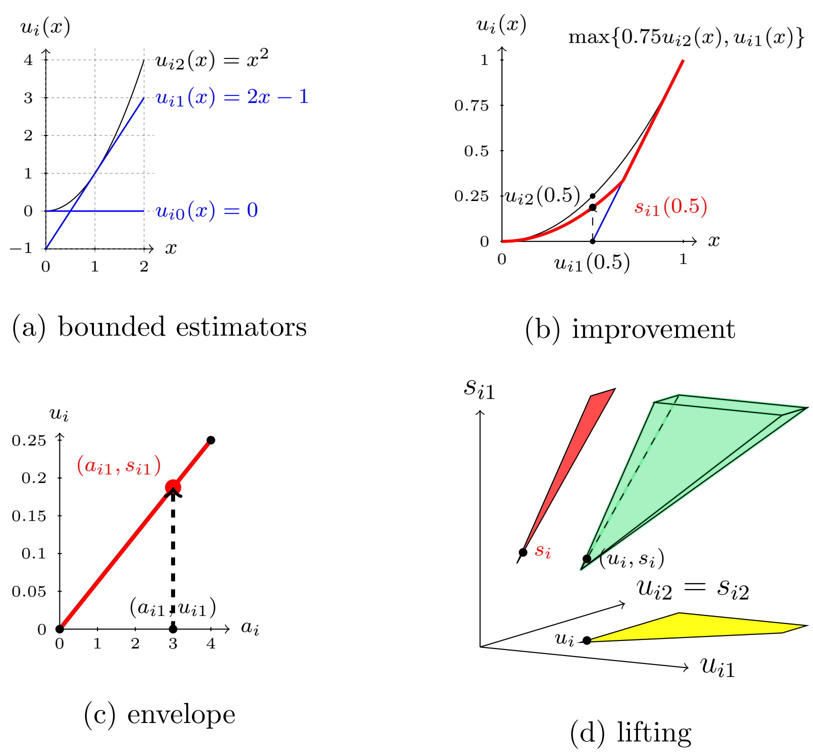


Fig. 2.1.: Illustration of the lifting procedure

Example 2.3.1 *Consider x^2 over the interval $[0, 2]$, and define*

$$u_i(x) = \big(u_{i0}(x), u_{i1}(x), u_{i2}(x)\big)$$

where $u_{i0}(x) := 0$, $u_{i1}(x) := 2x - 1$ and $u_{i2}(x) := x^2$, which are bounded from above by $a_{i0} = 0$, $a_{i1} = 3$ and $a_{i2} = 4$ respectively (see Figure 2.1a). Next, we *derive the largest underestimator* $s_{ij}(x)$ *bounded from above by* a_{ij} *by taking convex combinations of provided underestimators. More specifically, for all* j, $s_{ij}(x) := \max\{\sum_{j'=0}^{2} \lambda'_j u_{ij'}(x) \mid \sum_{j'=0}^{2} \lambda'_j a_{ij'} = a_{ij}\}$*. In Figure 2.1b, we tighten* $u_{i1}(x)$ *to* $s_{i1}(x) = \max\{0.75u_{i2}(x), u_{i1}(x)\}$*, because* $0.75u_{i2}(x) + 0.25u_{i0}(x) = 0.75u_{i2}(x)$ *is an underestimator which is bounded from above by* $0.75a_{i2} + 0.25a_{i0} = 3$*. We let* $s_{i0}(x) = u_{i0}(x)$ *and* $s_{i2}(x) = u_{i2}(x)$*.*

We will find it useful to visualize the evaluation of $s_i(x)$ *as depicted in Figure 2.1c. Consider* $x = 0.5$ *and let* $u_i = u_i(0.5)$*. In order to evaluate* $s_{i1}(x)$ *at* 0.5*, we compute* $\text{conc}_{[0,2]}(\xi)(a; u_i)$ *at* $a = a_{i1} = 3$*, where* $\xi(a; u_i)$ *is a univariate discrete function whose graph consists of the points,* $\{(a_{i0}, u_{i0}), (a_{i1}, u_{i1}), (a_{i2}, u_{i2})\}$*, which are depicted as black nodes in Figure 2.1c. In this way, the construction of the envelope of this univariate function lifts* $u_i = u_i(0.5) = (0, 0, 0.25)$ *to* $(u_i, s_i) = \big(u_i(0.5), s_i(0.5)\big) = \big((0, 0, 0.25), (0, 0.1875, 0.25)\big)$*. The obtained pair* (u_i, s_i) *belongs to* $PQ_i = \{u_{i0} = 0,\ u_{i1} \le s_{i1},\ u_{i2} = s_{i2},\ s_i \in Q_i\}$*, where* $Q_i := \big\{s_{i0} = 0,\ 0 \le \frac{s_{i2}-s_{i1}}{1} \le \frac{s_{i1}-s_{i0}}{3} \le 1\big\}$*; see Figure 2.1d. Observe, that in order to be able to draw a 3-D figure, we do not show axes for the variables* u_{i0} *and* s_{i0} *which are fixed to* 0 *and exploit* $u_{in} = s_{in}$ *to depict both these variables on the same axis.*

We now formally introduce the lifting operation illustrated in Example 2.3.1. Given a point $\bar{u} = (\bar{u}_1, \dots, \bar{u}_d) \in P_1 \times \cdots \times P_d$, for each $i \in \{1, \dots, d\}$, we lift $\bar{u}_i \in P_i$ to a point $(\bar{u}_i, \bar{s}_i)$, where $\bar{s}_i$ is an optimal solution of the following LP:

$$\begin{aligned} \min_{s} \quad & \sum_{j=0}^{n} s_{ij}, \\ \text{s.t.} \quad & \bar{u}_i \le s_i \\ & s_i \in Q_i. \end{aligned} \tag{2.17}$$

By Lemma 2.3.2, the feasible region of the linear program (2.17) is non-empty. We will show that (2.17) has an unique optimal solution and propose an algorithm, Algorithm 1, to solve (2.17) in $\mathcal{O}(n)$ operations. This algorithm relies on representing points in P_i as discrete univariate functions in the following way. With a point $u_i \in P_i$, we associate a discrete univariate function $\xi(a; u_i) : [a_{i0}, a_{in}] \mapsto \mathbb{R}$ as follows:

$$\xi(a; u_i) = \begin{cases} u_{ij} & a = a_{ij} \quad \text{for } j \in \{0, \ldots, n\} \\ -\infty & \text{otherwise.} \end{cases} \tag{2.18}$$

Moreover, let $\hat{\xi}(a; u_i) : [a_{i0}, a_{in}] \mapsto \mathbb{R}$ be the piecewise-linear interpolation of $\xi(a; u_i)$ such that $\hat{\xi}(a; u_i) = \xi(a; u_i)$ for $a \in \{a_{i0}, \ldots, a_{in}\}$ and, for all $j = 1, \ldots, n$, the restriction of $\hat{\xi}(a; u_i)$ to $[a_{ij-1}, a_{ij}]$ is linear. Next, we characterize points in the simplex Q_i as a family of univariate concave functions.

Lemma 2.3.4 *Given a point $s_i \in Q_i$, the extension $\hat{\xi}(a; s_i) : [a_{i0}, a_{in}] \mapsto \mathbb{R}$ is a non-decreasing concave function such that $\hat{\xi}(a; s_i) \leq a$ for $a \in [a_{i0}, a_{in}]$ and $\hat{\xi}(a_{i0}; s_i) = a_{i0}$. On the other hand, if a concave function $\psi : [a_{i0}, a_{in}] \mapsto \mathbb{R}$ satisfies*

$$\psi(a) \leq a \quad \text{for } a \in [a_{i0}, a_{in}], \qquad \psi(a_{i0}) = a_{i0}, \qquad \text{and } \psi \text{ is non-decreasing} \tag{2.19}$$

then $s_i := \big(\psi(a_{i0}), \ldots, \psi(a_{in})\big)$ belongs to Q_i.

Proof We start by proving the first part. Let $s_i \in Q_i$. By Lemma 2.3.1, there exists an unique non-negative vector $\lambda_i \in \mathbb{R}^{n+1}$ such that $s_i = \sum_{j=0}^{n} \lambda_{ij} v_{ij}$, where v_{ij} is defined as in (4.16). Then, there exists a $j' \in \{1, \ldots, n\}$ and $\gamma \in [0, 1]$ such that

$$\begin{aligned} \hat{\xi}(a; s_i) &= (1-\gamma)\xi(a_{ij'-1}; s_i) + \gamma \xi(a_{ij'}; s_i) \\ &= (1-\gamma) \sum_{j=0}^{n} \lambda_{ij} \xi(a_{ij'-1}; v_{ij}) + \gamma \sum_{j=0}^{n} \lambda_{ij} \xi(a_{ij'}; v_{ij}) \\ &= \sum_{j=0}^{n} \lambda_{ij} \big((1-\gamma)\xi(a_{ij'-1}; v_{ij}) + \gamma \xi(a_{ij'}; v_{ij})\big) \\ &= \sum_{j=0}^{n} \lambda_{ij} \hat{\xi}(a; v_{ij}), \end{aligned}$$

where the first equality is because for $a \in [a_{i0}, a_{in}]$, there exists $j' \in \{1, \ldots, n\}$ and $\gamma \in [0, 1]$ such that $a = (1-\gamma)a_{ij'-1} + \gamma a_{ij'}$ and $\hat{\xi}(a; s_i)$ is linear over $[a_{ij'-1}, a_{ij'}]$, the second equality is by linearity of $\xi(a; s_i)$ with respect to s_i, the third equality is by rearrangement of terms, and last equality holds because $\hat{\xi}(a; s_i)$ is linear in a over $[a_{ij'-1}, a_{ij'}]$. Since, for all $j \in \{0, \dots, n\}$, $\hat{\xi}(a; v_{ij}) \le a$ for $a \in [a_{i0}, a_{in}]$, $\hat{\xi}(a_{i0}; v_{ij}) = a_{i0}$, $\hat{\xi}(a; v_{ij})$ is a non-decreasing concave function, and these properties are closed under convex combination, it follows that $\hat{\xi}(a, s_i)$ follows these properties as well.

Assume that $\psi(a)$ is a univariate concave function satisfying (2.19). Then, we define $s_i := \big(\psi(a_{i0}), \dots, \psi(a_{in})\big)$ and let $z_i := Z_i(s_i)$, where Z_i is as defined in (4.17). By the discussion preceding Lemma 2.3.1, to prove that $s_i \in Q_i$, it suffices to show that $z_i \in \Delta_i$. Clearly, we have $z_{ij} \ge 0$ because ψ is non-decreasing. Moreover, we have $z_{i1} = \frac{s_{i1} - s_{i0}}{a_{i1} - a_{i0}} \le 1$ because $\psi(a_{i0}) = a_{i0}$ and $\psi(a_{i1}) \le a_{i1}$. Finally, we show that the concavity of ψ implies that $z_{in} \le \dots \le z_{i1}$. Let $j' \in \{1, \dots, n-1\}$, and let $\psi(a) \le L(a)$ be a supergradient inequality of ψ at $a_{ij'}$. It follows readily that

$$
\begin{aligned}
z_{ij'+1} &= \frac{\psi(a_{ij'+1}) - \psi(a_{ij'})}{a_{ij'+1} - a_{ij'}} \le \frac{L(a_{ij'+1}) - L(a_{ij'})}{a_{ij'+1} - a_{ij'}} \\
&= \frac{L(a_{ij'}) - L(a_{ij'-1})}{a_{ij'} - a_{ij'-1}} \le \frac{\psi(a_{ij'}) - \psi(a_{ij'-1})}{a_{ij'} - a_{ij'-1}} \\
&= z_{ij'},
\end{aligned}
$$

where first equality and last equality follow by definition, first inequality holds because $\psi(a_{ij'+1}) \le L(a_{ij'+1})$ and $\psi(a_{ij'}) = L(a_{ij'})$, second equality is the linearity of $L(a)$, second inequality holds because $\psi(a_{ij'}) = L(a_{ij'})$ and $\psi(a_{ij'-1}) \le L(a_{ij'-1})$. ■

Now, we present Algorithm 1 to lift a point in P_i to another in PQ_i. To lift, the algorithm constructs the concave envelope of the discrete one-dimensional function defined in (2.18). Formally, given a point $\bar{u}_i \in P_i$, let $\xi(a; \bar{u}_i)$ be a discrete function associated with the point $\bar{u}_i$. The graph of $\xi(a; \bar{u}_i)$ consists of $(n+1)$ points $(a_{i0}, \bar{u}_{i0}), \dots, (a_{in}, \bar{u}_{in})$ in $\mathbb{R}^2$, which are sorted using the first coordinate. The concave envelope of $\xi(a; \bar{u}_i)$ over $[a_{i0}, a_{in}]$ can be found using two-dimensional convex hull algorithms (see [49, 50]). In particular, since points of the graph of $\xi(a; \bar{u}_i)$ are sorted in terms of the first coordinate, Graham scan [50] takes $\mathcal{O}(n)$ to derive the envelope.

Proposition 2.3.1 *Given $\bar{u}_i \in P_i$ Algorithm 1 returns the unique optimal solution $\bar{s}_i$ of (2.17) in $\mathcal{O}(n)$.*

Algorithm 1 Lifting procedure

procedure LIFTING($\bar{u}_i$)

    construct function $\xi(a; \bar{u}_i)$

    apply Graham scan to obtain $\mathrm{conc}(\xi)(a; \bar{u}_i)$

    $(\bar{s}_{i0}, \ldots, \bar{s}_{in}) \leftarrow (\mathrm{conc}(\xi)(a_{i0}; \bar{u}_i), \ldots, \mathrm{conc}(\xi)(a_{in}; \bar{u}_i))$

    return $\bar{s}_i$.

end procedure

Proof We first show that $\bar{s}_i$ is a feasible solution to (2.17). Clearly, $\bar{u}_i \leq \bar{s}_i$ because $\xi(a; \bar{u}_i) \leq \operatorname{conc}(\xi)(a; \bar{u}_i)$. Next, we show that $\operatorname{conc}(\xi)(a; \bar{u}_i)$ satisfies the three conditions in (2.19), and, therefore, by the second result in Lemma 2.3.4, $\bar{s}_i \in Q_i$. First, $\operatorname{conc}(\xi)(a_{i0}; \bar{u}_i) = \xi(a_{i0}; \bar{u}_i) = \bar{u}_{i0} = a_{i0}$. Second, observe that $\xi(a; \bar{u}_i) \leq a$ since, for all i and j, $\bar{u}_{ij} \leq a_{ij}$. This implies that $\operatorname{conc}(\xi)(a; \bar{u}_i) \leq a$ because by definition $\operatorname{conc}(\xi)(a; \bar{u}_i)$ is the smallest concave overestimator of $\xi(a; \bar{u}_i)$ over $[a_{i0}, a_{in}]$. Last, we show the monotonicity of $\operatorname{conc}(\xi)(a; \bar{u}_i)$. Observe that, for every $a \in [a_{i0}, a_{in}]$, $\operatorname{conc}(\xi)(a; \bar{u}_i) \leq \bar{u}_{in} = \xi(a_{in}, \bar{u}_i) = \operatorname{conc}(\xi)(a_{in}; \bar{u}_i)$, where first inequality holds because $\bar{u}_{ij} \leq \bar{u}_{in}$ implies that $\xi(a; \bar{u}_i) \leq \bar{u}_{in}$. Consider two points a', a'' such that $a' < a'' < a_{in}$. Let $\lambda \in [0, 1]$ such that $a'' = (1 - \lambda)a' + \lambda a_{in}$. Observe that $\operatorname{conc}(\xi)(a''; \bar{u}_i) \geq (1 - \lambda) \operatorname{conc}(\xi)(a'; \bar{u}_i) + \lambda \operatorname{conc}(\xi)(a_{in}; \bar{u}_i) \geq \operatorname{conc}(\xi)(a'; \bar{u}_i)$, where the first inequality is by concavity of $\operatorname{conc}(\xi)$ and the second inequality is because $\operatorname{conc}(\xi)(a'; \bar{u}_i) \leq \operatorname{conc}(\xi)(a_{in}; \bar{u}_i)$.

Next, we prove by contradiction that $\bar{s}_i$ is the optimal solution of (2.17). Suppose that s'_i is a feasible solution so that $\sum_{j=0}^{n} s'_{ij} \leq \sum_{j=0}^{n} \bar{s}_{ij}$ and $s'_i \neq \bar{s}_i$. As $s'_i \in Q_i$ it follows from the first statement of Lemma 2.3.4 that $\hat{\xi}(a; s'_i)$ is a concave function over the interval $[a_{i0}, a_{in}]$. Moreover, we have $\xi(a; \bar{u}_i) \leq \hat{\xi}(a; s'_i)$ because $\bar{u}_i \leq s'_i$. In other words, $\hat{\xi}(a; s'_i)$ is a concave overestimator of $\xi(a; \bar{u}_i)$ over $[a_{i0}, a_{in}]$. By the hypothesis, $\sum_{j=0}^{n} \hat{\xi}(a_{ij}; s'_i) = \sum_{j=0}^{n} s'_{ij} \leq \sum_{j=0}^{n} \bar{s}_{ij} = \sum_{j=0}^{n} \operatorname{conc}(\xi)(a_{ij}; \bar{u}_i)$. Since $s'_i \neq \bar{s}_i$, there exists a $j' \in \{0, \ldots, n\}$ such that $\hat{\xi}(a_{ij'}; s'_i) = s'_{ij'} < \bar{s}_{ij} = \operatorname{conc}(\xi)(a_{ij'}; \bar{u}_i)$, contradicting that $\operatorname{conc}(\xi)(a; \bar{u}_i)$ is the smallest concave overestimator of $\xi(a; \bar{u}_i)$ over $[a_{i0}, a_{in}]$. ∎

Now, we study the following set

$$PQ' := \Big\{(u, s) \;\Big|\; s_i = \big(\operatorname{conc}(\xi)(a_{i0}; u_i), \ldots, \operatorname{conc}(\xi)(a_{in}; u_i)\big),\ i = 1, \ldots, d\Big\}. \tag{2.20}$$

Let $(\bar{u}, \bar{s}) \in PQ'$. We will recover a mapping from $\bar{u}$ to $\bar{s}$ that is implicit in the construction of $\operatorname{conc}(\xi)$. In order to do so, observe that the envelope, $\operatorname{conc}(\xi)(\cdot; \bar{u}_i)$, gives a representation of $(a_{ij}, \bar{s}_{ij})$ as a convex combination of points in $\big\{(a_{i0}, \bar{u}_{i0}), \ldots, (a_{in}, \bar{u}_{in})\big\}$.

Let $J = (J_1, \ldots, J_d)$ be a d-tuple such that $\{0, n\} \subseteq J_i \subseteq \{0, \ldots, n\}$. We define a linear map, $\Gamma_J : \mathbb{R}^{d \times (n+1)} \mapsto \mathbb{R}^{d \times (n+1)}$, as follows:

$$\begin{aligned} &\tilde{u}_{ij} = u_{ij} && \text{for } i \in \{1, \ldots, d\} \quad j \in J_i, \\ &\tilde{u}_{ij} = \gamma_{ij} u_{il(i,j)} + (1 - \gamma_{ij}) u_{ir(i,j)} && \text{for } i \in \{1, \ldots, d\} \quad j \notin J_i, \end{aligned} \tag{2.21}$$

where $l(i,j) := \max\{j' \in J_i \mid j' \leq j\}$, $r(i,j) := \min\{j' \in J_i \mid j' \geq j\}$, and $\gamma_{ij} = (a_{ir(i,j)} - a_{ij})/(a_{ir(i,j)} - a_{il(i,j)})$ for $j \notin J_i$. We will show that $\bar{s} = \Gamma_{J'}(\bar{u})$ for $J' = (J'_1, \ldots, J'_d)$ such that $J'_i := \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. Geometrically, we identify a face of Q in which the point $\bar{s}$ lies. In Section 2.4.2, this characterization plays an important role in establishing a polynomial time equivalence of facet generations between a vector of functions over P and over Q. Recall that, by Lemma 2.3.1, $\text{vert}(Q_i) = \{v_{i0}, \ldots, v_{in}\}$. For a d-tuple $J = (J_1, \ldots, J_d)$ such that $\{0, n\} \subseteq J_i \subseteq \{0, \ldots, n\}$, let $F_J := F_{1J(1)} \times \cdots \times F_{dJ(d)}$, where $F_{iJ(i)} := \text{conv}(\{v_{ij'} \mid j' \in J_i\})$. It follows readily that $F_{iJ(i)}$ is a face of the simplex Q_i and can also be described as the set of points of Q_i, which satisfy the following at equality:

$$\frac{s_{ij+1} - s_{ij}}{a_{ij+1} - a_{ij}} \leq \frac{s_{ij} - s_{ij-1}}{a_{ij} - a_{ij-1}} \qquad \text{for } j \notin J_i. \tag{2.22}$$

By Lemma 2.3.1, given any $j \notin J_i$, all points in Q_i, and therefore in $\text{vert}(Q_i)$, satisfy $\frac{s_{ij+1} - s_{ij}}{a_{ij+1} - a_{ij}} \leq \frac{s_{ij} - s_{ij-1}}{a_{ij} - a_{ij-1}}$. The vertices in $\{v_{ij'} \mid j' \in J_i\}$ satisfy these inequalities at equality while the remaining do not.

Proposition 2.3.2 *Let $J = (J_1, \ldots, J_d)$ be a d-tuple such that $\{0, n\} \subseteq J_i \subseteq \{0, \ldots, n\}$, and let Γ_J be a linear map defined in (2.21). Then, we have $\big(\Gamma_J(u), s\big) \in PQ$ for every $(u, s) \in PQ$. Moreover, inequality $\Gamma_J(s) \leq s$ defines the face F_J. Last, if $(\bar{u}, \bar{s}) \in PQ'$ then $\bar{s} = \Gamma_{J'}(\bar{u})$ and $\bar{s} \in F_{J'}$, where $J' = (J'_1, \ldots, J'_d)$ such that $J'_i := \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$.*

Proof We start by showing that $\big(\Gamma_J(u), s\big) \in PQ$ for every $(u, s) \in PQ$. Let $(u, s) \in PQ$ and define $\tilde{u} := \Gamma_J(u)$. Clearly, for $i \in \{1, \ldots, d\}$ and $j \in \{0, n\}$, we have $\tilde{u}_{ij} = u_{ij} = s_{ij}$, where first equality holds because $\{0, n\} \subseteq J_i$ and second equality holds because $(u, s) \in PQ$ implies that $u_{i0} = s_{i0}$ and $u_{in} = s_{in}$. Moreover, for all i and j, $a_{i0} \leq u_{ij'}$ for every $j' \in \{0, \ldots, n\}$ implies that $a_{i0} \leq \tilde{u}_{ij}$. Last, for $i \in$

$\{1,\ldots,d\}$, $\tilde{u}_i \le s_i$ follows because, for $j \in \{0,\ldots,n\}$, we have $\tilde{u}_{ij} \le \mathrm{conc}(\xi)(a_{ij};u_i) \le \hat{\xi}(a_{ij};s_i) = s_{ij}$, where first inequality holds because the point $(a_{ij},\tilde{u}_{ij})$ is expressible as a convex combination of the hypograph of $\xi(a;u_i)$, second inequality holds because, by $(u_i,s_i) \in PQ_i$ and Lemma 2.3.4, $\hat{\xi}(a;s_i)$ is a concave overestimator of $\xi(a;u_i)$, thus, of $\mathrm{conc}(\xi)(a;u_i)$, and the equality is by definition.

Next, we show that inequality $s \ge \Gamma_J(s)$ defines the facet F_J of Q. Since, for every $s \in Q$, $(s,s) \in PQ$, the validity of $s \ge \Gamma_J(s)$ over Q follows from the first result. Then, we need to show that $s = \Gamma_J(s)$ for $s \in \mathrm{vert}(F_J)$ and $s \ne \Gamma_J(s)$ for $s \in \mathrm{vert}(Q) \setminus \mathrm{vert}(F_J)$. Let $s' = (s'_1,\ldots,s'_d) \in \mathrm{vert}(Q)$. Observe that s' satisfies $s \ge \Gamma_J(s)$ at equality if and only if, for $i \in \{1,\ldots,d\}$, the associated extension function $\hat{\xi}(a;s'_i)$ is linear over $[a_{i\bar{j}_i}, a_{i\hat{j}_i}]$ for every $0 \le \bar{j}_i \le \hat{j}_i \le n$ such that $\bar{j}_i, \hat{j}_i \in J_i$ and $J_i \cap \{\bar{j}_i+1,\ldots,\hat{j}_i-1\} = \emptyset$. Observe that $\hat{\xi}(a;v_{ij})$ is linear over intervals that do not contain a_{ij}. If, for each i, there exists $j'_i \in J_i$ such that $s'_i = v_{ij'_i}$ then $\hat{\xi}(a;v_{ij'})$ is linear over $[a_{i\bar{j}_i}, a_{i\hat{j}_i}]$ because j'_i, being in J_i, does not belong to $\{\bar{j}_i+1,\ldots,\hat{j}_i-1\}$. Therefore, the point s' satisfies $s \ge \Gamma_J(s)$ at equality. On the other hand, assume that for some $i \in \{1,\ldots,d\}$, there exists $j'' \notin J_i$, so that $s'_i = v_{ij''}$. Since $\{0,n\} \subseteq J_i$ and $j'' \notin J_i$, there exist $\bar{j}_i, \hat{j}_i \in J_i$ such that $\bar{j}_i < j'' < \hat{j}_i$, $\hat{\xi}(a;v_{ij''_i})$ is not linear over $[a_{i\bar{j}_i}, a_{i\hat{j}_i}]$, and $s' \ne \Gamma_J(s')$.

Last, we prove the third result. Let $(\bar{u},\bar{s}) \in PQ'$. Then, we define $(\bar{z}_1,\ldots,\bar{z}_d) = \big(Z_1(\bar{s}_1),\ldots,Z_d(\bar{s}_d)\big)$, where Z_i is defined as in (4.17). Because $\bar{s}_i \in Q_i$, $0 \le \bar{z}_{in} \le \cdots \le \bar{z}_{i1} \le \bar{z}_{i0} = 1$. We will show that that $\bar{z}_{ij+1} = \bar{z}_{ij}$ for all i and $j \notin J'_i$. Thus, by (2.22), $\bar{s} \in F_{J'}$, and, therefore, $\bar{s} = \Gamma_{J'}(\bar{s}) = \Gamma_{J'}(\bar{u})$, where first equality holds by the second result and second equality holds by the definition of $\Gamma_{J'}$ and $\bar{s}_{ij} = \bar{u}_{ij}$ for all i and $j \in J'_i$. Now, we show that $\bar{z}_{ij+1} = \bar{z}_{ij}$ for all $j \notin J'_i$. By definition, there are $\bar{j}$, $\hat{j}$ and $\gamma \ge 0$ such that $\bar{s}_{ij} = \gamma\bar{u}_{i\bar{j}} + (1-\gamma)\bar{u}_{i\hat{j}}$. Since $j \notin J'_i$, we can assume that $0 \le \bar{j} < j < \hat{j} \le n$. If not, assume wlog that $j = \hat{j}$. Then, $0 = a_{i\hat{j}} - a_{ij} = \gamma(a_{i\hat{j}} - a_{i\bar{j}})$ implies that either $\gamma = 0$ or $a_{i\hat{j}} = a_{ij} = a_{i\bar{j}}$. In either case, $\bar{s}_{ij} = \bar{u}_{ij}$, contradicting that $j \notin J_i'$. Then, it follows that $(a_{ij}, \bar{s}_{ij})$, $(a_{i\bar{j}}, \bar{u}_{i\bar{j}})$, and $(a_{i\hat{j}}, \bar{u}_{i\hat{j}})$ are collinear. Let $L(a)$ be the function passing through these points. It follows that:

$$\begin{aligned} L(a_{ij}) &\geq \gamma' \operatorname{conc}(\xi)(a_{ij-1}, \bar{u}_i) + (1-\gamma') \operatorname{conc}(\xi)(a_{ij+1}, \bar{u}_i) \\ &\geq \gamma' L(a_{ij-1}) + (1-\gamma') L(a_{ij+1}) \\ &= L(a_{ij}), \end{aligned}$$

where $\gamma' = \frac{a_{ij+1} - a_{ij}}{a_{ij+1} - a_{ij-1}}$. The first inequality is because $L(a_{ij}) = \operatorname{conc}(\xi)(a_{ij}, \bar{u}_i)$ and $\operatorname{conc}(\xi)$ is concave, second inequality is because $\gamma' \in [0,1]$ and concavity of $\operatorname{conc}(\xi)$ implies that $\operatorname{conc}(\xi)(a_{ij-1}, \bar{u}_i) \geq L(a_{ij-1})$ and $\operatorname{conc}(\xi)(a_{ij+1}, \bar{u}_i) \geq L(a_{ij+1})$, and the equality is because of linearity of L. Therefore, equality holds throughout and $\bar{s}_{ij} = L(a_{ij}) = \gamma' \operatorname{conc}(\xi)(a_{ij-1}, \bar{u}_i) + (1-\gamma') \operatorname{conc}(\xi)(a_{ij+1}, \bar{u}_i) = \gamma' \bar{s}_{ij-1} + (1-\gamma')\bar{s}_{ij+1}$, thereby showing that $\bar{z}_{ij+1} = \bar{z}_{ij}$. ■

Let $\langle \alpha, s\rangle + \beta\phi + b \geq 0$ be a valid inequality generated for $\operatorname{conv}(\Phi^Q)$. We propose an algorithm, Algorithm 2, to generate another valid inequality $\langle \alpha', s\rangle + \beta\phi + b \geq 0$ for $\operatorname{conv}(\Phi^Q)$ such that $\alpha'_{ij} \leq 0$ for all i and $j \notin \{0, n\}$ and $\langle \alpha, s\rangle \geq \langle \alpha', s\rangle$ for every $s \in Q$. In words, given a valid inequality for $\operatorname{conv}(\Phi^Q)$, Algorithm 2 generates a valid inequality which dominates the original one over Φ^Q and also satisfies the sign condition discussed above. In the next proposition, we show the correctness of Algorithm 2 and discuss its complexity.

Proposition 2.3.3 *Let $\langle \alpha, s\rangle + \beta\phi + b \geq 0$ be a valid inequality for $\operatorname{conv}(\Phi^Q)$. Algorithm 2 generates an inequality $\langle \alpha', s\rangle + \beta\phi + b \geq 0$ which is valid for $\operatorname{conv}(\Phi^Q)$ such that $\alpha'_{ij} \leq 0$ for all i and $j \notin \{0, n\}$. Moreover, for each $s \in Q$, $\langle \alpha', s\rangle \leq \langle \alpha, s\rangle$ and there exists $\tilde{s} \in Q$ such that, for all i, $\tilde{s}_{in} = s_{in}$ and $\langle \alpha, \tilde{s}\rangle = \langle \alpha', s\rangle$. The procedure takes $\mathcal{O}(dn)$ time.*

Proof In Step 5, we initialize J_i^+ as a stack of indices j such that $\alpha_{ij} > 0$ and in Steps 9 and 10, we initialize a queue that contains all the indices in $\{0, \ldots, n\}$. In Step 14 of Algorithm 2, we pick $j \in J_i^+$ and remove it from the queue in Step 23. Indices are added to J_i^+ only in Step 20 and the added index belongs to $\{1, \ldots, n-1\}$. We

Algorithm 2 Sign Procedure

procedure $\textsc{Sign}(\alpha)$
  for i from 1 to d **do**
    for j from 1 to $n-1$ **do**
      if $\alpha_{ij} > 0$ **then**
        $\text{push}(j, J_i^+)$
      end if
    end for
    for j from 0 to n **do**
      $\text{prev}(j) = j - 1$
      $\text{succ}(j) = j + 1$
      $\alpha'_{ij} = \alpha_{ij};$
    end for
    while $J_i^+ \neq \emptyset$ **do**
      $j = \text{pop}(J_i^+)$
      $\vartheta_{ij} = \frac{a_{i\text{succ}(j)} - a_{ij}}{a_{i\text{succ}(j)} - a_{i\text{prev}(j)}}$
      for (j', ϱ) in $\big[\big(\text{prev}(j), \vartheta_{ij}\big), \big(\text{succ}(j), 1 - \vartheta_{ij}\big)\big]$ **do**
        $\text{neg} \leftarrow (\alpha_{ij'} \leq 0);$
        $\alpha'_{ij'} = \alpha'_{ij'} + \varrho\alpha'_{ij};$
        if $\text{neg} = \text{true}$ and $\alpha'_{ij'} > 0$ and $0 < j' < n$ **then**
          $\text{push}(j', J_i^+)$
        end if
      end for
      $\text{prev}\big(\text{succ}(j)\big) = \text{prev}(j)$; $\text{succ}\big(\text{prev}(j)\big) = \text{succ}(j);$
      $\text{prev}(j) = -1$; $\text{succ}(j) = n + 1;$
      $\alpha'_{ij} = 0;$
    end while
  end for
  return α'.
end procedure

argue that at the beginning of each iteration of the while loop starting at Step 13, if an index $j \in J_i^+$ then $\alpha'_{ij} > 0$, otherwise $\alpha'_{ij} \leq 0$. Further, no index in J_i^+ is repeated, and if $j \in J_i^+$ then j is also in the queue. This is certainly true at first iteration after initialization ends at Step 12. At Step 20, j' is added only if $\alpha'_{ij'} > 0$ and was negative previously and, therefore, not already in J_i^+. Moreover, j' is in the queue, being either the predecessor or successor of an index j in the queue. At Step 18, it follows that $\alpha'_{ij'}$ can only increase because Step 16 guarantees that $\varrho \in [0, 1]$ since $\vartheta_{ij} \in [0, 1]$ and $\alpha_{ij} > 0$ because j was just removed from J_i^+. Therefore, existing indices in J_i^+ continue to satisfy the invariant regarding positive coefficients. Moreover, if $\alpha_{ij'}$ turns positive at Step 18, it is added to J_i^+ in Step 20. In Step 23, j is removed from the queue, but its only copy in J_i^+ was removed already at Step 14. Since the size of the queue reduces by one in each iteration of the while loop starting at Step 13 and $\{0, n\}$ remain in the queue throughout, the loop executes at most $n - 1$ iterations. Since the outer for loop starting at Step 2 executes d iterations, the complexity of the algorithm is $\mathcal{O}(dn)$. Further, α'_{ij}, for any j not in queue, is set to zero at Step 25 and never updated at Step 18 because j has already been removed from the queue and its only copy was removed from J_i^+ at Step 14. Observe that at termination J_i^+ is empty. Any remaining index j in the queue is such that $\alpha_{ij} \leq 0$ and any index j outside the queue is such that $\alpha_{ij} = 0$.

We only need to establish the correctness of the i^{th} iteration of the outer for loop starting at Step 2. Let α^k be the α' at the start of k^{th} iteration of the while-loop spanning Steps 13-26, and, for notational convenience, assume $\alpha^1 = \alpha$. It follows easily that $\langle \alpha^k, s\rangle \leq \langle \alpha^{k-1}, s\rangle$, because $s \in Q$ implies, by Lemma 2.3.1, that $s_{ij} \geq \vartheta_{ij} s_{i\text{prev}(j)} + (1 - \vartheta_{ij}) s_{i\text{succ}(j)}$. Therefore, $\langle \alpha', s\rangle \leq \langle \alpha, s\rangle$. Let $(s, \phi) \in \Phi^Q$ and define $\tilde{s}^1 = s$. At the k^{th} iteration, we define $\tilde{s}^k$ such that, for $\text{prev}(j) < j' < \text{succ}(j)$, $\tilde{s}^k_{ij'} = \vartheta_{ij'} \tilde{s}^{k-1}_{i\text{prev}(j)} + (1 - \vartheta_{ij'}) \tilde{s}^{k-1}_{i\text{succ}(j)}$ and $\tilde{s}^k_{ij'} = \tilde{s}^{k-1}_{ij'}$ otherwise. We write the elements in the queue at the beginning of k^{th} iteration as H^k. Observe that, for $j' \notin H^k$, $\hat{\xi}(a; \tilde{s}^k_i)$ is linear at $a_{ij'}$. Since $\tilde{s}^k_i$ and $\tilde{s}^{k-1}_i$ differ only at indices not in H^k, it follows easily that $\tilde{s}^k_{ij'} = s_{ij'}$ for $j' \in H^k$. We show that $\langle \alpha^{l+1}, \tilde{s}^k\rangle = \langle \alpha^l, \tilde{s}^k\rangle$ for

$1 \le l \le k-1$. In the l^{th} iteration of the while loop, assume that, during Step 14, j' is popped from J_i^+, $\text{prev}(j') = \bar{j}$, and $\text{succ}(j') = \hat{j}$. Let $\vartheta = \frac{a_{i\hat{j}} - a_{ij'}}{a_{i\hat{j}} - a_{i\bar{j}}}$. Since the queue becomes smaller with iterations and $l \le k-1$, it follows that $j' \notin H^k$. Then, $\langle \alpha^{l+1}, \tilde{s}^k \rangle = \langle \alpha^l, \tilde{s}^k \rangle - \alpha^l_{ij'} \big(\tilde{s}^k_{ij'} - \vartheta \tilde{s}^k_{i\bar{j}} - (1-\vartheta)\tilde{s}^k_{i\hat{j}} \big)$, where, by the definition of $\tilde{s}^k$, $\{\bar{j}+1, \ldots, \hat{j}-1\} \notin H^k$, $\hat{\xi}(a; \tilde{s}^k)$ is linear between $[a_{i\bar{j}}, a_{i\hat{j}}]$ and, so, the term in the parenthesis on right-hand-side is zero. Since It follows from second result of Proposition 2.3.2 that $\tilde{s}^k \in Q$. Since $\{0, n\} \subseteq H^k$, it follows that $\tilde{s}^k_{in} = s_{in}$. Moreover, $\langle \alpha^k, \tilde{s}^k \rangle = \langle \alpha^k, s \rangle$ since $\tilde{s}^k_{ij} = s_{ij}$ for all $j \in H^k$ and $\alpha^k_{ij} = 0$ for $j \notin H^k$. It follows by considering the last iteration, say r, of the while-loop that there exists $\tilde{s}^r \in Q$ such that $\tilde{s}^r_{in} = s_{in}$, $\alpha' = \alpha^r$, and $\langle \alpha^r, s \rangle = \langle \alpha^r, \tilde{s}^r \rangle = \langle \alpha, \tilde{s}^r \rangle$. Moreover, since $\tilde{s}^r \in Q$ and $\phi = \phi(s_{1n}, \ldots, s_{dn}) = \phi(\tilde{s}^r_{1n}, \ldots, \tilde{s}^r_{dn})$, by assumption, $\langle \alpha, \tilde{s}^r \rangle + \beta\phi + b \ge 0$, thereby proving that $\langle \alpha', s \rangle + \beta\phi + b \ge 0$. ■

Given $(\bar{u}, \bar{s}) \in PQ'$ and a valid inequality $\langle \alpha, s \rangle + \beta\phi + b \ge 0$ for $\text{conv}(\Phi^Q)$ satisfying the sign condition guaranteed by Algorithm 2, the following lemma generates an inequality $\langle \alpha', u \rangle + \beta\phi + b \ge 0$ valid for $\text{conv}(\Phi^P)$ so that $\langle \alpha, \bar{s} \rangle = \langle \alpha', \bar{u} \rangle$.

Lemma 2.3.5 *Let $\langle \alpha, s \rangle + \beta\phi + b \ge 0$ be an inequality valid for* $\text{conv}(\Phi^Q)$. *Let $(\bar{u}, \bar{s}) \in PQ'$ and define $J = (J_1, \ldots, J_d)$ such that $J_i = \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. Consider a linear function $\langle \alpha', \cdot \rangle$ such that $\langle \alpha', u \rangle = \big\langle \alpha, \Gamma_J(u) \big\rangle$, where Γ_J is the linear transformation defined in (2.21). If, for all $i \in \{1, \ldots, d\}$ and $j \notin \{0, n\}$, α_{ij} is non-positive then, for every $(u, s) \in PQ$, $\langle \alpha, s \rangle \le \langle \alpha', u \rangle$ and equality is attained at $(\bar{u}, \bar{s})$. Moreover, the inequality $\langle \alpha', u \rangle + \beta\phi + b \ge 0$ is valid for* $\text{conv}(\Phi^P)$.

Proof Assume that $\alpha_{ij} \le 0$ for all i and $j \notin \{0, n\}$. We first observe that, for every $(u, s) \in PQ$, $\langle \alpha, s \rangle \le \big\langle \alpha, \Gamma_J(u) \big\rangle = \langle \alpha', u \rangle$, where the inequality holds because $\alpha_{ij} \le 0$ for all i and $j \notin \{0, n\}$ and, by the first result in Proposition 2.3.2, $(\Gamma_J(u), s) \in PQ$, and the equality holds by the definition of α' in the statement of the result. In particular, $\langle \alpha, \bar{s} \rangle = \big\langle \alpha, \Gamma_J(\bar{u}) \big\rangle = \langle \alpha', \bar{u} \rangle$, where first equality holds because, by the third result in Proposition 2.3.2, $\bar{s} = \Gamma_J(\bar{u})$.

Now, we show that $\langle \alpha', u\rangle + \beta\phi + b \geq 0$ is valid for $\mathrm{conv}(\Phi^P)$. Observe that the inequalities, $\langle \alpha', u\rangle \geq \langle \alpha, s\rangle$ and $\langle \alpha, s\rangle + \beta\phi + b \geq 0$, are valid for $\mathrm{conv}(\Phi^{PQ})$, where validity of the latter inequality follows since it is assumed to be valid for $\mathrm{conv}(\Phi^Q)$, which, by Lemma 2.3.2, equals $\mathrm{proj}_{(s,\phi)}\big(\mathrm{conv}(\Phi^{PQ})\big)$. This implies that $\langle \alpha', u\rangle + \beta\phi + b \geq 0$ is valid for $\mathrm{conv}(\Phi^{PQ})$ and, hence, for $\mathrm{conv}(\Phi^P)$ since it does not depend on s and, by Lemma 2.3.2, $\mathrm{conv}(\Phi^P) = \mathrm{proj}_{(u,\phi)}\big(\mathrm{conv}(\Phi^{PQ})\big)$. ■

Now, we show that polynomial time separability of $\mathrm{conv}(\Phi^Q)$ implies that of $\mathrm{conv}(\Phi^P)$.

Theorem 2.3.2 *The separation problem of* $\mathrm{conv}(\Phi^P)$ *can be solved in* $\mathcal{O}(dn)$ *time besides a call to the separation oracle for* $\mathrm{conv}(\Phi^Q)$*.*

Proof Let $(\bar{u}, \bar{\phi}) \in \mathbb{R}^{d\times(n+1)+1}$. We assume that $\bar{u} \in P$ because, if not, we can separate $\bar{u}$ from P in $\mathcal{O}(dn)$ time by finding a facet-defining inequality of P that is violated at $\bar{u}$. Let $\bar{s} := (\bar{s}_1, \ldots, \bar{s}_d)$, where $\bar{s}_i$ is the point returned by Algorithm 1 when $\bar{u}_i$ is provided as input. Observe that this step takes $\mathcal{O}(dn)$ time. Clearly, we have $(\bar{u}, \bar{s}) \in PQ'$, where PQ' is defined in (2.20). Now, we call the separation oracle for $\mathrm{conv}(\Phi^Q)$. If the oracle asserts that $(\bar{s}, \bar{\phi}) \in \mathrm{conv}(\Phi^Q)$ then, by equality in (2.15), $(\bar{u}, \bar{s}, \bar{\phi}) \in \mathrm{conv}(\Phi^{PQ})$, and thus, by Lemma 2.3.2, $(\bar{u}, \bar{\phi}) \in \mathrm{conv}(\Phi^P)$.

Now, suppose that $(\bar{s}, \bar{\phi}) \notin \mathrm{conv}(\Phi^Q)$. In this case, we will derive a hyperplane that separates $(\bar{u}, \bar{\phi})$ from $\mathrm{conv}(\Phi^P)$. We call the separation oracle for $\mathrm{conv}(\Phi^Q)$, with $(\bar{s}, \bar{\phi})$ as input. The oracle returns a hyperplane $\langle \alpha, s\rangle + \beta\phi + b = 0$ such that, for all $(s, \phi) \in \mathrm{conv}(\Phi^Q)$, $\langle \alpha, s\rangle + \beta\phi + b \geq 0$, whereas $\langle \alpha, \bar{s}\rangle + \beta\bar{\phi} + b < 0$. Using Proposition 2.3.3, we assume that $\alpha_{ij} \leq 0$ for all i and $j \notin \{0, n\}$. Otherwise, we can satisfy this sign requirement by executing, Algorithm 2, with complexity $\mathcal{O}(dn)$, using the generated inequality. We utilize Lemma 2.3.5 to derive the inequality $\langle \alpha', u\rangle + \beta\phi + b \geq 0$, which is valid for $\mathrm{conv}(\Phi^P)$. Observe that α' can easily be computed in $\mathcal{O}(dn)$ time. Given arbitrary $(u, \phi) \in \mathrm{conv}(\Phi^P)$, it follows that $\langle \alpha', \bar{u}\rangle + \beta\bar{\phi} + b = \langle \alpha, \bar{s}\rangle + \beta\bar{\phi} + b < 0 \leq \langle \alpha', u\rangle + \beta\phi + b$, where the first equality is by construction of α' and Lemma 2.3.5, first inequality is guaranteed by the separation oracle for $\text{conv}(\Phi^Q)$, and the last inequality follows directly Lemma 2.3.5. ■

Next, we explore the strength of cuts that are generated by the procedure described in Theorem 2.3.2. We will only discuss the strength of valid cuts for the hypograph of $\text{conc}_P(\phi)(u)$ since a similar argument applies for $\text{conv}_P(\phi)(u)$. Recall that we say an inequality $\phi \leq \langle \alpha, u\rangle + b$, valid for $\text{conv}(\Phi^P)$, is *tight* at a given point $\bar{u} \in P$ if $\text{conc}_P(\phi)(\bar{u}) = \langle \alpha, \bar{u}\rangle + b$.

Proposition 2.3.4 *Assume that the concave envelope,* $\text{conc}_P(\phi)(u)$*, is closed. Given a polynomial time separation oracle for* $\text{conc}_Q(\phi)(s)$ *which yields tight cuts, there exists a polynomial time separation algorithm for* $\text{conc}_P(\phi)(u)$ *which generates tight cuts.*

Proof Let $\bar{u} \in P$ and define $\bar{s} = (\bar{s}_1, \ldots, \bar{s}_d)$, where $\bar{s}_i$ is the point returned by Algorithm 1 when $\bar{u}_i$ is provided as the input. Suppose that the separation oracle generates a valid inequality $\phi \leq \langle \alpha, s\rangle + b$ of $\text{conc}_Q(\phi)(s)$, which is tight at $\bar{s}$. We assume without loss of generality that $\alpha_{ij} \leq 0$ for all i and $j \notin \{0, n\}$ since otherwise we apply Algorithm 2 to generate a new inequality which is tight at $\bar{s}$ and satisfies the sign requirement. Let $\phi \leq \langle \alpha', u\rangle + b$ be the inequality obtained using Lemma 2.3.5. Then,

$$\text{conc}_Q(\phi)(\bar{s}) = \langle \alpha, \bar{s}\rangle + b = \langle \alpha', \bar{u}\rangle + b \geq \text{conc}_P(\phi)(\bar{u}) \geq \text{conc}_P(\phi)(\bar{s}) \geq \text{conc}_Q(\phi)(\bar{s}),$$

where first equality is because $\phi \leq \langle \alpha, s\rangle + b$ is tight at $\bar{s}$, second equality holds because, by third result in Proposition 2.3.2, $\bar{s} = \Gamma_J(\bar{u})$ and, by Lemma 2.3.5, $\langle \alpha', \bar{u}\rangle = \big\langle \alpha, \Gamma_J(\bar{u})\big\rangle$, the first inequality holds because, by Lemma 2.3.5, $\phi \leq \langle \alpha', u\rangle + b$ is valid for $\text{conc}_P(\phi)(u)$, second inequality holds because, by Lemma 2.2.1, the closedness of $\text{conc}_P(\phi)(u)$ implies that it is non-increasing in u_{ij} for all i and $j \notin \{0, n\}$ and $(\bar{u}, \bar{s}) \in PQ$ implies that $\bar{u} \leq \bar{s}$ and $\bar{u}_{ij} = \bar{s}_{ij}$ for all i and $j \in \{0, n\}$, and the last inequality follows because $Q \subseteq P$ and thus $\text{conc}_P(\phi)(s) \geq \text{conc}_Q(\phi)(s)$ for every $s \in Q$. Therefore, equalities hold throughout. In particular, we obtain that $\langle \alpha', \bar{u}\rangle + b = \text{conc}_P(\phi)(\bar{u}) = \text{conc}_P(\phi)(\bar{s}) = \text{conc}_Q(\phi)(\bar{s})$. ■

We argued in the proof of Proposition 2.3.4 that, given a pair $(\bar{u}, \bar{s}) \in PQ'$, we have $\text{conc}_P(\phi)(\bar{u}) = \text{conc}_Q(\phi)(\bar{s})$. We denote by $\mathcal{L}(u_i)$ the feasible region of (2.17) with u_i as the given input. It is easy to see that, for every $u \in T(\bar{u}, \bar{s}) := \{u \mid \bar{u} \le u \le \bar{s}\}$, we have $(u, \bar{s}) \in PQ'$ because, for $i \in \{1, \dots, d\}$, $\mathcal{L}(u_i) \subseteq \mathcal{L}(\bar{u}_i)$ and $\bar{s}_i \in \mathcal{L}(u_i)$. Therefore, the following result follows from the proof of Proposition 2.3.4.

Corollary 2.3.1 *Assume that the concave envelope $\text{conc}_P(\phi)(u)$ is closed. Given $(\bar{u}, \bar{s}) \in PQ'$, we have $\text{conc}_P(\phi)(u) = \text{conc}_Q(\phi)(\bar{s})$ for every $u \in T(\bar{u}, \bar{s})$, where $T(\bar{u}, \bar{s}) = \{u \mid \bar{u} \le u \le \bar{s}\}$.*

2.3.3 Application in factorable programming

In this subsection, we show that the factorable programming scheme can be improved by considering a special case of the problem treated in Sections 2.2 and 2.3. We consider the case when ϕ is a bilinear term and P is defined with $d = n = 2$.

Theorem 2.3.3 *Let $f_1^L \le a_1 \le f_2^U$ and $f_2^L \le a_2 \le f_2^U$. Then, consider the set:*

$$P = \left\{ f_1^L \le u_1 \le \min\{f_1, a_1\}, f_1 \le f_1^U, f_2^L \le u_2 \le \min\{f_2, a_2\}, f_2 \le f_2^U \right\}.$$

The following overestimators inequalities are valid for the epigraph of $f_1 f_2$ over P:

$$f_1 f_2 \le \min \left\{ \begin{array}{l} r_1 := f_2^L f_1 + f_1^U f_2 - f_1^U f_2^L \\ r_2 := (f_2^L - a_2) u_1 + (a_1 - f_1^U) u_2 + a_2 f_1 + f_1^U f_2 - a_1 f_2^L \\ r_3 := (f_2^L - f_2^U) u_1 + a_1 f_2 + f_2^U f_1 - a_1 f_2^L \\ r_4 := (f_1^L - f_1^U) u_2 + a_2 f_1 + f_1^U f_2 - f_1^L a_2 \\ r_5 := (a_2 - f_2^U) u_1 + (f_1^L - a_1) u_2 + f_2^U f_1 + a_1 f_2 - f_1^L a_2 \\ r_6 := f_1 f_2^U + f_1^L f_2 - f_1^L f_2^U \end{array} \right\}.$$

Proof To show that r_2 is a valid overestimator, observe that $f_1 f_2 = r_2 - (a_1 - u_1)(u_2 - f_2^L) - (f_1 - u_1)(a_2 - u_2) - (f_2 - u_2)(f_1^U - f_1) \le r_2$. Similarly, it follows that r_3, r_4, and r_5 are overestimators because $f_1 f_2 = r_3 - (a_1 - u_1)(f_2 - f_2^L) - (f_1 -$

$u_1)(f_2 - f_2^U) \le r_3$, $f_1 f_2 = r_4 - (f_1 - f_1^L)(a_2 - u_2) - (f_1^U - f_1)(f_2 - u_2) \le r_4$, and $f_1 f_2 = r_5 - (u_1 - f_1^L)(a_2 - u_2) - (a_1 - u_1)(f_2 - u_2) - (f_1 - u_1)(f_2^U - f_2) \le r_5$. ■

Observe that Theorems 2.1.1 and 2.3.3 use a slightly different notation to denote P than the rest of Sections 2.2-2.4. In particular, we hide the subscript j of u_{ij}, as it is unnecessary when there is a single underestimator. We also drop the subscript j of s_{ij} in the foregoing discussion. In [42], relying heavily on the equivalence results established in Section 2.3, we provide an explicit representation of $\text{conv}(\Phi^P)$ for a setup that significantly generalizes the one introduced in Theorems 2.1.1 and 2.3.3. When specialized to the case treated here, *i.e.*, $n = d = 2$ and $\phi = f_1 f_2$, the result shows that $\text{conv}(\Phi^P)$ is obtained as an intersection of the sets described in Theorems 2.1.1 and 2.3.3. We now use the results in Section 2.3 to show that one of the inequalities obtained in Theorem 2.3.3 is a facet of $\text{conv}(\Phi^P)$. Extending this argument to each inequality given in the results, one can easily show that the inequalities in Theorems 2.1.1 and 2.3.3 describe $\text{conv}(\Phi^P)$.

Consider, for example, the inequality $f_1 f_2 \le r_3$ and a point $p := (\bar{u}_1, \bar{f}_1, \bar{u}_2, \bar{f}_2) = (a_1, 0.5a_1 + 0.5f_1^U, 0.5f_2^L + 0.5a_2, 0.25f_2^L + 0.75f_2^U)$ that belongs to P. Algorithm 1 maps this point to $q := (\bar{s}_1, \bar{f}_1, \bar{s}_2, \bar{f}_2)$, a point in Q, with $\bar{s}_1 := \bar{u}_1$ and $\bar{s}_2 := \frac{f_2^U - a_2}{f_2^U - f_2^L} f_2^L + \frac{a_2 - f_2^L}{f_2^U - f_2^L}(0.25f_2^L + 0.75f_2^U) = 0.25f_2^L + 0.75a_2$. Let $g_3 := (f_2^L - f_2^U)s_1 + a_1 f_2 + f_2^U f_1 - a_1 f_2^L$. The inequality $f_1 f_2 \le g_3$ is tight at the extreme points of Q given by $(f_1^L, f_1^L, f_2^L, f_2^L)$, (a_1, a_1, f_2^L, f_2^L), (a_1, a_1, a_2, a_2), (a_1, a_1, a_2, f_2^U), and (a_1, f_1^U, a_2, f_2^U), where for each point, the variables are ordered as in (s_1, f_1, s_2, f_2). Now, observe that q can be written as a convex combination of the above extreme points with multipliers 0, 0.25, 0, 0.25, and 0.5 respectively. Since the validity of $f_1 f_2 \le g_3$ can be shown in the same manner as that of $f_1 f_2 \le r_3$, it follows that $f_1 f_2 \le g_3$ is a facet-defining inequality of $\text{conc}_Q(\phi)$ that is tight at q. Since the coefficient of s_1 in g_3 is already non-positive, we do not need to invoke Algorithm 2 to satisfy this sign-condition. Then, as in Lemma 2.3.5, we replace each s_i with its defining expression in terms of u_i to obtain an inequality that is valid for $\text{conc}_P(\phi)$ and defines one of its facet. We provide more detail and illustrate the ideas involved. Normally, we would replace s_1 (resp. s_2) with

u_1 (resp. $\frac{f_2^U - a_2}{f_2^U - f_2^L} f_2^L + \frac{a_2 - f_2^L}{f_2^U - f_2^L} f_2$). However, since s_2 does not appear in the expression defining g_3, we obtain r_3 simply by substituting u_1 for s_1 in g_3, and thus derive the inequality, $f_1 f_2 \geq r_3$, that is valid for $\text{conc}_P(\phi)$. Using the points tight in Q, it follows easily that this inequality also defines a facet for $\text{conc}_P(\phi)$. We next show that it is also tight at p, the point that was initially chosen for its derivation, thereby, demonstrating the more general fact shown in the proof of Proposition 2.3.4. Consider the point $r := (\bar{s}_1, \bar{f}_1, f_2^L, \bar{f}_2)$ that is a convex combination of tight points, as seen from the expression, $r = 0.25(a_1, a_1, f_2^L, f_2^L) + 0.25(a_1, a_1, f_2^L, f_2^U) + 0.5(a_1, f_1^U, f_2^L, f_2^U)$. Here, the points in the rhs belong to P, are tight, and were obtained from the tight points in Q by substituting the s_2 coordinate with f_2^L. Then, since $p = \frac{1}{3}r + \frac{2}{3}q$, we have expressed p as a convex combination of tight points and shown that $f_1 f_2 \leq r_3$ defines a facet for $\text{conc}_P(\phi)$ and is tight at p.

Remark 2.3.3 *Consider the bilinear product $f_1(x)f_2(x)$ and assume $f_1(x) \leq o_1(x)$ yields an overestimator, $o_1(x)$. Then, Theorems 2.1.1 and 2.3.3 can be used by replacing f_1 with $-f_1$. In this case, the more involved transformation discussed in (2.5) is not necessary. Nevertheless, (2.5) is useful if besides the overestimator, $o_1(x)$, we also have an underestimator, $u_1(x)$, available for $f_1(x)$ and wish to exploit both estimators in the construction of cuts.*

2.4 Extensions

In this section, we consider a *vector* of functions $\theta : \mathbb{R}^d \mapsto \mathbb{R}^k$ over the polytope P and their graph

$$\Theta^P := \big\{(u, \theta) \bigm| \theta = \theta(u_{1n}, \ldots, u_{dn}),\ u \in P\big\}.$$

We refer to $\text{conv}(\Theta^P)$ as the *simultaneous hull* of Θ^P and are interested in solving the separation problem associated with this set. We, similarly, define $\Theta^Q := \big\{(s, \theta) \bigm| \theta = \theta(s_{1n}, \ldots, s_{dn}),\ s \in Q\big\}$ and generalize the results from Sections 2.3.2 to this setting. More specifically, we will show that, given a polynomial time separation oracle for conv(Θ^Q), the separation problem for conv(Θ^P) can also be solved in polynomial time. We will also prove a sharper result when conv(Θ^Q) is a polytope, such as is the case when θ is a vector of multilinear functions (see Corollary 2.7 in [51]). For this setting, we will assume that we have access to a polynomial time oracle that generates facet-defining cuts for a family of lower-dimensional polytopes of conv(Θ^Q), using which we will generate facet-defining cuts for conv(Θ^P) in polynomial time.

2.4.1 Polynomial time equivalence of separations for simultaneous hulls

In this subsection, we show that proof techniques for Theorem 2.3.2 can be generalized to establish polynomial time equivalence of separations between conv(Θ^P) and conv(Θ^Q). First, Lemma 2.3.2 and the equality in (2.15) can be easily generalized to the context of conv(Θ^P) using a similar argument. Let Θ^{PQ} be the graph of θ over polytope PQ, that is, $\Theta^{PQ} := \bigl\{(u,s,\theta) \bigm| \theta = \theta(s_{1n},\dots,s_{dn}),\ (u,s)\in PQ\bigr\}$.

Lemma 2.4.1 *We have* $\mathrm{proj}_{(u,\theta)}(\Theta^{PQ}) = \Theta^P$ *and* $\mathrm{proj}_{(s,\theta)}(\Theta^{PQ}) = \Theta^Q$*. Moreover,* $\mathrm{conv}(\Theta^P) = \bigl\{(u,s,\theta) \bigm| (s,\theta)\in \mathrm{conv}(\Theta^Q), (u,s)\in PQ\bigr\}$.

Now, observe that Algorithm 2 can be applied to modify coefficients of a valid inequality for conv(Θ^Q).

Proposition 2.4.1 *Given a valid inequality* $\langle\alpha,s\rangle + \langle\beta,\theta\rangle + b \ge 0$ *for* conv(Θ^Q)*, Algorithm 2 generates a valid inequality* $\langle\alpha',s\rangle + \langle\beta,\theta\rangle + b \ge 0$ *for* conv(Θ^Q) *so that* $\alpha'_{ij}\le 0$ *for all* i *and* $j\notin\{0,n\}$ *and* $\langle\alpha',s\rangle \le \langle\alpha,s\rangle$ *for every* $s\in Q$.

Proof Let α' be the vector returned by Algorithm 2 when α is given as input. Then, for any $(s,\theta)\in\Theta^Q$, there exists $\tilde{s}\in Q$ such that $\langle\alpha,s\rangle \ge \langle\alpha',s\rangle = \langle\alpha,\tilde{s}\rangle \ge -\langle\beta,\theta\rangle - b$, where the first inequality, first equality, and the existence of $\tilde{s}$ follow from Proposition 2.3.3. The second inequality follows because $(\tilde{s},\theta)\in\Theta^Q$ and $\langle\alpha,s\rangle + \langle\beta,\theta\rangle + b \ge 0$ is assumed to be valid for conv(Θ^Q). ■

Theorem 2.4.1 *The separation problem of* conv(Θ^P) *can be solved in polynomial time given a polynomial separation oracle for* conv(Θ^Q).

Proof The proof is similar to that of Theorem 2.3.2. We construct $\bar{s}$ using Algorithm 1 with $\bar{u}$ as input. If $(\bar{s}, \bar{\theta}) \in \text{conv}(\Theta^Q)$ then, since $(\bar{u}, \bar{s}) \in PQ$, Lemma 2.4.1 shows that $(\bar{u}, \bar{\theta}) \in \text{conv}(\Theta^P)$. If $(\bar{s}, \bar{\theta}) \notin \text{conv}(\Theta^Q)$, we use the separation oracle of $\text{conv}(\Theta^Q)$ and Algorithm 2 and Proposition 2.4.1 to obtain an inequality $\langle \alpha, s\rangle + \langle \beta, \theta\rangle + b \geq 0$ valid for $\text{conv}(\Theta^Q)$ that separates $(\bar{s}, \bar{\theta})$ from $\text{conv}(\Theta^Q)$. Then, we use the transformation of Lemma 2.3.5 to obtain α' and observe that, for all $(u, \theta) \in \text{conv}(\Theta^P)$ and $(u, s) \in PQ$, $\langle \alpha', u\rangle + \langle \beta, \theta\rangle + b \geq \langle \alpha, s\rangle + \langle \beta, \theta\rangle + b \geq 0$. Since, by Lemma 2.3.5, $\langle \alpha', \bar{u}\rangle = \langle \alpha, \bar{s}\rangle < -\langle \beta, \bar{\theta}\rangle - b$, the inequality is not satisfied at $(\bar{u}, \bar{\theta})$ and, thus, separates $(\bar{u}, \bar{\theta})$ from $\text{conv}(\Phi^P)$. ■

2.4.2 Polynomial time equivalence of facet generations for simultaneous hulls

We start by formally defining a family of polytopes of the form Q that result when subsets of the outer-approximators of $f_i(x)$ are considered. Let $\mathcal{J}$ be a collection of d-tuples defined as follows:

$$\mathcal{J} := \Big\{ (J_1, \ldots, J_d) \;\Big|\; \{0, n\} \subseteq J_i \subseteq \{0, \ldots, n\} \;\forall i \in \{1, \ldots, d\} \Big\}.$$

Consider $J = (J_1, \ldots, J_d) \in \mathcal{J}$. For any $y \in \mathbb{R}^{d\times(n+1)}$, let $y_J := (y_{1J(1)}, \ldots, y_{dJ(n)})$, where $y_{iJ(i)}$ are the components of y_i corresponding to the index J_i. Let $\bar{J} := (\bar{J}_1, \ldots, \bar{J}_d)$, where $\bar{J}_i$ is the complement of J_i, *i.e.*, $\bar{J}_i = \{0, \ldots, n\} \setminus J_i$. Using these definitions, we can now write, up to reordering of variables, that $y = (y_J, y_{\bar{J}})$. Let $a = (a_1, \ldots, a_d)$ be a vector in $\mathbb{R}^{d\times(n+1)}$ so that a_i is strictly increasing for every $i \in \{1, \ldots, d\}$. For any $J = (J_1, \ldots, J_d) \in \mathcal{J}$, we define $Q_J := Q_{1J(1)} \times \cdots \times Q_{dJ(d)}$, where $Q_{iJ(i)}$ is the simplex defined in (4.16) with parameter $a_{iJ(i)} \in \mathbb{R}^{|J(i)|}$. Consider the graph of θ over Q_J defined as $\Theta^{Q_J} := \Big\{ (s_J, \theta) \;\Big|\; \theta = \theta(s_{1n}, \ldots, s_{dn}),\; s_J \in Q_J \Big\}$. In the following, we will show that the *facet generation problem* of $\text{conv}(\Theta^P)$ can be solved in polynomial time, given a facet generation oracle for $\text{conv}(\Theta^{Q_J})$ for $J \in \mathcal{J}$. By a facet generation problem for a polyhedron S, we mean that in addition to the separation problem of S, we return a hyperplane that contains S and does not contain x if $x \notin \mathrm{aff}(S)$, or return a facet-defining inequality of S that is not satisfied at x.

Now, we associate with $\mathcal{J}$ a family of faces of Q. Namely, for each $J \in \mathcal{J}$,

$$F_J := F_{1J(1)} \times \cdots \times F_{dJ(d)},$$

where $F_{iJ(i)} := \mathrm{conv}\big(\{v_{ij} \mid j \in J_i\}\big)$ and $v_{ij} = (a_{i0}, \ldots, a_{ij-1}, a_{ij}, \ldots, a_{ij})$ for all i and j. Similarly, we consider the graph of θ over F_J defined as $\Theta^{F_J} := \big\{(s, \theta) \bigm| \theta = \theta(s_{1n}, \ldots, s_{dn}),\ s \in F_J\big\}$. In the next result, we will provide the invertible affine isomorphism relating points in polytopes $\mathrm{conv}(\Theta^{Q_J})$ with those in $\mathrm{conv}(\Theta^{F_J})$. Recall that two polytopes $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$ are *affinely isomorphic* if there is an affine map $f : \mathbb{R}^m \mapsto \mathbb{R}^n$ that is a bijection between the points of the two sets. Let $f : \mathbb{R}^m \mapsto \mathbb{R}^k$ and X be a convex subset of $\mathbb{R}^m$. For any valid inequality $\langle \alpha, x\rangle + \langle \beta, \mu\rangle + b \geq 0$ of the convex hull of $\big\{(x, \mu) \mid \mu = f(x),\ x \in X\big\}$, we shall denote by $T_f^{(\alpha,\beta,b)}(X)$ the face of $\mathrm{conv}\big(\big\{(x,\mu) \bigm| \mu = f(x),\ x \in X\big\}\big)$ defined by the valid inequality.

Lemma 2.4.2 *Assume that* $\mathrm{conv}(\Theta^Q)$ *is a polytope. Let* $J = (J_1, \ldots, J_d) \in \mathcal{J}$ *and let* $\langle \alpha, s\rangle + \langle \beta, \theta\rangle + b \geq 0$ *be a valid inequality of* $\mathrm{conv}(\Theta^Q)$ *so that* $\alpha_{\bar{J}} = 0$*. Then, the face* $T_\theta^{(\alpha,\beta,b)}(F_J)$ *of* $\mathrm{conv}(\Theta^{F_J})$ *is affinely isomorphic to the face* $T_\theta^{(\alpha_J,\beta,b)}(Q_J)$ *of* $\mathrm{conv}(\Theta^{Q_J})$*.*

Proof Let $J = (J_1, \ldots, J_d) \in \mathcal{J}$. Since F_J is a face of Q and $\mathrm{conv}(\Theta^Q)$ is a polytope, it follows readily that $\mathrm{conv}(\Theta^{F_J})$ is a polytope. We first show that $\mathrm{conv}(\Theta^{F_J})$ is affinely isomorphic to $\mathrm{conv}(\Theta^{Q_J})$. Consider an affine map $A : s_J \mapsto t$ such that $t_{ij} = s_{ij}$ for $j \in J_i$ and $t_{ij} = (1-\gamma_{ij})s_{il(i,j)} + \gamma_{ij}s_{ir(i,j)}$ for $j \notin J_i$, where $l(i,j) = \max\{j' \in J_i \mid j' \leq j\}$, $r(i,j) = \min\{j' \in J_i \mid j' \geq j\}$, and $\gamma_{ij} = \frac{a_{ij} - a_{il(i,j)}}{a_{ir(i,j)} - a_{il(i,j)}}$. It follows from second result in Proposition 2.3.2 that A maps the polytope Q_J into the face F_J. The inverse of A is defined as $s \mapsto s_J$ and maps the face F_J into the simplex Q_J. This is because, for any $s_i \in \mathrm{vert}(F_{iJ(i)})$, there exists a $k \in J_i$ such that $s_{ij} = \min\{a_{ij}, a_{ik}\}$ for all $j \in J_i$. Thus, $s_{iJ(i)} \in \mathrm{vert}\big(Q_{iJ(i)}\big)$. Consider the affine transformation, Π, defined as

$(s_J, \theta) \mapsto \big(A(s_J), \theta\big)$ and its inverse, Π^{-1}, $(s, \theta) \mapsto (s_J, \theta)$. Note that, in calling the projection operation as an inverse of Π, we are interpreting Π as a transformation into the affine hull of Φ^{F_J} rather than into Φ^Q. In other words, $\Pi\Theta^{Q_J} = \Theta^{F_J}$ and $\Pi^{-1}\Theta^{F_J} = \Theta^{Q_J}$. Therefore, $\operatorname{conv}(\Theta^{F_J}) = \operatorname{conv}(\Pi\Theta^{Q_J}) = \Pi \operatorname{conv}(\Theta^{Q_J})$ and, similarly, $\operatorname{conv}(\Theta^{Q_J}) = \Pi^{-1} \operatorname{conv}(\Theta^{F_J})$. It follows that $\operatorname{conv}(\Theta^{F_J})$ is affinely isomorphic to $\operatorname{conv}(\Theta^{Q_J})$.

Now, let $\langle \alpha, s\rangle + \langle \beta, \theta\rangle + b \geq 0$ be a valid inequality for $\operatorname{conv}(\Theta^Q)$ so that $\alpha_{\bar{J}} = 0$. Then, the validity of the inequality $\langle \alpha_J, s_J\rangle + \langle \beta, \theta\rangle + b \geq 0$ for $\operatorname{conv}(\Theta^{Q_J})$ follows since $\alpha_{\bar{J}} = 0$. Clearly, the corresponding faces, $T_\theta^{(\alpha,\beta,b)}(F_J)$ and $T_\theta^{(\alpha_J,\beta,b)}(Q_J)$, are affinely isomorphic under the mapping Π. ■

As a consequence of Proposition 2.4.1, we obtain the following monotonic property for a non-vertical facet-defining inequality of $\operatorname{conv}(\Theta^Q)$ using an argument similar to that in the proof of Lemma 2.2.1. A inequality $\langle \alpha, s\rangle + \langle \beta, \theta\rangle + b \geq 0$ is non-vertical if $\beta \neq 0$.

Lemma 2.4.3 *Assume that* $\operatorname{conv}(\Theta^Q)$ *is a polytope. Let* $\langle \alpha, s\rangle + \langle \beta, \theta\rangle + b \geq 0$ *be a non-vertical face-defining inequality of* $\operatorname{conv}(\Theta^Q)$. *Then,* $\alpha_{ij} \leq 0$ *for all* i *and* $j \notin \{0, n\}$.

Assume that $\operatorname{conv}(\Theta^Q)$ is a polytope and observe that $\operatorname{conv}(\Theta^{PQ})$ is a polytope since, by the second part of Lemma 2.4.1, $\operatorname{conv}(\Theta^{PQ}) = \big\{(u, s, \theta) \bigm| (s, \theta) \in \operatorname{conv}(\Theta^Q),\ (u, s) \in PQ\big\}$. Consequently, by the first part of Lemma 2.4.1, $\operatorname{conv}(\Theta^P)$ is a polytope. For any $J \in \mathcal{J}$, F_J is a face of Q and, so, $\operatorname{conv}(\Theta^{F_J})$ is a face of $\operatorname{conv}(\Theta^Q)$. By Lemma 2.4.3, if $\langle \alpha_J, s_J\rangle + \langle \beta, \theta\rangle + b \geq 0$ is a non-vertical facet-defining inequality of $\operatorname{conv}(\Theta^{Q_J})$ then $\alpha_{ij} \leq 0$ for all i and $j \in J_i \setminus \{0, n\}$.

Theorem 2.4.2 *Assume that* $\operatorname{conv}(\Theta^Q)$ *is a polytope. The facet generation problem of* $\operatorname{conv}(\Theta^P)$ *can be solved in polynomial time if there exists a polynomial time facet generation oracle of* $\operatorname{conv}(\Theta^{Q_J})$ *for every* $J \in \mathcal{J}$.

Proof Let $(\bar{u}, \bar{\theta}) \in \mathbb{R}^{d\times(n+1)+k}$ and assume that $\bar{u} \in P$. Let $\bar{s}$ be the point returned by Algorithm 1 with $\bar{u}$ as input, and define $J = \{J_1, \ldots, J_d\}$, where $J_i := \{j \mid$

$\bar{u}_{ij} = \bar{s}_{ij}\}$. We assume that $(\bar{s}, \bar{\theta}) \notin \Theta^Q$; otherwise $(\bar{u}, \bar{\theta}) \in \Theta^P$ as in the proof of Theorem 2.4.1. We only consider the case when the facet-generation oracle produces a facet-defining inequality, $\langle \alpha_J, s_J\rangle + \langle \beta, \theta\rangle + b \geq 0$, of $\mathrm{conv}(\Theta^{Q_J})$ with $\langle \alpha_J, \bar{s}_J\rangle + \langle \beta, \bar{\theta}\rangle + b < 0$. Instead, if the oracle returns a hyperplane which contains Θ^{Q_J}, a similar proof without the need for the point $(\tilde{s}_J, \tilde{\theta})$, we define later, can be constructed easily. By Lemma 2.4.3, $\alpha_{ij} \leq 0$ for all i and $j \notin J_i \backslash \{0, n\}$. Let $\tilde{\alpha}_J = \alpha_J$ and $\tilde{\alpha}_{\bar{J}} = 0$. Since $\langle \tilde{\alpha}, \bar{u}\rangle + \langle \beta, \bar{\theta}\rangle + b < 0$, we only need to show that the inequality $\langle \tilde{\alpha}, u\rangle + \langle \beta, \theta\rangle + b \geq 0$ is facet defining for $\mathrm{conv}(\Theta^P)$.

To prove the validity of the inequality, we consider a point $(u, s, \theta) \in \Theta^{PQ}$ and observe that $\langle \tilde{\alpha}, u\rangle \geq \langle \tilde{\alpha}, s\rangle = \langle \tilde{\alpha}_J, s_J\rangle \geq -\langle \beta, \theta\rangle - b$, where the first inequality is because $\tilde{\alpha}_{ij}(u_{ij} - s_{ij}) \geq 0$, the first equality is because $\tilde{\alpha}_{\bar{J}} = 0$, and the second inequality is because of the validity of $\langle \tilde{\alpha}_J, s_J\rangle + \langle \beta, \theta\rangle + b \geq 0$ for Θ^{Q_J}. Therefore, the inequality is valid for Θ^{PQ}, and hence, by Lemma 2.4.1, for Θ^P.

For simplicity of notation, for any set S we abbreviate $T_\theta^{(\tilde{\alpha},\beta,b)}(S)$ as $T(S)$. We will show that $\dim\bigl(T(P)\bigr) = \dim\bigl(\Theta^P\bigr) - 1$, and thus, conclude that $(\tilde{\alpha}, \beta, b)$ defines a facet of $\mathrm{conv}(\Theta^P)$. We start by constructing a subset H of $T(P)$ so that, for $i' \in \{1, \ldots, d\}$ and $j' \notin J_i$, there exist two points, $(\hat{s}, \hat{\theta})$ and $(\check{s}, \check{\theta})$, in H which differ only in coordinate corresponding to $s_{i'j'}$, that is, $(\hat{s}, \hat{\theta}) - (\check{s}, \check{\theta}) = (\delta e_{i'j'}, 0)$ where $\delta \neq 0$ and $e_{i'j'}$ is the j'^{th} principal vector in the i'^{th} subspace. Consider $(\hat{s}_J, \hat{\theta}) \in \mathrm{ri}\bigl(T(Q_J)\bigr)$ and observe that $\hat{s}_J \in \mathrm{ri}(Q_J)$. If not, one of the inequalities defining Q_J is tight at all points in $T(Q_J)$ contradicting that (α_J, β, b) defines a non-vertical facet of $\mathrm{conv}(\Theta^{Q_J})$. Next, we extend the point $\hat{s}_J$ to the point $\hat{s}$ of F_J using the transformation A defined as in the proof of Lemma 2.4.2. By Lemma 2.4.2, $\hat{s} \in \mathrm{ri}(F_J)$ and $(\hat{s}, \hat{\theta}) \in \mathrm{ri}\bigl(T(F_J)\bigr)$. It follows that $\hat{s}_{ij} > a_{i0}$ for all i and $j \neq 0$. Moreover, there exist a set of points $\bigl\{(s^k, \theta^k)\bigr\}_{k\in K} \subseteq T(F_J) \cap \Theta^{F_J}$ and convex multipliers $\lambda \in \mathbb{R}^{|K|}$ such that $(\hat{s}, \hat{\theta}) = \sum_{k\in K} \lambda_k (s^k, \theta^k)$. Now, let $\check{s}^k$ be a point so that $\check{s}^k_{ij} = s^k_{ij}$ for $(i, j) \neq (i', j')$ and $\check{s}^k_{ij} = a_{i0}$ otherwise. Since $\tilde{\alpha}_{i'j'} = 0$ and $j' \neq n$, it follows that, for all $k \in K$, $(\check{s}^k, \theta^k) \in T(P) \cap \Theta^P$. Then, we define $(\check{s}, \check{\theta}) := \sum_k \lambda_k (\check{s}^k, \theta^k)$. Therefore, by construction, $(\check{s}, \check{\theta}) \in T(P)$ and $(\hat{s}, \hat{\theta}) \in T(F_J) \subseteq T(P)$, and $(\hat{s}, \hat{\theta}) - (\check{s}, \check{\theta}) = \big((\hat{s}_{i'j'} - a_{i'0})e_{i'j'}, 0\big) \neq 0$.

Now, we argue that if $(s, \theta) \in \Theta^Q$ then, for $\delta_{ij} \in \mathbb{R}$, $\big(s + \sum_{i=1}^{d} \sum_{j \notin \{0,n\}} e_{ij}\delta_{ij}, \theta\big) \in \operatorname{aff}\big(T(P) \cup (A(\tilde{s}_J), \tilde{\theta})\big)$. We first prove the case when $\delta_{ij} = 0$ for all i and $j \neq \{0, n\}$. It follows from Lemma 2.3.1 that the point $(s_J, \theta) \in \Theta^{Q_J}$. Thus, there exists a point $(\tilde{s}_J, \tilde{\theta}) \in \Theta^{Q_J}$, not dependent on (s_J, θ), such that (s_J, θ) can be expressed as an affine combination of points in $T(Q_J) \cup (\tilde{s}_J, \tilde{\theta})$. Let Π be an affine mapping so that $\Pi(s_J, \theta) = \big(A(s_J), \theta\big)$. Then,

$$\Pi(s_J, \theta) \in \Pi\big(\operatorname{aff}\big(T(Q_J) \cup (\tilde{s}_J, \tilde{\theta})\big)\big) = \operatorname{aff}\big(T(F_J) \cup (A(\tilde{s}_J), \tilde{\theta})\big) \subseteq \operatorname{aff}\big(T(P) \cup (A(\tilde{s}_J), \tilde{\theta})\big),$$

where first inclusion holds because $(s_J, \theta) \in \operatorname{aff}\big(T(Q_J) \cup (\tilde{s}_J, \tilde{\theta})\big)$, the equality holds by Lemma 2.4.2 and by the fact that A, as an affine transformation, commutes with affine combinations, and the containment holds due to $T(F_J) \subseteq T(P)$. Because the point (s, θ) differs from $\big(A(s_J), \theta\big)$ only in coordinates corresponding to variables of the type s_{ij} for $i \in \{1, \ldots, d\}$ and $j \notin J_i$, it follows from the existence of H that (s, θ) can be expressed as an affine combination of points in $T(P) \cup \big(A(\tilde{s}_J), \tilde{\theta}\big)$. Next, we prove the general case. Let $i'' \in \{1, \ldots, d\}$ and $j'' \notin \{0, n\}$. Then, consider the point (s'', θ''), where $s''_i = v_{i0}$ for $i \neq i''$, $s''_{i''} = v_{i''j''}$ (see Lemma 2.3.1 for the definition of v_{ij}), and $\theta'' = \theta(s''_{1n}, \ldots, s''_{dn})$. Clearly, there exists $\epsilon > 0$ so that two distinct points, (s'', θ'') and $(s'' - \epsilon e_{i''j''}, \theta'')$, belong to Θ^Q and are, as shown above, expressible as affine combinations of points in $T(P)$ and $\big(A(\tilde{s}_J), \tilde{\theta}\big)$. Since these points differ only in the variable $s_{i''j''}$, it follows that we can change the variable $s_{i''j''}$ for any $(s, \theta) \in \Theta^Q$ arbitrarily while remaining in the affine hull of $T(P) \cup \big(A(\tilde{s}_J), \tilde{\theta}\big)$. In other words, the affine hull of Θ^P is parallel to each u_{ij} coordinate for all i and $j \notin \{0, n\}$.

Last, let $(\dot{u}, \dot{\theta}) \in \Theta^P$ and define $(\dot{s}, \dot{\theta})$ as a point so that $\dot{s}_i$ is the point returned by Algorithm 1 when $\dot{u}_i$ is provided as input. Then, it follows that $(\dot{s}, \dot{\theta}) \in \Theta^Q$. Since $\dot{u}_{ij} = \dot{s}_{ij}$ for all $i \in \{1, \ldots, d\}$ and $j \in \{0, n\}$, it follows readily that $(\dot{u}, \dot{\theta})$ is expressible as an affine combination of $T(P) \cup \big(A(\tilde{s}_J), \tilde{\theta}\big)$. This shows that $\dim\big(T(P)\big) = \dim(\Theta^P) - 1$. ∎

Corollary 2.4.1 *Assume that* $\mathrm{conc}_Q(\phi)(s)$ *is a polyhedral function. Further assume that there is an oracle that for any* $J \in \mathcal{J}$ *and a point in* Q *generates a facet-defining inequality of* Θ^Q *that is tight at that point. Then, there exists a facet generation algorithm for* $\mathrm{conc}_P(\phi)(u)$ *which generates tight cuts in* $\mathcal{O}(dn)$ *besides a call to the above facet generation oracle.*

Proof Let $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_d)$ and $\bar{s} = (\bar{s}_1, \ldots, \bar{s}_d)$, where $\bar{s}_i$ is returned by Algorithm 1 when it is provided with $\bar{u}_i$ as input. Let $J = (J_1, \ldots, J_d)$, where $J_i = \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. The oracle returns a facet-defining inequality $\phi \leq \langle \alpha_J, s_J \rangle + b$ of $\mathrm{conc}_{Q_J}(\phi)(s_J)$ that is tight at $\bar{s}_J$. Let $\tilde{\alpha}$ be a vector in $\mathbb{R}^{d(n+1)}$ such that $\tilde{\alpha}_J = \alpha_J$ and $\tilde{\alpha}_{\bar{J}} = 0$. Then, by Theorem 2.4.2, $\phi \leq \langle \tilde{\alpha}, s \rangle + b$ is facet-defining inequality for $\mathrm{conc}_P(\phi)(u)$. The proof is complete by observing that $\mathrm{conc}_P(\phi)(\bar{u}) = \mathrm{conc}_Q(\phi)(\bar{s}) = \mathrm{conc}_{F_J}(\phi)(\bar{s}) = \langle \tilde{\alpha}, \bar{s} \rangle + b = \langle \tilde{\alpha}, \bar{u} \rangle + b$, where first equality holds by Corollary 2.3.1, second equality is because $\bar{s} \in F_J$, third equality is because of the assumed tightness property of the oracle and the affine isomorphism of Θ^{Q_J} and Θ^{F_J} shown in Lemma 2.4.2. The last equality follows because $\tilde{\alpha}_{\bar{J}} = 0$ and $\bar{u}_J = \bar{s}_J$. ∎

2.5 Conclusions

In this chapter, we tightened the factorable relaxation by proposing a new relaxation framework for mixed-integer nonlinear programs. The framework gives the first structured approach to relax composite functions using the inner-function structure. This is achieved by relaxing the outer-function over a polytope P that encapsulates information about the inner-functions implicit in their estimators. The structure of P is relatively complex in that its extreme points grow exponentially with the number of estimators, even when the number of inner-functions is fixed. Instead, we devised a fast combinatorial algorithm to solve the separation problem over P using an oracle to separate over Q, a much simpler subset of P. For vertex-generated outer-functions, with a fixed number of inner functions, we gave a tractable polyhedral representation for the convex hull over P. When the outer-function is a bilinear term, and each inner-function has one estimator, we developed closed-form expressions for new valid inequalities, generalizing the factorable programming scheme. More specifically, if the inner functions have n_1 and n_2 estimators, we derived $4n_1n_2 + 2n_1 + 2n_2$ inequalities besides the four McCormick inequalities. The new relaxations do not introduce variables beyond those used in the factorable scheme. In Chapter 3, we consider specially structured outer-functions for which convexification of the graph over Q is tractable, and using our results here, devise tractable algorithms for convexification over P. If the convex hull of the graph of outer-function over Q is polyhedral and there is a facet-generation oracle for this graph and some of its subsets, we constructed a facet-generation separation algorithm for this graph over P. Finally, we generalized our results to the setting involving a vector of outer-functions.

3. TRACTABLE RELAXATIONS OF COMPOSITE FUNCTIONS

The sole example of composite relaxations in Chapter 2 for which explicit inequalities are available concerns the product of two bounded functions each furnished with an underestimator. In this chapter, we describe ways in which we extend these results beyond the above example setting. First, we treat a larger class of outer-functions, in particular, those that are supermodular and concave-extendable over Q. Second, we allow arbitrarily many estimators for each function. Formally, for a composite function with d inner-functions, each equipped with n estimators, we devise an algorithm that generates, whenever possible, a facet-defining inequality in $\mathcal{O}(dn \log d)$ time to separate a given point from the hypograph of the outer-function over Q. The number of facet-defining inequalities of this hypograph is $\binom{dn}{n,n,\dots,n}$, which, by Stirling's approximation, grows asymptotically as fast as $\frac{d^{dn+\frac{1}{2}}}{(2n\pi)^{\frac{d-1}{2}}}$, or exponentially with respect to d and n. Though numerous, since these inequalities are generated using a fast combinatorial separation algorithm, they can be derived iteratively, with little computational overhead, to cut off infeasible regions from MINLP relaxations. The geometric structure of these inequalities relates to a certain triangulation of Q, which provides many insights. We show that a result of [8] regarding separability of concave envelopes can be generalized to our setting. We also show that even when additional estimators are available, any inequalities generated using fewer estimators are still facet-defining. We specialize our results to the example setting of Chapter 2, proving that the 12 inequalities therein yield the convex hull of the graph of the bilinear product over P; the polytope which was modeled using one estimator for each inner function. Third, we extend our algorithm to allow simultaneous separation of a vector of composite functions, each with an outer-function that is supermodular and concave-extendable over Q. Fourth, we consider infinitely many estimators for each inner-function, assuming additionally that the outer-function is convex when all but one of its arguments are fixed. We show that, in this case, the composite relaxation arises as the solution of an optimal transport problem [38]. The reduction proceeds by expressing each inner-function as the expectation of a random variable that is completely determined by its estimators. When the outer-function is also supermodular, we provide an explicit integral formula, whose evaluation gives a closed-form expression for the composite relaxation.

3.1 Problem setup and geometric structure

Let $\phi \circ f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$ be a composite function denoted as $\phi \circ f(x) = \phi\bigl(f(x)\bigr)$, where $f : X \mapsto \mathbb{R}^d$ is a vector of *bounded* functions over X. We shall write $f(x) := \bigl(f_1(x), \dots, f_d(x)\bigr)$ and refer to $f(x)$ as inner functions and $\phi(\cdot)$ as the outer function. In Chapter 2, we proposed a framework to relax the graph of $\phi \circ f$, that is, $\mathrm{gr}(\phi \circ f) = \bigl\{(x, \phi) \bigm| \phi = \phi\bigl(f(x)\bigr),\ x \in X\bigr\}$. In this section, we review these ideas and relate our setting to some relevant convexification results. Throughout this chapter, we shall denote the convex hull of set S by $\mathrm{conv}(S)$, the convex (resp. concave) envelope of $f(x)$ over S by $\mathrm{conv}_S(f)$ (resp. $\mathrm{conc}_S(f)$), the projection of a set S to the space of x variables by $\mathrm{proj}_x(S)$, the extreme points of S by $\mathrm{vert}(S)$, the dimension of the affine hull of S by $\dim(S)$, and the relative interior of S by $\mathrm{ri}(S)$.

3.1.1 A relaxation framework for composite functions

Let $n \in \mathbb{Z}$. Let $u : \mathbb{R}^m \mapsto \mathbb{R}^{d\times(n+1)}$ be a vector of bounded functions denoted as $u(x) = \bigl(u_1(x), \dots, u_d(x)\bigr)$ and let $a = (a_1, \dots, a_d)$ be a vector in $\mathbb{R}^{d\times(n+1)}$. For all i and $x \in X$, assume that $f_i(x) \in [a_{i0}, a_{in}]$ and that u and a satisfy the following inequalities:

$$a_{i0} < \cdots < a_{in}, \quad u_{ij}(x) \leq \min\{f_i(x), a_{ij}\}, \quad u_{i0}(x) = a_{i0}, \quad u_{in}(x) = f_i(x).$$

We review some basic ideas regarding the structure of P. First, in Chapter 2, the polytope $P := \prod_{i=1}^{d} P_i$, where

$$P_i = \left\{ u_i \in \mathbb{R}^{n+1} \;\middle|\; \begin{array}{l} u_{ij} \leq u_{in} \text{ and } a_{i0} \leq u_{ij} \leq a_{ij} \\ u_{i0} = a_{i0},\ a_{i0} \leq u_{in} \leq a_{in} \end{array} \right\}, \tag{3.1}$$

was introduced as an *abstraction* of underestimators for the inner-functions $f(x)$. The polytope P_i introduces a variable u_{ij} for each underestimator $u_{ij}(x)$, where u_{in} is a special underestimator that equals $f_i(x)$. Each $u_{ij}(x)$ underestimates the corresponding inner-function $u_{in}(x)$ and is bounded from above by a_{ij}. Thus, $u_{ij} \leq u_{in}$ (resp. $u_{ij} \leq a_{ij}$) models that $u_{ij}(x)$ underestimates $f_i(x)$ (resp. a_{ij}). If a_{i0} is a lower bound for u_{in}, then we are allowed to impose $a_{i0} \leq u_{ij}$. Second, our assumption that each function f_i has n underestimators is without loss of generality. Third, if some of the estimators are overestimators, Chapter 2 gives an affine transformation to reduce the treatment to one involving only underestimators. We denote by Φ^P the graph of the outer function $\phi(u_{1n}, \ldots, u_{dn})$ over P, given formally as

$$\Phi^P = \left\{ (u, \phi) \in \mathbb{R}^{d\times(n+1)} \times \mathbb{R} \;\middle|\; \phi = \phi(u_{1n}, \ldots, u_{dn}),\ u \in P \right\}.$$

The following result shows how the hypograph of $\phi \circ f$ can be relaxed using variables $(x, u_{\cdot n}, \phi)$, where $u_{\cdot n} := (u_{1n}, \ldots, u_{dn})$.

Theorem 3.1.1 (Theorem 2.2.1) *Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a continuous function and let $f : \mathbb{R}^m \mapsto \mathbb{R}^d$ be a vector of functions, each of which is bounded over $X \subseteq \mathbb{R}^m$. If $\big(a, u(x)\big)$ satisfies (3.1.1) then $\operatorname{hyp}(\phi \circ f) \subseteq \operatorname{proj}_{(x,\phi)}(R)$, where $\operatorname{hyp}(\phi \circ f)$ is the hypograph of $\phi \circ f$ over X and*

$$R := \left\{ (x, u_{\cdot n}, \phi) \;\middle|\; \phi \leq \operatorname{conc}_P(\phi)(u),\ u(x) \leq u,\ u_{\cdot n} = f(x),\ x \in X \right\}.$$

If the graph of $f(x)$, expressed using the constraints $u_{\cdot n} = f(x)$ and $x \in X$ in the definition of R, is outer-approximated with a convex set, and, for $j \neq n$, $u_{ij}(x)$ is convex, we obtain a convex relaxation of $\operatorname{hyp}(\phi \circ f)$. Moreover, $\operatorname{conc}_P(\phi)$ is non-increasing in u_{ij}, for all $i \in \{1, \ldots, d\}$ and $j \notin \{0, n\}$. Substituting $u_{ij}(x)$ for such

u_{ij} variables and dropping $u_{ij} \geq u_{ij}(x)$ projects the u_{ij} variables out of the convex relaxation.

The idea behind Theorem 3.1.1 is that the constraints in P are satisfied by underestimators and, as such, inequalities valid for $\text{conv}(\Phi^P)$ are also valid for the underestimators. Since variables u_{ij}, for $j \neq n$, are eventually replaced with their defining function, $u_{ij}(x)$, the relaxation, just like the factorable one, uses the original variables x and an introduced variable u_{in} for each inner function $f_i(x)$.

In this chapter, we will solve the *facet-generation problem* of $\text{conc}_P(\phi)$, assuming that ϕ is supermodular and concave-extendable over P. Under these conditions, the hypograph of $\text{conc}_P(\phi)$ is a polyhedron. By the facet-generation problem of a full-dimensional polyhedron S, we mean that, given a vector y, we establish that either $y \in S$ or find a facet-defining inequality of S that is not satisfied by y. The facet-generation problem for $\text{conc}_P(\phi)$ is, in general, NP-Hard because P includes, as a special case, the unit hypercube and ϕ can be any bilinear function. Nonetheless, on the positive side, Chapter 2 showed that the facet-generation problem for $\text{conc}_P(\phi)$ is tractable if the facet-generation problem for $\text{conc}_Q(\phi)$ and some of its faces is tractable. The polytope Q, which is the domain of the latter function $\text{conc}_Q(\phi)$, is a subset of P that we will describe shortly. Simultaneously, we will also review other results relevant to devising separation algorithms for $\text{conc}_P(\phi)$.

Let $a = (a_1, \ldots, a_d)$ be a vector in $\mathbb{R}^{d\times(n+1)}$ so that each subvector a_i is strictly increasing. Then, $Q := \prod_{i=1}^{d} Q_i$, where Q_i is the simplex in $\mathbb{R}^{n+1}$, whose extreme points are given as follows:

$$v_{ij} := (a_{i0}, \ldots, a_{ij}, a_{ij}, \ldots, a_{ij}) \qquad \text{for } j = 0, \ldots, n. \tag{3.2}$$

We will consider projections of Q defined using a d-tuple of index sets. Let $J = (J_1, \ldots, J_d) \in \mathcal{J}$, where each J_i belongs to the collection:

$$\mathcal{J} := \big\{ (J_1, \ldots, J_d) \;\big|\; \{0, n\} \subseteq J_i \subseteq \{0, \ldots, n\} \; \forall i \in \{1, \ldots, d\} \big\}. \tag{3.3}$$

As a succinct notation, for any $y \in \mathbb{R}^{d\times(n+1)}$, we let $y_J := (y_{1J(1)}, \dots, y_{dJ(n)})$, where $y_{iJ(i)}$ consists of coordinates of y_i from the index-set J_i. We can then write, up to reordering of variables, that $y = (y_J, y_{\bar{J}})$. Define

$$Q_J := Q_{1J(1)} \times \cdots \times Q_{dJ(d)}, \tag{3.4}$$

where $Q_{iJ(i)}$ is the simplex defined in (3.2) with the parameter vector $a_{iJ(i)} \in \mathbb{R}^{|J(i)|}$ and observe that Q_J is a projection of Q to the coordinates contained in the d-tuple J.

Given a point $(\bar{u}, \bar{\phi})$, the separation algorithm for $\text{conc}_P(\phi)$ constructs another point $(\bar{s}, \bar{\phi}) \in \mathbb{R}^{d\times(n+1)+1}$. This point is then separated from $\text{conc}_Q(\phi)$ using the separation oracle for one of its faces. We first describe how $(\bar{s}, \bar{\phi})$ is obtained from $(\bar{u}, \bar{\phi})$. With each point $u_i \in \mathbb{R}^{n+1}$, we associate a discrete univariate function $\xi(a; u_i) : [a_{i0}, a_{in}] \mapsto \mathbb{R}$ defined so that $\xi(a; u_i) = u_{ij}$ if $a = a_{ij}$ for some $j \in \{0, \dots, n\}$ and $\xi(a; u_i) = -\infty$ otherwise. Then, $\bar{s} = (\bar{s}_1, \dots, \bar{s}_d)$, where

$$\bar{s}_i = \big(\text{conc}(\xi)(a_{i0}; \bar{u}_i), \dots, \text{conc}(\xi)(a_{in}; \bar{u}_i)\big),$$

and $\text{conc}(\xi)(\cdot; u_i)$ is the concave envelope of $\xi(\cdot; u_i)$ over $[a_{i0}, a_{in}]$. So, $\bar{u}_i$ (resp. $\bar{s}_i$) is the vector of values of $\xi(\chi; u_i)$ (resp. $\text{conc}(\xi)(\chi; u_i)$) generated by sequentially setting χ to the values in $(a_{i0}, \dots, a_{in})$. For each $\bar{u} \in P$, the corresponding $\bar{s}$ is captured in the set:

$$PQ' := \big\{(u, s) \;\big|\; (u_1, \dots, u_d) \in P,\ s_i := \big(\text{conc}(\xi)(a_{i0}; u_i), \dots, \text{conc}(\xi)(a_{in}; u_i)\big)\ \forall i\big\}. \tag{3.5}$$

Since $\bar{u}$ and $\bar{s}$ are related via a concave envelope construction, we can lift $\bar{u}$ to its unique lifting $(\bar{u}, \bar{s}) \in PQ'$ using a two-dimensional convex hull algorithm, such as Graham scan [50], that, for each of the d discrete univariate functions $\xi(a; u_i)$ finds its envelope in $\mathcal{O}(n)$ time. Given $\bar{u} \in P$, we will find a face of Q containing $\bar{s}$ and, consequently, identify which face of $\text{conc}_Q(\phi)$ is to be separated from $(\bar{s}, \bar{\phi})$. Observe that for any $\bar{u}_i \notin Q_i$, $\bar{u}_i$ violates certain facet-defining inequalities of Q_i. As the next result shows, these inequalities define facets, whose intersection yields a face of Q_i containing $\bar{s}_i$. Together, for all i, these faces define the face of interest of Q, which is described using d-tuples of index sets. Thus, we associate with a d-tuple, $J = (J_1, \ldots, J_d) \in \mathcal{J}$, the following face F_J of Q:

$$F_J := F_{1J(1)} \times \cdots \times F_{dJ(d)},$$

where $F_{iJ(i)} := \operatorname{conv}\bigl(\{v_{ij} \mid j \in J_i\}\bigr)$ and, for all i and j, $v_{ij} = (a_{i0}, \ldots, a_{ij-1}, a_{ij}, \ldots, a_{ij})$. Clearly, $F_{iJ(i)}$ is a face of Q_i because Q_i is a simplex whose vertices form a superset of those of $F_{iJ(i)}$. With J, we also associate a linear map, $\Gamma_J : \mathbb{R}^{d\times(n+1)} \mapsto \mathbb{R}^{d\times(n+1)}$ that maps a $u \in P$ to $\tilde{u} \in P$ as follows:

$$\tilde{u}_{ij} = u_{ij} \quad \text{for } j \in J_i \qquad \text{and} \qquad \tilde{u}_{ij} = (1-\gamma_{ij})u_{il(i,j)} + \gamma_{ij}u_{ir(i,j)} \quad \text{for } j \notin J_i, \tag{3.6}$$

where $l(i,j) := \max\{j' \in J_i \mid j' \leq j\}$, $r(i,j) := \min\{j' \in J_i \mid j' \geq j\}$, and, for $j \notin J_i$, $\gamma_{ij} = (a_{ij} - a_{ir(i,j)})/(a_{ir(i,j)} - a_{il(i,j)})$. In other words, $\tilde{u}_{ij}$ is obtained by restricting the domain of $\xi(a; u_i)$, a function of a, to $a_{iJ(i)}$ and then linearly interpolating the function at a_{ij} for $j \notin J_i$.

Proposition 3.1.1 (Proposition 2.3.2) *Let $(\bar{u}, \bar{s}) \in PQ'$ and $J = (J_1, \ldots, J_d)$ be defined so that $J_i := \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. Then, $\bar{s} = \Gamma_J(\bar{u})$. The set Q satisfies the inequalities $s \geq \Gamma_J(s)$. Then F_J is the face of Q defined by the inequalities $s \leq \Gamma_J(s)$. Moreover, $\bar{s} \in F_J$.*

Given $(\bar{u}, \bar{\phi})$, we call the facet-generation oracle of $\operatorname{conc}_{Q_J}(\phi)(s_J)$ to generate an inequality $\phi \leq \langle \alpha_J, s_J\rangle + b$ of $\operatorname{conc}_{Q_J}(\phi)(s_J)$, which is tight at $\bar{s}$, *i.e.*, $\operatorname{conc}_{Q_J}(\phi)(\bar{s}_J) = \langle \alpha_J, \bar{s}_J\rangle + b$. Let $\tilde{\alpha}$ be a vector in $\mathbb{R}^{d\times(n+1)}$ such that $\tilde{\alpha}_J = \alpha_J$ and $\tilde{\alpha}_{\bar{J}} = 0$. By Corollary (2.3.1), the inequality $\phi \leq \bigl\langle (\alpha_J, 0), (s_J, s_{\bar{J}})\bigr\rangle + b$ defines a facet of $\operatorname{conc}_P(\phi)$ that is tight at $\bar{u}$. We summarize the above discussion for later use.

Proposition 3.1.2 *Assume that $\operatorname{conc}_Q(\phi)(s)$ is a polyhedral function. Given $\bar{u} \in P$, the unique point $(\bar{u}, \bar{s}) \in PQ'$ can be found in $\mathcal{O}(dn)$ time. Let $J = (J_1, \ldots, J_d)$, where $J_i := \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. Then, $\bar{s} \in F_J$ and $\bar{s}_J \in Q_J$. If $\phi \leq \langle \alpha_J, s_J\rangle + b$ is a facet-defining inequality for $\operatorname{conc}_{Q_J}(s)$ that is tight at $\bar{s}_J$ then the inequality $\phi \leq \bigl\langle (\alpha_J, 0), (u_J, u_{\bar{J}})\bigr\rangle + b$ defines a facet of $\operatorname{conc}_P(\phi)$ that is tight at $\bar{u}$.*

3.1.2 Supermodularity and staircase triangulation

This chapter considers supermodular and concave-extendable functions over a Cartesian product of simplices, Q. Such functions have a concave envelope that is closely related to certain triangulations of Q. In this subsection, we explore these connections. Before we begin, we formally define concave-extendability of functions [47] and triangulations of polyhedral domains [52]; both are prevalent notions in convexification literature.

Definition 3.1.1 ([16]) *A function $g : D \mapsto \mathbb{R}$, where D is a polytope, is said to be concave-extendable (resp. convex-extendable) from $X \subseteq D$ if the concave (resp. convex) envelope of $g(x)$ is determined by X only, that is, the concave envelope of g and $g|_X$ over P are identical, where $g|_X$ is the restriction of g to X that is defined as:*

$$g|_X = \begin{cases} g(x) & x \in X \\ -\infty & \textit{otherwise.} \end{cases}$$

Definition 3.1.2 (Triangulation [52]) *Let $D \subseteq \mathbb{R}^n$. A set of polyhedra $\mathcal{R} := \{R_1, \ldots, R_r\}$ forms a* polyhedral subdivision *of D if $D = \bigcup_{i=1}^{r} R_i$ and $R_i \cap R_j$ is a (possibly empty) face of both R_i and R_j. Moreover, if each R_i is a simplex, then $\mathcal{R}$ is a* triangulation *of D.*

The non-vertical facets of a polyhedral function, when projected, divide the domain into polyhedral sets, which form a polyhedral subdivision; a subdivision that can be further refined into a triangulation. Thus, if the concave envelope of $\phi(\cdot)$ is polyhedral over Q, there is a triangulation $\mathcal{R}$ of Q such that the concave envelope affinely interpolates each simplex R_i of this triangulation (Theorem 2.4 in [16]). By affine interpolation, we mean that the function value at any point $s \in R_i$ is obtained as the affine combination of function values at $\text{vert}(R_i)$. Therefore, the concave envelope of $\phi(\cdot)$ over Q, when polyhedral, is uniquely described by the triangulation $\mathcal{R}$ of Q.

We will eventually be interested in extending the domain of the concave envelope outside of Q. To do so, we will use the following construction, which extends the domain of $\text{conc}_Q(\phi)$ to $\text{aff}(Q)$, a set that contains P. We describe this construction for a generic function $\chi : D \mapsto \mathbb{R}$, whose domain, D, is a subset of $\mathbb{R}^n$ and this domain is assumed to be endowed with a triangulation $\mathcal{R}$. Define $\chi^{R_i}(x) : \text{aff}(D) \mapsto \mathbb{R}$ as the unique affine function that satisfies $\chi^{R_i}(x) = \chi(x)$ for all $x \in \text{vert}(R_i)$. Moreover, define $h(x)$ so that, for all $x \in R_i \in \mathcal{R}$, $h(x) = \chi^{R_i}(x)$ and assume that it is concave. Now, to extend $h(x)$ to $\text{aff}(D)$, we consider another function $\chi^{\mathcal{R}}(x)$ defined as $\min_i \chi^{R_i}(x)$ and show that it matches $h(x)$ over D. If not, there exists some $(i, j) \in \{1, \ldots, r\}^2$ and an x such that although $x \in R_i$, $\chi^{R_i}(x) > \chi^{R_j}(x)$. Now, pick $y \in \text{int}(R_j)$ and a sufficiently small $\epsilon > 0$ so that $y + \epsilon(x - y) \in R_j$. Then, as the following argument shows, $h(x)$ cannot be concave, violating our assumption:

$$\begin{aligned} h(y) + \epsilon\big(h(x) - h(y)\big) &= \chi^{R_j}(y) + \epsilon\big(\chi^{R_i}(x) - \chi^{R_j}(y)\big) \\ &> \chi^{R_j}(y) + \epsilon\big(\chi^{R_j}(x) - \chi^{R_j}(y)\big) \\ &= \chi^{R_j}\big(y + \epsilon(x - y)\big) = h\big(y + \epsilon(x - y)\big), \end{aligned} \tag{3.7}$$

where the first equality is by the definition of h, the first inequality is because $\chi^{R_i}(x) > \chi^{R_j}(x)$ and $\epsilon > 0$, the second equality is because χ^{R_j} is affine, and the last equality is by the definition of h.

We turn our attention now to a specific triangulation of a product of simplices, referred to as the *staircase triangulation*. Let $S_i \subseteq \mathbb{R}^n$ be a simplex so that $\text{vert}(S_i) = \{\nu_{i0}, \ldots, \nu_{in}\}$ and let $S := \prod_{i=1}^d S_i$ be the Cartesian product of these simplices. The extreme points of S, which are $\prod_{i=1}^d \{\nu_{i0}, \ldots, \nu_{in}\}$, can then be depicted on the grid $\mathcal{G}$ given by $\{0, \ldots, n\}^d$. More specifically, the extreme point $(\nu_{ij_i})_{i=1}^d$ will be associated with the grid-point $\{j_i\}_{i=1}^d$. The coordinates of the extreme point are recovered from those of the grid-point using *grid-labels*, markers that label coordinate j along direction i as ν_{ij}. We remark that although grid-labels depend on the specific geometry of S, the grid only depends on the number of simplices and the dimension of each simplex.

A *monotone staircase* is a set of $dn+1$ ordered points $(p_0, \ldots, p_{dn})$, where $p_i \in \mathcal{G}$ for all $i \in \{0, \ldots, dn\}$ and the sequence satisfies the following properties, (i) $p_0 = (0, \ldots, 0)$ and (ii) for all $i \in \{1, \ldots, dn\}$, $p_i - p_{i-1} = e_k$, where $k \in \{1, \ldots, d\}$ and e_k denotes the k^{th} principal vector in $\mathbb{R}^d$. We refer to movement from p_{i-1} to p_i as the i^{th} move. Since $p_i \in \mathcal{G}$ for all i, by property (ii) there are exactly n moves in each coordinate direction. Moreover, $p_{dn} = (n, \ldots, n)$. Thus, the monotone staircase is a lattice path of monotonically increasing points in $\mathbb{Z}^n$ from $(0, \ldots, 0)$ to $(n, \ldots, n)$, hence resembling a staircase, where each step is of possibly different height. The staircase can be specified succinctly as a vector $\pi = (\pi_1, \ldots, \pi_{dn})$, where $\pi_i \in \{1, \ldots, d\}$ is the coordinate direction of the i^{th} move. Thus, we will refer to such vector π as *movement* vector in the grid $\mathcal{G}$. Given a vector π, we will often need to track where the k^{th} move leaves us on the grid. This is obtained using the transformation Π, which is defined as $\Pi(\pi, k) := p_0 + \sum_{j=1}^{k} e_{\pi_j}$, where e_i is the i^{th} principal vector in $\mathbb{R}^d$. The corresponding staircase can then be recovered as $\big(\Pi(\pi, k)\big)_{k=0}^{dn}$. In Figure 3.1, we see the set of all monotone staircases on the 4×3 gird.

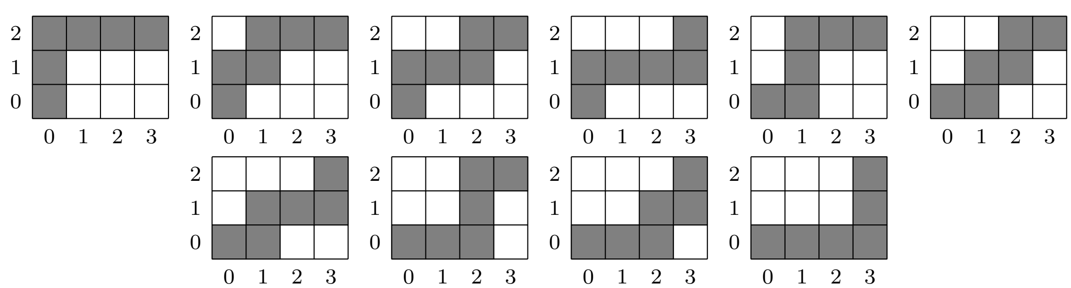


Fig. 3.1.: Monotone staircases in the 4×3 grid.

As described before, with each grid point $p_k = (j_i)_{i=1}^{d}$, we associate the extreme point of S, $(\nu_{ij_1}, \ldots, \nu_{dj_d})$, which we denote as $\operatorname{ext}(S, p_k)$. The set of all extreme points associated with the grid-points along a staircase π describe a simplex Ξ_π. In particular, if we let $p_k = \Pi(\pi, k)$ then $\Xi_\pi := \operatorname{conv}\big(\bigcup_{k=0}^{dn} \operatorname{ext}(S, p_k)\big)$. That this set defines a simplex follows from the affine independence of $\operatorname{ext}(S, p_k)$ for $k \in \{0, \ldots, dn\}$, which in turn follows from linear independence of difference vectors $\operatorname{ext}\big(S, p_{k+1}\big) -$

$\text{ext}\big(S, p_k\big)$ for $k = 0, \ldots, dn-1$. In particular, the difference vectors for a move k along i^{th} and another move k' along i' coordinate, where $i' \neq i$, are linearly independent because these vectors are non-zero along different variables. On the other hand, the difference vectors for moves, all of which are along the i^{th} coordinate, are linearly independent because they are of the form $(0_1, \ldots, 0_{i-1}, \nu_{ij+1} - \nu_{ij}, 0_{i+1}, \ldots, 0_d)$, where $0_{i'}$ is the zero vector in the subspace of variables defining $S_{i'}$ and $\nu_{ij+1} - \nu_{ij}$ are the difference vectors between adjacent extreme points of S_i.

Definition 3.1.3 (Staircase triangulation [52]**)** *The set of all monotone staircases in the grid $\mathcal{G}$ defines a staircase triangulation of $\prod_{i=1}^{d} S_i$. Each monotone staircase, defined by the movement vector π, yields a simplex, $\Xi(\pi)$, in this triangulation.*

We argue that the staircase triangulation is a triangulation of S. Observe that the $\mathbb{R}^{n\times n}$ matrix, $M_i := (\nu_{i1} - \nu_{i0}, \ldots, \nu_{in} - \nu_{i0})$, is invertible because S_i is a full-dimensional simplex. Then, we consider the affine mapping that maps an $x \in \mathbb{R}^n$ to $UM_i^{-1}(x - \nu_{i0})$, where $U \in \mathbb{R}^{n\times n}$ is an upper-triangular matrix of all ones. Under this transformation, the simplex S_i maps to $\Lambda_i := \{z_i \in \mathbb{R}^n \mid 0 \le z_{in} \le \cdots \le z_{i1} \le 1\}$. Given a point $s = (s_1, \ldots, s_n) \in S$, we obtain $z = (z_1, \ldots, z_n)$, where $z_i = UM_i^{-1}(s_i - \nu_{i0})$. Then, we sort the coordinates of the vector z in a non-increasing order so that if $z_{ij} = z_{ij'}$ for some i and $j' > j$, we place z_{ij} ahead of $z_{ij'}$ in the ordering. Now, for any $k \in \{1, \ldots, dn\}$, if z_{ij} is the k^{th} order-statistic, we define $\Theta(k) = \big(\Theta_1(k), \Theta_2(k)\big) := (i, j)$. Given a movement vector π, Θ can be recovered using $\Theta_1(k) = \pi_k$, and $\Theta_2(k) = \big|\{j \mid \pi(j) = \pi(k), 1 \le j \le k\}\big|$. Then, we verify that z (resp. s) belongs to the simplex whose extreme points form the monotone staircase defined by the movement vector $\pi' = \big(\Theta_1(k)\big)_{k=1}^{dn}$, *i.e.*, $\text{conv}\big(\bigcup_{k=0}^{dn} \text{ext}\big(\Lambda, \Pi(\pi', k)\big)\big)$ (resp. $\text{conv}\big(\bigcup_{k=0}^{dn} \text{ext}(S, \Pi(\pi', k))\big)$). We denote $\text{ext}\big(\Lambda, \Pi(\pi', k)\big)$ as p_k. Note that $p_0 = (0_1, \ldots, 0_d)$, where 0_i is a zero vector in the space of z_i variables and $p_k = p_{k-1} + e_{\Theta(k)}$, where $e_{\Theta(k)}$ is a unit vector in the direction of $z_{\Theta(k)}$. Then,

$$z = (1 - z_{\Theta(1)})p_0 + \sum_{k=1}^{dn-1} \big(z_{\Theta(k)} - z_{\Theta(k+1)}\big)p_k + z_{\Theta(dn)}p_{dn}. \tag{3.8}$$

Thus, each point z belongs to some simplex in the staircase triangulation. Moreover, points that belong to two simplices have at least two consistent orderings of their coordinates, ensuring some of them are equal, implying that the point belongs to a common face of both simplices. Thus, monotone staircases triangulate S.

If a function $f : [0,1]^n \mapsto \mathbb{R}$ is *supermodular* when restricted to $\{0,1\}^n$ and concave extendable from $\{0,1\}^n$ then $\mathrm{conc}_{[0,1]^n}(f)(x)$ coincides with the *Lovász extension* of f [16, 53]. The Lovász extension interpolates the staircase triangulation of $[0,1]^n$, which can be regarded as a Cartesian product of simplices. In this setting, the corresponding triangulation is referred to as *Kuhn's triangulation*. We relate the concave envelope over a product of simplices to that of certain subsets of $[0,1]^n$, also treated in [16].

Definition 3.1.4 ([54]) *A function $\eta(x) : S \subseteq \mathbb{R}^n \mapsto \mathbb{R}$ is said to be supermodular if $\eta(x' \vee x'') + \eta(x' \wedge x'') \geq \eta(x') + \eta(x'')$ for all x', $x'' \in S$. Here, $x' \vee x''$ denotes the component-wise maximum and $x' \wedge x''$ denotes the component-wise minimum of x' and x'' and S is assumed to be a lattice, that is, $x' \vee x''$ and $x' \wedge x''$ belong to S whenever x' and x'' belong to S.*

Proposition 3.1.3 *Let $\Lambda_i = \{z_i \mid 0 \leq z_{i1} \leq \cdots \leq z_{in} \leq 1\}$ and $\Lambda = \prod_{i=1}^{d} \Lambda_i$. Then,* $\mathrm{vert}(\Lambda)$ *defines a lattice. Let $\eta : \Lambda \mapsto \mathbb{R}$ be a supermodular function when restricted to* $\mathrm{vert}(\Lambda)$. *Then, the concave envelope of η over Λ is given by the staircase triangulation of Λ, i.e.* $\mathrm{conc}_{\Lambda}(\eta)(x) = \eta^{\mathcal{S}}(x)$ *for every $x \in \Lambda$, where $\mathcal{S}$ is the staircase triangulation of Λ.*

Proof Since $\mathrm{vert}(\Lambda_i)$ forms a chain and $\mathrm{vert}(\Lambda)$ is a Cartesian product of these chains, $\mathrm{vert}(\Lambda)$ forms a lattice. Using Corollary 3.4 in [16], the concave envelope of η over Λ is obtained by affine interpolation of simplices obtained by ordering z in specific linear orders in a manner consistent with how they are ordered within each Λ_i. The so linearly ordered z belongs to one of the simplices in the staircase triangulation of Λ. ∎

Although detecting whether a function is supermodular is NP-Hard [55], there are important special cases where this property can be readily detected [54]. A particularly useful result establishes supermodularity of a composition of functions.

Lemma 3.1.1 (Lemma 2.6.4 in [54]) *Consider a lattice X and let $K = \{1, \ldots, k\}$. For $i \in K$, let $f_i(x)$ be increasing supermodular (resp. submodular) functions on X and Z_i be convex subsets of $\mathbb{R}$. Assume $Z_i \supseteq \{f_i(x) \mid x \in X\}$. Let $g(z_1, \ldots, z_k, x)$ be supermodular in $(z_1, \ldots, z_k, x)$ on $\prod_{i=1}^{k} Z_i \times X$. If for $i \in K$, $\bar{z}_{i'} \in Z_{i'}$ for $i' \in K \setminus \{i\}$, and $\bar{x} \in X$, $g(\bar{z}_1, \ldots, \bar{z}_{i-1}, z_i, \bar{z}_{i+1}, \ldots, \bar{z}_k, \bar{x})$ is increasing (resp. decreasing) and convex in z_i on Z_i then $g\big(f_1(x), \ldots, f_k(x), x\big)$ is supermodular on X.*

By choosing $g(z_1, \ldots, z_k, x)$ appropriately as $z_1 z_2 \cdots z_k$ or $-z_1 z_2 \cdots z_k$, it follows that a product of nonnegative, increasing (resp. decreasing) supermodular functions is also nonnegative increasing (resp. decreasing) and supermodular; see Corollary 2.6.3 in [54]. Also, it follows trivially that a conic combination of supermodular functions is supermodular.

3.2 On finitely many estimators for inner functions

In this section, we devise a facet generation algorithm, Algorithm 3, that separates $\operatorname{conc}_P(\phi)$ assuming ϕ satisfies some assumptions and takes $\mathcal{O}(dn \log d)$ time. As a result, we discover various interesting properties of the envelope $\operatorname{conc}_P(\phi)$. In Section 3.2.2, we develop a decomposition result that applies to a class of bilinear functions and show that for a function ϕ from this class, $\operatorname{conc}_Q(\phi)$ is obtained by aggregating concave envelopes of each bilinear term, and thus generalize a result of [8]. Last, we use Algorithm 3 to solve the facet generation problem of a vector of functions over P.

3.2.1 Tractable concave envelopes

In this subsection, we will, under previously stated technical conditions on ϕ, characterize, in closed form, the concave envelope of $\phi(s_{1n}, \ldots, s_{dn})$ over Q and solve the facet generation problem of $\text{conc}_P(\phi)(u)$ in $\mathcal{O}(dn \log d)$ time. Specializing to a multilinear function ϕ, we find that $\text{conc}_P(\phi)$, the envelope over P, is obtained simply by extending the domain, from Q to P, of certain inequalities that describe $\text{conc}_Q(\phi)$, the envelope over Q. Further specializing our result to a bilinear term ϕ with only on non-trivial underestimator for each inner function, we recover Theorems 2.1.1 and 2.3.3 and show that the inequalities describe the convex hull of graph of ϕ in this setting.

Lemma 2.3.1 in [56] shows that for each $i = \{1, \ldots, d\}$, the simplex Q_i can be expressed as:

$$0 \le \frac{s_{in} - s_{in-1}}{a_{in} - a_{in-1}} \le \cdots \le \frac{s_{i1} - s_{i0}}{a_{i1} - a_{i0}} \le 1, \quad s_{i0} = a_{i0}.$$

Observe that, because of the last equality, Q_i is not full-dimensional. Nevertheless, we can mimic the construction following Definition 3.1.3 using the above representation of Q_i. Here, for an $s_i \in Q_i$ we define $z_i \in \Delta_i := \{z \mid 0 \le z_{in} \le \cdots \le z_{i1} \le z_{i0} = 1\}$ as follows:

$$z_{i0} = 1 \qquad \text{and} \qquad z_{ij} = \frac{s_{ij} - s_{ij-1}}{a_{ij} - a_{ij-1}} \quad \text{for } j = 1, \ldots, n. \tag{3.9}$$

We refer to the above transformation as Z_i so that $z_i = Z_i(s_i)$. The inverse of Z_i is then defined as:

$$s_{ij} = a_{i0} z_{i0} + \sum_{k=1}^{j} (a_{ik} - a_{ik-1}) z_{ik} \quad \text{for } j = 0, \ldots, n. \tag{3.10}$$

Unlike the treatment following Definition 3.1.3, here z has $(n+1)d$ indices. To address this discrepancy, we impose additional condition on how z is sorted. We require that the first d entries, when z is sorted in non-increasing order, be z_{i0}, $i = 1, \ldots, d$. Then, if the $(d+k)^{\text{th}}$ variable in this order is z_{ij}, we associate with $\pi_k = i$, that is a movement which steps from $j-1$ to j along the i^{th} direction. Thus, movement vector π describes the simplex of the staircase triangulation that contains s.

The transformation, $\big(Z_1(s_1), \ldots, Z_n(s_n)\big)$, which maps $Q := \prod_{i=1}^d Q_i$ to $\Delta := \prod_{i=1}^d \Delta_i$, has a few properties that will be useful in our development. First, being an affine transformation, it maps vertices of Q to those of Δ. Recall that $\mathrm{vert}(Q_i) = \{v_{i0}, \ldots, v_{in}\}$, where $v_{ij} = (a_{i0}, \ldots, a_{ij-1}, a_{ij}, \ldots, a_{in})$. After the transformation, the vertex v_{ij} maps to $\zeta_{ij} = \sum_{j'=0}^{j} e_{ij'}$, where $e_{ij'}$ is the j' principal vector in the space spanned by variables $(z_{i0}, \ldots, z_{in})$. Clearly, $\mathrm{vert}(\Delta_i)$ form a chain, where $\zeta_{i0} \leq \cdots \leq \zeta_{in}$. Consequently, we show that $\mathrm{vert}(Q_i)$ also form a chain, where the vertices are ordered as $v_{i0} \leq \cdots \leq v_{in}$. This is because (3.10) is an increasing mapping. More specifically, if $z_i, z'_i \in \Delta_i$ such that $z_i \geq z'_i$, then $z_{i0} = 1$ and $a_{ik} - a_{ik-1} > 0$ imply that $s_i = Z_i^{-1}(z_i) \geq Z_i^{-1}(z'_i) = s'_i$. Then, the definition of $\vee$ and $\wedge$ as coordinate-wise maximum and minimum and the observation that $\mathrm{vert}(Q_i)$ form a chain together imply

$$Z_i^{-1}(\zeta_{ij} \vee \zeta_{ij'}) = Z_i^{-1}(\zeta_{ij}) \vee Z_i^{-1}(\zeta_{ij'}) \quad Z_i^{-1}(\zeta_{ij} \wedge \zeta_{ij'}) = Z_i^{-1}(\zeta_{ij}) \wedge Z_i^{-1}(\zeta_{ij'}) \tag{3.11}$$

$$Z_i(s_{ij} \vee s_{ij'}) = Z_i(s_{ij}) \vee Z_i(s_{ij'}) \quad Z_i(s_{ij} \wedge s_{ij'}) = Z_i(s_{ij}) \wedge Z_i(s_{ij'}), \tag{3.12}$$

In other words, Z_i^{-1} (resp. Z) distributes over $\vee$ and $\wedge$ as long as the arguments are vertices of Δ_i (resp. Q_i). The vertex $(v_{ij_i})_{i=1}^d$ maps to $(\zeta_{ij_i})_{i=1}^d$ under Z_i and will be represented by the same grid-point $(j_i)_{i=1}^d$ on $\mathcal{G}$.

It follows that $\mathrm{vert}(Q)$ and $\mathrm{vert}(\Delta)$ form isomorphic lattices, *i.e.*, Cartesian product of d chains each with $n+1$ elements. As such $\mathrm{vert}(Q)$ and $\mathrm{vert}(\Delta)$ satisfy the conditions in Definition 3.1.4 and that allows us to define supermodular functions over these domains. We exploit the isomorphism of the two lattices to relate a supermodular function ψ over $\mathrm{vert}(Q)$ to another function η which is supermodular over $\mathrm{vert}(\Delta)$.

Lemma 3.2.1 *A function $\psi : \mathrm{vert}(Q) \mapsto \mathbb{R}$ is supermodular over $\mathrm{vert}(Q)$ if and only if the function $\eta : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ defined by $\eta(z) = \psi(Z^{-1}(z))$ is supermodular over $\mathrm{vert}(\Delta)$.*

Proof We show that $\eta(z)$ is supermodular over $\mathrm{vert}(\Delta)$ when $\psi(s)$ is supermodular over $\mathrm{vert}(Q)$. Consider z' and z'' in $\mathrm{vert}(\Delta)$ and observe that:

$$\begin{aligned}\eta(z') + \eta(z'') &= \psi(Z^{-1}(z')) + \psi(Z^{-1}(z'')) \\ &\le \psi\big(Z^{-1}(z') \vee Z^{-1}(z'')\big) + \psi\big(Z^{-1}(z') \wedge Z^{-1}(z'')\big) \\ &= \psi\big(Z^{-1}(z' \vee z'')\big) + \psi\big(Z^{-1}(z' \wedge z'')\big) \\ &= \eta(z' \vee z'') + \eta(z' \wedge z''),\end{aligned}$$

where the first quality is from the definition of η, the first inequality follows from the supermodularity of ψ over $\mathrm{vert}(Q)$, the second equality is implied by equalities in (3.11), and the last equality holds by definition of η. The converse follows similarly by observing that $\eta(Z(s)) = \psi(Z^{-1}(Z(s))) = \psi(s)$ and utilizing (3.12) instead of (3.11). ■

Theorem 3.2.1 *Assume that $\phi(s_{1n}, \ldots, s_{dn})$ is concave-extendable from* $\mathrm{vert}(Q)$ *and is supermodular when restricted to the lattice set* $\mathrm{vert}(Q)$*. Let $\mathcal{S}$ be the staircase triangulation of Q. Then,* $\mathrm{conc}_Q(\phi)(s) = \phi^{\mathcal{S}}(s)$ *for each $s \in Q$, where $\phi^{\mathcal{S}}(s)$ interpolates ϕ over vertices of a simplex in $\mathcal{S}$ that contains s.*

Proof By Lemma 3.2.1, the function $\eta(z) = \phi\big((Z_1^{-1}(z_1))_n, \ldots, (Z_d^{-1}(z_d))_n\big)$ is supermodular over $\mathrm{vert}(\Delta)$ since ϕ is supermodular when restricted to the lattice set $\mathrm{vert}(Q)$. It is also concave extendable over $\mathrm{vert}(\Delta)$ since the affine transformation that maps $\{(s, \phi) \mid \phi \le \phi(s_{1n}, \ldots, s_{dn}), s \in Q\}$ to $\{(z, \eta) \mid \eta \le \eta(z_{1n}, \ldots, z_{dn}), z \in \Delta\}$ preserves this property.

By Proposition 3.1.3, the concave envelope of η over Δ is determined by the staircase triangulation $\mathcal{K}$ of Δ. The same triangulation also gives the concave envelope of ϕ over Q because the two hypographs, that of η over Δ and of ϕ over Q, are related via the affine transformation Z. It is just that, for a monotone staircase specified by movement vector π, the simplex in the triangulation of Δ is $\mathrm{conv}\big(\bigcup_{k=1}^{dn} \mathrm{ext}(\Delta, \Pi(\pi, k))\big)$ while that in the triangulation of Q is $\mathrm{conv}\big(\bigcup_{k=1}^{dn} \mathrm{ext}(Q, \Pi(\pi, k))\big)$. ■

Remark 3.2.1 *We show that supermodularity of ϕ over the lattice set* $\text{vert}(Q)$ *follows from its supermodularity over $\prod_{i=1}^{d}\{a_{i0},\dots,a_{in}\}$, which in turn follows from its supermodularity over $\prod_{i=1}^{d}[a_{i0},a_{in}]$. That $\prod_{i=1}^{d}\{a_{i0},\dots,a_{in}\}$ is a lattice follows easily, but, this lattice is also a projection of lattice of points in* $\text{vert}(Q)$ *onto the space of $(s_{1n},\dots,s_{dn})$ variables since the operators $\vee$ and $\wedge$ find maximum (resp. minimum) along each component independently. It therefore suffices to check supermodularity of ϕ over $\prod_{i=1}^{d}\{a_{i0},\dots,a_{in}\}$ because ϕ depends only one $(s_{1n},\dots,s_{dn})$. Then, since $\prod_{i=1}^{d}\{a_{i0},\dots,a_{in}\}\subseteq\prod_{i=1}^{d}[a_{i0},a_{in}]$, if ϕ is shown to be the supermodular over $\prod_{i=1}^{d}[a_{i0},a_{in}]$, its supermodularity over* $\text{vert}(Q)$ *follows.*

Next, we explicitly derive the inequalities that interpolate the function ϕ over the extreme points of each simplex in the staircase triangulation. We first recall connections between monotone staircases in the grid and simplices of the triangulation as it applies in this context. Let Ω be the set of movement vectors that define monotone staircases over the grid $\mathcal{G}$ given by $\{0,1,\dots,n\}^d$. For $\pi\in\Omega$, the k^{th} extreme point $\text{ext}(Q,\Pi(\pi,k))$ will be denoted as $\mathcal{V}(\pi,k)$. The corresponding simplex $\text{conv}\big(\mathcal{V}(\pi,0),\dots,\mathcal{V}(\pi,dn)\big)$ will be denoted as Υ_π and the triangulation $\{\Upsilon_\pi\}_{\pi\in\Omega}$ as Υ. In addition, we define $m(i,j)=k$ if $\pi(k)=i$ and $j=\sum_{k'\le k}\mathbb{1}\big(\pi(k')=\pi(k)\big)$, *i.e.*, for a pair (i,j), $m(i,j)$ returns k if the k^{th} movement is the j^{th} step in coordinate direction i. Observe that $\mathcal{V}\big(\pi,m(i,j)\big)-\mathcal{V}\big(\pi,m(i,j)-1\big)=(0_1,\dots,0_{i-1},v_{ij}-v_{ij-1},0_{i+1},\dots,0_d)$, where 0_k is the zero vector in the space of s_k variables. If $\mathcal{V}(\pi,k)=(s_1,\dots,s_d)$, we denote $(s_{1n},\dots,s_{dn})$ by $\mathcal{V}_{\cdot n}(\pi,k)$. Let $\langle\alpha^\pi,s\rangle+\beta^\pi$ be the unique affine function so that, for all i, $\alpha^\pi_{i0}=0$ and $\phi(s_{1n},\dots,s_{dn})=\langle\alpha^\pi,s\rangle+\beta^\pi$ for $s\in\text{vert}(\Upsilon_\pi)$. Then, to derive expressions for α^π,β^π observe that:

$$\begin{aligned}\phi\Big(\mathcal{V}_{\cdot n}\big(\pi,m(i,j)\big)\Big)&-\phi\Big(\mathcal{V}_{\cdot n}\big(\pi,m(i,j)-1\big)\Big)\\&=\Big\langle\alpha^\pi,\mathcal{V}_{\cdot n}\big(\pi,m(i,j)\big)-\mathcal{V}_{\cdot n}\big(\pi,m(i,j)-1\big)\Big\rangle\\&=\langle\alpha^\pi_i,v_{ij}-v_{ij-1}\rangle=(a_{ij}-a_{ij-1})\sum_{j'=j}^{n}\alpha^\pi_{ij'}.\end{aligned}$$

Let $\vartheta(\pi,i,j)=\frac{\phi(\mathcal{V}_{\cdot n}(\pi,m(i,j)))-\phi(\mathcal{V}_{\cdot n}(\pi,m(i,j)-1))}{a_{ij}-a_{ij-1}}$. Then, by differencing the equations and fitting the equation at $\mathcal{V}(\pi,0)$, we obtain the following explicit formulae:

$$\alpha_{ij}^{\pi}=\begin{cases}0 & j=0\\ \vartheta(\pi,i,j)-\vartheta(\pi,i,j+1) & 1\le j<n\\ \vartheta(\pi,i,n) & j=n\end{cases}$$
$$b^{\pi}=\phi\big(\mathcal{V}_{\cdot n}(\pi,0)\big)-\big\langle\alpha^{\pi},\mathcal{V}(\pi,0)\big\rangle. \tag{3.13}$$

By construction, the inequality $\phi\le\langle\alpha^{\pi},s\rangle+b^{\pi}$ is tight over

$$\Upsilon_{\pi}:=\operatorname{conv}\big(\mathcal{V}(\pi,0),\ldots,\mathcal{V}(\pi,dn)\big),$$

where, by tight, we mean that, for $s\in\Upsilon_{\pi}$, $\operatorname{conc}_Q(\phi)(s)=\langle\alpha^{\pi},s\rangle+b^{\pi}$. More generally, the tight set for a valid inequality $f\le\langle\alpha,x\rangle+b$ of a function $f:X\mapsto\mathbb{R}$ will represent the set $\{x\in X\mid\operatorname{conc}_X(f)(x)=\langle\alpha,x\rangle+b\}$, and will be denoted as $T_f^{(\alpha,b)}(X)$. So, we can succinctly express our conclusion regarding the tight set of $\phi\le\langle\alpha^{\pi},s\rangle+b^{\pi}$ as $\Upsilon_{\pi}\subseteq T_{\phi}^{(\alpha^{\pi},b^{\pi})}(Q)$. Although (3.13) describes the coefficients of the interpolating inequalities in the general case, we remark the following special case, where the coefficient α_{ij}^{π} becomes zero.

Remark 3.2.2 *Let ϕ be a multilinear function and consider a movement vector π such that the j^{th} move along coordinate i is adjacent to the $j+1^{th}$ move along this coordinate, i.e., $m(i,j+1)=m(i,j)+1$. Then, because the function $\phi(s_{1n},\ldots,s_{dn})$ is affine when all but s_{in} is fixed, it follows that $\vartheta(\pi,i,j)=\vartheta(\pi,i,j+1)$, which implies by (3.13) that $\alpha_{ij}^{\pi}=0$.*

Now, we use Theorem 3.2.1 to compute, at a given point $\bar{s}\in Q$, a non-vertical facet-defining inequality of the hypograph of $\operatorname{conc}_Q(\phi)(s)$. To this end, it suffices to find a movement vector π so that $\bar{s}$ belongs to the corresponding simplex Υ_{π} and then to compute the function $\phi^{\Upsilon_{\pi}}(s)$ using (3.13), where $\phi^{\Upsilon_{\pi}}(s)$ is the affine interpolating function tight at $\mathcal{V}(\pi,0),\ldots,\mathcal{V}(\pi,dn)$. As shown, in our discussion following Definition 3.1.3, that a simple sorting of the coordinates of $\bar{z}:=Z(\bar{s})$ reveals this staircase. In our context, $\bar{z}_{i0}=1$ for all i and recall that to derive π we ignore

Algorithm 3 Facet-Generation over Q

procedure Facet-Generation($\bar{s}$)

  $\bar{z} \leftarrow Z(\bar{s})$;

  BeginSort

    sort $\bar{z}$ to find a movement vector π so that $\bar{s} \in \Upsilon_\pi$;

    $\bar{z}_{i0}$ are sorted before $\bar{z}_{ij}$ for $j \neq 0$;

    if, for any $j \geq 1$, $z_{ij} = z_{ij+1}$ then they are adjacent in the sorted order;

  EndSort

  compute an affine function $\phi^{\Upsilon_\pi}(s) = \langle \alpha^\pi, s \rangle + b^\pi$ by using equation (3.13);

  return (α^π, b^π).

end procedure

the ordering of these coordinates assuming they are placed first in the sorted order. Then, if the $d+k$ largest coordinate of $\bar{z}$ is $\bar{z}_{ij}$, we let $\Theta(k) = (i,j)$ and define $\pi = \big(\Theta_1(1), \ldots, \Theta_1(dn)\big)$. Slightly adjusting (3.8), we can express $\bar{s}$ as a convex combination of $\mathcal{V}(\pi, 0), \ldots, \mathcal{V}(\pi, dn)$ as follows:

$$\bar{s} = \big(1 - z_{\Theta(1)}\big)\mathcal{V}(\pi, 0) + \sum_{k=1}^{dn-1} \big(z_{\Theta(k)} - z_{\Theta(k+1)}\big)\mathcal{V}(\pi, k) + z_{\Theta(dn)}\mathcal{V}(\pi, dn).$$

Corollary 3.2.1 *Assume that $\phi(s_{1n}, \ldots, s_{dn})$ is concave-extendable from* $\mathrm{vert}(Q)$ *and is supermodular when restricted to the vertices of Q. Given a point $\bar{s} \in Q$, Algorithm 3 takes $\mathcal{O}(dn \log d)$ operations to find a non-vertical facet-defining inequality of* $\mathrm{conc}_Q(\phi)$ *which is tight at $\bar{s}$.*

Proof The correctness of Algorithm 3 is due to Theorem 3.2.1 and because Algorithm 3 identifies a π such that $\bar{s} \in \Upsilon_\pi$. The time complexity is $\mathcal{O}(dn \log d)$ because the computation of Z takes $\mathcal{O}(dn)$ time and d sorted lists each of size n can be merged in $\mathcal{O}(dn \log d)$ time using the d-way merge sort algorithm (see 5.4.1 in [57]). ■

The inequality obtained using Theorem 3.2.1 is facet-defining for $\mathrm{conc}_Q(\phi)(s)$ since it interpolates ϕ over the extreme points of a simplex Υ_π. Moreover, when $\bar{s}$ belongs to a face of Q, this inequality describes a facet of the hypograph of ϕ restricted to this face, a property that can be exploited as shown in [56] to develop facet-defining inequalities over P. Recall that for $J = (J_1, \ldots, J_d) \in \mathcal{J}$, where $\mathcal{J}$ is a collection of d-tuples defined as in (3.3), we defined $F_J := \prod_{i=1}^d F_{iJ(i)}$ as a face of Q, where $F_{iJ(i)}$ is defined as the convex hull of $\{v_{ij} \mid j \in J_i\}$. We can also describe the face $F_{iJ(i)}$ as the set of points of Q_i which satisfy the following facet-defining constraints of Q at equality:

$$\frac{s_{ij+1} - s_{ij}}{a_{ij+1} - a_{ij}} \leq \frac{s_{ij} - s_{ij-1}}{a_{ij} - a_{ij-1}} \qquad \text{for } j \notin J_i. \tag{3.14}$$

Corollary 3.2.2 *Assume $\bar{s} \in F_J$, and, when $\bar{s}$ is input, let (α^π, b^π) be the pair generated by Algorithm 3. If $\phi(s_{1n}, \ldots, s_{dn})$ is supermodular when restricted to the vertices of Q and concave-extendable from* $\mathrm{vert}(Q)$ *then (α^π, b^π) defines a non-vertical* *facet of* $\text{conc}_{F_J}(\phi)(s)$ *and the corresponding inequality is tight at* $\bar{s}$*. Moreover, if* $\phi(s_{1n},\ldots,s_{dn})$ *is a multilinear function then, for all* $j\notin J_i$*,* $\alpha^\pi_{ij}=0$*.*

Proof As $\bar{s}\in F_J$ it follows from (3.14) that $\bar{z}_{ij+1}=\frac{\bar{s}_{ij+1}-\bar{s}_{ij}}{a_{ij+1}-a_{ij}}=\frac{\bar{s}_{ij}-\bar{s}_{ij-1}}{a_{ij}-a_{ij-1}}=\bar{z}_{ij}$ for all i and $j\notin J_i$. Therefore, the sorting in Algorithm 3 guarantees that the movement vector π is such that for all i and $j\notin J_i$, the $j+1^{\text{st}}$ move along coordinate i follows immediately after the j^{th} move. This implies that

$$\text{for } i\in\{1,\ldots,d\} \text{ and } j\notin J_i \qquad m(i,j)+1=m(i,j+1). \tag{3.15}$$

Therefore, when ϕ is multilinear, the last statement in the result follows from Remark 3.2.2.

Under the assumption on ϕ, it follows from Corollary 3.2.1 that the inequality $\phi\le\langle\alpha^\pi,s\rangle+b^\pi$ is valid for $\text{conc}_Q(\phi)(s)$, and, thus, also valid for $\text{conc}_{F_J}(\phi)(s)$. We will show that $\dim\bigl(T_\phi^{(\alpha^\pi,b^\pi)}(F_J)\bigr)=\dim(F_J)$. Clearly, we have $\dim\bigl(T_\phi^{(\alpha^\pi,b^\pi)}(F_J)\bigr)\le\dim(F_J)$. Now, consider the simplex Υ_π defined by the movement vector π. It follows readily that $\Upsilon_\pi\cap F_J\subseteq T_\phi^{(\alpha^\pi,b^\pi)}(F_J)$ since $\text{conc}_Q(\phi)(s)=\text{conc}_{F_J}(\phi)(s)$ for every $s\in F_J$. Thus, the proof is complete if we can show that $\dim(F_J)\le\dim\bigl(\text{vert}(\Upsilon_\pi)\cap F_J\bigr)$, where $\text{vert}(\Upsilon_\pi)=\{\mathcal{V}(\pi,0),\ldots,\mathcal{V}(\pi,dn)\}$. Since $\mathcal{V}(\pi,k-1)\le\mathcal{V}(\pi,k)$ for all $k=1,\ldots,dn$, it follows from (3.15) that for all i and $j\notin J_i$, the only grid point where the grid-label along i^{th} coordinate is v_{ij} is $\Pi\bigl(\pi,m(i,j)\bigr)$ with the corresponding point $\mathcal{V}(\pi,m(i,j))$. In other words, for all i and $j\notin J_i$, $\text{vert}(\Upsilon_\pi)\cap\{s\mid s_i=v_{ij}\}=\mathcal{V}\bigl(\pi,m(i,j)\bigr)$ and, thus

$$\text{vert}(\Upsilon_\pi)\setminus\Bigl\{\mathcal{V}\bigl(\pi,m(i,j)\bigr)\Bigm| i=1,\ldots,d,\ j\notin J_i\Bigr\}\subseteq F_J.$$

This implies that $|\,\text{vert}(\Upsilon_\pi\cap F_J)|\ge dn+1-\sum_{i=1}^d|\bar{J}_i|=\dim(F_J)+1$, where $\bar{J}_i=\{0,\ldots,n\}\setminus J_i$. Since points in $\text{vert}(\Upsilon_\pi)$ are affinely independent, we conclude that $\dim(F_J)\le\dim\bigl(\text{vert}(\Upsilon_\pi)\cap F_J\bigr)\le\dim\bigl(T_\phi^{(\alpha^\pi,b^\pi)}(F_J)\bigr)$. ■

Next, we show that the facet generation problem of $\text{conc}_P(\phi)(u)$ can be solved in $\mathcal{O}(dn\log d)$. To prove this result, we need the following result shown in Chapter 2 which relates the concave envelope over the face F_J to the envelope over the projection

Q_J, where Q_J is defined in (3.4). Recall that two sets $C \subseteq \mathbb{R}^c$ and $D \subseteq \mathbb{R}^d$ are *affinely isomorphic* if there is an affine map $f : \mathbb{R}^c \mapsto \mathbb{R}^d$ that is a bijection between the points of the two sets. Consider an affine map $A : s_J \mapsto \tilde{s}$ defined as

$$\tilde{s}_{ij} = s_{ij} \quad \text{for } j \in J_i \quad \text{and} \quad \tilde{s}_{ij} = (1 - \gamma_{ij})s_{il(i,j)} + \gamma_{ij} s_{ir(i,j)} \quad \text{for } j \notin J_i, \tag{3.16}$$

where $l(i,j) = \max\{j' \in J_i \mid j' \leq j\}$, $r(i,j) = \min\{j' \in J_i \mid j' \geq j\}$, and $\gamma_{ij} = \frac{a_{ij} - a_{il(i,j)}}{a_{ir(i,j)} - a_{il(i,j)}}$. The inverse of A is defined as a map which transforms s to s_J.

Lemma 3.2.2 (Lemma 2.4.2) *Assume that* $\operatorname{conc}_Q(\phi)$ *is a polyhedral function. Let* $J = (J_1, \ldots, J_d) \in \mathcal{J}$*. Then,* $\operatorname{conc}_{F_J}(\phi)(s) = \operatorname{conc}_{Q_J}(\phi)(s_J)$ *for every* $s \in F_J$*. Let* $\phi \leq \langle \alpha, s\rangle + b$ *be a valid inequality of* $\operatorname{conc}_Q(\phi)(s)$ *so that* $\alpha_{\bar{J}} = 0$*. Then, the two tight sets,* $T_\phi^{(\alpha,b)}(F_J)$ *and* $T_\phi^{(\alpha_J,b)}(Q_J)$ *are affinely isomorphic under the affine map* A *defined in (3.16).*

Theorem 3.2.2 *Assume that* $\phi(s_{1n}, \ldots, s_{dn})$ *is concave-extendable from* $\operatorname{vert}(Q)$ *and is supermodular when restricted to the vertices of* Q*. Let* $\bar{u} \in P$*. Then, a facet-defining inequality of* $\operatorname{conc}_P(\phi)$ *which is tight at* $\bar{u}$ *can be found in* $\mathcal{O}(dn \log d)$ *operations.*

Proof Let $\bar{u} \in P$ and let $\bar{s}$ be the unique point so that $(\bar{u}, \bar{s}) \in PQ'$, which can be found in $\mathcal{O}(nd)$ and where PQ' is defined as in (3.5). Define $J = (J_1, \ldots, J_d)$, where $J_i := \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. It follows from Proposition 3.1.1 that $\bar{s} \in F_J$. Given $\bar{s}$ as input, let (α^π, b^π) denote the pair generated by Algorithm 3. Now, we derive an inequality $\phi \leq \langle \alpha', s\rangle + b'$ defined so that $\langle \alpha', s\rangle + b' = \langle \alpha^\pi, \Gamma_J(s)\rangle + b^\pi$, where $\Gamma_J : s \in \mathbb{R}^{d(n+1)} \mapsto \tilde{s} \in \mathbb{R}^{d(n+1)}$ is a linear map defined in (3.6). We will show that $\phi \leq \langle \alpha', s\rangle + b'$ defines a facet of $\operatorname{conc}_{F_J}(\phi)(s)$ which is tight at $\bar{s}$. Then, since $\alpha'_{\bar{J}} = 0$, it follows from Lemma 3.2.2 that (α'_J, b') defines a non-vertical facet of $\operatorname{conc}_{Q_J}(\phi)$ which is tight at $\bar{s}_J$. Therefore, by Proposition 3.1.2, (α', b') defines a non-vertical facet of $\operatorname{conc}_P(\phi)(u)$ which is tight at $\bar{u}$.

We now show that $\phi \leq \langle \alpha', s\rangle + b'$ defines a facet of $\operatorname{conc}_{F_J}(\phi)(s)$ tight at $\bar{s}$. The validity of the inequality for $\operatorname{conc}_{F_J}(\phi)(x)$ follows because for every $s \in F_J$

$$\operatorname{conc}_{F_J}(\phi)(s) \leq \langle \alpha^\pi, s\rangle + b^\pi = \langle \alpha^\pi, \Gamma_J(s)\rangle + b^\pi = \langle \alpha', s\rangle + b', \tag{3.17}$$

where the first inequality holds by the validity of $\phi \leq \langle \alpha^\pi, s\rangle + b^\pi$ for $\text{conc}_{F_J}(\phi)(s)$, first equality holds because, by Proposition 3.1.1, $s \in F_J$ implies $s = \Gamma_J(s)$, and the second equality is by the definition of (α', b'). Now, the proof is complete because, by Corollary 3.2.2, the first inequality in (3.17) is satisfied at equality for $\dim(F_J) + 1$ affinely independent points in F_J, and, in particular, for the point $\bar{s}$. ■

Next, we specialize our study to the case when the outer function is multilinear. Let $\mathcal{S}$ be the staircase triangulation of Q. Recall that $\phi^{\mathcal{S}} : \text{aff}(Q) \mapsto \mathbb{R}$ is obtained by extending the affine interpolation function of ϕ over the affine hull of Q. In Theorem 3.2.1, we argued that, under the assumed conditions on ϕ, $\phi^{\mathcal{S}}(s) = \text{conc}_Q(\phi)(s)$ for every $s \in Q$. Next, we show that if a multilinear function $\phi(s_{1n}, \ldots, s_{dn})$ is supermodular over $\text{vert}(Q)$ then, for every $u \in P$, $\phi^{\mathcal{S}}(u) = \text{conc}_P(\phi)(u)$. The point to note here is that, for the multilinear case, the concave envelope over P requires no other non-vertical inequalities beyond those needed to describe the concave envelope over Q, a result that does not hold in general for concave-extendable, supermodular functions.

Corollary 3.2.3 *Assume that the function $\phi(s_{1n}, \ldots, s_{dn})$ is multilinear and supermodular when restricted to vertices of Q. Let $\mathcal{S}$ be the staircase triangulation of Q. Then, for every $u \in P$, $\text{conc}_P(\phi)(u) = \phi^{\mathcal{S}}(u)$.*

Proof Let $\bar{u} \in P$. Then, as in Proposition 3.1.2, compute $(\bar{u}, \bar{s}) \in PQ'$ and define $J = (J_1, \ldots, J_d)$ so that $J_i = \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. Since ϕ is multilinear, it is concave-extendable from $\text{vert}(Q)$. Moreover, ϕ is supermodular when restricted to $\text{vert}(Q)$. Therefore, we may construct a facet-defining inequality using Algorithm 3, whose output will be denoted as the pair (α^π, b^π). Then, by Proposition 3.1.2, $\bar{s} \in F_J$ and, by Corollary 3.2.2, the inequality $\phi \leq \langle \alpha^\pi, s\rangle + b^\pi$ defines a facet of $\text{conc}_{F_J}(\phi)$ such that for all i and $j \notin J_i$, $\alpha^\pi_{ij} = 0$. Moreover, by Corollary 3.2.1, this inequality is tight at $\bar{s}$. Then, by Lemma 3.2.2, $\phi \leq \langle \alpha^\pi_J, s_J\rangle + b^\pi$ is a facet-defining inequality of $\text{conc}_{Q_J}(\phi)$ that is tight at $\bar{s}_J$ and, by Proposition 3.1.2, $\phi \leq \langle \alpha^\pi, u\rangle + b^\pi$ is a facet-defining inequality of $\text{conc}_P(\phi)$ that is tight at $\bar{u}$. ■

We have shown that the Algorithm 3 can be used to generate facet-defining inequalities for $\operatorname{conc}_P(\phi)$, where P is a product of polytopes defined in (3.1), each of which depends on underestimators for an inner-function. Assume that we construct another polytope P' using a subset of the underestimators used to define P. Then, since P' is a projection of P and projection commutes with convexification, it follows readily that that $\operatorname{conc}_P(\phi)$ projects to $\operatorname{conc}_{P'}(\phi)$. Therefore, any inequalities valid for $\operatorname{conc}_{P'}(\phi)$ are also valid for $\operatorname{conc}_P(\phi)$. However, we will show that a stronger property holds. The facet-defining inequalities also define facets of $\operatorname{conc}_P(\phi)$. Towards this end, we introduce some notation to describe a projection of P obtained by selecting a subset of uderestimators. This subset is specified using a d-tuple $J = (J_1, \ldots, J_d) \in \mathcal{J}$, where each tuple specifies which underestimators are selected. We denote the corresponding projection of P as P_J which is now the Cartesian product $P_{1J(1)} \times \cdots \times P_{dJ(d)}$, where $P_{iJ(i)}$ is the polytope defined in (3.1) using underestimators $u_{iJ(i)}$ with a vector of bounds $a_{iJ(i)}$.

Theorem 3.2.3 *Assume* $\operatorname{conc}_P(\phi)$ *is a polyhedral function and* $\phi \leq \langle \alpha_J, u_J \rangle + b$ *is a facet-defining inequality of* $\operatorname{conc}_{P_J}(\phi)(u_J)$*. Then,* $\phi \leq \langle \alpha, u \rangle + b$ *is a facet-defining inequality of* $\operatorname{conc}_P(\phi)(u)$*, where* $\alpha := (\alpha_J, 0)$*.*

Proof Since $\phi \leq \langle \alpha_J, u_J \rangle + b$ is a facet-defining inequality of $\operatorname{conc}_{P_J}(\phi)(u_J)$ and $a_{i0} < a_{in}$, there is a point $\bar{u} \in P$ with $\operatorname{conc}_{P_J}(\phi)(\bar{u}_J) = \langle \alpha_J, \bar{u}_J \rangle + b$ such that, for $i \in \{1, \ldots, d\}$ and $j \neq 0$, $a_{i0} < \bar{u}_{ij}$. We assume without loss of generality that $\bar{u} = A(\bar{u}_J)$ because we can use the linear transformation A defined in (3.16) to lift points from P_J to P and affine maps commute with convexification. Now, let $i' \in \{1, \ldots, d\}$ and $j' \notin J_{i'}$, and consider the face $P' := \{u \in P \mid u_{i'j'} = a_{i0}\}$. We construct $\hat{u} \in P'$ that matches $\bar{u}$ except that $\hat{u}_{i'j'} \neq \bar{u}_{i'j'}$. Such a $\hat{u}$ can be obtained using the same argument as above where $\hat{u} = B(\bar{u}_J)$ and B is defined similarly to A, except that $\hat{u}_{i'j'} = a_{i'0}$. It follows that $e_{i'j'}$ is in the affine hull of $T_\phi^{(\alpha,b)}(P)$. Then,

$$\begin{aligned}\dim(P) \geq \dim\bigl(T_\phi^{(\alpha,b)}(P)\bigr) &\geq \dim\bigl(T_\phi^{(\alpha_J,b)}(P_J)\bigr) + \sum_{i=1}^{d}\bigl(n - |J_i|\bigr)\\ &= \dim(P_J) + \sum_{i=1}^{d}\bigl(n - |J_i|\bigr) = \dim(P),\end{aligned}$$

where the first inequality is because $T_\phi^{(\alpha,b)}(P) \subseteq P$, second inequality is because in our argument above the choice of (i', j') was arbitrary except that $j' \notin J_{i'}$, the first equality is because (α_J, b) defines a facet of $\mathrm{conc}_{PJ}(\phi)$, and the second equality is by the definition of P. Therefore, equalities holds throughout and, in particular, $\dim(P) = \dim\bigl(T_\phi^{(\alpha,b)}(P)\bigr)$. ■

We have described a way to develop inequalities for composite functions as long as the outer-function is supermodular and concave-extendable. To extend the applicability of this result, we now turn our attention to a particular linear transformation that can be used to convert some functions that are not ordinarily supermodular into supermodular functions. This transformation is well-studied when the domain of the function is $\{0, 1\}^d$, a special case of $\mathrm{vert}(Q)$. In this case, the transformation, often referred to as switching, chooses a set $D \subseteq \{1, \ldots, d\}$ and considers a new function $\phi'(x_1, \ldots, x_d)$ defined as $\phi(y_1, \ldots, y_d)$, where $y_i = (1 - x_i)$ if $i \in D$ and $y_i = x_i$ otherwise. We will now generalize this switching operation to Q. To do so, we will need permutations σ_i of $\{0, \ldots, n\}$ for each $i \in \{1, \ldots, d\}$. We use the permutation σ_i to define an affine transformation that maps v_{ij} to $v_{i\sigma_i(j)}$. Let P^{σ_i} be a permutation matrix in $\mathbb{R}^{(n+1)\times(n+1)}$ such that, for all (i', j'), $P^{\sigma_i}_{i'j'} = 1$ when $i' = \sigma(j')$ and zero otherwise. Then, the affine transformation associated with σ_i is given by $A^{\sigma_i} = Z_i^{-1} \circ UP^{\sigma_i}U^{-1} \circ Z_i$, where $\circ$ denotes the composition operator and U is an upper triangular matrix of all ones. We let $A^\sigma(s) := \bigl(A^{\sigma_1}(s_1), \ldots, A^{\sigma_d}(s_d)\bigr)$. We will particularly be interested in the case where $\sigma_i = \{n, \ldots, 0\}$ for $i \in T$ and

$\sigma_i = \{0, \ldots, n\}$ otherwise. In this case, we denote $A^\sigma(s)$ by $s(T)$. Clearly, for $i \notin T$, $s(T)_i = s_i$. To compute $s(T)_i$ where $i \in T$, we use the following expression

$$s(T)_{ij} = a_{i0} + \sum_{k=1}^{j} (a_{ik} - a_{ik-1})(1 - z_{in+1-k}) \quad \text{for } j = 0, \ldots, n, \tag{3.18}$$

where z denotes $Z(s)$. Then, we define $\phi(T)(s_1, \ldots, s_d) = \phi\big(s(T)_{1n}, \ldots, s(T)_{dn}\big)$ and we say that $\phi(T)$ is obtained from ϕ by switching T. It follows easily that $\operatorname{conc}_Q(\phi)(s) = \operatorname{conc}_Q\big(\phi(T)\big)\big(s(T)\big)$. (Similar conclusions can be easily drawn for arbitrary permutations σ, where $\phi(\sigma)(s)$ is defined as $\phi\big(A^\sigma(s)_{1n}, \ldots, A^\sigma(s)_{dn}\big)$. In this case, $\operatorname{conc}_Q(\phi)(s) = \operatorname{conc}_Q\big(\phi(\sigma)\big)\big((A^\sigma)^{-1}(s)\big)$.) More specifically, if the switched function $\phi(T)$ is supermodular when restricted to the vertices of Q then $\operatorname{conc}_Q(\phi)(s)$ is determined by the *switched* staircase triangulation specified by T, whose grid-labels are obtained by labelling coordinates directions, for $i \in T$, as they were, and, for $i \notin T$, in a reversed order $v_{in}, \ldots, v_{i0}$. Then, for any movement vector π in the grid given by $\{0, \ldots, n\}^d$, the corresponding simplex is defined as $\operatorname{conv}\big(\bigcup_{k=0}^{dn} \operatorname{ext}\big(Q(T), \Pi(\pi, k)\big)\big)$, where $Q(T) := \{s(T) \mid s \in Q\}$. The following result records the above construction for later use.

Corollary 3.2.4 *Assume function $\phi(s_{1n}, \ldots, s_{dn})$ is concave-extendable from* $\operatorname{vert}(Q)$*. Let T be a subset of $\{1, \ldots, d\}$. If $\phi(T)(s_1, \ldots, s_d)$ is supermodular when restricted to* $\operatorname{vert}(Q)$ *then* $\operatorname{conc}_Q(\phi)(s)$ *is determined by the switched staircase triangulation specified by T.*

We now consider a special case that was studied in Chapter 2 and used to improve factorable programming. The case setting requires that the outer-function ϕ is a bilinear term and each inner function has only one non-trivial underestimator.

Corollary 3.2.5 (Theorem 2.1.1 and Theorem 2.3.3) *Let $a_{i0} \leq a_{i1} \leq a_{i2}$ for $i = 1, 2$, and define $P := \big\{(u, f) \mid a_{i0} \leq u_i \leq \min\{f_i, a_{i1}\},\ f_i \leq a_{i2},\ i = 1, 2\big\}$. Then,* *non-vertical facet-defining inequalities of the convex hull of* $\big\{(u,f) \mid \phi = f_1 f_2,\ (u,f) \in P\big\}$ *are given as follows:*

$$\phi \geq e_1 := a_{22}f_1 + a_{12}f_2 - a_{12}a_{22},$$
$$\phi \geq e_2 := (a_{22}-a_{21})u_1 + (a_{12}-a_{11})u_2 + a_{21}f_1 + a_{11}f_2 + a_{11}a_{21} - a_{11}a_{22} - a_{12}a_{21},$$
$$\phi \geq e_3 := (a_{22}-a_{20})u_1 + a_{20}f_1 + a_{11}f_2 - a_{11}a_{22},$$
$$\phi \geq e_4 := (a_{12}-a_{10})u_2 + a_{21}f_1 + a_{10}f_2 - a_{12}a_{21},$$
$$\phi \geq e_5 := (a_{21}-a_{20})u_1 + (a_{11}-a_{10})u_2 + a_{20}f_1 + a_{10}f_2 - a_{11}a_{21},$$
$$\phi \geq e_6 := a_{10}f_2 + a_{20}f_1 - a_{10}a_{20},$$
$$\phi \leq r_1 := a_{20}f_1 + a_{12}f_2 - a_{12}a_{20},$$
$$\phi \leq r_2 := (a_{20}-a_{21})u_1 + (a_{11}-a_{12})u_2 + a_{21}f_1 + a_{12}f_2 - a_{11}a_{20},$$
$$\phi \leq r_3 := (a_{20}-a_{22})u_1 + a_{11}f_2 + a_{22}f_1 - a_{11}a_{20},$$
$$\phi \leq r_4 := (a_{10}-a_{12})u_2 + a_{21}f_1 + a_{12}f_2 - a_{10}a_{21},$$
$$\phi \leq r_5 := (a_{21}-a_{22})u_1 + (a_{10}-a_{11})u_2 + a_{22}f_1 + a_{11}f_2 - a_{10}a_{21},$$
$$\phi \leq r_6 := f_1 a_{22} + a_{10}f_2 - a_{10}a_{22}.$$

Proof Let $\phi(f_1,f_2) = f_1 f_2$. We verify that the set of inequalities, $\phi \geq e_i$, $i = 1,\ldots,6$, defines the set of non-vertical facets of the epigraph of $\mathrm{conv}_P(\phi)\big((u_1,f_1),(u_2,f_2)\big)$. Let $v_{i0} = (a_{i0}, a_{i0})$, $v_{i1} = (a_{i1}, a_{i1})$, $v_{i2} = (a_{i1}, a_{i2})$, and define $Q_i := \mathrm{conv}(\{v_{i0}, v_{i1}, v_{i2}\})$ for $i = 1,2$. For $T = \{2\}$, we have

$$\phi(T)\big((u_1,f_1),(u_2,f_2)\big) := f_1\left(a_{20}+(a_{21}-a_{20})\Big(1-\frac{f_2-u_2}{a_{22}-a_{21}}\Big)+(a_{22}-a_{21})\Big(1-\frac{u_2-a_{20}}{a_{21}-a_{20}}\Big)\right),$$

where we have used (3.18) to derive the term after f_1 on the RHS of the above expression. Using the above expression, $\phi(T)(v_{1j_1}, v_{2j_2}) = a_{1j_1}a_{2(2-j_2)}$ for $j_1, j_2 \in \{0,1,2\}$. It follows that $\phi(T)$ is submodular when restricted to vertices of Q since $a_{i0} \leq a_{i1} \leq a_{i2}$ for $i = 1,2$. Let $\{\pi^1,\ldots,\pi^6\}$ be the set of movement sequences in $\mathbb{Z}^2$ from $(0,0)$ to $(2,2)$, where $\pi^1 := (1,1,2,2)$, $\pi^2 := (1,2,1,2)$, $\pi^3 := (1,2,2,1)$, $\pi^4 := (2,1,1,2)$,

$\pi^5 := (2,1,2,1)$, and $\pi^6 := (2,2,1,1)$. The set of movement sequences defines the switched staircase triangulation $\{\Upsilon_{\pi^1}(T),\ldots,\Upsilon_{\pi^6}(T)\}$, where for $i \in \{1,\ldots,6\}$

$$\Upsilon_{\pi^i}(T) = \operatorname{conv}\left(\left\{(v_{1j_1}, v_{2(2-j_2)}) \;\middle|\; (j_1,j_2) = (0,0) + \sum_{p=1}^{k} e_{\pi^i_p},\ k = 0,\ldots,4\right\}\right),$$

(see Figure 3.2 for the grid representation of the triangulation). Since the bilinear

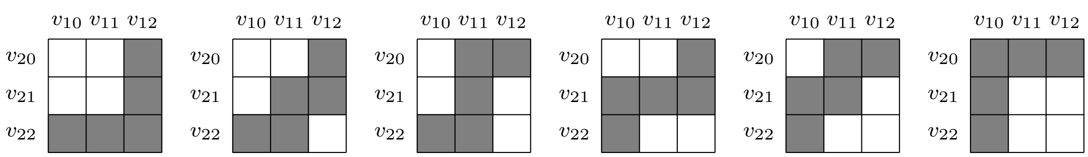


Fig. 3.2.: Gird representation of switched staircase triangulation

term is obviously convex-extendable from $\operatorname{vert}(Q)$, it follows from Corollary 3.2.4 that the convex envelope of ϕ over Q is determined by the switched staircase triangulation $\{\Upsilon_{\pi^1}(T),\ldots,\Upsilon_{\pi^6}(T)\}$. Moreover, each function e_i affinely interpolates f_1f_2 over the extreme points of simplex $\Upsilon_{\pi^i}(T)\big)$. As such, each inequality $\phi \geq e_i$, where $i \in \{1,\ldots,6\}$, describes a non-vertical facet of the epigraph of $\operatorname{conv}_Q(\phi)(s)$, and the result follows from Corollary 3.2.3.

A similar argument can be used to show that, for each $i \in \{1,\ldots,6\}$, $\phi \leq r_i$ defines a non-vertical facet-defining of the hypograph of $\operatorname{conc}_P(\phi)(u)$. This case does not require switching since the bilinear term is already supermodular. ∎

3.2.2 On the strength of termwise relaxation of bilinear functions

In this subsection, we consider a weighted graph $G = (V,E)$ with node set $V = \{1,\ldots,d\}$ and edge set E. With this graph, we associate a bilinear function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ defined as $\phi(s_{1n},\ldots,s_{dn}) = \sum_{e\in E} c_e \prod_{i\in e} s_{in}$, where, by $i \in e$, we mean that edge e is incident with node i. We assume that an edge exists only if the corresponding weight $c_e \neq 0$. We call an edge positive if $c_e > 0$ and negative if $c_e < 0$. We shall denote the graph of ϕ over Q as ϕ^Q and we will study whether its convex hull

$$\operatorname{conv}(\phi^Q) = \left\{(s,\phi) \;\middle|\; \operatorname{conv}_Q(\phi)(s) \leq \phi \leq \operatorname{conc}_Q(\phi)(s), s \in Q\right\}$$

is obtained by a simple relaxation, one obtained by convexifying each bilinear term separately. To formally define the latter relaxation, let $\phi_e(s)$ denote the bilinear term $c_e \prod_{i\in e} s_{in}$ associated with an edge $e \in E$. Then, we construct the *termwise-relaxation* of ϕ^Q by underestimating $\phi(\cdot)$ with $\sum_{e\in E} \operatorname{conv}_Q(\phi_e)(s)$ and overestimating it with $\sum_{e\in E} \operatorname{conc}_Q(\phi_e)(s)$, where each term in the summation could be obtained by using Corollary 3.2.4, or more specifically, when there is one non-trivial underestimator for each inner function, using Corollary 3.2.5. Succinctly, the termwise relaxation is defined as follow:

$$\Psi := \Big\{(s,\phi) \;\Big|\; \sum_{e\in E} \operatorname{conv}_Q(\phi_e)(s) \le \phi \le \sum_{e\in E} \operatorname{conc}_Q(\phi_e)(s), s \in Q\Big\}.$$

Clearly, Ψ is convex superset of ϕ^Q and therefore also a superset of $\operatorname{conv}\big(\phi^Q\big)$. We show that, if the graph G satisfies some conditions, $\operatorname{conv}(\phi^Q)$ coincides with Ψ. Since the sign for all c_e can be reversed, it suffices to consider the equivalence $\sum_{e\in E} \operatorname{conc}_Q(\phi_e)(s) = \operatorname{conc}_Q(\phi)(s)$. We assume without loss of generality, by transforming s_{ij} to $\frac{s_{ij}-a_{i0}}{a_{in}-a_{i0}}$ for all i and j, if necessary, that for all i, $a_{i0} = 0$ and $a_{in} = 1$.

We call an edge $e \in E$ is positive if $c_e > 0$ and negative otherwise. A (signed) graph is said to be *balanced* if every cycle has an even number of negative edges (see [58]). It is shown in Theorem 3 of [58] that a graph is balanced if and only if the vertex set $V(G)$ can be partitioned into subsets T_1 and T_2 so that each positive edge of G connects two nodes from the same subset and each negative edge connects two nodes from different subsets. We will argue, by switching the variables which correspond to one of the partitioned subsets, that we can transform ϕ into a supermodular function. For $s \in Q$, let $s_{\cdot n}$ denote the vector $(s_{1n}, \ldots, s_{dn})$, and, for an edge $e = (i,j) \in E$, let s_{en} denote the vector (s_{in}, s_{jn}).

Lemma 3.2.3 *Consider a graph G and let ϕ be a bilinear function defined by the graph G. There exists a subset T of V so that, for $s(T)$ as defined in (3.18), the function $\phi(T)(s_{1n}, \ldots, s_{dn}) = \phi(s(T)_{1n}, \ldots, s(T)_{dn})$ is supermodular when restricted to $\operatorname{vert}(Q)$ if and only if graph G is balanced.*

Proof Assume $G = (V, E)$ is a balanced graph. Then, using Theorem 3 in [58], we partition V into subsets T_1 and T_2 such that positively signed edges connect nodes of the same subset and the negatively signed edges connect nodes of the opposite subset. Then, to show that $\phi(T_1)(\cdot)$ is supermodular over $\mathrm{vert}(Q)$, it suffices to show that, for each edge e, $\phi_e(T_1)(s_{en})$ is supermodular over $\mathrm{vert}(Q)$. By (3.18) it follows that, for $i \in T_1$, $s_i \geq s'_i$ and $s(T_1)_i \leq s'(T_1)_i$ whenever $z_i \geq z'_i$. Since $\phi_e(s_{en})$ is supermodular when $e \in T_1$ or $e \in T_2$ and submodular otherwise, it follows that, for each edge e, $\phi_e(T_1)(s_{en})$ is supermodular.

We now show the converse, *i.e.*, there does not exist a T such that $\phi(T)$ is supermodular when restricted to $\mathrm{vert}(Q)$. Since the graph is not balanced, there exists a cycle that contains an odd number of negative edges. Since this cycle leaves and enters T an even number of times, it follows that there is a negative edge either contained in T or in its complement. Let this edge be $e := (k, l)$. Assume without loss of generality that $k, l \in T$ as the other case is similar. Consider vertices v' and v'' corresponding to grid points $(j'_i)_{i=1}^d$ and $(j''_i)_{i=1}^d$, where we assume that $j'_k = j''_l = 1$, $j'_l = j''_k = 2$, and $j'_i = j''_i$ otherwise. Then, it follows that

$$\begin{aligned}&\phi(T)(v'_{\cdot n} \vee v''_{\cdot n}) + \phi(T)(v'_{\cdot n} \wedge v''_{\cdot n}) - \phi(T)(v'_{\cdot n}) - \phi(T)(v''_{\cdot n})\\&= c_e\big(a_{kn-2}a_{ln-2} + a_{kn-1}a_{ln-1} - a_{kn-2}a_{ln-1} + a_{kn-1}a_{ln-2}\big) < 0,\end{aligned}$$

where the inequality follows from the supermodularity of the bilinear product and $c_e < 0$. Therefore, it follows $\phi(T)$ is not supermodular. ■

In Theorem 3.2.4, we show that the balanced graphs are exactly the ones for which the termwise relaxation Ψ coincides with $\mathrm{conv}(\phi^Q)$. To prove this result, we need the following lemma. Recall that we say that the concave envelope of a function f is determined by a triangulation $\mathcal{K} = \{K_1, \ldots, K_r\}$ if the concave envelope of f over $\bigcup_{i=1}^r K_i$ is $\min_{i=1}^r \chi^{K_i}(s)$, where χ^{K_i} is the affine function interpolating $\big(v, f(v)\big)$ for all $v \in \mathrm{vert}(K_i)$.

Lemma 3.2.4 ([51]) *Consider a function $f : \mathrm{vert}(D) \mapsto \mathbb{R}$ so that $f(s) = \sum_{j=1}^m f_j(s)$, where D is a polytope. If concave envelopes of $f_j(s)$, $j = 1, \ldots, m$, are determined* *by the same triangulation $\mathcal{K}$ of D then* $\text{conc}_D(f)(s) = \sum_{j=1}^m \text{conc}_D(f_j)(s)$ *for every* $s \in D$*. Moreover, if a common triangulation does not generate concave envelopes of* f_j *for all* j*, then there exists* $s \in D$ *such that* $\text{conc}_D(f) < \sum_{j=1}^m \text{conc}_D(f)$.

Proof Let $\mathcal{K} = \{K_1, \ldots, K_r\}$. Define $f^{\mathcal{K}}(s) := \min_{i=1}^r \chi_j^{K_i}(s)$ (resp. $f_j^{\mathcal{K}}(s) := \min_{i=1}^r \chi_j^{K_i}(s)$), where χ^{K_i} (resp. $\chi_j^{K_i}$) affinely interpolates $\big(v, f(v)\big)$ (resp. $\big(v, f_j(v)\big)$ for all $v \in \text{vert}(K_i)$. Let $s \in K_i$. Since $\mathcal{K}$ is the triangulation of D, it follows that for some $\lambda \geq 0$ such that $\sum_v \lambda_v = 1$

$$\begin{aligned}\sum_{j=1}^m \text{conc}_D(f_j)(s) \geq \text{conc}_D(f)(s) \geq \chi^{K_i}(s) = \sum_{v\in\text{vert}(K_i)} \lambda_v \chi^{K_i}(v) = \sum_{v\in\text{vert}(K_i)} \lambda_v f(v)\\ = \sum_{v\in\text{vert}(K_i)} \lambda_v \Big(\sum_{j=1}^m f_j(v)\Big) = \sum_{j=1}^m \sum_{v\in\text{vert}(K_i)} \lambda_v f_j(v) = \sum_{j=1}^m \chi_j^{K_i}(s)\\ \geq \sum_{j=1}^m \text{conc}_D(f_j)(s),\end{aligned}$$

where the first inequality is because $\sum_{j=1}^m \text{conc}_D(f_j)$ is a concave overestimator of f, the second inequality is because of Jensen's inequality and $\text{conc}_D(f)$ is a concave function, the first equality is by definition of $\chi^{K_i}(s)$, the second equality is because $\chi^{K_i}(v) = f(v)$ for all $v \in \text{vert}(K_i)$, the third equality is because of definition of f_j, the fourth equality is by interchanging the order of summation, the last equality is by the definition of $\chi_j^{K_i}(s)$ and $f_j(v) = \chi_j^{K_i}(v)$, and the last inequality is because $\text{conc}(f_j)(x) = \min_{i=1}^r \chi^{K_i}(s)$. Therefore, equality holds throughout and $\text{conc}_D(f)(s) = \chi^{K_i}(s) = \sum_{j=1}^m \text{conc}_D(f_j)(s)$.

Now, we consider the case when a common triangulation does not exist. Let $\mathcal{K} = \{K_1, \ldots, K_r\}$ be the triangulation associated with the concave envelope of f and let j be such that the concave envelope of f_j is not associated with $\mathcal{K}$. Clearly, $\text{conc}_D(f_j)(s) \geq \chi_j^{K_i}(s)$ for all $s \in K_i$. But, there must exist an i and an $s \in K_i$ such that $\text{conc}_D(f_j)(s) > \chi_j^{K_i}(s)$. Otherwise, as shown in (3.7) $\text{conc}_D(f_j)(s) = \min_{i=1}^r \chi_j^{K_i}(s)$, which contradicts the assertion that the concave envelope of f_j is not associated with the triangulation $\mathcal{K}$. Let $s = \sum_{v\in \text{vert}(K_i)} v\lambda_v$ express s as a convex combination of vertices of K_i. It follows that

$$\begin{aligned}\text{conc}_D(f)(s) &= \sum_{v\in\text{vert}(K_i)} f(v)\lambda_v = \sum_{v\in\text{vert}(K_i)} \sum_{j'=1}^{m} f_{j'}(v)\lambda_v \\ &= \sum_{j'=1}^{m} \sum_{v\in\text{vert}(K_i)} f_{j'}(v)\lambda_v = \sum_{j'=1}^{m} \chi_{j'}^{K_i}(s) < \sum_{j'=1}^{m} \text{conc}_D(f_{j'})(s),\end{aligned}$$

where the first equality is because $\mathcal{K}$ is the triangulation associated with $\text{conc}_D(f)$, the second equality is by definition of f, the third equality is by interchanging the summations, the fourth equality is by the definition of $\chi_j^{K_i}(s)$ and the strict inequality is because for $j' \neq j$, $\text{conc}_D(f_{j'})(s) \geq \chi_{j'}^{K_i}(s)$ and we have chosen s so that $\text{conc}_D(f_j)(s) > \chi_j^{K_i}(s)$. ■

Theorem 3.2.4 *Consider a graph G and a bilinear function ϕ defined on G. Then, $\sum_{e\in E} \text{conc}_Q(\phi_e)(s) = \text{conc}_Q(\phi)(s)$ if and only if G is balanced.*

Proof Suppose that graph G is balanced. Then, we show that $\sum_{e\in E} \text{conc}_Q(\phi_e)(s) = \text{conc}_Q(\phi)(s)$. By Lemma 3.2.3, there exists a subset T of V such that, for all $e \in E$, $\phi_e(T)(s_{1n}, \ldots, s_{dn})$ is supermodular when restricted to $\text{vert}(Q)$. By Theorem 1.2 in [9], for all $e \in E$, ϕ_e is concave-extendable from $\text{vert}(Q)$. By Corollary 3.2.4, $\text{conc}_Q(\phi_e)(s)$ is determined by the same switched staircase triangulation for all $e \in E$. So, by Lemma 3.2.4, we conclude that $\sum_{e\in E} \text{conc}_Q(\phi_e)(s) = \text{conc}_Q(\phi)(s)$.

Now, suppose that G is not balanced. We construct a point $s \in Q$ so that $\text{conc}_Q(\phi)(s) < \sum_{e\in E} \text{conc}_Q(\phi_e)(s)$. Let $\bar{s}_i = \frac{1}{2}(0, a_{i1}, \ldots, a_{in-1}, 1)$ for all $i = 1, \ldots, d$. Then, we obtain

$$\begin{aligned}\text{conc}_Q(\phi)(\bar{s}_1, \ldots, \bar{s}_d) &= \text{conc}_{[0,1]^d}\, \phi\Big(\frac{1}{2}, \ldots, \frac{1}{2}\Big) \\ &< \sum_{e\in E} \text{conc}_{[0,1]^2}\, \phi_e\Big(\frac{1}{2}, \frac{1}{2}\Big) = \sum_{e\in E} \text{conc}_Q(\phi)(\bar{s}),\end{aligned}$$

where first and last equality hold by Lemma 3.2.2 and strict inequality follows from Theorem 4 in [8]. (Alternately, the existence of a point that satisfies the strict inequality follows from Lemma 3.2.4, and that $\left(\frac{1}{2}, \ldots, \frac{1}{2}\right)$ is such a point is a consequence of strict supermodularity of the bilinear term). ■

The hypercube $[0,1]^d$ arises as a special case of Q where $n=1$ and the variables $(s_{10},\dots,s_{d0})$ are projected out. In this case, Theorem 3.2.4 recovers the results of [8] and [21] regarding when McCormick envelopes [5] applied termwise suffice to obtain the concave envelope of a bilinear function ϕ over $[0,1]^d$.

Corollary 3.2.6 (Theorem 4 [8] and Theorem 3.10 [21]) *Consider the graph G associated with a bilinear function $\phi : [0,1]^d \mapsto \mathbb{R}$. Then, the termwise relaxation of the hypograph of $\phi(x)$ over $[0,1]^d$ coincides with the hypograph of $\operatorname{conc}_P(\phi)(u)$ if and only if every cycle in G has an even number of negative edges.*

3.2.3 Tractable simultaneous convex hull

We now extend our results to simultaneous convexification of a vector of functions $\theta : \mathbb{R}^d \mapsto \mathbb{R}^\kappa$. Consider the hypograph of $\theta : \mathbb{R}^d \mapsto \mathbb{R}^\kappa$ over a polytope $P := P_1 \times \cdots \times P_d$ defined as

$$\Theta^P := \bigl\{(u,\theta) \in \mathbb{R}^{d\times(n+1)} \times \mathbb{R}^\kappa \bigm| \theta \le \theta(u_{1n},\dots,u_{dn}),\ u \in P\bigr\},$$

where P_i is the polytope defined in (3.1). For $k \in \{1,\dots,\kappa\}$, let $\Theta_k^P := \bigl\{(u,\theta) \bigm| \theta_k \le \theta_i(u_{1n},\dots,u_{dn}),\ u \in P\bigr\}$ be the hypograph of θ_k over P. Since $\Theta^P \subseteq \bigcap_{k=1}^\kappa \operatorname{conv}(\Theta_k^P)$, it follows that $\operatorname{conv}(\Theta^P)$ is a subset of $\bigcap_{k=1}^\kappa \operatorname{conv}(\Theta_k^P)$, where the former will be referred to as the *simultaneous* convex hull of Θ^P, while the latter as the *individual* convex hull of Θ^P. Clearly, it is often the case that $\operatorname{conv}(\Theta^P) \subsetneq \bigcap_{k=1}^\kappa \operatorname{conv}(\Theta_k^P)$. Nevertheless, we will characterize conditions for which the simultaneous hull of Θ^P coincides with the individual hull of Θ^P.

Theorem 3.2.5 *If concave envelopes of θ_k, $k=1,\dots,\kappa$, over Q are determined by the same triangulation $\mathcal{K}$ then $\operatorname{conv}(\Theta^P) = \bigcap_{k=1}^\kappa \operatorname{conv}(\Theta_k^P)$.*

Proof Clearly, $\operatorname{conv}\bigl(\Theta^P\bigr) \subseteq \bigcap_{k=1}^\kappa \operatorname{conv}\bigl(\Theta_k^P\bigr)$ because $\Theta^P \subseteq \bigcap_{k=1}^\kappa \operatorname{conv}\bigl(\Theta_k^P\bigr)$ and the latter set is convex. To show $\bigcap_{k=1}^\kappa \operatorname{conv}\bigl(\Theta_k^P\bigr) \subseteq \operatorname{conv}\bigl(\Theta^P\bigr)$, we consider a point $(\bar u,\bar\theta) \in \bigcap_{k=1}^\kappa \operatorname{conv}\bigl(\Theta_k^P\bigr)$, *i.e.*, $(\bar\theta_1,\dots,\bar\theta_d) \le \bigl(\operatorname{conc}_P(\theta_1)(\bar u),\dots,\operatorname{conc}_P(\theta_d)(\bar u)\bigr)$. Then, we lift $\bar{u}$ to the unique point $\bar{s}$ so that $(\bar{u}, \bar{s}) \in PQ'$, and define $J = (J_1, \ldots, J_d)$, where $J_i := \{j \mid \bar{u}_{ij} = \bar{s}_{ij}\}$. Let $D = (D_1, \ldots, D_d)$ so that $D_i \subseteq \bar{J}_i := \{0, \ldots, n\} \setminus J_i$. Let $\bar{u}^D$ be a point of P so that, for $i \in \{1, \ldots, d\}$, $\bar{u}^D_{ij} = \bar{s}_{ij}$ if $j \notin D_i$ and $\bar{u}^D_{ij} = \bar{u}_{ij}$ otherwise. Since $\bar{u}^{\bar{J}} = \bar{u}$ where $\bar{J}$ denotes $(\bar{J}_1, \ldots, \bar{J}_d)$, the proof is complete if we show by induction on $|D| := \sum_{i=1}^d |D_i|$ that there is a set $M_D \subseteq \mathrm{vert}(P)$ and λ_v, for $v \in M_D$, that are independent of k and, for $k \in \{1, \ldots, \kappa\}$, satisfy

$$\big(\bar{u}^D, \mathrm{conc}_P(\theta_k)(\bar{u}^D)\big) = \sum_{v \in M_D} \lambda_v \big(v, \theta_k(v_{\cdot n})\big), \quad \sum_{v \in M_D} \lambda_v = 1, \quad \lambda \geq 0, \tag{3.19}$$

where $v_{\cdot n}$ denotes $(v_{1n}, \ldots, v_{dn})$. For the base case with $|D| = 0$, we have $\bar{u}^D = \bar{s} \in Q$. Since $\mathcal{K}$ is assumed to be the common triangulation that determines $\mathrm{conc}_Q(\theta_k)$ for every $k \in \{1, \ldots, \kappa\}$, there exists $K \in \mathcal{K}$ and convex multipliers λ_v such that, for $k \in \{1, \ldots, \kappa\}$, $\big(\bar{u}^D, \mathrm{conc}_Q(\theta_k)(\bar{u}^D)\big) = \sum_{v \in \mathrm{vert}(K)} \lambda_v \big(v, \theta_k(v_{\cdot n})\big)$. Therefore, the base case is established because $\mathrm{vert}(K) \subseteq \mathrm{vert}(Q) \subseteq \mathrm{vert}(P)$, and, by Corollary 2.3.1, we have that, for every $k \in \{1, \ldots, \kappa\}$, $\mathrm{conc}_P(\theta_k)(\bar{u}^D) = \mathrm{conc}_Q(\theta_k)(\bar{u}^D)$. For the inductive step, consider $D = (D_1, \ldots, D_d)$ so that $D_i \subseteq \bar{J}_i$ and assume that the result holds for any tuple D' such that $|D'| < |D|$. Since $|D| \neq 0$, there is a pair (i', j') so that $j' \in D_{i'}$. Let $D' := (D'_1, \ldots, D'_d)$ so that $D'_i = D_i$ if $i \neq i'$ and $D'_{i'} = D_{i'} \setminus \{j'\}$. Consider an affine mapping A so that, for $i \neq i'$ and $j \neq j'$, $A(u)_{ij} = u_{ij}$ while $A(u)_{i'j'} = a_{i'0}$. Define $\gamma := \frac{u_{i'j'} - a_{i'0}}{s_{i'j'} - a_{i'0}}$. It follows from the induction hypothesis that there exists a set $M_{D'} \subseteq \mathrm{vert}(P)$ and convex multipliers λ_v, one for each $v \in M_{D'}$, so that, for $k \in \{1, \ldots, \kappa\}$, $\big(\bar{u}^{D'}, \mathrm{conc}_P(\theta_k)(\bar{u}^{D'})\big) = \sum_{v \in M^{D'}} \lambda_v \big(v, \theta_k(v_{\cdot n})\big)$. Then, for $k \in \{1, \ldots, \kappa\}$,

$$\begin{aligned}
\big(\bar{u}^D, \mathrm{conc}_P(\theta_k)(\bar{u}^D)\big) &= \gamma\big(\bar{u}^{D'}, \mathrm{conc}_P(\theta_k)(\bar{u}^{D'})\big) + (1-\gamma)\big(A(\bar{u}^{D'}), \mathrm{conc}_P(\theta_k)(\bar{u}^{D'})\big) \\
&= \gamma \sum_{v \in M_{D'}} \lambda_v \big(v, \theta_k(v_{\cdot n})\big) + (1-\gamma) \sum_{v \in M_{D'}} \lambda_v \big(A(v), \theta_k(v_{\cdot n})\big) \\
&= \gamma \sum_{v \in M_{D'}} \lambda_v \big(v, \theta_k(v_{\cdot n})\big) + (1-\gamma) \sum_{v \in M_{D'}} \lambda_v \Big(A(v), \theta_k\big(A(v_{\cdot n})\big)\Big),
\end{aligned}$$

where the first equality holds as, by Corollary 2.3.1, $\mathrm{conc}_P(\theta_k)(\bar{u}^D) = \mathrm{conc}_P(\theta_k)(\bar{u}^{D'})$, the second equality follows from the induction hypothesis because $|D'| < |D|$ and exploits that convexification commutes with affine transformation, and the last equality is because $j' \neq n$. The induction step is established by observing that, for any $u \in \text{vert}(P)$, $A(u) \in \text{vert}(P)$. ■

Corollary 3.2.7 *Assume that, for every $k \in \{1, \ldots, \kappa\}$, the function $\theta_k(s_{1n}, \ldots, s_{dn})$ is concave-extendable from $\text{vert}(Q)$ and is supermodular when restricted to the vertices of Q. Then, $\text{conv}\bigl(\Theta^P\bigr) = \bigcap_{k=1}^{\kappa} \text{conv}\bigl(\Theta^P_k\bigr)$. Moreover, the facet generation problem of $\text{conv}(\Theta^P)$ can be solved in $\mathcal{O}(\kappa dn \log d)$.*

Proof It follows from Theorem 3.2.5 that $\text{conv}\bigl(\Theta^P\bigr) = \bigcap_{k=1}^{\kappa} \text{conv}\bigl(\Theta^P_k\bigr)$ because, by Theorem 3.2.1, concave envelopes of θ_k, $k = 1, \ldots, \kappa$, are determined by a triangulation of Q, that does not depend on k. Now, we argue that the facet generation problem of $\text{conv}(\Theta^P)$ can be solved by separating $\text{conv}(\Theta^P_k)$ individually. Let $(\bar{u}, \bar{\theta}) \in \mathbb{R}^{d\times(n+1)} \times \mathbb{R}^{\kappa}$. Then, for each $k \in \{1, \ldots, \kappa\}$, we call the procedure in Theorem 3.2.2 to solve the facet generation problem of $\text{conv}(\Theta^P_k)$. If $(\bar{u}, \bar{\theta}) \in \bigcap_{k=1}^{\kappa} \text{conv}(\Theta^P_k)$ then, as shown above, $(\bar{u}, \bar{\theta}) \in \text{conv}(\Theta^P)$. Otherwise, without loss of generality, we assume that $(\bar{u}, \bar{\theta}) \notin \text{conv}(\Theta^P_1)$ and the procedure outputs a facet-defining inequality $\theta_1 \leq \langle \alpha, u\rangle + b$ of $\text{conc}_P(\theta_1)$ that is violated by $(\bar{u}, \bar{\theta})$. As $\text{conv}(\Theta^P) \subseteq \bigcap_{k=1}^{\kappa} \text{conv}(\Theta^P_k)$, this inequality is valid for $\text{conv}(\Theta^P)$. To complete the proof, we will show that it defines a facet of $\text{conv}(\Theta^P)$. Let $T := \bigl\{(u, \theta) \in \text{conv}(\Theta^P) \bigm| \theta_1 = \langle \alpha, u\rangle + b\bigr\}$ and let $M := \bigl\{u \in \text{vert}(P) \bigm| \langle \alpha, u\rangle + b = \theta_1(u_{\cdot n})\bigr\}$. Observe that $\bigl\{(u, \theta) \bigm| u \in M,\ \theta_1 = \theta_1(u_{\cdot n}),\ \theta_k \geq \theta_k(u_{\cdot n}) \text{ for } k = 2, \ldots, \kappa\bigr\}$ is a subset of T. Therefore, $\dim\bigl(\text{aff}(T)\bigr) \geq \dim(M) + \kappa - 1 = \dim(P) + \kappa - 1 \geq \dim(\Theta^P) - 1$, where the equality holds because $\theta_1 \leq \langle \alpha, u\rangle + b$ defines a facet of Θ_1. Thus, T is a facet of $\text{conv}(\Theta^P)$. ■

3.3 Infinitely many estimators for inner functions

Sections 3.1 to 3.2 considered composite functions and described a way to relax them while exploiting *finitely* many estimators for each inner function. A natural follow-up question is to understand the limiting relaxation, one obtained using *infinitely* many estimators for each inner function. We will explore the structure of this relaxation in Section 3.3. To begin, we review some basic concepts from probability theory that we later use to characterize the structure of limiting relaxation.

We consider real-valued random variables with support in $[0,1]$. Each such random variable A_i induces a probability measure on $\mathbb{R}$ which can be described by its (cumulative) distribution function F_i, that is: $F_i(a_i) = \Pr\{A_i \leq a_i\}$ for $a_i \in (-\infty, \infty)$. The expectation of random variable A_i is

$$\mathbb{E}[A_i] = \int_{-\infty}^{\infty} a_i \mathrm{d}F_i(a_i).$$

Let $\mathcal{F}$ be the set of all distribution functions with support in $[0,1]$. Any distribution function $F_i \in \mathcal{F}$ has three properties; it is non-decreasing, right-continuous, and ranges from 0 to 1 with $F_i(0) = 0$ and $F_i(1) = 1$. Conversely, any function satisfying these three properties is a distribution function for some random variable whose support is a measurable subset of $[0,1]$. Then, $\mathcal{F}$ is a convex subset of $\mathcal{B}$, where the latter set denotes the convex cone whose elements are all bounded nonnegative univariate functions on $\mathbb{R}$. The convexity of $\mathcal{F}$ follows because the three properties charactering functions in $\mathcal{F}$ are closed under convex combinations and any function satisfying these properties belongs to $\mathcal{F}$. The extreme set of $\mathcal{F}$, denoted as $\mathrm{ext}(\mathcal{F})$, is the set of distribution functions with Dirac measures over $[0,1]$, *i.e.*, $\mathrm{ext}(\mathcal{F}) := \left\{H_{\delta(a)} \;\middle|\; a \in [0,1]\right\}$, where $\delta(a)$ denotes the Dirac measure at point a and $H_{\delta(a)}$ denotes the corresponding distribution function. The right-continuity of a non-decreasing function implies that the function is continuous except possibly at a finite or countable set of points where the graph of the distribution function has a vertical gap.

Due to the vertical gaps, a distribution function F_i does not always have an inverse. To circumvent this issue, a generalized inverse is used instead that is defined for any $\lambda \in [0,1]$ as follows:

$$F_i^{-1}(\lambda) := \min\left\{a_i \in [0,1] \;\middle|\; F_i(a_i) \geq \lambda\right\}.$$

The generalized inverse, $F_i^{-1}(\lambda)$, is non-decreasing and left-continuous on $[0,1]$. Like the distribution function F_i, the generalized inverse function, F_i^{-1}, can have at most countably many jumps where if it fails to be continuous. Observe that $F_i^{-1}(\lambda) \leq a_i$ if and only if $\lambda \leq F_i(a_i)$. This is because for any (a_i, λ), if $F(a_i) \geq \lambda$, it follows by minimization and feasibility of a_i in the definition of $F_i^{-1}(\lambda)$ that $F_i^{-1}(\lambda) \leq a_i$. Moreover, if $F_i^{-1}(\lambda) \leq a_i$, there exists a decreasing sequence a_i^k and $a_i^k \to a_i' \leq a_i$ such that $F_i(a_i^k) \geq \lambda$. Then, it follows that $F_i(a_i) \geq F_i(a_i') \geq \lambda$, where the first inequality is because $a_i \geq a_i'$ and the second is because F_i is right-continuous.

For distribution functions $F_1, \ldots, F_d \in \mathcal{F}$, we define $\Pi(F_1, \ldots, F_d)$ as the set of all joint distribution functions on $\mathbb{R}^d$ whose marginals are $F_1, \ldots, F_d$. Therefore, a distribution function F belongs to $\Pi(F_1, \ldots, F_d)$ if and only if it satisfies the following properties. First, F is non-decreasing in each variable. Second, F is right-continuous in the sense that $\lim_{\delta \to 0^+} F(a_1+\delta, \ldots, a_d+\delta) = F(a_1, \ldots, a_d)$. Third, $F(0, \ldots, 0) = 0$ and $F(1, \ldots, 1) = 1$. Finally, for each i and $a_i \in (-\infty, \infty)$, $F(\infty, \ldots, a_i, \ldots, \infty) = F_i(a_i)$. Let $\mathcal{B}^d$ denote the convex cone of of bounded nonnegative functions on $\mathbb{R}^d$. Then, $\Pi(F_1, \ldots, F_d)$ is a convex subset of $\mathcal{B}^d$ because the above properties are closed under taking convex combinations and any functions satisfying these properties belong to $\Pi(F_1, \ldots, F_d)$. To clarify the joint distribution function, we add it as a subscript to the expectation operator so that, for a continuous function $\phi : [0,1]^d \mapsto \mathbb{R}$, its expectation under $F \in \Pi(F_1, \ldots, F_d)$ is denoted as

$$\mathbb{E}_F\big[\phi(A_1, \ldots, A_d)\big] := \int_{\mathbb{R}^d} \phi(a_1, \ldots, a_d) \mathrm{d}F(a_1, \ldots, a_d),$$

where $(A_1, \ldots, A_d)$ follows F, which will be denoted as $(A_1, \ldots, A_d) \sim F$. Since the random variables $A_1, \ldots, A_d$ are assumed to have supports in $[0,1]$ and ϕ is continuous, it follows that, for all $F \in \Pi(F_1, \ldots, F_d)$, the expectation $\mathbb{E}_F\big[\phi(A_1, \ldots, A_d)\big]$ is finite.

3.3.1 Envelope characterization via optimal transport

Consider a composite function $\phi \circ f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$ defined as $\phi \circ f(x) = \phi(f(x))$, where $f : \mathbb{R}^m \mapsto \mathbb{R}^d$ is a vector of bounded functions over X and $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ is a continuous function. We assume, without loss of generality, that, for every $x \in X$, $f(x) \in [0,1]^d$ and $\phi : [0,1]^d \mapsto \mathbb{R}$ since otherwise if, for every $x \in X$, $f_i^L \leq f(x) \leq f_i^U$, we can define $\bar{f}_i(x) = (f_i(x) - f_i^L)/(f_i^U - f_i^L)$ as the i^{th} inner-function and $\bar{\phi}(\bar{f}) = \phi\big((f_1^U - f_1^L)\bar{f}_1 + f_1^L, \ldots, (f_1^U - f_d^L)\bar{f}_d + f_d^L\big)$ as the outer-function. For each point (x, f), where $x \in X$ and $f = f(x)$, in Section 3.3.2 we will use underestimating functions of $f_i(x)$ to derive a marginal distribution function $F_i \in \mathcal{F}$. This marginal distribution will be such that the expected value of the corresponding random variable A_i, denoted as $\mathbb{E}_{F_i}[A_i]$, equals f_i. Consequently, to relax the hypograph of the composite function $\phi(f(x))$, it will suffice to over-estimate $\phi\big(\mathbb{E}_{F_1}(A_1), \ldots, \mathbb{E}_{F_d}(A_d)\big)$. We now discuss how this will be achieved. For notational brevity, we extend the outer-function ϕ to define $\tilde{\phi}$ so that, for any $(F_1, \ldots, F_d) \in \mathcal{F}^d$,

$$\tilde{\phi}(F_1, \ldots, F_d) = \phi\big(\mathbb{E}_{F_1}[A_1], \ldots, \mathbb{E}_{F_d}[A_d]\big),$$

where $A_i \sim F_i$. To see the functional $\tilde{\phi}$ as an extension of ϕ from $[0,1]^d$ to $\mathcal{F}^d$, we map an $f \in [0,1]^d$ into $\mathcal{F}^d$ as $\big(H_{\delta(f_1)}, \ldots, H_{\delta(f_d)}\big)$, where $H_{\delta(f_i)}$, as defined before, is the distribution function with its mass concentrated at f_i. In this subsection, we will derive the concave envelope of $\tilde{\phi}$ over its domain $\mathcal{F}^d$, that is we will find the lowest concave overestimator of the extension $\tilde{\phi}$ over $\mathcal{F}^d$. This envelope will be denoted as $\operatorname{conc}_{\mathcal{F}^d}(\tilde{\phi})$. More specifically, we show that when the outer-function ϕ satisfies certain conditions, $\operatorname{conc}_{\mathcal{F}^d}(\tilde{\phi})$ is the solution to an optimal transport problem [38]. This solution can be derived explicitly when ϕ satisfies some additional requirements.

Before characterizing the concave envelope of $\tilde{\phi}$ over $\mathcal{F}^d$, we discuss how this setting relates to the discrete case. In Section 3.1 and 3.2, we introduced a sequence of mappings $(x, f) \mapsto u \mapsto s \mapsto z$, where the first map evaluated underestimators, second map was defined by constructing two-dimensional concave envelopes, and the third map was via an affine transformation Z. It is the z-space that is intimately related to $\mathcal{F}^d$. More specifically, let $z = (z_1, \ldots, z_d)$, where $z_i = (z_{i0}, \ldots, z_{in}) \in \Delta_i$. Let F_i be defined as 1 at a_{in} and above, $1 - z_{ik}$ in $[a_{ik-1}, a_{ik})$ for $k \in \{1, \ldots, n\}$, and 0 below a_{i0}. Since the mapping from z to $(F_1, \ldots, F_d)$ is affine, results regarding concave envelopes over Δ translate to those about $\operatorname{conc}_{\mathcal{F}^d}(\tilde{\phi})(F_1, \ldots, F_d)$. The following treatment will generalize the above constructions, allowing for more general distributions that are not necessarily supported at a finite set of discrete points.

We start by making an observation assuming $d = 1$. If ϕ is a univariate convex function, by Jensen's inequality, $\tilde{\phi}(F_1) = \phi\big(\mathbb{E}_{F_1}[A_1]\big) \leq \mathbb{E}_{F_1}\big[\phi(A_1)\big]$. We now extend this idea to the multidimensional case. We will show in Lemma 3.3.1 that, as long as, ϕ is convex in each argument when other arguments are fixed, there exists a joint distribution $F \in \Pi(F_1, \ldots, F_d)$ so that

$$\tilde{\phi}(F_1, \ldots, F_d) \leq \mathbb{E}_F[\phi(A_1, \ldots, A_d)]. \tag{3.20}$$

The above inequality immediately implies that

$$\operatorname{conc}_{\mathcal{F}^d}\big(\tilde{\phi}\big)(F_1, \ldots, F_d) = \operatorname{conc}_{\mathcal{F}^d}\big(\tilde{\phi}|_{\operatorname{ext}(\mathcal{F}^d)}\big)(F_1, \ldots, F_d), \tag{3.21}$$

i.e., it suffices to restrict $\tilde{\phi}$ to the extreme points for $\mathcal{F}^d$ for the purpose of constructing $\operatorname{conc}_{\mathcal{F}^d}\big(\tilde{\phi}\big)$. To see this, observe that the left hand side in (3.21) is at least as large as the right hand side. We now argue the converse relationship. Observe that

$$\begin{aligned} \tilde{\phi}(F_1, \ldots, F_d) \leq \mathbb{E}_F\big[\phi(A_1, \ldots, A_d)\big] &= \int \tilde{\phi}(H_{\delta(a_1)}, \ldots, H_{\delta(a_d)}) \mathrm{d}F(a) \\ &\leq \operatorname{conc}_{\mathcal{F}^d}\big(\tilde{\phi}|_{\operatorname{ext}(\mathcal{F}^d)}\big)(F_1, \ldots, F_d), \end{aligned} \tag{3.22}$$

where the first inequality holds by the hypothesis (3.20), the equality is by definition of $\tilde{\phi}$. The last inequality holds by the concavity of $\operatorname{conc}_{\mathcal{F}^d}(\tilde{\phi}|_{\operatorname{ext}(\mathcal{F}^d)})$ and that $H_{\delta(a_i)}$ are the extreme points of $\mathcal{F}$, since $F_i(a_i) = \displaystyle\int H_{\delta(b_i)}(a_i) \mathrm{d}F_i(b_i)$. Then, it follows that the converse $\operatorname{conc}_{\mathcal{F}^d}(\tilde{\phi})(F_1, \ldots, F_d) \leq \operatorname{conc}_{\mathcal{F}^d}(\tilde{\phi}|_{\operatorname{ext}(\mathcal{F}^d)})(F_1, \ldots, F_d)$ holds. The relation (3.21) shows that $\operatorname{conc}_{\mathcal{F}^d}(\tilde{\phi})$ is the lowest concave extension of $\tilde{\phi}$ restricted to $\operatorname{ext}(\mathcal{F}^d)$. We now return to establish that the inequality (3.20) holds under certain hypothesis on the structure of ϕ.

Lemma 3.3.1 *Let $A_1, \ldots, A_d$ be independent random variables. If $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ is a continuous function which is convex in each argument when other arguments are fixed then $\tilde{\phi}(F_1, \ldots, F_d) \leq \mathbb{E}\big[\phi(A_1, \ldots, A_d)\big]$, where the equality is attained when the function ϕ is multilinear.*

Proof For any index set I, we will denote the joint distribution of $\{A_i \mid i \in I\}$ as F_I. The set of integers $\{i, \ldots, j\}$ will be denoted $[i, j]$ so that the joint distribution of $\{A_i, \ldots, A_d\}$ will be written as $F_{[i,j]}$. We prove the inequality in the statement of the result by induction on d. The base case $d = 1$ follows from Jensen's inequality as was remarked earlier. For the inductive step, we have:

$$\begin{aligned}
\tilde{\phi}(F_1, \ldots, F_d) &\leq \mathbb{E}_{F_{[1,d-1]}}\Big[\phi\Big(A_1, \ldots, A_{d-1}, \mathbb{E}_{F_d}\big[A_d\big]\Big)\Big] \\
&= \mathbb{E}_{F_{[1,d-1]}}\Big[\phi\Big(A_1, \ldots, A_{d-1}, \mathbb{E}_{F_d}\big[A_d \mid A_1, \ldots, A_{d-1}\big]\Big)\Big] \\
&= \mathbb{E}_{F_{[1,d-1]}}\Big[\phi\Big(\mathbb{E}_{F_d}\big[A_1, \ldots, A_{d-1}, A_d \mid A_1, \ldots, A_{d-1}\big]\Big)\Big] \\
&\leq \mathbb{E}_{F_{[1,d-1]}}\Big[\mathbb{E}_{F_d}\Big[\phi\big(A_1, \ldots, A_d\big) \,\Big|\, A_1, \ldots, A_{d-1}\Big]\Big] \\
&= \mathbb{E}_{F_{[1,d]}}\Big[\phi\big(A_1, \ldots, A_d\big)\Big],
\end{aligned}$$

where the first inequality is by induction hypothesis, the first equality is by the independence of A_i, the second equality is because $\mathbb{E}[A_i \mid A_i] = A_i$, the second inequality is due to Jensen's inequality, and the last equality holds because of the law of iterated expectations. The proof is complete by observing that each inequality becomes an equality if ϕ is linear when all but one of its arguments are fixed. ∎

Next, for an arbitrary $(F_1, \ldots, F_d) \in \mathcal{F}^d$, consider the following optimization problem

$$\hat{\phi}(F_1, \ldots, F_d) = \sup\Big\{\mathbb{E}_F\big[\phi(A_1, \ldots, A_d)\big] \,\Big|\, (F_1, \ldots, F_d) \in \mathcal{F}^d,\ F \in \Pi(F_1, \ldots, F_d)\Big\}, \tag{3.23}$$

where we recall that $\Pi(F_1, \ldots, F_d)$ denotes the set of joint distributions with $F_1, \ldots, F_d$ as marginals. The optimization problem (3.23) is called the *multivariate Monge-Kantorovich* problem on the real line (Section 2 in [59]). We next argue that (3.23) is a reformulation for the right hand side of (3.21).

Proposition 3.3.1 *For* $(F_1, \ldots, F_d) \in \mathcal{F}^d$*,* $\hat{\phi}(F_1, \ldots, F_d) = \mathrm{conc}_{\mathcal{F}^d}\big(\tilde{\phi}|_{ext(\mathcal{F}^d)}\big)(F_1, \ldots, F_d)$*. Moreover, if* ϕ *is continuous and convex in each argument when other arguments are fixed,* $\hat{\phi}(F_1, \ldots, F_d) = \mathrm{conc}_{\mathcal{F}^d}\big(\tilde{\phi}\big)(F_1, \ldots, F_d)$*.*

Proof Let $(F_1, \ldots, F_d) \in \mathcal{F}^d$. Then, $\hat{\phi}(F_1, \ldots, F_d) \leq \mathrm{conc}\big(\tilde{\phi}|_{\mathrm{ext}(\mathcal{F}^d)}\big)(F_1, \ldots, F_d)$ because, as we argued in (3.22) that for $F \in \Pi(F_1, \ldots, F_d)$, $\mathbb{E}_F[\phi(A_1, \ldots, A_d)] \leq$

$\operatorname{conc}\big(\tilde{\phi}|_{\operatorname{ext}(\mathcal{F}^d)}\big)(F_1, \ldots, F_d)$. To prove the converse, it suffices to show that $\hat{\phi}$ is concave in $\mathcal{F}^d$ because, for $(F_1, \ldots, F_d) \in \operatorname{ext}(\mathcal{F}^d)$, each $F_i = H_{\delta(a_i)}$ for some $a_i \in [0, 1]$ and it follows by considering the multidimensional Dirac distribution at $(a_1, \ldots, a_d)$ that $\hat{\phi}(F_1, \ldots, F_d) \geq \tilde{\phi}(F_1, \ldots, F_d)$. To see that $\hat{\phi}$ is concave, let $(F_1, \ldots, F_d)$ and $(G_1, \ldots, G_d)$ be two points in $\mathcal{F}^d$ and let α be chosen to satisfy $0 \leq \alpha \leq 1$. Then, for any $F \in \Pi(F_1, \ldots, F_d)$ and $G \in \Pi(G_1, \ldots, G_d)$, let $A \sim F$, $B \sim G$, and $C \sim \alpha F + (1-\alpha)G$. We have,

$$\begin{aligned}\hat{\phi}\big(\alpha(F_1, \ldots, F_d) + (1-\alpha)(G_1, \ldots, G_d)\big) &\geq \mathbb{E}_{(\alpha F + (1-\alpha)G)}[\phi(C_1, \ldots, C_d)] \\ &= \alpha \mathbb{E}_F[\phi(A_1, \ldots, A_d)] + (1-\alpha)\mathbb{E}_G[\phi(B_1, \ldots, B_d)],\end{aligned}$$

where the inequality holds because $\alpha F + (1-\alpha)G$ is a feasible solution to (3.23) at $\alpha(F_1, \ldots, F_d) + (1-\alpha)(G_1, \ldots, G_d)$, and the equality holds because expectation of a mixture distribution is the mixture of the expectations under distributions being mixed. Since the inequality holds for every (F, G) in $\Pi(F_1, \ldots, F_d) \times \Pi(G_1, \ldots, G_d)$, it also holds for the supremum of $\alpha \mathbb{E}_F\big[\phi(A_1, \ldots, A_d)\big] + (1-\alpha)\mathbb{E}_G\big[\phi(B_1, \ldots, B_d)\big]$ over $\Pi(F_1, \ldots, F_d) \times \Pi(G_1, \ldots, G_d)$. Therefore, $\hat{\phi}\big(\alpha(F_1, \ldots, F_d) + (1-\alpha)(G_1, \ldots, G_d)\big) \geq \alpha\hat{\phi}(F_1, \ldots, F_d) + (1-\alpha)\hat{\phi}(G_1, \ldots, G_d)$, showing that $\hat{\phi}$ is concave.

The second statement follows from the first statement since Lemma 3.3.1 implies (3.21). ∎

As a result, when the inequality in (3.20) is satisfied, the functional $\hat{\phi}(F_1, \ldots, F_d)$ coincides with the lowest concave overestimator of $\tilde{\phi}$ over $\mathcal{F}^d$. Then, to compute $\hat{\phi}$ at $(F_1, \ldots, F_d)$, we need to solve the multivariate Monge-Kantorovich problem (3.23). The latter problem has an explicit solution under certain conditions (see, for example, Theorem 5 in [60]).

Theorem 3.3.1 ([61] and Theorem 5 in [60]) *Let $F_1, \ldots, F_d$ be d probability distribution functions on the real line and let $F^*(a) = \min_i F_i(a_i)$. Then, for any continuous supermodular function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$,*

$$\sup_{F \in \Pi(F_1, \ldots, F_d)} \int \phi \mathrm{d}F = \int \phi \mathrm{d}F^*$$

if $\phi \leq \varphi$ for some continuous function φ such that $\int \varphi \mathrm{d}T$ is finite and constant for all $F \in \Pi(F_1, \ldots, F_d)$.

We apply this result in our setting to obtain an explicit integral representation of the optimal functional $\hat{\phi}$.

Theorem 3.3.2 *If $\phi : [0,1]^d \mapsto \mathbb{R}$ is a continuous supermodular function then*

$$\hat{\phi}(F_1, \ldots, F_d) = \mathbb{E}_{F^*}[\phi(A_1, \ldots, A_d)] = \int_0^1 \phi\big(F_1^{-1}(\lambda), \ldots, F_n^{-1}(\lambda)\big)\mathrm{d}\lambda,$$

where $F^*(a) = \min\{F_1(a_1), \ldots, F_d(a_d)\}$.

Proof Let $(F_1, \ldots, F_d) \in \mathcal{F}^d$. Since $A_1, \ldots, A_d$ have supports in $[0,1]$ and the function ϕ is continuous, $\mathbb{E}_F[\phi(A_1, \ldots, A_d)]$ is finite for all $F \in \Pi(F_1, \ldots, F_d)$. We choose $\varphi(a)$ to be a constant c defined as $\max_{a' \in [0,1]^d} |\phi(a')|$. By definition, $\phi(f) \leq c$ and $\int c\mathrm{d}F = c$ is finite for all $F \in \Pi(F_1, \ldots, F_d)$. It follows from Theorem 3.3.1 that we have $\hat{\phi}(F_1, \ldots, F_d) = \mathbb{E}_{F^*}[\phi(A_1, \ldots, A_d)]$. Now, consider a random variable U that is uniformly distributed over $[0,1]$ and observe that, for all $(a_1, \ldots, a_d) \in \mathbb{R}^d$,

$$\begin{aligned}\Pr\big(F_1^{-1}(U) \leq a_1, \ldots, F_d^{-1}(U) \leq a_d\big) &= \Pr(U \leq F_1(a_1), \ldots, U \leq F_d(a_d)) \\ &= \min\big\{F_1(a_1), \ldots, F_d(a_d)\big\},\end{aligned}$$

where the first equality is because $F_i^{-1}(U) \leq a_i$ if and only if $U \leq F_i(a_i)$ and the second equality is because U is uniformly distributed. In other words, the distribution function of the random vector $\big(F_1^{-1}(U), \ldots, F_d^{-1}(U)\big)$ is $F^*(a)$. Let $A^* = (A_1^*, \ldots, A_d^*) \sim F^*$, and, for $\delta = 2^{-m}$, define

$$\begin{aligned}M_k &= \Pr\big\{(A_1^*, \ldots, A_d^*) \bigm| k\delta \leq \phi(A_1^*, \ldots, A_d^*) < (k+1)\delta\big\} \\ &= \Pr\Big\{\big(F_1^{-1}(U), \ldots, F_d^{-1}(U)\big) \Bigm| k\delta \leq \phi\big(F_1^{-1}(U), \ldots, F_d^{-1}(U)\big) < (k+1)\delta\Big\}.\end{aligned}$$

Then, it follows that

$$\int \phi \mathrm{d}F^* = \lim_{m \to \infty} \sum_{k=0}^{2^m - 1} k\delta M_k = \mathbb{E}_U\Big[\phi\big(F_1^{-1}(U), \ldots, F_d^{-1}(U)\big)\Big],$$

where both equalities follow from the piecewise approximations of ϕ where it is replaced with $k\delta$ whenever it evaluates to a value in the range $[k\delta, (k+1)\delta)$ and the

Dominated Convergence Theorem (see Theorem 16.4 in [62]), which applies because of the existence of φ. ■

Similar to Corollary 3.2.4, Theorem 3.3.2 can be used to characterize $\hat{\phi}$ for functions ϕ that become supermodular when their domain is transformed affinely, by using an operation such as the switching operation. Recall that the function $\phi(T)$, obtained by switching the domain of $\phi : [0,1]^d \mapsto \mathbb{R}$, is described using a set $T \subseteq \{1, \ldots, d\}$ so that $\phi(T)(f) = \phi(f(T))$, where $f(T)_i = 1 - f_i$ if $i \in T$ and $f(T)_i = f_i$ otherwise. We define a marginal distribution $F_i(T)$ so that

$$F_i(T)(b) = \Pr\{1 - A_i \leq b\} \text{ for } i \in T \text{ and } F_i(T)(b) = \Pr\{A_i \leq b\} \text{ otherwise.} \quad (3.24)$$

Then, it follows that for $i \in T$

$$\begin{aligned} F_i(T)^{-1}(\lambda) &= \min\bigl\{b \bigm| F_i(T)(b) \geq \lambda\bigr\} = \min\Bigl\{b \Bigm| 1 - \sup_{a<1-b} F_i(a) \geq \lambda\Bigr\} \\ &= 1 - \max\Bigl\{d \Bigm| 1 - \sup_{a<d} F_i(a) \leq 1 - \lambda\Bigr\} \\ &= 1 - \sup\{a \mid F_i(a) \leq 1 - \lambda\}. \end{aligned}$$

The following result explicitly characterizes $\hat{\phi}$ when $\phi(T)$, instead of ϕ, is supermodular.

Corollary 3.3.1 *If $\phi(T)$ is a continuous function and there exists a $T \subseteq \{1, \ldots, d\}$ so that $\phi(T)$ is supermodular then*

$$\hat{\phi}(F_1, \ldots, F_d) = \int_0^1 \phi(T)\bigl(F_1(T)^{-1}(\lambda), \ldots, F_d(T)^{-1}(\lambda)\bigr)\mathrm{d}\lambda.$$

Proof We will show that $\hat{\phi}(F_1, \ldots, F_d)$ equals $\widehat{\phi(T)}\bigl(F_1(T), \ldots, F_d(T)\bigr)$ and then, the result follows directly from Theorem 3.3.2 using supermodularity of $\phi(T)$. To show $\hat{\phi}(F_1, \ldots, F_d) \leq \widehat{\phi(T)}\bigl(F_1(T), \ldots, F_d(T)\bigr)$, consider $F \in \Pi(F_1, \ldots, F_d)$, and let $A \sim F$. Define a random vector A' so that, for $S \in [0,1]^d$, $\Pr\{A' \in S\} := \Pr\bigl\{A \in \{a(T) \mid a \in S\}\bigr\}$. Then, let F' be the cumulative distribution function of A', *i.e.*, $F'(a) = \Pr\bigl\{A \in$

$\{a(T) \mid a \in [0, a]\}\big\}$. For $\delta = 2^{-m}$, let $S_k := \big\{a \in [0,1]^d \bigm| k\delta \le \phi(a) < (k+1)\delta\big\}$, and thus,

$$\begin{aligned} M_k :&= \Pr\big\{A \in S_k\big\} \\ &= \Pr\big\{A' \in \{a(T) \mid a \in S_k\}\big\} \\ &= \Pr\big\{A' \bigm| k\delta \le \phi(T)(A') < (k+1)\delta\big\}, \end{aligned} \tag{3.25}$$

where the second equality holds by the definition of A', and the third equality holds because $\phi(T)\big(a(T)\big) = \phi(a)$. Therefore,

$$\int \phi \mathrm{d}F = \lim_{m\to\infty} \sum_{k=0}^{2^m-1} k\delta M_k = \int \phi(T)\mathrm{d}F' \le \widehat{\phi(T)}\big(F_1(T), \ldots, F_d(T)\big),$$

where the two equalities follow from the Dominated Convergence Theorem (see Theorem 16.4 in [62]) and the first and last equalities in (3.25), and the inequality holds because, using (3.24), the marginal distribution of A'_i is given by $\Pr\{A'_i \le a_i\} = F_i(T)(a_i)$. Hence, $\hat{\phi}(F_1, \ldots, F_d) \le \widehat{\phi(T)}\big(F_1(T), \ldots, F_d(T)\big)$. Since the reverse inequality follows by a similar argument, the proof is complete. ∎

3.3.2 Composite relaxations via random variables

In this subsection, we will assume that, for each inner function, an underestimator is available and is parametrized by its real-valued upper bound a_i. For a given $x \in X$, the underestimator will vary with bound a_i and will be used to derive the marginal cumulative distribution function F_i used in Section 3.3.1. Then, we will use Theorem 3.3.2 to construct the composite relaxation. To relate the marginal distributions to the underestimating functions, we will find it useful to work with an alternate characterization of a distribution function with support over $[0, 1]$ in terms of a concave function on the real line. To derive this function, we truncate the associated random variable to lie below a bound and study how the expectation varies with this bound. Formally, we define $E : \mathbb{R} \times \mathcal{F}$ as:

$$E(a_i, F_i) = \mathbb{E}_{F_i}\big[\min\{A_i, a_i\}\big] \qquad \text{for } a_i \in \mathbb{R} \text{ and } F_i \in \mathcal{F}, \tag{3.26}$$

where $A_i \sim F_i$. We will write $E_{F_i}(a_i)$ (resp. $E_{a_i}(F_i)$) when we wish to convey that F_i (resp. a_i) is fixed. It is the right derivative of $E_{F_i}(a_i)$ that relates to the cumulative distribution F_i. Recall that the left and right derivative of a univariate function $s(a)$ are defined as $s'_-(a) = \lim_{\delta \nearrow 0} \frac{s(a+\delta)-s(a)}{\delta}$ and $s'_+(a) = \lim_{\delta \searrow 0} \frac{s(a+\delta)-s(a)}{\delta}$.

Lemma 3.3.2 (Theorem 1 in [63]) *For a distribution function $F_i \in \mathcal{F}$, the univariate function $E_{F_i}(a_i)$ is non-decreasing concave such that*

$$E_{F_i}(0) = 0,\ E_{F_i}(1) = \mathbb{E}[A_i],\ (E_{F_i})'_-(a_i) = 1 \text{ for } a_i \leq 0,\ (E_{F_i})'_+(a_i) = 0 \text{ for } a_i \geq 1. \tag{3.27}$$

The distribution function F_i can be recovered from E_{F_i} using $F_i(a_i) = 1 - (E_{F_i})'_+(a_i)$, where $(E_{F_i})'_+(x) := \lim_{x' \searrow x} \frac{E_{F_i}(x') - E_{F_i}(x)}{x'-x}$, the right derivative of E_{F_i}. On the other hand, any concave function $s_i(a_i)$ on $\mathbb{R}$ with the properties that

$$s_i(0) = 0,\ s_i(1) = \text{a finite value},\ (s_i)'_-(a_i) = 1 \text{ for } a_i \leq 0,\ (s_i)'_+(a_i) = 0 \text{ for } a_i \geq 1 \tag{3.28}$$

is $E_{F_i}(a_i)$ for some cdf $F_i \in \mathcal{F}$.

Proof For a distribution function $F_i \in \mathcal{F}$, the univariate function $E_{F_i}(a_i)$ is clearly non-decreasing. It is concave because, for $a'_i, a''_i \in \mathbb{R}$ and $\alpha \in [0,1]$, $\mathbb{E}_{F_i}[\min\{A_i, \alpha a'_i + (1-\alpha)a''_i\}] \leq \alpha \mathbb{E}_{F_i}[\min\{A_i, a'_i\}] + (1-\alpha)\mathbb{E}_{F_i}[\min\{A_i, a''_i\}]$, where the inequality holds by the concavity of $\min\{a_i, a'_i\}$ in a'_i. Moreover, $E_{F_i}(0) = \displaystyle\int \min\{a_i, 0\} \mathrm{d}F_i = 0$ and $E_{F_i}(1) = \displaystyle\int \min\{a_i, 1\} \mathrm{d}F_i = \mathbb{E}_{F_i}[A_i]$. In addition, for $a_i \leq 0$

$$\lim_{\delta \searrow 0} \frac{E_{F_i}(a_i) - E_{F_i}(a_i - \delta)}{\delta} = \lim_{\delta \searrow 0} \frac{a_i - (a_i - \delta)}{\delta} = 1;$$

and for $a_i \geq 1$

$$\lim_{\delta \searrow 0} \frac{E_{F_i}(a_i + \delta) - E_{F_i}(a_i)}{\delta} = \lim_{\delta \searrow 0} \frac{\mathbb{E}[A_i] - \mathbb{E}[A_i]}{\delta} = 0.$$

Therefore, it follows from Theorem 1 in [63] that $F_i(a_i)$ can be recovered using $1 - (E_{F_i})'_+(a_i)$.

Now, let $s_i(a_i)$ be a concave function satisfying (3.28). It follows from $(s_i)'_-(0) = 1$, the concavity of s_i, and $s(0) = 0$ that $s_i(a_i) \le a_i$. Similarly, it follows from $s_i(1) = c$ for some constant c, $(s_i)'_+(1) = 0$, and the concavity of s_i that $s_i(a_i) \le c$ for all a_i. Finally, since $(s_i)'_-(a_i) = 1$ for $a_i \le 0$, it follows that, for $b_i < 0$, $s_i(b_i) = s_i(b_i) - s_i(0) = \displaystyle\int_0^{b_i} (s_i)'_-(a_i) = b_i$. Therefore, $\lim_{a_i \to -\infty}(a_i - s_i(a_i)) = 0$. Thus, by Theorem 1 in [63], there exists a distribution function F_i with support over $\mathbb{R}$ such that $E_{F_i}(a_i) = s_i(a_i)$. We are done if we show that $F_i(b_i) = 0$ for $b_i < 0$ and $F_i(b_i) = 1$ for $b_i > 1$. Assume that there exists $b_i < 0$ such that $F_i(b_i) > 0$. Then, $s_i(0) = \displaystyle\int \min\{a_i, 0\} \mathrm{d}F_i < 0$, a contradiction. Similarly, suppose that $F_i(1) < 1$. This case also leads to a contradiction as follows, $\lim_{a_i \to \infty} s_i(a_i) = \displaystyle\int a_i \mathrm{d}F_i > \int \min\{a_i, 1\} \mathrm{d}F_i = c \ge \lim_{a_i \to \infty} s_i(a_i)$, where the last inequality follows because $s_i(a_i) \le c$ for all a_i. ■

Lemma 3.3.2 establishes a connection between certain univariate concave functions and distribution functions. We now relate these univariate fun

functions with certain underestimators of the inner functions. Formally, we consider a function $s_i : W \times \mathbb{R} \mapsto \mathbb{R}$ such that, for $(x, f, a_i) \in W \times \mathbb{R}$,

$$s_i(x, f, a_i) \in \Big[\min\big\{a_i, f_i^L + \big(f_i - f_i^L\big)\mathbb{1}_{a_i \ge f_i^U}\big\},\ \min\big\{a_i, f_i\big\}\Big], \tag{3.29}$$

where W outer-approximates the graph of inner function $f(x)$, and $\mathbb{1}_{\text{clause}}$ is one if the clause is true and 0 otherwise. For any (x, f, a_i), the range of values for $s_i(x, f, a_i)$ is non-empty because $f_i^L + \big(f_i - f_i^L\big)\mathbb{1}_{a_i \ge f_i^U} \le f_i$. In other words, (3.29) requires that, for $a_i \in [0, 1]$, $s_i(x, f, a_i)$ underestimates $\min\{f_i, a_i\}$ over W, for $a_i \le 0$, $s_i(x, f, a_i)$ is a_i, and, for $a_i \ge 1$, the function coincides with f_i. We construct one such function in the following remark.

Remark 3.3.1 *Let W_i be a convex outerapproximation of the graph of inner function $f_i(x)$. With each constant $a_i \in \mathbb{R}$, associate a set $S_i(a_i) := \big\{(x, f_i, \rho_i) \bigm| \rho_i \ge \min\{a_i, f_i\},\ (x, f_i) \in W_i\big\}$. Define $s_i : W_i \times \mathbb{R} \mapsto \mathbb{R}$ so that, for any $a_i \in \mathbb{R}$,*

$$s_i\big(x, f_i, a_i\big) := \inf\Big\{\rho_i \Bigm| (x, f_i, \rho_i) \in \operatorname{conv}\big(S_i(a_i)\big)\Big\}. \tag{3.30}$$

To see that the function s_i satisfies the requirements in (3.29), consider a constant $a_i \in \mathbb{R}$. If $a_i \leq 0$ then $\min\{a_i, f_i\}$ equals a_i, the set $S_i(a_i)$ is convex, and $s_i(x, f_i, a_i) = a_i$. Similarly, if $a_i \geq 1$, $\min\{a_i, f_i\}$ equals f_i, $S_i(a_i)$ is convex, and $s_i(x, f_i, a_i) = f_i$. If $0 \leq a_i \leq 1$ then $0 \leq s_i(x, f_i, a_i) \leq \min\{a_i, f_i\}$, where the first inequality holds because, for each $(x, f_i, \rho_i) \in S_i(a_i)$, $0 \leq \min\{a_i, f_i\} \leq \rho_i$, and the second inequality holds because, for every $(x, f_i) \in W_i$, $\bigl(x, f_i, \min\{a_i, f_i\}\bigr) \in \operatorname{conv}\bigl(S_i(a_i)\bigr)$. Furthermore, $s_i(x, f_i, a_i)$ is a convex function over W_i; see Theorem 5.3 in [64]. In fact, for any fixed $a_i \in \mathbb{R}$, $s_i(x, f_i, a_i)$ is the convex envelope of the function $\min\{a_i, f_i\}$ over W_i. In contrast, for any fixed $(x, f_i) \in W_i$, $s_i(x, f_i, a_i)$ is a concave function in a_i. To see this, consider two distinct points a_i', a_i'' in $\mathbb{R}$, and define $\tilde{a}_i := \lambda a_i' + (1-\lambda)a_i''$ for some $\lambda \in (0, 1)$. Since $\lambda s_i(x, f_i, a_i') + (1-\lambda)s_i(x, f_i, a_i'')$ (resp. $s_i(x, f_i, \tilde{a}_i)$) is a convex underestimator (resp. convex envelope) of $\min\{f_i, \tilde{a}_i\}$ over W_i, it follows that $\lambda s_i(x, f_i, a_i') + (1-\lambda)s_i(x, f_i, a_i'') \leq s_i(x, f_i, \tilde{a}_i)$.

In the following example, we consider the quadratic term, derive the underestimator (3.30) explicitly, and illustrate that this underestimator, treated as a function of a_i, is concave and, via the transformation discussed in Lemma 3.3.2, yields a distribution function.

Example 3.3.1 *Consider the quadratic term x_1^2 over the interval $[0, 2]$. Here, the quadratic term varies over $[0, 4]$, while in our formal treatment, we have assumed that $f_i \in [0, 1]$. However, as we discussed before, this does not pose any issues since an affine transformation of the function can be used to normalize any bounded range to $[0, 1]$. In our current setting, we could use $\frac{1}{4}x_1^2$ as the inner function instead of x_1^2. The function $s_1(x_1, x_1^2, a_1)$ defined in (3.30) can be computed explicitly as follows*

$$s_1(x_1, x_1^2, a_1) = \begin{cases} a_1 & a_1 \leq 0 \\ \bigl(-4 + 2\sqrt{-a_1 + 4}\bigr)(2 - x_1) + a_1 & 0 \leq 2 - \sqrt{-a_1 + 4} \leq x_1 \leq 2, \\ x_1^2 & \textit{otherwise.} \end{cases} \tag{3.31}$$

Figure 3.3a illustrates the function at $a_1 = 3$, where we see that this function is the largest convex underestimator of x_1^2 over $[0, 2]$ bounded by 3. In contrast, Figure 3.3b depicts $s_1(x_1, x_1^2, a_1)$ as a function of a_1 at $x_1 = 1$ and it is easily verified that this function is concave and satisfies the requirements in (3.28). Therefore, by Lemma 3.3.2, $(s_1)'_+(1, 1, a_1)$ is a survival function (1− distribution function) for a random variable with support in $[0, 4]$ and is depicted in Figure 3.3c. For general a_1, the right derivative of $s_1(x_1, x_1^2, a_1)$ with respect to a_1 is as follows:

$$(s_1)'_+(x_1, x_1^2, a_1) = \begin{cases} 1 & a_1 < 0, \\ 1 - \frac{2-x_1}{\sqrt{-a_1+4}} & 0 \le 2 - \sqrt{-a_1+4} < x_1 \le 2, \\ 0 & \textit{otherwise.} \end{cases} \tag{3.32}$$

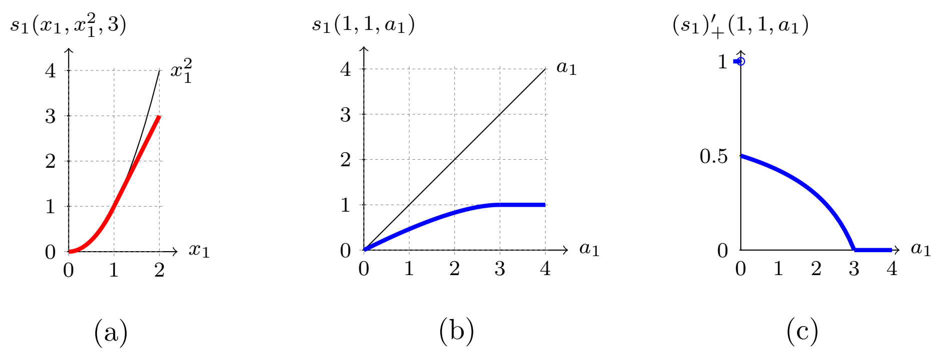


Fig. 3.3.: (a) the largest convex underestimator of x_1^2 over $[0, 2]$ that is bounded from above by 3. (b) a concave function which satisfies (3.28). (c) the right derivative of $s_1(1, 1, a_1)$ is a survival function.

In the following result, we formally relate the underestimator $s_i(x, f, a_i)$ with a distribution function. Assume that we are given a function $s_i(x, f, a_i)$ that satisfies (3.29). We denote by $(s_i)'_+(x, f, a_i)$ the right derivative of $s_i(x, f, a_i)$ with respect to a_i. We will fix $(x, f) \in W$ and characterize $1 - (s_i)'_+(x, f, a_i)$ as a distribution function, which, for notational brevity, we denote as $S_i^{x,f}(a_i)$.

Proposition 3.3.2 *Let $s_i : W \times \mathbb{R}$ be a function satisfying (3.29), and define $S_i^{x,f}(a_i) := 1-(s_i)'_+(x,f,a_i)$. If, for any given $(x,f) \in W$, $s_i(x,f,a_i)$ is concave in a_i then $S_i^{x,f}$ is a distribution function so that $\int a_i \mathrm{d}S_i^{x,f}(a_i) = f_i$. Moreover, $s_i(x,f,a_i) = E_{S_i^{x,f}}(a_i)$.*

Proof It is easy to verify that function $s_i(x,f,a_i)$ that is concave in a_i and satisfies the requirement in (3.29) also satisfies the four conditions in (3.28). In particular, since $s_i(x,f,a_i)$ equals a_i (resp. f_i) when $a_i \le 0$ (resp. $a_i \ge 1$), it follows that $(s_i)'_-(x,f,a_i)$ (resp. $(s_i)'_+(x,f,a_i)$) equals 1 (resp. 0). Then, by the second part of Lemma 3.3.2, there exists a distribution function $F_i \in \mathcal{F}$ such that $s_i(x,f,a_i) = E_{F_i}(a_i)$. By the first part of Lemma 3.3.2, $F_i(a_i) = 1-(E_{F_i})'_+(a_i) = 1-(s_i)'_+(x,f,a_i) = S_i^{x,f}(a_i)$. Moreover, $\int a_i \mathrm{d}S_i^{x,f}(a_i) = \lim_{a_i \to +\infty} s_i(x,f,a_i) = f_i$, where the first equality is because $s(x,f,a_i)$ equals $E_{S_i^{x,f}}(a_i)$, which in turn approaches the RHS as $a_i \to \infty$, and the second equality follows directly from (3.29). ■

Equipped with Proposition 3.3.1 and Proposition 3.3.2, we are ready to derive the limiting relaxation for the hypograph of $\phi \circ f$ as follows. For each $i \in \{1,\dots,d\}$, let $S_i^{x,f}(a_i)$ be a function defined as in Proposition 3.3.2. Then, if the outer-function ϕ satisfies (3.20), for example, as in Lemma 3.3.1, if ϕ is continuous and convex in each argument when other arguments are fixed, we obtain that, for every $(x,f) \in W$,

$$\phi(f) = \tilde{\phi}(S_1^{x,f},\dots,S_d^{x,f}) \le \hat{\phi}\big(S_1^{x,f},\dots,S_d^{x,f}\big),$$

where the first equality holds by Proposition 3.3.2 and the definition of $\tilde{\phi}$, and the inequality holds because, by Proposition 3.3.1, $\hat{\phi}$ is the lowest concave overestimator of $\tilde{\phi}$ over $\mathcal{F}^d$. We will show that the limiting relaxation $\hat{\phi}(S_1^{x,f},\dots,S_d^{x,f})$ has a convex representation in the space of variables (x,f). We illustrate the ideas on an example before providing a formal discussion.

Example 3.3.2 *Consider $x_1^2x_2^2$ over the rectangle $[0,2]^2$. We use Proposition 3.3.2 to derive distribution functions from underestimators of x_i^2. For underestimator $s_i(x,x_i^2,a_i)$ given in (3.31), $1-(s_i)'_+(x,x_i^2,a_i)$ is easily computed using the right derivative in (3.32). For notational brevity, let $S_i^{x_i}(a_i)$ denote $1-(s_i)'_+(x,x_i^2,a_i)$*

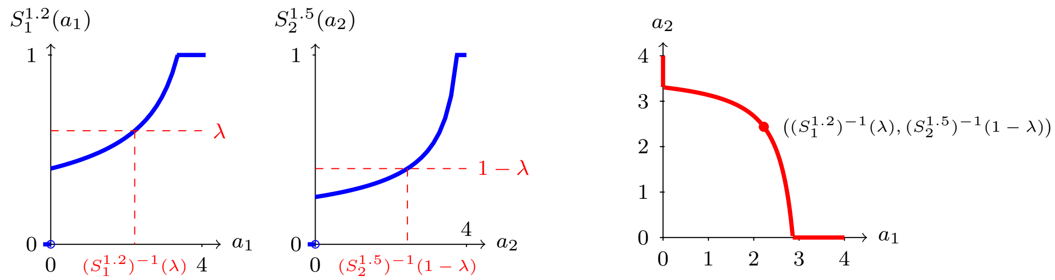
since it depends only on the i^{th} coordinate of x. For any $x_i \in [0,2]$, it follows from Proposition 3.3.2 that $S_i^{x_i}$ is a distribution function of a random variable $A_i^{x_i}$ such that $\mathbb{E}[A_i^{x_i}] = x_i^2$. Then, in this example setting, our construction is essentially derived from the following argument:

$$\begin{aligned}
x_1^2x_2^2 &= \mathbb{E}[A_1^{x_1}]\mathbb{E}[A_2^{x_2}] = \mathbb{E}_D[A_1^{x_1}A_2^{x_2}] \qquad \textit{where } D(a_1,a_2) := S_1^{x_1}(a_1)S_2^{x_2}(a_2) \\
&\geq \inf_G \bigl\{\mathbb{E}_G[A_1^{x_1}A_2^{x_2}] \bigm| G \in \Pi(S_1^{x_1}, S_2^{x_2})\bigr\} \\
&= \mathbb{E}_{G^*}[A_1^{x_1}A_2^{x_2}] \qquad \textit{where } G^*(a_1,a_2) := \max\{0, S_1^{x_1}(a_1) + S_2^{x_2}(a_2) - 1\} \\
&= \mathbb{E}_U\Bigl[\bigl(S_1^{x_1}\bigr)^{-1}(U)\bigl(S_2^{x_2}\bigr)^{-1}(1-U)\Bigr],
\end{aligned} \tag{3.33}$$

where the first equality is because $\mathbb{E}[A_i^{x_i}] = x_i^2$, the second equality is by constructing D by coupling $A_1^{x_1}$ and $A_2^{x_2}$ independently, the first inequality holds because the product distribution D has $S_1^{x_1}$ and $S_2^{x_2}$ as marginals. The third equality holds because G^* is feasible to the optimization problem on the LHS, and because, for two marginals $S_1^{x_1}, S_2^{x_2} \in \mathcal{F}$, $\Pr\{(A_1^{x_1} > a_1) \cup (A_2^{x_2} > a_2)\} \leq \Pr\{A_1^{x_1} > a_1\} + \Pr\{A_2^{x_2} > a_2\}$. This implies that, for $G \in \Pi(S_1^{x_1}, S_2^{x_2})$, $G(a_1,a_2) \geq G^*(a_1,a_2)$, where the RHS is known as Hoeffding-Fréchet lower bound [65, 66], and, thus, $\mathbb{E}_G[A_1^{x_1}A_2^{x_2}] \geq \mathbb{E}_{G^*}[A_1^{x_1}A_2^{x_2}]$ by [?] since the bilinear term is a correlation affine function. The third equality also follows from the more general result in Corollary 3.3.1 choosing either $T = \{1\}$ or $T = \{2\}$. The last equality holds because the distribution function of the random vector $\bigl((S_1^{x_1})^{-1}(U), (S_2^{x_2})^{-1}(1-U)\bigr)$ is G^*. We depict in Figure 3.4a the marginal distributions $S_1^{1.2}$ and $S_2^{1.5}$. For a given $U = \lambda$, we compute their inverse values to locate a point on the curve that is the locus of support points for G^*; see Figure 3.4b. We evaluate the last term in (3.33) by integrating the function value at points on this curve to derive the following limiting composite relaxation:

$$\begin{aligned}
x_1^2x_2^2 &\geq \max\biggl\{0, \int_{1-\frac{x_1}{2}}^{\frac{x_2}{2}} \Bigl(\frac{4\lambda^2 - 4 + 4x_1 - x_1^2}{\lambda^2}\Bigr)\Bigl(\frac{4\lambda^2 - 8\lambda + 4x_2 - x_2^2}{(1-\lambda)^2}\Bigr)\mathrm{d}\lambda\biggr\} \\
&= \max\biggl\{0, 2\ln\biggl[\frac{x_1x_2}{(2-x_1)(2-x_2)}\biggr] - 4\cdot(x_1+x_2-2)\cdot\bigl((3-x_1)\cdot(2-x_2)-x_1\bigr)\biggr\},
\end{aligned}$$

whose convexity, although not directly apparent from the resulting formula, is a consequence of Corollary 3.3.2 proved later.



(a) marginal distributions

(b) support of optimal transport plan

Fig. 3.4.: Underestimating $x_1^2 x_2^2$ at the point $x = (1.2, 1.5)$ via optimal transport

To show that the the limiting relaxation is convex in (x, f), we will need certain monotonicity properties of the optimal functional $\hat{\phi}$. To this end, we review order relations over $\mathcal{F}$. Let $F_i, G_i \in \mathcal{F}$ be two distribution functions. We say $F_i \preceq G_i$ if

$$E_{F_i}(a_i) \leq E_{G_i}(a_i) \quad \text{for all } a_i \in \mathbb{R} \quad \text{and} \quad E_{F_i}(a_i) = E_{G_i}(a_i) \quad \text{as } a_i \to \infty. \tag{3.34}$$

If $A_i \sim F_i$ and $B_i \sim G_i$, then the order $F_i \preceq G_i$ can equivalently be written as $\mathbb{E}\big[\min\{A_i, a_i\}\big] \leq \mathbb{E}\big[\min\{B_i, a_i\}\big]$ for all $a_i \in \mathbb{R}$ and $\mathbb{E}[A_i] = \mathbb{E}[B_i]$. This ordering is typically referred to as the *concave order* of two random variables (see Theorem 3.A.1 in [67]), which can equivalently be defined as follows. A_i is said to be smaller than B_i in the *concave order* if $\mathbb{E}[\psi(A_i)] \leq \mathbb{E}[\psi(B_i)]$ for all concave functions $\psi : \mathbb{R} \mapsto \mathbb{R}$, provided the expectations exist and are equal. Next, we consider another order which is defined by dropping the second requirement in (3.34). Namely, we say $F_i \prec G_i$ if

$$E_{F_i}(a) \leq E_{G_i}(a_i) \quad \text{for } a_i \in \mathbb{R},$$

or, equivalently, $\mathbb{E}\big[\min\{A_i, a_i\}\big] \leq \mathbb{E}\big[\min\{B_i, a_i\}\big]$ for all $a_i \in \mathbb{R}$. This order is the *increasing concave order* of two random variables A_i and B_i, and is equivalently defined by requiring, $\mathbb{E}[\psi(A_i)] \leq \mathbb{E}[\psi(B_i)]$ for all increasing concave function $\psi : \mathbb{R} \mapsto \mathbb{R}$ (see Theorem 4.A.2 in [67]). Given two tuples of distribution functions $(F_1, \ldots, F_d)$ and $(G_1, \ldots, G_d)$, we say $(F_1, \ldots, F_d) \preceq (G_1, \ldots, G_d)$ (resp. $(F_1, \ldots, F_d) \prec (G_1, \ldots, G_d)$) if, for each $i \in \{1, \ldots, d\}$, $F_i \preceq G_i$ (resp. $F_i \prec G_i$). To prove the monotonicity of $\hat{\phi}$, we also need the following technical lemma.

Lemma 3.3.3 (Theorem 1 in [68]) *Let $\psi(\lambda, u_1, \ldots, u_d)$ be a continuous function mapping from $[0,1] \times \mathbb{R}^d$ to $\mathbb{R}$. Then, we have*

$$\int_0^1 \psi(\lambda, \eta_1, \ldots, \eta_d)\mathrm{d}\lambda \leq \int_0^1 \psi(\lambda, \gamma_1, \ldots, \gamma_d)\mathrm{d}\lambda$$

for each system of non-decreasing bounded univariate functions η_i, γ_i, $i = 1, \ldots, d$, such that

$$\int_0^p \eta_i(\lambda)\mathrm{d}\lambda \geq \int_0^p \gamma_i(\lambda)\mathrm{d}\lambda \quad 0 \leq p \leq 1 \text{ and } \int_0^1 \eta_i(\lambda)\mathrm{d}\lambda = \int_0^1 \gamma_i(\lambda)\mathrm{d}\lambda, \tag{3.35}$$

if and only if the function ψ is convex in u_i when the other arguments are fixed, supermodular over $\mathbb{R}^d$ when λ is fixed, and

$$\int_0^\delta \Big(\psi(p+\delta+\lambda, u) - \psi(p+\delta+\lambda, u-he_i) + \psi(p+\lambda, u-he_i) - \psi(p+\lambda, u)\Big)\mathrm{d}\lambda \geq 0 \tag{3.36}$$

for all $0 \leq p \leq 1-2\delta$, $\delta > 0$, $h \geq 0$, $i = 1, \ldots, d$, where e_i is the i^{th} principal vector in $\mathbb{R}^d$.

Proposition 3.3.3 *Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a continuous function which is convex in each argument when the other arguments are fixed, and assume that there exists a $T \subseteq \{1, \ldots, d\}$ so that $\phi(T)$ is supermodular. If $(F_1, \ldots, F_d) \preceq (G_1, \ldots, G_d)$ then $\hat{\phi}(F_1, \ldots, F_d) \geq \hat{\phi}(G_1, \ldots, G_d)$. The weaker condition $(F_1, \ldots, F_d) \prec (G_1, \ldots, G_d)$ suffices to show $\hat{\phi}(F_1, \ldots, F_d) \geq \hat{\phi}(G_1, \ldots, G_d)$ if $\phi(T)$ is also non-increasing in each argument.*

Proof Assume that $(F_1, \ldots, F_d) \preceq (G_1, \ldots, G_d)$. We will show that $\hat{\phi}(F_1, \ldots, F_d) \geq \hat{\phi}(G_1, \ldots, G_d)$. Let $\eta_i(\lambda) := F_i(T)^{-1}(\lambda)$ and define $\psi(\lambda, \eta_1, \ldots, \eta_d) = \phi(T)(\eta_1, \ldots, \eta_d)$. Observe that the hypothesis on ψ in Lemma 3.3.3 is satisfied by this definition since $\phi(T)$ is independent of λ and is assumed to be supermodular and convex in each argument when the others are fixed. Moreover, since the first argument of ψ is ignored, the condition (3.36) holds trivially. Now, we show that (3.35) is satisfied with our definition. Since, for any univariate function ψ, $F_i \preceq G_i$ implies that $\mathbb{E}[\psi(1-A_i)] \leq \mathbb{E}[\psi(1-B_i)]$ where $A_i \sim F_i$ and $B_i \sim G_i$, it follows that $F_i(T) \preceq G_i(T)$, where

$F_i(T)$ and $G_i(T)$ are defined as in (3.24). It follows from Theorem 3.A.5 in [67] that $F_i(T) \preceq G_i(T)$ if and only if

$$\int_0^p F_i(T)^{-1}(\lambda)\mathrm{d}\lambda \le \int_0^p G_i(T)^{-1}(\lambda)\mathrm{d}\lambda \quad \text{for } p \in [0,1],$$

with equality achieved at $p = 1$. Last, observe that $F_i(T)^{-1}(\lambda)$ and $G_i(T)^{-1}(\lambda)$ are non-decreasing. Therefore, it follows from Lemma 3.3.3 that

$$\int_0^1 \phi(T)(F_1(T)^{-1}(\lambda),\dots,F_d(T)^{-1}(\lambda))\mathrm{d}\lambda \ge \int_0^1 \phi(T)(G_1(T)^{-1}(\lambda),\dots,G_d(T)^{-1}(\lambda))\mathrm{d}\lambda.$$

Hence, by Corollary 3.3.1, we conclude that $\hat{\phi}(F_1,\dots,F_d) \ge \hat{\phi}(G_1,\dots,G_d)$.

Assume now that ϕ is non-increasing. We prove that $\hat{\phi}(F_1,\dots,F_d) \ge \hat{\phi}(G_1,\dots,G_d)$ under the weaker condition $(F_1,\dots,F_d) \prec (G_1,\dots,G_d)$. Clearly, $F_i \prec G_i$ implies $F_i(T) \prec G_i(T)$. Then, it follows from Theorem 4.A.6 in [67] that $F_i \prec G_i$ if and only if there exists a distribution function $D_i \in \mathcal{F}$ such that $F_i(T)^{-1}(\lambda) \le D_i^{-1}(\lambda)$ for all $\lambda \in [0,1]$ and $D_i \preceq G_i(T)$. Therefore,

$$\begin{aligned}\int_0^1 \phi(T)\big(F_1(T)^{-1}(\lambda),\dots,F_d(T)^{-1}(\lambda)\big)\mathrm{d}\lambda &\ge \int_0^1 \phi(T)\big(D_1^{-1}(\lambda),\dots,D_d^{-1}(\lambda)\big)\mathrm{d}\lambda \\ &\ge \int_0^1 \phi(T)\big(G_1(T)^{-1}(\lambda),\dots,G_d(T)^{-1}(\lambda)\big)\mathrm{d}\lambda,\end{aligned}$$

where the first inequality holds because $\phi(T)$ is non-increasing and, for every $i \in \{1,\dots,d\}$ and $\lambda \in [0,1]$, $F_i(T)^{-1}(\lambda) \le D_i^{-1}(\lambda)$, and second inequality was established above. Thus, the result follows from Corollary 3.3.1. ■

Now, we derive the limiting composite relaxation. Our construction will be based on three key ideas: (a) for each (x,f), the underestimator will be used as in Proposition 3.3.2 to derive marginal distributions $S_1^{x,f},\dots,S_d^{x,f}$; (b) under the technical condition (3.20), Proposition 3.3.1 will be used to relax $\phi(f)$ by $\hat{\phi}(S_1^{x,f},\dots,S_d^{x,f})$; (c) $\hat{\phi}(S_1^{x,f},\dots,S_d^{x,f})$ will be further relaxed to $\hat{\phi}(F_1,\dots,F_d)$, where $(F_1,\dots,F_d) \preceq (S_1^{x,f},\dots,S_d^{x,f})$ using Proposition 3.3.3 without sacrificing the quality of the relaxation. We present these ideas formally in the next result, where we also prove the convexity of the resulting relaxation.

Theorem 3.3.3 *Let $\phi \circ f$ be a composite function, where $\phi : [0,1]^d \mapsto \mathbb{R}$ is a continuous function which is convex in each argument when other arguments are fixed, and $f : \mathbb{R}^m \mapsto [0,1]^d$ is a vector of functions over a subset X of $\mathbb{R}^m$. For each $i \in \{1,\dots,d\}$, let $s_i : W \times \mathbb{R} \mapsto \mathbb{R}$ be a function which satisfies (3.29) and is concave in a_i. Define $S_i^{x,f}(a_i) := 1 - (s_i)'_+(x,f,a_i)$. Then, $\mathrm{proj}_{(x,\phi)}(R)$ is a relaxation of hypograph of $\phi \circ f$, where:*

$$R := \left\{ (x,f,\phi,F_1,\dots,F_d) \,\middle|\, \begin{aligned} &\phi \leq \hat{\phi}(F_1,\dots,F_d),\ (F_1,\dots,F_d) \in \mathcal{F}^d \\ &(x,f) \in W,\ S_i^{x,f} \preceq F_i \end{aligned} \right\}, \tag{3.37}$$

Moreover,

1. *if, for each fixed a_i, $s_i(x,f,a_i)$ is convex in (x,f) and W is convex then R is convex;*

2. *if ϕ is a supermodular function then*

$$\mathrm{proj}_{(x,f,\phi)}(R) = \left\{ (x,f,\phi) \,\middle|\, \begin{aligned} &(x,f) \in W \\ &\phi \leq \int_0^1 \phi\Big(\big(S_1^{x,f}\big)^{-1}(\lambda),\dots,\big(S_d^{x,f}\big)^{-1}(\lambda)\Big)\mathrm{d}\lambda \end{aligned} \right\}; \tag{3.38}$$

3. *if ϕ is supermodular and non-increasing in each argument and, for fixed $a_i \in \mathbb{R}$, $s_i(x,f(x),a_i)$ is convex in x then we obtain a convex relaxation of the hypograph of $\phi \circ f$:*

$$\left\{ (x,\phi) \,\middle|\, x \in X,\ \phi \leq \int_0^1 \phi\Big(\big(S_1^{x,f(x)}\big)^{-1}(\lambda),\dots,\big(S_d^{x,f(x)}\big)^{-1}(\lambda)\Big)\mathrm{d}\lambda \right\}. \tag{3.39}$$

Proof We first prove that $\mathrm{hyp}(\phi \circ f) \subseteq \mathrm{proj}_{(x,\phi)}(R)$, where $\mathrm{hyp}(\phi \circ f)$ denotes the hypograph of $\phi \circ f$. Let $(x,\phi) \in \mathrm{hyp}(\phi \circ f)$. Define $f := f(x)$ and observe that $(x,f) \in W$ because W is a relaxation of $\mathrm{gr}(f)$. Moreover, let $F_i := S_i^{x,f}$. By Proposition 3.3.2, F_i is the distribution function of a random variable A_i such that $\mathbb{E}[A_i] = f_i$. By $(x,\phi) \in \mathrm{hyp}(\phi \circ f)$, $f = f(x)$, $F_i = S_i^{x,f}$, and Proposition 3.3.1, we obtain $\phi \leq \phi\big(f(x)\big) = \phi(f) = \tilde{\phi}(S_1^{x,f},\dots,S_d^{x,f}) = \tilde{\phi}(F_1,\dots,F_d) \leq \hat{\phi}(F_1,\dots,F_d)$.

Therefore, we conclude that $(x, f, \phi, F_1, \ldots, F_d) \in R$. In other words, $\text{hyp}(\phi \circ f) \subseteq \text{proj}_{(x,\phi)}(R)$.

We now show that under the three conditions in the theorem statement the claimed structure for the relaxation holds. We begin with Condition 1. Assume that, for each a_i, $s_i(x, f, a_i)$ is convex in (x, f) and W is convex. To show that R is convex, it suffices to argue that constraint $S_i^{x,f_i} \preceq F_i$ defines a convex set because W is assumed to be convex and, by Proposition 3.3.1, the functional $\hat{\phi}$ is concave over $\mathcal{F}^d$. By definition, $S_i^{x,f} \preceq F_i$ is equivalent to

$$E_{a_i}\big(S_i^{x,f}\big) \leq E_{a_i}(F_i) \quad \forall a_i \in \mathbb{R} \text{ and } \int a_i \mathrm{d}S_i^{x,f}(a_i) = \int a_i \mathrm{d}F_i(a_i). \tag{3.40}$$

For each $a_i \in \mathbb{R}$, by Proposition 3.3.2, we have $E_{a_i}\big(S_i^{x,f}\big) - E_{a_i}(F_i) = s_i(x, f, a_i) - E_{a_i}(F_i)$. Since we assumed that, for each a_i, $s_i(x, f, a_i)$ is a convex function, the result follows if $E_{a_i}(F_i)$ is a concave function for each a_i. The latter follows because, for $\alpha \in [0, 1]$ and $F_i, G_i \in \mathcal{F}$, $E_{a_i}\big(\alpha F + (1-\alpha)G\big) = \int \min\{a_i', a_i\}\mathrm{d}\big(\alpha F_i(a_i') + (1-\alpha)G_i(a_i')\big) = \alpha E_{a_i}(F) + (1-\alpha)E_{a_i}(G)$. The second equation in (3.40) defines a convex set because, by Proposition 3.3.2, $\int a_i \mathrm{d}S_i^{x,f}(a_i) = f_i$, and, for $F_i, G_i \in \mathcal{F}$ so that $\int a_i \mathrm{d}F_i = \int a_i \mathrm{d}G_i = f_i$, $\int a_i \mathrm{d}\big(\alpha F_i(a_i) + (1-\alpha)G_i(a_i)\big) = \alpha \int a_i \mathrm{d}F_i(a_i) + (1-\alpha)\int a_i \mathrm{d}G_i(a_i) = f_i$. It follows that the set described in (3.40) is convex in the space of $(x, f, F_1, \ldots, F_d)$ variables.

Next, we prove Condition 2. Let

$$R' := \big\{(x, f, \phi) \bigm| (x, f) \in W,\ \phi \leq \hat{\phi}(S_1^{x,f}, \ldots, S_1^{x,f})\big\},$$

which, by Theorem 3.3.2, is the set in the RHS of (3.38). To show $R' \subseteq \text{proj}_{(x,f,\phi)}(R)$, we consider a point $(x, f, \phi) \in R'$ and define $(F_1, \ldots, F_d) = (S_1^{x,f}, \ldots, S_d^{x,f})$. Then, by Proposition 3.3.2, we have $\int a_i \mathrm{d}F_i(a_i) = f_i$. Since $S_i^{x,f} \preceq F_i$ holds trivially, it follows that $(x, f, \phi, F_1, \ldots, F_d) \in R$, showing that $R' \subseteq \text{proj}_{(x,f,\phi)}(R)$. To prove that $\text{proj}_{(x,f,\phi)}(R) \subseteq R'$, we consider a point $(x, f, \phi, F_1, \ldots, F_d)$ of R and show that $(x, f, \phi) \in R'$. It follows readily that $\phi \leq \hat{\phi}(F_1, \ldots, F_d) \leq \hat{\phi}\big(S_1^{x,f}, \ldots, S_d^{x,f}\big)$, where second inequality holds because $\big(S_1^{x,f}, \ldots, S_d^{x,f}\big) \preceq (F_1, \ldots, F_d)$ and, by Proposition 3.3.3, the functional $\hat{\phi}$ is non-increasing under the order $\preceq$. Since $(x, f) \in W$, we conclude that $(x, f, \phi) \in R'$ and $\text{proj}_{(x,f,\phi)}(R) \subseteq R'$.

Last, we prove Condition 3. Assume that $s_i(x, f(x), a_i)$ is convex in x. We start by showing $\text{proj}_{(x,\phi)}(\tilde{R})$ is a convex relaxation of the hypograph of $\phi \circ f$, where

$$\tilde{R} := \left\{(x, \phi, F_1, \dots, F_d) \;\middle|\; \phi \leq \hat{\phi}(F_1, \dots, F_d),\ (F_1, \dots, F_d) \in \mathcal{F}^d,\ S_i^{x,f(x)} \prec F_i\right\}. \tag{3.41}$$

For this proof, we let $W := \text{gr}(f)$. Then, we observe that $\text{hyp}(\phi \circ f) \subseteq \text{proj}_{(x,\phi)}(R) \subseteq \text{proj}_{(x,\phi)}(\tilde{R})$, where the first containment was shown above and the second containment holds because, by Proposition 3.3.1, the constraint $S_i^{x,f(x)} \preceq F_i$ is equivalent to $S_i^{x,f(x)} \prec F_i$ and $f_i(x) = \displaystyle\int a_i \mathrm{d}F_i$. The convexity of $\tilde{R}$ follows because $\hat{\phi}$ is concave over $\mathcal{F}^d$, and the constraint $S_i^{x,f(x)} \prec F_i$ is convex because it can be imposed using $s_i\big(x, f(x), a_i\big) \leq E_{a_i}(F_i)$ for every $a_i \in \mathbb{R}$. This defines a convex set because, for every $a_i \in \mathbb{R}$, $s_i\big(x, f(x), a_i\big)$ is convex and $E_{a_i}(F_i)$ was shown to be linear in the proof of Condition 1. Now, let R'' be the set defined by (3.39). We will show $\text{proj}_{(x,\phi)}(\tilde{R}) = R''$. We have $R'' = \text{proj}_{(x,\phi)}(R) \subseteq \text{proj}_{(x,\phi)}(\tilde{R})$, where the first equality holds by Condition 2 and $W := \text{gr}(f)$, and the second equality holds because $R \subseteq \tilde{R}$. Now, to prove $\text{proj}_{(x,\phi)}(\tilde{R}) \subseteq R''$, we consider a point $(x, f, \phi, F_1, \dots, F_d)$ of $\tilde{R}$ and show $(x, \phi) \in R''$. It follows readily that $\phi \leq \hat{\phi}(F_1, \dots, F_d) \leq \hat{\phi}\big(S_1^{x,f(x)}, \dots, S_d^{x,f(x)}\big)$, where the second inequality holds by $\big(S_1^{x,f(x)}, \dots, S_d^{x,f(x)}\big) \prec (F_1, \dots, F_d)$ and because Proposition 3.3.3 shows that $\hat{\phi}$ is non-increasing in $\prec$ under the assumed properties of ϕ. Thus, by Theorem 3.3.2, $(x, \phi) \in R''$ and, therefore, $\text{proj}_{(x,\phi)}(\tilde{R}) \subseteq R''$. ■

Remark 3.3.2 *We remark that $\hat{\phi}\big(S_1^{x,f}, \dots, S_d^{x,f}\big)$ coincides with the composite function $\phi \circ f$ when the underestimating function $s_i(x, f, a_i)$ equals* $\min\{f_i(x), a_i\}$. *To see this, consider an $x \in X$ and let $f = f(x)$. It follows that, for $i = 1, \dots, d$, $s_i(x, f, a_i) = a_i$ if $a_i \leq f_i$ and $s_i(x, f, a_i) = f_i$ otherwise. Therefore, $S_i^{x,f}(a_i) = 0$ if $a_i \leq f_i$ and* 1 *otherwise. In other words, $S_i^{x,f}$ corresponds to the distribution function of a Dirac measure with all its mass at f_i. Therefore, the only joint distribution,* *feasible in the optimal transport formulation (3.23), is the distribution of a Dirac measure with all its mass at* $(f_1, \ldots, f_d)$. *In other words,* $\hat{\phi}(S_1^{x,f}, \ldots, S_d^{x,f}) = \phi(f)$.

Observe that since the locus of points over which the integral (3.38) or (3.39) is taken is independent of the function ϕ, (3.38) or (3.39) can be used to simultaneously treat a vector of functions θ_k, $k \in \{1, \ldots, \kappa\}$. Using arguments similar to those in Conditions 2 and 3 of Theorem 3.3.3, we can extend the result to treat functions that become supermodular after switching. We record this result for its use in applications such as Example 3.3.2.

Corollary 3.3.2 *Let* R *be the set defined in (3.37). Assume the same setup as Theorem 3.3.3 except that the assumed properties on* ϕ *apply to* $\phi(T)$, *where* T *is some subset of* $\{1, \ldots, d\}$. *Then, Condition 2 applies with the definition of* $\operatorname{proj}_{(x,f,\phi)}(R)$ *replaced with:*

$$\operatorname{proj}_{(x,f,\phi)}(R) = \left\{ (x, f, \phi) \,\middle|\, \begin{array}{l} (x, f) \in W \\ \phi \leq \displaystyle\int_0^1 \phi(T)\Big(\big(S_1^{x,f}\big)(T)^{-1}(\lambda), \ldots, \big(S_d^{x,f}\big)(T)^{-1}(\lambda)\Big) \mathrm{d}\lambda \end{array} \right\}.$$

Similarly, Condition 3 applies, where the relaxation is replaced with:

$$\left\{ (x, \phi) \,\middle|\, x \in X,\ \phi \leq \int_0^1 \phi(T)\Big(\big(S_1^{x,f(x)}\big)(T)^{-1}(\lambda), \ldots, \big(S_d^{x,f(x)}\big)(T)^{-1}(\lambda)\Big) \mathrm{d}\lambda \right\}.$$

3.4 Conclusions

In this chapter, we developed new tractable relaxations for composite functions. Our relaxations leverage the composite relaxation framework proposed in Chapter 2 that involves convexifying the outer-function over a polytope P. The polytope P encodes the structure of inner-functions using n estimators for each function. The structure of P generalizes that of a hypercube; the set used in factorable relaxations for a similar purpose and derived using bounds on the inner-function. Although convexifying general outer-functions over P is NP-Hard, we showed that when the outer-function is supermodular and concave-extendable, its concave envelope over P

is determined by the staircase triangulation of a subset Q of P. Using this result, we found exponentially many inequalities describing the concave envelope of the outer-function over P. Since the polyhedral subdivision of P is invariant with the outer-function, we could convexify simultaneously the hypograph of a vector of composite functions. We also derived various inequalities regarding the structure of inequalities for the special case where the outer-function is multilinear.

We extended our results to the case with infinitely many estimators for each inner-function, by assuming that the outer-function is convex in each argument. For this extension, we described a marginal distribution for each inner-function by considering how underestimating function varies as a function of its upper bound. We then reformulated the concave envelope construction to an optimal transport problem and showed that the problem has an explicit solution when the outer-function is supermodular. Moreover, when the outer-function is non-decreasing, we exploited monotonicity properties of the explicit solution for the optimal transport problem with respect to a certain stochastic order to show that, as long as the underestimating functions were convex, we can derive a convex relaxation for the composite function in the space of the original problem variables.

4. EXTRACTING STRUCTURE FROM EXPRESSION TREES FOR DISCRETE AND CONTINUOUS RELAXATIONS OF MINLPS

In Chapters 2 and 3, we extracted a polytope P and its subset Q from expression trees to construct composite relaxations. In this chapter, we show that composite relaxations also lay a foundation for improving relaxation hierarchies and standard MIP relaxations.

First, we present a family of inequalities, called *staircase inequalities*, which is obtained by telescoping composite function $\phi \circ f$ in a certain way so that, under certain conditions, each summand can be overestimated using a simple operation. Staircase inequalities can be used to derive the Lovász extension of set function [53] and a tractable relaxation for composite functions in Chapter 3. These staircase inequalities are not implied by standard reformulation-linearization technique (RLT) [39]. Furthermore, recursively using staircase inequalities and linearization yield RLT inequalities and more.

We also argue that the composite relaxations are well-suited for constructing MIP relaxations via a discretization scheme. Our proposed MIP relaxations are obtained by reinterpreting the incremental formulation [35] and combining it with the composite relaxations, and are provable to be tighter than the standard ones in [36, 37]. Further improved MIP relaxations can be constructed using staircase inequalities. For composite functions with discrete domains and univariate inner-functions, we obtain ideal logarithmic MIP formulations for their graphs. In contrast, the state-of-the-art formulation from [40] is not ideal. By exploiting our results for discrete domains, we show that, for certain compositions of univariate functions, we can construct a sequence of polyhedral relaxations that converge, in the limit, to the concave envelope.

4.1 Staircase expansions and stepwise relaxations

Consider a composite function $\phi \circ f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$ defined as $\phi \circ f(x) = \phi\bigl(f(x)\bigr)$. We shall write $f(x) := \bigl(f_1(x), \ldots, f_d(x)\bigr)$ and refer to $f(\cdot)$ as *inner-functions* while $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ will be referred to as the *outer-function*. In this section, we will present a procedure to derive overestimators for composite function $\phi \circ f$. More specifically, we telescope composite function $\phi \circ f$ in a certain way so that, under certain conditions, each summand can be overestimated using a simple operation. As an illustration, we apply this technique to derive the Lováze extension of set function [53] and a tractable relaxation for composite functions in Chapter 3. We will assume that for every $x \in X$ we have $f(x) \in [f^L, f^U]$ for some vectors f^L and f^U in $\mathbb{R}^d$. We further assume, without loss of generality, that $[f^L, f^U] = [0,1]^d$ since otherwise we can treat $\bar{f}_i(x) = \bigl(f_i(x) - f_i^L\bigr)/(f_i^U - f_i^L)$ as an inner-function and $\bar{\phi}(\bar{f}) = \phi\bigl((f_1^U - f_1^L)\bar{f}_1 + f_1^L, \ldots, (f_d^U - f_d^L)\bar{f}_d + f_d^L\bigr)$ as an outer-function.

Let $(n_1, \ldots, n_d) \in \mathbb{Z}^d$, and let $u : \mathbb{R}^m \mapsto \mathbb{R}^{\sum_{i=1}^d (n_i+1)}$ be a vector of functions so that

$$u(x) = \bigl(u_1(x), \ldots, u_d(x)\bigr) \qquad \text{and} \qquad \bigl(u_{1n_1}(x), \ldots u_{dn_d}(x)\bigr) = \bigl(f_1(x), \ldots, f_d(x)\bigr).$$

Throughout this chapter, we introduce a vector of variables u to represent $u(x)$, and we assume without loss of generality and for notational simplicity that $n(1) = \cdots = n(d) = n$. For $(j_1, \ldots, j_d) \in \{0, \ldots, n\}^d$, a subvector $(u_{1j_1}, \ldots, u_{dj_d})$ of u can be conveniently represented on a grid $\mathcal{G}$. More precisely, we define the grid $\mathcal{G}$ on a d-dimensional space where each coordinate takes one of the values in $\{0, \ldots, n\}$. Then, a subvector $(u_{ij_i})_{i=1}^d$ can be depicted as the point $\{j_i\}_{i=1}^d$ on the grid. In other words, if we label the grid marker j along the coordinate direction i as u_{ij}, we can read off the labels to find the corresponding subvector of u. We shall call these labels *grid-labels*, which depend on u, while the grid depends only on the number of subspaces and the dimension of each subspace. Given a grid point $p_i = (j_i)_{i=1}^d$, we denote the corresponding subvector $(u_{1j_1}, \ldots, u_{dj_d})$ of u by (u, p_k). Essentially, (u, p_k) reads off the grid labels, derived from u, for the grid-point p_k to obtain the corresponding subvector.

A *lattice path* in $\mathcal{G}$ from $(0,\ldots,0)$ to $(n,\ldots,n)$ is a sequence of points $p_0,\ldots,p_\tau$ in $\mathbb{Z}^d$ such that $p_0=(0,\ldots,0)$ and $p_k=(n,\ldots,n)$. In particular, a *monotone staircase* is a lattice path in $\mathbb{Z}^d$ of length $dn+1$ such that, for all $t\in\{1,\ldots,dn\}$, $p_k-p_{k-1}=e_i$, where e_i is the i^{th} principal vector in $\mathbb{R}^d$. We refer to the movement p_{k-1} to p_k as the k^{th} move. Clearly, there are exactly n moves in each coordinate direction. Thus, the staircase can be specified succinctly as a *movement vector* $\pi=(\pi_1,\ldots,\pi_{dn})$ where $\pi_k\in\{1,\ldots,d\}$ is the coordinate direction of k^{th} move. We will often need to track where the k^{th} move leaves us on the grid. This is obtained using the transformation Π, which is defined as $\Pi(\pi,k):=p_0+\sum_{j=1}^k e_{\pi(j)}$. Given a move π, the corresponding staircase can be recovered as $\big(\Pi(\pi,k)\big)_{k=0}^{dn}$, and thus the corresponding subvectors of u are denoted by $\big\{\big(u,\Pi(\pi,k)\big)\big\}_{k=0}^{dn}$. The *staircase expansion* of a function $\phi:\mathbb{R}^d\mapsto\mathbb{R}$ under a movement π is a function $\mathcal{D}^\pi(\phi):\mathbb{R}^{d\times(n+1)}\mapsto\mathbb{R}$ defined as follows:

$$\mathcal{D}^\pi(\phi)(u):=\phi\big(u,\Pi(\pi,0)\big)+\sum_{k=1}^{dn}\Big(\phi\big(u,\Pi(\pi,k)\big)-\phi\big(u,\Pi(\pi,k-1)\big)\Big). \tag{4.1}$$

Lemma 4.1.1 *Consider a composite function $\phi\circ f: X\subseteq\mathbb{R}^m\mapsto\mathbb{R}^d$, and let $u:\mathbb{R}^m\mapsto\mathbb{R}^{d\times(n+1)}$ so that, for $i\in\{1,\ldots,d\}$, $u_{in}(x)=f_i(x)$. Then, for any staircase π we have*

$$\phi\big(f(x)\big)=\mathcal{D}^\pi(\phi)\big(u(x)\big)\le\phi\big(u(x),\Pi(\pi,0)\big)+\sum_{k=1}^{dn}\tilde{\mathcal{D}}^{(\pi,k)}(\phi)\big(u(x)\big)\quad \textit{for } x\in X,$$

provided that $\mathcal{D}^{(\pi,k)}(\phi)\big(u(x)\big)\le\tilde{\mathcal{D}}^{(\pi,k)}(\phi)\big(u(x)\big)$ is valid over X.

Proof Let $x\in X$, and let $u:=u(x)$. Then, the equality in the result holds by observing that

$$\phi\big(f(x)\big)=\phi(u_{1n},\ldots,u_{dn})=\phi\big(u,\Pi(\pi,dn)\big)=\mathcal{D}^\pi(\phi)(u),$$

where the first equality holds by the hypothesis that for $i\in\{1,\ldots,d\}$ $u_{in}(x)=f_i(x)$, the second equality holds because $\Pi(\pi,dn)=(n,\ldots,n)$, and the last equality holds because the sum in (4.1) telescopes, leaving the term $\phi\big(u,\Pi(\pi,dn)\big)$. ∎

In the next example we use the idea in Lemma 4.1.1 to derive a relaxation for the product of two functions.

Example 4.1.1 (Theorem 2.3.3) *Consider the product of two functions $f_1(x)f_2(x)$ over a convex set $X \subseteq \mathbb{R}^m$, where for $x \in X$ we assume that $f_i^L \leq f_i(x) \leq f_i^U$. For $i = 1, 2$, assume that there exist constants f_i^L, f_i^U and a_i, and a function $u_i : X \mapsto \mathbb{R}$ so that for $x \in X$ we have*

$$f_i^L \leq a_i \leq f_i^U \quad f_i^L \leq f_i(x) \leq f_i^U \quad u_i(x) \leq f_i(x) \quad \textit{and} \quad u_i(x) \leq a_i. \tag{4.2}$$

We only derive one of inequalities in Theorem 2.3.3 since others can be derived in a similar way. Let $u_i^+(x)$ denote $\max\{u_i(x), f_i^L\}$. Then, it follows that

$$\begin{aligned}
f_1(x)f_1(x) &= f_1^L f_2^L + \big(u_1^+(x) - f_1^L\big)f_2^L + u_1^+(x)\big(u_2^+(x) - f_2^L\big) + \big(f_1(x) - u_1^+(x)\big)u_2^+(x) \\
&\quad + f_1(x)\big(f_2(x) - u_2^+(x)\big) \\
&\leq f_1^L f_2^L + \big(u_1^+(x) - f_1^L\big)f_2^L + a_1\big(u_2^+(x) - f_2^L\big) + \big(f_1(x) - u_1^+(x)\big)a_2 \\
&\quad + f_1^U\big(f_2(x) - u_2^+(x)\big) \\
&= (f_2^L - a_2)u_1^+(x) + (a_1 - f_2^U)u_2^+(x) + a_2 f_1(x) + f_1^U f_2(x) - a_1 f_2^L \\
&\leq (f_2^L - a_2)u_1(x) + (a_1 - f_2^U)u_2(x) + a_2 f_1(x) + f_1^U f_2(x) - a_1 f_2^L
\end{aligned}$$

where the first equality holds because the right hand side telescopes to $f_1(x)f_2(x)$, the first inequality holds because inequalities in (4.2) imply that

$$\begin{aligned}
\big(a_1 - u_1^+(x)\big)\big(u_2^+(x) - f_2^L\big) &\geq 0, \\
\big(f_1(x) - u_1^+(x)\big)\big(a_2 - u_2^+(x)\big) &\geq 0, \\
\big(f_1^U - f_1(x)\big)\big(f_2(x) - u_2^+(x)\big) &\geq 0,
\end{aligned}$$

the second equality holds by rearrangement, and the last inequality is implied by $u_i^+(x) \geq u_i(x)$ and inequalities in (4.2).

Motivated by the Example 4.1.1, we now consider a particular system of estimating functions for the inner-function $f(\cdot)$. Namely, let $\big(u(x), a(x)\big)$ be a pair of vectors of functions mapping from $\mathbb{R}^{d\times(n+1)}$ to $\mathbb{R}$ so that for $x \in X$

$$\begin{aligned} &0 \le a_{i0}(x) \le \ldots \le a_{in}(x) \le 1, \\ &u_{ij}(x) \le \min\big\{f_i(x), a_{ij}(x)\big\} \qquad \text{for all } j \in \{0,\ldots,n\}, \\ &u_{i0}(x) = a_{i0}(x) \quad u_{in}(x) = f_i(x). \end{aligned} \tag{4.3}$$

The first requirement states that $u_{ij}(\cdot)$ is an underestimator for $f_i(\cdot)$, which is bounded from above by bounding function $a_{ij}(\cdot)$. The second requirement requires that elements of $a_i(\cdot)$ are ordered in a non-decreasing order. The last requirement says that the first underestimator $u_{i0}(\cdot)$ and last underestimator $u_{in}(\cdot)$ is the smallest bounding function $a_{i0}(\cdot)$ and the inner function $f_i(\cdot)$ itself. In Section 4.4, we will provide tractable procedures for constructing a pair satisfying (4.3).

Now, we exploit ordering relationships in (4.3) to relax each step different function $\mathcal{D}^{(\pi,k)}\big(u(x)\big)$. For a subset I of $\{1,\ldots,d\}$, we define a replacement operator χ^I : $(u,a) \in \mathbb{R}^{d\times(n+1)} \times \mathbb{R}^{d\times(n+1)} \mapsto t \in \mathbb{R}^{d\times(n+1)}$ so that $t_i = u_i$ if $i \in I$ and $t_i = a_i$ otherwise. For a function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ and for a staircase π in the grid given by $\{0,\ldots,n\}^d$, let $\mathcal{B}^\pi(\phi) : \mathbb{R}^{n\times(n+1)} \times \mathbb{R}^{n\times(n+1)} \mapsto \mathbb{R}$ be a function defined as follows:

$$\mathcal{B}^\pi(\phi)(u,a) = \phi\big(u, \Pi(\pi,0)\big) + \sum_{k=1}^{dn} \mathcal{B}^{(\pi,k)}(\phi)(u,a), \tag{4.4}$$

where $\mathcal{B}^{(\pi,k)}(\phi)(u,a) := \mathcal{D}^{(\pi,k)}(\phi)\big(\chi^{\pi(k)}(u,a)\big)$ is a function obtained from $\mathcal{D}^{(\pi,k)}(\phi)(u)$ by replacing all arguments u_i by a_i but the π_k^{th} coordinate direction. Next, we show that the supermodularity of $\phi(\cdot)$ implies that $(\phi \circ f)(x) \le \mathcal{B}^\pi(\phi)\big(u(x), a(x)\big)$ for every $x \in X$ and that $\mathcal{B}^\pi(\phi)(\cdot,\cdot)$ processes certain monotone properties.

Definition 4.1.1 ([54]) *A function $\eta(x) : S \subseteq \mathbb{R}^n \mapsto \mathbb{R}$ is said to be supermodular if $\eta(x' \vee x'') + \eta(x' \wedge x'') \ge \eta(x') + \eta(x'')$ for all x', $x'' \in S$. Here, $x' \vee x''$ denotes the component-wise maximum and $x' \wedge x''$ denotes the component-wise minimum of x' and x'' and assume that $x' \vee x''$ and $x' \wedge x''$ belong to S whenever x' and x'' in S.*

Although detecting whether a function is supermodular is NP-Hard [55], there are important special cases where this property can be readily detected [54]. For example, a product of nonnegative, increasing (decreasing) supermodular functions is also nonnegative increasing (decreasing) and supermodular; see Corollary 2.6.3 in [54]. Also, it follows trivially that a conic combination of supermodular functions is supermodular.

Theorem 4.1.1 *Consider a composite function $\phi \circ f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$, and consider three pairs $\big(u(x), a(x)\big)$, $\big(u'(x), a(x)\big)$ and $\big(u(x), a'(x)\big)$ which satisfy the requirement (4.3) such that $u(x) \leq u'(x)$ and $a(x) \leq a'(x)$. If the outer-function $\phi(\cdot)$ is supermodular over $[0,1]^d$ then, for each staircase π, we have $\phi\big(f(x)\big) \leq \mathcal{B}^\pi(\phi)\big(u'(x), a(x)\big) \leq \mathcal{B}^\pi(\phi)\big(u(x), a(x)\big)$ for every $x \in X$.*

Proof Let $\big(u(x), a(x)\big)$ and $\big(u'(x), a(x)\big)$ be two pairs satisfying (4.3) such that $u(x) \leq u'(x)$. Then, we start with showing that $\mathcal{B}^\pi(\phi)\big(u'(x), a(x)\big) \leq \mathcal{B}^\pi(\phi)\big(u(x), a(x)\big)$. Let $x \in X$, and define $u := u(x)$, $u' := u'(x)$, and $a := a(x)$. For $i \in \{1, \ldots, d\}$, define $\{t_{i1}, \ldots, t_{in}\} := \big\{t \bigm| \pi(t) = i\big\}$ so that $t_{i1} \leq \ldots \leq t_{in}$. It follows that

$$
\begin{aligned}
&\mathcal{B}^\pi(\phi)(u', a) \\
&= \phi\big(u', \Pi(\pi, 0)\big) + \sum_{i=1}^{d} \sum_{j=1}^{n} \Big(\phi\big(\chi^i(u', a), \Pi(\pi, t_{ij})\big) - \phi\big(\chi^i(u', a), \Pi(\pi, t_{ij} - 1)\big)\Big) \\
&= \phi\big(u', \Pi(\pi, 0)\big) + \sum_{i=1}^{d} \Big(-\phi\big(\chi^i(u', a), \Pi(\pi, t_{i1} - 1)\big) + \\
&\quad \sum_{j=1}^{n-1} \Big(\phi\big(\chi^i(u', a), \Pi(\pi, t_{ij})\big) - \phi\big(\chi^i(u', a), \Pi(\pi, t_{i(j+1)} - 1)\big)\Big) + \phi\big(\chi^i(u', a), \Pi(\pi, t_{in})\big)\Big) \\
&\leq \phi\big(u, \Pi(\pi, 0)\big) + \sum_{i=1}^{d} \Big(-\phi\big(\chi^i(u, a), \Pi(\pi, t_{i1} - 1)\big) + \\
&\quad \sum_{j=1}^{n-1} \Big(\phi\big(\chi^i(u, a), \Pi(\pi, t_{ij})\big) - \phi\big(\chi^i(u, a), \Pi(\pi, t_{i(j+1)} - 1)\big)\Big) + \phi\big(\chi^i(u, a), \Pi(\pi, t_{in})\big)\Big) \\
&= \mathcal{B}^\pi(\phi)(u, a).
\end{aligned}
\tag{4.5}
$$

To see that the inequality holds when $\phi(\cdot)$ is supermodular, we observe that $\big(\chi^i(u', a), \Pi(\pi, t_{i1}-1)\big) = \big(\chi^i(u, a), \Pi(\pi, t_{i1}-1)\big)$ since $\big(\Pi(\pi, t_{i1}-1)\big)_i = 0$ and $u_{i0} = u'_{i0}$, that $\big(\chi^i(u', a), \Pi(\pi, t_{in})\big) = \big(\chi^i(u, a), \Pi(\pi, t_{in})\big)$ since $(\Pi(\pi, t_{in}))_i = n$ and $u_{in} = u'_{in}$, and that

$$
\begin{aligned}
&\big(\chi^i(u', a), \Pi(\pi, t_{ij})\big) \wedge \big(\chi^i(u, a), \Pi(\pi, t_{i(j+1)} - 1)\big) = \big(\chi^i(u, a), \Pi(\pi, t_{ij})\big),\\
&\big(\chi^i(u', a), \Pi(\pi, t_{ij})\big) \vee \big(\chi^i(u, a), \Pi(\pi, t_{i(j+1)} - 1)\big) = \big(\chi^i(u', a), \Pi(\pi, t_{i(j+1)} - 1)\big),
\end{aligned}
$$

where two equalities hold because $u \leq u'$, $a_{i0} \leq \ldots \leq a_{in}$, and $\Pi(\pi, t_{ij}) \leq \Pi(\pi, t_{i(j+1)}-1)$.

Next, consider a vector of underestimators $s(x)$ of $f(x)$ so that for each $i \in \{1, \ldots, d\}$ and $j \in \{0, \ldots, n\}$, $s_{ij}(x) = \min\big\{f_i(x), a_{ij}(x)\big\}$. We will show that $(\phi \circ f)(x) \leq \mathcal{B}^\pi(\phi)\big(s(x), a(x)\big)$, thus, together with the above result, yielding $(\phi \circ f)(x) \leq \mathcal{B}^\pi(\phi)\big(u(x), a(x)\big)$ since the pair $\big(s(x), a(x)\big)$ also satisfies (4.3) and $u(x) \leq s(x)$. Let $x \in X$, and define $a = a(x)$ and $s = s(x)$. By Definition 4.1, we have $(\phi \circ f)(x) = \mathcal{D}^\pi(\phi)(s)$, and need to show $\mathcal{D}^\pi(\phi)(s) \leq \mathcal{B}^\pi(\phi)(s, a)$. By the Definition 4.4, it suffices to show that, for $k \in \{1, \ldots, dn\}$, $\mathcal{D}^{(\pi,k)}(\phi)(s) \leq \mathcal{B}^{(\pi,k)}(\phi)(s, a)$, which can be established by observing that

$$
\begin{aligned}
\mathcal{D}^{(\pi,k)}(\phi)(s) &:= \phi\big(s, \Pi(\pi, k)\big) - \phi\big(s, \Pi(\pi, k-1)\big)\\
&\leq \phi\big(\chi^{\pi(k)}(s, a), \Pi(\pi, k)\big) - \phi\big(\chi^{\pi(k)}(s, a), \Pi(\pi, k-1)\big) =: \mathcal{B}^{\pi,k}(\phi)(s, a),
\end{aligned}
$$

where the inequality holds because $\phi(\cdot)$ is assumed to be supermodular, and $s_{i0} \leq \cdots \leq s_{in}$, $s_{ij} \leq a_{ij}$, and $\Pi(\pi, k-1) \leq \Pi(\pi, k)$ imply that

$$
\begin{aligned}
&\big(s, \Pi(\pi, k)\big) \wedge \big(\chi^{\pi(k)}(s, a), \Pi(\pi, k-1)\big) = \big(s, \Pi(\pi, k-1)\big),\\
&\big(s, \Pi(\pi, k)\big) \vee \big(\chi^{\pi(k)}(s, a), \Pi(\pi, k-1)\big) = \big(\chi^{\pi(k)}(s, a), \Pi(\pi, k)\big).
\end{aligned}
$$

The proof is complete. ■

Remark 4.1.1 *In this remark we treat the case when overestimators for inner-functions $f(\cdot)$ are provided. More specifically, we show that overestimators of $f(\cdot)$ can be transformed into underestimators and thus our treatment in Theorem 4.1.1 is without loss of generality. For $i \in \{1, \ldots, d\}$, let (A_i, B_i) be a partition of $\{0, \ldots, n\}$ so that*

$\{0,n\} \subseteq A_i$, and let $\big(u(x),a(x)\big)$ be a pair of vectors of functions from $\mathbb{R}^{d\times(n+1)}$ to R so that for $x \in X$:

$$\begin{aligned}
&0 \leq a_{i0}(x) \leq \ldots \leq a_{in}(x) \leq 1,\\
&\text{for each } j \in A_i: \; u_{ij}(x) \leq \min\big\{f_i(x), a_{ij}(x)\big\},\\
&\text{for each } j \in B_i: \; \max\big\{f_i(x), a_{ij}(x)\big\} \leq u_{ij}(x),\\
&u_{i0}(x) = a_{i0}(x) \quad u_{in}(x) = f_i(x).
\end{aligned} \tag{4.6}$$

Consider an affine transformation $L : (u,a) \in \mathbb{R}^{d\times(n+1)+d\times(n+1)} \mapsto (\tilde{u},\tilde{a}) \in \mathbb{R}^{d\times(n+1)+d\times(n+1)}$ defined as follows:

$$\tilde{u}_{ij} = u_{ij} \quad \tilde{a}_{ij} = a_{ij} \quad \text{for } j \in A_i \quad \text{and} \quad \tilde{u}_{ij} = a_{ij} - u_{ij} + u_{in} \quad \tilde{a}_{ij} = a_{ij} \quad \text{for } j \in B_i. \tag{4.7}$$

Clearly, the transformation L is invertible. More specifically, given a vector $(t,b) \in \mathbb{R}^{d\times(n+1)+d\times(n+1)}$ the transformation $(u,a) := L^{-1}(t,b)$ is given by $a = b$, and $u_{ij} = t_{ij}$ for all i and $j \in A_i$ and $u_{ij} = b_{ij} - t_{ij} + t_{in}$ for all i and $j \in B_i$. Observe that for each pair $\big(u(x),a(x)\big)$ satisfying (4.6) the transformed pair $T\big(u(x),a(x)\big)$ satisfies (4.3). Thus, by Theorem 4.1.1 we obtain that

$$(\phi \circ f)(x) \leq \mathcal{B}^{\pi}(\phi)\Big(T\big(u(x),a(x)\big)\Big) \quad \text{for } x \in X.$$

provided that $\phi(\cdot)$ is supermodular over $[0,1]^d$.

To derive overestimators for a composite function $\phi \circ f$, the Theorem 4.1.1 requires the outer-function $\phi(\cdot)$ to be supermodular over $[0,1]^d$, which contains the range of inner-functions $f(\cdot)$. Next, we consider a particular linear transformation that can be used to make the outer-function supermodular. This transformation, referred to as switching, chooses a set $D \subseteq \{1,\ldots,d\}$ and considers a new function $\phi(D) : [0,1]^d \mapsto \mathbb{R}$ defined as $\phi(D)(f_1,\ldots,f_d) = \phi\big(f(D)\big)$, where $f(D)_i = 1 - f_i$ if $i \in D$ and $f(D)_i = f_i$ otherwise. Moreover, for a given pair $\big(u(x),a(x)\big)$ satisfying (4.3), we define a new pair $\big(\tilde{u}(x), \tilde{a}(x)\big)$ so that for $i \notin T$ $\tilde{u}_i(x) = u_i(x)$ and $\tilde{a}_i(x) = a_i(x)$, and otherwise

$$\tilde{u}_{ij}(x) = \big(1 - a_{in-j}(x)\big) - \big(1 - u_{in-j}(x)\big) + \big(1 - u_{in}(x)\big) \text{ and } \tilde{a}_{ij}(x) = 1 - a_{in-j}(x) \quad \forall j \tag{4.8}$$

It turns out that $\phi \circ f$ can be expressed as $\phi(T)\big(\bar{u}_{1n}(x), \ldots, \bar{u}_{dn}(x)\big)$. Moreover, the new pair $\big(\tilde{u}(x), \tilde{a}(x)\big)$ satisfies the requirement in (4.3) as the pair $\big(u(x), a(x)\big)$ is assumed to satisfy (4.3). Therefore, the next result follows directly from Theorem 4.1.1.

Corollary 4.1.1 *Consider a composite function $\phi \circ f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$, and consider a pair $\big(u(x), a(x)\big)$ satisfying (4.3). If $f(T)$ is supermodular for some $T \subseteq \{0, \ldots, d\}$ then, for each staircase π, we have*

$$(\phi \circ f)(x) \leq \mathcal{B}^{\pi}\big(\phi(T)\big)\big(\tilde{u}(x), \tilde{a}(x)\big) \quad \textit{for every } x \in X,$$

where the pair $\big(\tilde{u}(x), \tilde{a}(x)\big)$ is defined as in (4.8).

Example 4.1.2 (Theorem 2.1.1) *Consider the same setup as Example 4.1.1. We will derive one of inequalities in [56] since others can be derived using similar argument. Let $\tilde{f}_2(x) := f_2^U - f_2(x)$, $\tilde{f}_2^L := 0$, and $\tilde{f}_2^U := f_2^U - f_2^L$. In addition, let $\tilde{u}_2(x) := (f_2^U - a_2) - \big(f_2^U - u_2(x)\big) + \big(f_2^U - f_2(x)\big)$ and $\tilde{a}_2 := f_2^U - a_2$. Let $\tilde{u}_2^+(x)$ denote* $\max\big\{\tilde{u}_2(x), \tilde{f}_2^L\big\}$. *Then, it follows that*

$$\begin{aligned}
f_1(x)f_1(x) &= f_1(x)\big(f_2^U - \tilde{f}_2(x)\big) \\
&= f_1^L\big(f_2^U - \tilde{f}_2^L\big) + \big(u_1^+(x) - f_1^L\big)\big(f_2^U - \tilde{f}_2^L\big) + u_1^+(x)\Big(\big(f_2^U - \tilde{u}_2^+(x)\big) - \big(f_2^U - \tilde{f}_2^L\big)\Big) \\
&\quad + \big(f_1(x) - u_1^+(x)\big)\big(f_2^U - \tilde{u}_2^+(x)\big) + f_1(x)\Big(\big(f_2^U - \tilde{f}_2(x)\big) - \big(f_2^U - \tilde{u}_2^+(x)\big)\Big) \\
&\geq f_1^L\big(f_2^U - \tilde{f}_2^L\big) + \big(u_1^+(x) - f_1^L\big)\big(f_2^U - \tilde{f}_2^L\big) + a_1\Big(\big(f_2^U - \tilde{u}_2^+(x)\big) - \big(f_2^U - \tilde{f}_2^L\big)\Big) \\
&\quad + \big(f_1(x) - u_1^+(x)\big)\big(f_2^U - \tilde{a}_2\big) + f_1^U\Big(\big(f_2^U - \tilde{f}_2(x)\big) - \big(f_2^U - \tilde{u}_2^+(x)\big)\Big) \\
&= (f_2^U - a_2)u_1^+(x) + (f_1^U - a_1)u_2^+(x) + a_2 f_1(x) + a_1 f_2(x) + a_1 a_2 - a_1 f_2^U - f_1^U a_2 \\
&\geq (f_2^U - a_2)u_1(x) + (f_1^U - a_1)u_2(x) + a_2 f_1(x) + a_1 f_2(x) + a_1 a_2 - a_1 f_2^U - f_1^U a_2,
\end{aligned}$$

where the first inequality holds because the following inequalities are valid for $[f_1^L, f_1^U]\times [f_2^L, f_2^U]$

$$u_1^+(x) \leq a_1 \quad \tilde{f}^L \leq \tilde{u}_2^+(x) \quad u_1^+(x) \leq f_1(x) \quad \tilde{u}_2^+(x) \leq \tilde{a}_2 \quad f_1(x) \leq f_1^U \quad \tilde{u}_2(x) \leq \tilde{f}_2(x),$$

and the bilinear function $b_1(f_2^U - b_2)$ *is submodular over* $[f_1^L, f_1^U] \times [0, f_2^U - f_2^L]$.

4.1.1 Lovász extension of set function

Consider a function $f(\cdot)$ mapping from $\{0,1\}^d$ to $\mathbb{R}$. Then, $f(\cdot)$ can be viewed as a set function in the following way. We define $f'(\cdot)$ as $f'(K) = f\big(\sum_{j\in K} e_j\big)$ for a subset $K \subseteq \{1,\ldots,d\}$, where e_j is the j^{th} principal vector in $\mathbb{R}^d$. In [53], the author introduces an extension of the function f, which originally is defined on $0,1$ vectors, to all non-negative vectors, which is well-known as *Lovász extension* of set functions. Given any $x \in [0,1]^d$, there exists a permutation π of $\{1,\ldots,n\}$ such that $x_{\pi(1)} \geq \cdots \geq x_{\pi(d)}$. Then, x is expressible as follows $x = (1-x_{\pi(1)})v_0^\pi + \sum_{i=1}^{d-1}(x_{\pi(i)} - x_{\pi(i+1)})v_i^\pi + x_{\pi(n)}v_n^\pi$, where $v_0^\pi = (0,\ldots,0)$ and $v_i^\pi = v_{i-1}^\pi + e_{\pi(i)}$ for $i = 1,\ldots,d$. The Lovász extension of $f(x)$ is defined as

$$\hat{f}(x) = (1-x_{\pi(1)})f(v_0^\pi) + \sum_{i=1}^{d-1}(x_{\pi(i)} - x_{\pi(i+1)})f(v_i^\pi) + x_{\pi(n)}f(v_n^\pi).$$

In the following, we show that relaxation techniques in Theorem 4.1.1 yields an alternative derivation of Lovász extension of set function [53].

Our derivation can be divided into three steps. First, given $\pi \in \Omega$, the set of all permutations of $\{1,\ldots,d\}$, we express the function $f(\cdot)$ as a function $\mathcal{D}^\pi(f) : \{0,1\}^d \mapsto \mathbb{R}$ defined as follows:

$$\begin{aligned}\mathcal{D}^\pi(f)(x) &:= f(0) + \sum_{i=1}^{d}\Big(f(e_{\pi(1)}x_{\pi(1)} + \cdots + e_{\pi(i)}x_{\pi(i)}) \\ &\qquad\qquad - f(e_{\pi(1)}x_{\pi(1)} + \cdots + e_{\pi(i-1)}x_{\pi(i-1)})\Big) \qquad (4.9)\\ &=: f(0) + \sum_{i=1}^{d} \mathcal{D}^{(\pi,i)}(f)(x).\end{aligned}$$

Clearly, terms in the second expression of (4.9) telescopes, leaving $f(x)$. Next, we consider a function obtained from $\mathcal{D}^\pi(f)$ by replacing arguments $(x_{\pi(1)}, \dots, x_{\pi(i-1)})$ in $\mathcal{D}^{(\pi,i)}(f)(x)$ by the all-ones vector $(1, \dots, 1) \in \mathbb{R}^{i-1}$. Namely, let $\mathcal{B}^\pi(f) : \{0,1\}^d \mapsto \mathbb{R}^d$ be a function so that

$$\begin{aligned} \mathcal{B}^\pi(f)(x) &:= f\big(v_0^\pi\big) + \sum_{i=1}^d \Big(f\big(v_{i-1}^\pi + e_{\pi(i)} x_{\pi(i)}\big) - f\big(v_{i-1}^\pi\big) \Big) \\ &=: f(v_0^\pi) + \sum_{i=1}^d \mathcal{B}^{(\pi,i)}(f)(x). \end{aligned} \tag{4.10}$$

It follows readily that for every $x \in \{0,1\}^d$ we have $\mathcal{D}^\pi(f)(x) \leq \mathcal{B}^\pi(f)(x)$ since for each $i \in \{1, \dots, d\}$ the supermodularity of $f(\cdot)$ implies that $\mathcal{D}^{(\pi,i)}(f)(x) \leq \mathcal{B}^{(\pi,i)}(f)(x)$. Last, we extend $\mathcal{B}^\pi(f)(\cdot)$ from $\{0,1\}^d$ to $\mathbb{R}^d$ by affinely interpolating each univariate function $\mathcal{B}^{(\pi,i)}(\phi)$ over $\{0,1\}$, namely,

$$\hat{\mathcal{B}}^\pi(f)(x) = f(v_0^\pi) + \sum_{i=1}^d \big(f(v_i^\pi) - f(v_{i-1}^\pi)\big) x_{\pi(i)}. \tag{4.11}$$

In summary, if $f(\cdot)$ is supermodular then

$$f(x) = \mathcal{D}^\pi(f)(x) \leq \mathcal{B}^\pi(f)(x) = \hat{\mathcal{B}}^\pi(f)(x) \quad \text{for every } x \in \{0,1\}^d.$$

In the next result, we show that $\hat{\mathcal{B}}^\Omega(f)(x) := \min_{\pi \in \Omega} \hat{\mathcal{B}}^\pi(f)(x)$ coincides with the Lovász extension of the set function $f(\cdot)$ when $f(\cdot)$ is supermodular.

Corollary 4.1.2 (Proposition 4.1 in [53] and Theorem 3.3 in [16]) *Consider a set function $f : \{0,1\}^d \mapsto \mathbb{R}$ and its Lovász extension $\hat{f}(x)$. If $f(\cdot)$ is supermodular then $\hat{\mathcal{B}}^\Omega(f)(x) = \hat{f}(x)$ for every $x \in [0,1]^d$. Moreover, the concave envelope of $f(\cdot)$ over $[0,1]^d$ is given by $\hat{\mathcal{B}}^\Omega(f)(x)$.*

Proof Let $x \in [0,1]^d$, and let $\pi \in \Omega$ so that $x_{\pi(1)} \geq \cdots \geq x_{\pi(d)}$. Then, it follows readily that $\hat{f}(x) = \mathcal{B}^\pi(f)(x)$. If $\mathcal{B}^\pi(f)(x) = \mathcal{B}^\Omega(f)(x)$ then we can conclude that

$\hat{f}(x) = \mathcal{B}^{\Omega}(f)(x)$. Clearly, we have $\hat{\mathcal{B}}^{\pi}(f)(x) \geq \hat{\mathcal{B}}^{\Omega}(f)(x)$. To show that $\hat{\mathcal{B}}^{\pi}(f)(x) \leq \hat{\mathcal{B}}^{\Omega}(f)(x)$ we observe that for $\omega \in \Omega$,

$$\begin{aligned}
\hat{\mathcal{B}}^{\pi}(f)(x) &= (1 - x_{\pi(1)})f(v_0^{\pi}) + \sum_{i=1}^{d-1}(x_{\pi(i)} - x_{\pi(i+1)})f(v_i^{\pi}) + x_{\pi(n)}f(v_n^{\pi}) \\
&\leq (1 - x_{\pi(1)})\hat{\mathcal{B}}^{\omega}(f)(v_0^{\pi}) + \sum_{i=1}^{d-1}(x_{\pi(i)} - x_{\pi(i+1)})\hat{\mathcal{B}}^{\omega}(f)(v_i^{\pi}) + x_{\pi(n)}\hat{\mathcal{B}}^{\omega}(f)(v_n^{\pi}) \\
&= \hat{\mathcal{B}}^{\omega}(f)(x),
\end{aligned}$$

where the first equality holds by definition, the first inequality holds because $1 \geq x_{\pi(1)} \geq \cdots \geq x_{\pi(d)} \geq 0$ and for every $x \in \{0,1\}^d$ we have $f(x) \leq \hat{\mathcal{B}}^{\pi}(f)(x)$, and the second equality holds by the linearity of $\hat{\mathcal{B}}^{\omega}(f)(\cdot)$ and the convex combination representation of x in terms of $\{v_0^{\pi}, \ldots, v_n^{\pi}\}$.

Now, the proof is complete by observing that for every $x \in [0,1]^d$

$$\mathrm{conc}_{[0,1]^d}(f)(x) \leq \hat{\mathcal{B}}^{\Omega}(f)(x) = \hat{f}(x) \leq \mathrm{conc}_{[0,1]^d}(f)(x),$$

where the first inequality holds because $\hat{\mathcal{B}}^{\Omega}(f)$ is concave and overestimates the set function $f(\cdot)$, and the second inequality holds because $\hat{f}(x)$ is defined as convex combinations of $\{f(v_i)\}_{i=0}^d$ for some $\{v_i\}_{i=0}^d \subseteq \{0,1\}^d$. ■

To illustrate the above construction, we derive the concave envelope of the multilinear monomial over a rectangle in the positive orthant.

Example 4.1.3 (Theorem 1 in [11]) *Consider the multilinear monomial* $m(x) = \prod_{i=1}^d x_i$ *over a rectangle* $[x^L, x^U] := \prod_{i=1}^d [x_i^L, x_i^U]$, *where* $x_i^U > x_i^L \geq 0$ *for all* i. *Let* π *be a movement vector in the grid given by* $\{0,1\}^d$. *For* $i \in \{1, \ldots, d\}$ *define* $I(i) := \{\pi(i') \mid i' \leq i\}$ *as a subset of* $\{1, \ldots, d\}$. *It turns out that*

$$\begin{aligned}
m(x) = \prod_{i=1}^d (x_i^L + (x_i - x_i^L)) &= \prod_{i=1}^d x_i^L + \sum_{i=1}^d \Big(\prod_{j \in I(i) \setminus \pi(i)} x_j\Big) \cdot \Big(\prod_{j \notin I(i)} x_j^L\Big) \cdot (x_{\pi(i)} - x_{\pi(i)}^L) \\
&\leq \prod_{i=1}^d x_i^L + \sum_{i=1}^d \Big(\prod_{j \in I(i) \setminus \pi(i)} x_j^U\Big) \cdot \Big(\prod_{j \notin I(i)} x_j^L\Big) \cdot (x_{\pi(i)} - x_{\pi(i)}^L) \\
&:= M^{\pi}(x).
\end{aligned}$$

Therefore, $\mathrm{conc}_{[x^L,x^U]}(m)(x) \leq \min_{\pi \in \Omega} M^{\pi}(x)$.

Next, we extend constructions in (4.9), (4.10), and (4.11) to obtain the concave envelope of set functions over sets other than the standard hypercube. Consider a subset $\Delta := \prod_{i=1}^{d} \Delta_i$ of the hypercube $[0,1]^{d\times(n+1)}$, where $\Delta_i := \{z_i \in \mathbb{R}^{n+1} \mid 1 = z_{i0} \geq z_{i1} \geq \cdots \geq z_{in} \geq 0\}$. It can be verified that

$$\mathrm{vert}(\Delta_i) = \{\zeta_{ij}\}_{j=0}^{n}, \tag{4.12}$$

where ζ_{ij} denotes the vector $\sum_{j'=0}^{j} e_{ij'}$ and $e_{ij'}$ is the j' principal vector in the space spanned by variables $(z_{i0},\dots,z_{in})$. Therefore, $\mathrm{vert}(\Delta_i)$ forms a chain in $\mathbb{R}^{n+1}$, and thus $\mathrm{vert}(\Delta) = \prod_{i=1}^{d} \mathrm{vert}(\Delta_i)$ is a lattice set. Just as a subvector $(u_{ij_i})_{i=1}^{d}$ of $u \in \mathbb{R}^{d\times(n+1)}$, an extreme point $(\zeta_{ij_i})_{i=1}^{d}$ of Δ can also be conveniently represented on a grid $\mathcal{G}$ given by $\{0,\dots,n\}^n$. Here, we label as the extreme point ζ_{ij}, instead of u_{ij}, the grid marker j along the coordinate direction i. As a consequence of this re-labelling, an extreme point $(\zeta_{ij_i})_{i=1}^{d}$ of Δ is then depicted as the point $\{j_i\}_{i=1}^{d}$ on the grid. Given a grid point $p_k = (j_i)_{i=1}^{d}$, we denote the corresponding extreme point $(\zeta_{1j_1},\dots,\zeta_{dj_d})$ by $\mathrm{ext}(\Delta, p_k)$. Given a movement vector π in the grid $\mathcal{G}$, there is a simplex Υ_π, which is defined by $\mathrm{conv}\big(\bigcup_{k=0}^{dn} \mathrm{ext}(\Delta, \Pi(\pi,k))\big)$. The set of simplices $\{\Upsilon_\pi\}_{\pi\in\Omega}$ yields a subdivision Δ, referred as the *staircase triangulation* of Δ; see Chapter 3 for details.

Given a movement vector π in the grid given by $\{0,\dots,n\}^d$, and given a $k \in \{1,\dots,dn\}$, let $\Theta(\pi,k) := (\Theta_1(\pi,k), \Theta_2(\pi,k))$ so that

$$\Theta_1(\pi,k) = \pi(k) \quad \text{and} \quad \Theta_2(\pi,k) = \big|\big\{j \;\big|\; \pi(j) = \pi(k),\ 0 \leq j \leq k\big\}\big|. \tag{4.13}$$

For any function $\psi : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$, we define $\mathcal{D}^\pi(\psi) : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ so that

$$\begin{aligned}\mathcal{D}^\pi(\psi)(z) := \psi\big(&\mathrm{ext}(\Delta, \Pi(\pi,0))\big)\\ &+ \sum_{k=1}^{dn}\Big(\psi\big(\mathrm{ext}(\Delta,\Pi(\pi,0)) + e_{\Theta(\pi,1)}z_{\Theta(\pi,1)} + \cdots + e_{\Theta(\pi,k)}z_{\Theta(\pi,k)}\big)\\ &- \psi\big(\mathrm{ext}(\Delta,\Pi(\pi,0)) + e_{\Theta(\pi,1)}z_{\Theta(\pi,1)} + \cdots + e_{\Theta(\pi,k-1)}z_{\Theta(\pi,k-1)}\big)\Big).\end{aligned}$$

Clearly, $\mathcal{D}^\pi(\psi)(z)$ telescopes, leaving $\psi(z)$. If $\psi(\cdot)$ is supermodular, we obtain that $\mathcal{D}^\pi(\psi)(z) \leq \mathcal{B}^\pi(\psi)$, where

$$\begin{aligned}\mathcal{B}^\pi(\psi)(z) &:= \psi\big(\mathrm{ext}(\Delta, \Pi(\pi, 0))\big)\\ &\qquad + \sum_{k=1}^{dn}\Big(\psi\big(\mathrm{ext}(\Delta, \Pi(\pi, k-1)) + e_{\Theta(\pi,k)} z_{\Theta(\pi,k)}\big) - \psi\big(\mathrm{ext}(\Delta, \Pi(\pi, k-1))\big)\Big).\end{aligned}$$

Affinely interpolating each term in $\mathcal{B}^\pi(\psi)(z)$ over $[0, 1]$, we obtain a function $\hat{\mathcal{B}}^\pi(\psi) : \mathbb{R}^{d\times(n+1)} \mapsto \mathbb{R}$ defined as

$$\begin{aligned}\hat{\mathcal{B}}^\pi(\psi)(z) &:= \psi\big(\mathrm{ext}(\Delta, \Pi(\pi, 0))\big)\\ &\qquad + \sum_{k=1}^{dn}\Big(\psi\big(\mathrm{ext}(\Delta, \Pi(\pi, k))\big) - \psi\big(\mathrm{ext}(\Delta, \Pi(\pi, k-1))\big)\Big) z_{\Theta(\pi,k)}.\end{aligned} \tag{4.14}$$

Corollary 4.1.3 *The concave envelope of a supermodular function $\psi : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ coincides with $\min_{\pi\in\Omega} \hat{\mathcal{B}}^\pi(\psi)(z)$, where Ω is the set of movement vectors that define monotone staircases over the grid $\mathcal{G}$ given by $\{0, 1, ..., n\}^d$.*

Proof It follows readily that $\mathrm{conc}_\Delta(\psi)(z) \leq \min_{\pi\in\Omega} \hat{\mathcal{B}}^\pi(\psi)(z)$ for every $z \in \Delta$ because the latter is concave and for every $\pi \in \Omega$ and $z \in \mathrm{vert}(\Delta)$ we have $\psi(z) = \mathcal{D}^\pi(\psi)(z) \leq \mathcal{B}^\pi(\psi)(z) \leq \hat{\mathcal{B}}^\pi(\psi)(z)$, where the first equality follows by staircase expansion, the first inequality holds by the supermodularity of $\psi(\cdot)$ over $\mathrm{vert}(\Delta)$, and the last inequality holds by affine interpolation.

To show $\min_{\pi\in\Omega} \hat{\mathcal{B}}^\pi(\psi)(z) \leq \mathrm{conc}_\Delta(\psi)(z)$, we consider a point $\bar{z} \in \Delta$, and sort the coordinates of $\bar{z}$ assuming that $\bar{z}_{10}, \ldots, \bar{z}_{d0}$ are placed first the sorted order. Then, we define a move vector $\bar{\pi}$ such that $\bar{\pi}(k) = i$ if the $d + k$ largest coordinate of $\bar{z}$ is $\bar{z}_{ij}$. It can be verified that

$$\begin{aligned}\bar{z} &= (1 - \bar{z}_{\Theta(\bar{\pi},1)})\,\mathrm{ext}(\Delta, \Pi(\bar{\pi}, 1))\\ &\qquad + \sum_{k=1}^{dn-1}(\bar{z}_{\Theta(\bar{\pi},k)} - \bar{z}_{\Theta(\bar{\pi},k+1)})\,\mathrm{ext}(\Delta, \Pi(\bar{\pi}, k)) + \bar{z}_{\Theta(\bar{\pi},dn)}\,\mathrm{ext}(\Delta, \Pi(\bar{\pi}, dn))\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
\min_{\pi\in\Omega} \hat{\mathcal{B}}^{\bar{\pi}}(\psi)(\bar{z}) &\leq \hat{\mathcal{B}}^{\bar{\pi}}(\psi)(\bar{z}) \\
&= (1-\bar{z}_{\Theta(\bar{\pi},1)})\psi\big(\mathrm{ext}(\Delta,\Pi(\bar{\pi},1))\big) + \sum_{k=1}^{dn-1}(\bar{z}_{\Theta(\bar{\pi},k)} - \bar{z}_{\Theta(\bar{\pi},k+1)})\psi\big(\mathrm{ext}(\Delta,\Pi(\bar{\pi},k))\big) \\
&\quad + \bar{z}_{\Theta(\bar{\pi},dn)}\psi\big(\mathrm{ext}(\Delta,\Pi(\bar{\pi},dn))\big) \\
&\leq \mathrm{conc}_{\Delta}(\psi)(\bar{z}),
\end{aligned}$$

where the first equality holds by the linearity of $\hat{\mathcal{B}}(\psi)(\cdot)$ and, for $k \in \{1,\ldots,dn\}$, $\hat{\mathcal{B}}(\psi)\big(\mathrm{ext}(\Delta,\Pi(\bar{\pi},dn))\big) = \psi\big(\mathrm{ext}(\Delta,\Pi(\bar{\pi},dn))\big)$, and the last inequality holds by the concavity of $\mathrm{conc}_{\Delta}(\psi)$. ■

4.1.2 Tractable composite relaxations

In Chapter 2, we proposed a framework to relax the graph of a composite function $\phi\circ f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$, that is $\mathrm{gr}(\phi\circ f) := \big\{(x,\phi) \bigm| \phi = \phi\big(f_1(x),\ldots,f_d(x)\big),\ x \in X\big\}$. Given a pair $\big(u(x),a\big)$ of estimating functions of the inner-function $f(x)$ which satisfies requirements in (4.3), where $a := (a_1,\ldots,a_d)$ is a pre-specified vector in $\mathbb{R}^{d\times(n+1)}$ such that, for $i \in \{1,\ldots,d\}$, $a_{i0} < \cdots < a_{in}$, they encapsulated ordering relationship of the pair $\big(u(x),a\big)$ in a polytope $P := \prod_{i=1}^d P_i$, where

$$P_i = \left\{u_i \in \mathbb{R}^{n+1} \;\middle|\; a_{i0} \leq u_{ij} \leq \min\{a_{ij},u_{in}\},\ u_{i0} = a_{i0},\ a_{i0} \leq u_{in} \leq a_{in}\right\}. \qquad (4.15)$$

Here, the polytope P_i introduces a variable u_{ij} for each underestimator $u_{ij}(x)$, where $u_{i0}(x)$ (resp. $u_{in}(x)$) is a special underestimator that equals the lower bound a_{i0} (resp. $f(x)$). We denote by Φ^P the graph of the outer-function ϕ over polytope P, which is formally defined as

$$\Phi^P := \big\{(u,\phi) \in \mathbb{R}^{d\times(n+1)} \bigm| \phi = \phi(u_{1n},\ldots,u_{dn}),\ u \in P\big\}.$$

For a pair $\big(u(x),a\big)$ satisfying (4.3), the following result yields relations for the graph of $\phi\circ f$, which will be referred as *composite relaxation* of $\phi\circ f$.

Theorem 4.1.2 (Theorem 2.2.1) *Let $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ be a continuous function and let $f : \mathbb{R}^m \mapsto \mathbb{R}^d$ be a vector of functions, each of which is bounded over $X \subseteq \mathbb{R}^m$. If $\big(u(x), a\big)$ satisfies (4.3) then $\mathrm{gr}(\phi \circ f) \subseteq \mathrm{proj}_{(x,\phi)}(R)$, where*

$$R := \big\{(x, u, \phi) \bigm| (u, \phi) \in \mathrm{conv}(\Phi^P),\ u(x) \leq u,\ u_{\cdot n} = f(x),\ x \in X\big\}.$$

If the graph of $f(x)$, expressed using the constraints $u_{\cdot n} = f(x)$ and $x \in X$ in the definition of R, is outer-approximated with a convex set, and, for $j \neq n$, $u_{ij}(x)$ is convex, one obtains a convex relaxation of $\mathrm{gr}(\phi \circ f)$. Moreover, for each $i \in \{1, \ldots, d\}$ and $j \neq n$, substituting $u_{ij}(x)$ for u_{ij} projects out the u_{ij} variable out of the convex relaxation.

In order to derive the convex hull of Φ^P, we identified a subset Φ^Q of Φ^P defined as

$$\Phi^Q := \big\{(s, \phi) \bigm| \phi = \phi(s_{1n}, \ldots, s_{dn}),\ s \in Q\big\},$$

where $Q := \prod_{i=1}^d Q_i$ and Q_i is a simplex in $\mathbb{R}^{n+1}$ so that its extreme points are defined as follows:

$$v_{ij} = (a_{i0}, \ldots, a_{ij-1}, a_{ij}, \ldots, a_{ij}) \qquad \text{for } j \in \{0, \ldots, n\}. \tag{4.16}$$

Moreover, they proved that the convex hull of Φ^Q together with a small size of inequalities suffice to describe the convex hull of Φ^P in the space of (u, s, ϕ) variables.

Lemma 4.1.2 *The convex hull of Φ^P can be described in the space of (u, s, ϕ) variables as follows:*

$$\big\{(u, s, \phi) \bigm| \mathrm{conv}_Q(\phi)(s) \leq \phi \leq \mathrm{conc}_Q(\phi)(s),\ s \in Q,\ u \in P,\ u \leq s,\ u_{\cdot n} = s_{\cdot n}\big\},$$

where $\mathrm{conv}_Q(\phi)$ (resp. $\mathrm{conc}_Q(\phi)$) denote the convex (resp. concave) envelope of $\phi(s_{1n}, \ldots, s_{dn})$ over Q.

Proof The result follows directly from Lemma 2.3.2 and Lemma 2.3.3. ∎

Remark 4.1.2 *We remark that (convex) relaxations of the set Φ^Q lead to (convex) relaxations for the set Φ^P. More specifically, given a set $R^Q \subseteq \mathbb{R}^{d\times(n+1)+1+\tau}$ so that* $\mathrm{proj}_{(s,\phi)}(R^Q) \subseteq \Phi^Q$ *for some $\tau \geq 0$, we obtain that*

$$\begin{aligned}\Phi^P &= \bigl\{(u,\phi) \bigm| (s,\phi) \in \Phi^Q,\ u \in P,\ u \leq s,\ u_{\cdot n} = s_{\cdot n}\bigr\}\\ &\subseteq \bigl\{(u,\phi) \bigm| (s,\phi,y) \in R^Q,\ u \in P,\ u \leq s,\ u_{\cdot n} = s_{\cdot n}\bigr\},\end{aligned}$$

where the equality follows from Lemma 2.3.2, and the containment holds by the hypothesis about R^Q.

As a result of Lemma 4.1.2, we can focus on deriving the concave envelope of $\phi(s_{1n},\ldots,s_{dn})$ over Q since the convex envelope can be derived in a similar way. We start with introducing an invertible affine mapping from Q to Δ, which is defined as in (4.12). Consider the affine transformation $Z(s) = \bigl(Z_1(s_1),\ldots,Z_d(s_i)\bigr)$ which maps Q to Δ, where $Z_i : \mathbb{R}^{n_i+1} \mapsto \mathbb{R}^{n_i+1}$ is defined as follows

$$z_{i0} = 1 \qquad \text{and} \qquad z_{ij} = \frac{s_{ij} - s_{ij-1}}{a_{ij} - a_{ij-1}} \quad \text{for } j = 1,\ldots,n_i. \tag{4.17}$$

In addition, $Q = Z^{-1}(\Delta)$, where $Z^{-1}(z) = \bigl(Z_1^{-1}(z_1),\ldots,Z_d^{-1}(z_d)\bigr)$ and Z_i^{-1} is given by:

$$s_{ij} = a_{i0}z_{i0} + \sum_{k=1}^{j}(a_{ik} - a_{ik-1})z_{ik} \quad \text{for } j = 0,\ldots,n_i. \tag{4.18}$$

As a result, the simplex Q_i can be described by the following inequalities:

$$0 \leq \frac{s_{in} - s_{in-1}}{a_{in} - a_{in-1}} \leq \cdots \leq \frac{s_{i1} - s_{i0}}{a_{i1} - s_{i0}} \quad s_{i0} = a_{i0}. \tag{4.19}$$

Moreover, for any function $\eta : Q \mapsto \mathbb{R}$, we have that $\mathrm{conc}_Q(\eta)(s) = \mathrm{conc}_\Delta(\psi)\bigl(Z(s)\bigr)$, where $\psi(z) := \eta\bigl(Z^{-1}(z)\bigr)$.

Corollary 4.1.4 (Theorem 3.2.1) *Let $\eta : \mathrm{vert}(Q) \mapsto \mathbb{R}$, and define $\psi : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ so that $\psi(z) = \eta\bigl(Z^{-1}(z)\bigr)$. If $\eta(\cdot)$ is supermodular, $\mathrm{conc}_Q(\eta)(s) = \min_{\pi\in\Omega} \hat{\mathcal{B}}^\pi(\psi)\bigl(Z(s)\bigr)$ for every $s \in Q$. In particular, if $\phi(s_{1n},\ldots,s_{dn})$ is concave-extendable from $\mathrm{vert}(Q)$ and supermodular when restricted to $\mathrm{vert}(Q)$ then*

$$\mathrm{conc}_Q(\phi)(s) = \min_{\pi\in\Omega}\biggl\{\phi\bigl(a,\Pi(\pi,0)\bigr) + \sum_{k=1}^{dn}\Bigl(\phi\bigl(a,\Pi(\pi,k)\bigr) - \phi\bigl(a,\Pi(\pi,k-1)\bigr)\Bigr)z_{\Theta(k)}\biggr\},$$

where recall $\Theta(\pi,k)$ is defined as in (4.13) and z_{ij} is defined as in (4.17).

Proof This result follows directly from Corollary 4.1.3. ■

Example 4.1.4 *Consider a nonlinear function* $\sqrt{x_1 + x_2^2}$ *over* $[0,5] \times [0,2]$. *Let* $s_1(x) := (0, x_1)$, *and let* $s_2(x) := \big(0,\ \max\{\frac{3}{4}x_2^2, 2x_2 - 1\},\ x_2^2\big)$. *It turns out* $a_1 = (0,5)$ *and* $a_2 = (a_{20}, a_{21}, a_{22}) = (0,3,4)$ *is a vector of upper bounds for* $s_1(x)$ *and* $s_2(x)$ *over* $[0,5] \times [0,2]$, *respectively. Moreover, it can be verified that, for each* $x \in [0,5] \times [0,2]$, *the point* $s_i(x)$ *satisfies (4.19). Since* $\sqrt{s_{11} + s_{22}}$ *is submodular over* $[0,5] \times [0,4]$, *it follows from Corollary 4.1.4 that we obtain the following underestimator of* $\sqrt{x_1 + x_2^2}$ *over* $[0,5] \times [0,2]$:

$$
\max \left\{ \begin{aligned}
&0 + \left(\sqrt{5} - 0\right)\frac{s_{11}(x) - 0}{5 - 0} + \left(\sqrt{5+3} - \sqrt{5}\right)\frac{s_{21}(x) - 0}{3 - 0} \\
&\qquad\qquad\qquad\qquad + \left(\sqrt{5+4} - \sqrt{5+3}\right)\frac{s_{22}(x) - s_{21}(x)}{4 - 3} \\
&0 + \left(\sqrt{3} - 0\right)\frac{s_{21}(x) - 0}{3 - 0} + \left(\sqrt{5+3} - \sqrt{3}\right)\frac{s_{11}(x) - 0}{5} \\
&\qquad\qquad\qquad\qquad + \left(\sqrt{5+4} - \sqrt{5+3}\right)\frac{s_{22}(x) - s_{11}(x)}{4 - 3} \\
&0 + \left(\sqrt{3} - 0\right)\frac{s_{21}(x) - 0}{3 - 0} + \left(\sqrt{4} - \sqrt{3}\right)\frac{s_{22}(x) - s_{21}(x)}{4 - 3} \\
&\qquad\qquad\qquad\qquad + \left(\sqrt{5+4} - \sqrt{4}\right)\frac{s_{11}(x) - 0}{5 - 0}
\end{aligned} \right\},
$$

whose convexity can be easily verified.

To extend the applicability of Corollary 4.1.4, we now turn our attention to a particular linear transformation that can be used to convert some functions that are not ordinarily supermodular into supermodular functions. In particular, we will generalize the switching operation over $[0,1]^d$, which is used in Corollary 4.1.1, to Q. To do so, we will need permutations σ_i of $\{0, \dots, n\}$ for each $i \in \{1, \dots, d\}$. We use the permutation σ_i to define an affine transformation that maps v_{ij} to $v_{i\sigma_i(j)}$. Let P^{σ_i} be a permutation matrix in $\mathbb{R}^{(n+1)\times(n+1)}$ such that, for all (i', j'), $P^{\sigma_i}_{i'j'} = 1$ when $i' = \sigma(j')$ and zero otherwise. Then, the affine transformation associated with σ_i is given by

$$
A^{\sigma_i} = Z_i^{-1} \circ U P^{\sigma_i} U^{-1} \circ Z_i, \tag{4.20}
$$

where $\circ$ denotes the composition operator and U is an upper triangular matrix of all ones. We let $A^\sigma(s) := \bigl(A^{\sigma_1}(s_1), \ldots, A^{\sigma_d}(s_d)\bigr)$. Then, we define $\phi(\sigma)(s_1, \ldots, s_d) = \phi\bigl(A^\sigma(s)_{1n}, \ldots, A^\sigma(s)_{dn}\bigr)$ and say that $\phi(\sigma)$ is obtained from ϕ by switching with σ. It follows easily that $\mathrm{conc}_Q(\phi)(s) = \mathrm{conc}_Q\bigl(\phi(\sigma)\bigr)\bigl((A^\sigma)^{-1}(s)\bigr)$.

Corollary 4.1.5 *Let* $\eta : \mathrm{vert}(Q) \mapsto \mathbb{R}$ *be a function, and let* $\sigma := (\sigma_1, \ldots, \sigma_d)$*, where, for* $i \in \{1, \ldots, d\}$*,* σ_i *is a permutation of* $\{0, \ldots, n\}$*. Then, if* $\phi(\sigma)(s)$ *is supermodular when restricted to* $\mathrm{vert}(Q)$*, we have*

$$\mathrm{conc}_Q(\phi)(s) = \min_{\pi \in \Omega} \hat{B}^\pi\bigl(\psi(\sigma)\bigr)\Bigl(Z\bigl((A^\sigma)^{-1}(s)\bigr)\Bigr) \quad \text{for } s \in Q,$$

where $\psi(\sigma) : \Delta \mapsto \mathbb{R}$ *is defined as* $\psi(\sigma)(z) = \phi(\sigma)\bigl(Z^{-1}(z)\bigr)$*, and* A^σ *is defined as in (4.20).*

4.2 Connection to RLT relaxations of polynomial programs

Consider the feasible region X of a polynomial optimization in n variables defined as

$$X = \Bigl\{x \in \mathbb{R}^m \Bigm| g_i(x) \geq 0,\ i = 1, \ldots, k\Bigr\}.$$

Let $\gamma \geq \max_{i=1}^k \deg(g_i)$ be a given integer, where $\deg(p)$ denote the degree of a polynomial function $p(\cdot)$. To obtain a generic LP relaxation of X, the RLT procedure in [39, 69] *reformulates* X by generating implied constraints using distinct product forms:

$$g_1(x)^{\alpha_1} \cdots g_k(x)^{\alpha_k} \geq 0, \qquad \sum_{i=1}^k \alpha_i \leq \gamma. \tag{4.21}$$

After this, RLT *expands* the left-hand side of each resulting polynomial inequality so that it becomes a weighted sum of distinct monomials. Last, RLT *linearizes* the resulting polynomial inequalities by substituting a new variable y_α for each monomial term x^α, so as to obtain linear inequalities in terms of introduced y variables. The resulting LP relaxation is called the degree-γ RLT relaxation for X, and will be denoted as $\mathrm{RLT}_\gamma(X)$. In this section, we study the connection between the reformulation-linearization technique and the technique introduced in Section 4.1.

4.2.1 Improving RLT relaxations of polynomial functions

We first restrict our study to the graph of a polynomial a polynomial function $p : [x^L, x^U] \subseteq \mathbb{R}^m \mapsto \mathbb{R}$, that is, $\mathrm{gr}(p) := \bigl\{(x,\phi) \bigm| \phi = p(x),\ x^L \leq x \leq x^U\bigr\}$. For a given integer $\gamma \geq \deg(p)$, the typical degree-γ RLT relaxation for $\mathrm{gr}(p)$ is obtained by linearizing monomials in the following polynomial constraints with corresponding y variables:

$$\begin{aligned}
&\phi = p(x),\\
&(x_1 - x_1^L)^{\alpha_1} \cdots (x_d - x_d^L)^{\alpha_d} (x_1^U - x_1)^{\beta_1} \cdots (x_d^U - x_d)^{\beta_d} \geq 0 \quad \text{for } \sum_{i=1}^{d} (\alpha_i + \beta_i) \leq \gamma.
\end{aligned}$$

Now, we represent p as a composite function, that is, $p(x) = \phi\bigl(f_1(x), \ldots, f_d(x)\bigr)$ for some multilinear function $\phi(\cdot)$ and polynomial function $f_i(\cdot)$. Assume that there exists a pair $\bigl(u(x), a\bigr)$ satisfying requirement (4.3), where $u : \mathbb{R}^m \mapsto \mathbb{R}^{d\times(n+1)}$ is a vector of polynomials and $a := (a_1, \ldots, a_d)$ is a pre-specified vector in $\mathbb{R}^{d\times(n+1)}$ such that, for $i \in \{1, \ldots, d\}$, $a_{i0} < \cdots < a_{in}$. It turns out that the typical RLT relaxation of $\mathrm{gr}(p)$ can be improved through exploiting the following valid system for $\mathrm{gr}(p)$:

$$\begin{aligned}
&\phi = p(x), \quad x - x^L \geq 0, \quad x^U - x \geq 0,\\
&u_{in}(x) - u_{ij}(x) \geq 0, \quad a_{ij} - u_{ij}(x) \geq 0, \quad u_{in}(x) - a_{i0} \geq 0 \quad \forall i\ \forall j.
\end{aligned} \tag{4.22}$$

However, the following example shows that this improved RLT relaxation does not imply all inequalities obtained using the technique in Theorem 4.1.1.

Example 4.2.1 *Consider the graph* $G := \bigl\{(x,\phi) \bigm| \phi = x_1^2 x_2^2,\ x \in [0,2]^2\bigr\}$, *and consider the following system of valid inequalities for* G*:*

$$x_i \geq 0 \quad 2 - x_i \geq 0 \quad x_i^2 \geq 0 \quad 4 - x_i^2 \geq 0 \quad x_i^2 - (2x_i - 1) \geq 0 \quad 3 - (2x_i - 1) \geq 0. \tag{4.23}$$

Using Corollary 4.1.1 we obtain that

$$
\begin{aligned}
\phi &= x_1^2 x_2^2 \\
&= 0 \cdot 4 + 0 \cdot \Big(\min\big\{4, x_2^2 - (2x_2 - 1) + 3\big\} - 4\Big) \\
&\qquad + \big(\max\{0, 2x_1 - 1\} - 0\big) \cdot \min\big\{4, x_2^2 - (2x_2 - 1) + 3\big\} \\
&\qquad + \max\{0, 2x_1 - 1\} \cdot \Big(x_2^2 - \min\big\{4, x_2^2 - (2x_2 - 1) + 3\big\}\Big) \\
&\qquad + \big(x_1^2 - \max\{0, 2x_1 - 1\}\big) \cdot x_2^2 \\
&\geq 0 + 0 + \big(\max\{0, 2x_1 - 1\} - 0\big) \cdot 3 + 3 \cdot \Big(x_2^2 - \min\big\{4, x_2^2 - (2x_2 - 1) + 3\big\}\Big) + 0 \\
&\geq 3(2x_1 - 1) + 3\big((2x_2 - 1) - 3\big) \\
&= 6x_1 + 6x_2 - 15,
\end{aligned}
$$

where the second equality follows from a staircase expansion, and the first inequality holds due to the validity of $0 \leq \max\{0, 2x_1 - 1\} \leq \min\{x_1^2, 3\}$ *over* $[0, 2]$ *and that of* $\max\{x_2^2, 3\} \leq \min\big\{4, x_2^2 - (2x_2 - 1) + 3\big\} \leq 4$ *over* $[0, 2]$.

Notice that to obtain the above inequality, we exploit the fact that $0 \leq \max\{0, 2x_1 - 1\} \leq x_1^2$. *However, the reformulation-linearization technique fails to deal with those constraints involving piecewise functions* $\max\{0, 2x_1 - 1\}$. *Therefore, it can be expected that applying RLT over (4.23) fails to generate the inequality* $\phi \geq 6x_1 + 6x_2 - 15$. *Moreover, minimizing an affine function* $\phi - (6x_1 + 6x_2 - 15)$ *over the degree-*4 *RLT relaxation of (4.23) yields* -0.36, *while it has been shown that* $\phi - (6x_1 + 6x_2 - 15) \geq 0$ *is valid for* G.

The second idea of improving RLT relaxations of $\mathrm{gr}(p)$ relies on the construction in Theorem 4.1.2. By the Theorem 4.1.2, an extended formulation of $\mathrm{gr}(p)$ is given as follows

$$
\big\{(x, u, \phi) \bigm| (u, \phi) \in \Phi^P,\ u(x) \leq u,\ u_{\cdot n} = f(x),\ x^L \leq x \leq x^U\big\}, \tag{4.24}
$$

where $\Phi^P := \big\{(u, \phi) \bigm| \phi = \phi(u_{1n}, \ldots, u_{dn}),\ u \in P\big\}$ and $P := \prod_{i=1}^d P_i$ is defined as in (4.15). Therefore, a natural idea to construct relaxations for $\mathrm{gr}(p)$ is to replace the set Φ^P in (4.24) with its RLT relaxations.

Example 4.2.2 *Assume the same setup as Example 4.2.1. The polytope P, which models the ordering relationship among functions, is given by*

$$0 \le u_i \le \min\{3, f_i\} \quad \textit{and} \quad 0 \le f_i \le 4 \quad \textit{for } i = 1, 2.$$

It follows readily that the inequality $\phi \ge 6x_1 + 6x_2 - 15$ in Example 4.2.1 is implied by $u_i \ge 2x_i - 1$ and the following RLT type constraints:

$$\phi = f_1 f_2 \quad (f_1 - u_1)(f_2 - 0) \ge 0 \quad (3 - u_1)(4 - f_2) \ge 0.$$

However, the degree-2 RLT fails to yield the convex hull of $\big\{(f,u) \bigm| \phi = f_1 f_2,\ (f,u) \in P\big\}$.

Although RLT fails to generate $\operatorname{conv}(\Phi^P)$ in the d^{th}-level, we will show that, exploiting the structure of the Cartesian product of simplices Q defined as in (4.16), RLT is helpful in describing the convex hull $\operatorname{conv}(\Phi^P)$. We consider a set of points $\mathcal{M}^Q_{\mathcal{I}}$ satisfying all degree d multilinear monomial equations over the Cartesian product of simplices Q, that is ,

$$\mathcal{M}^Q_{\mathcal{I}} := \Big\{(s,m) \Bigm| s \in Q,\ m_{(I,e)} = \prod_{i \in I} s_{ie_i} \ \forall I \in \mathcal{I} \ \forall e \in E\Big\},$$

where $\mathcal{I}$ is a collection of subsets of $\{1,\ldots,d\}$, and E is a set of all d-tuples so that, for each $e := (e_1,\ldots,e_d) \in E$, we have $e_i \in \{0,\ldots,n\}$. The set $\mathcal{M}^Q_{\mathcal{I}}$ is related to another set of points $\mathcal{M}^\Lambda$ satisfying all degree-d multilinear monomial equations over the Cartesian product of standard simplices $\Lambda := \prod_{i=1}^d \Lambda_i$, that is,

$$\mathcal{M}^\Lambda := \Big\{(\lambda, w) \Bigm| \lambda \in \Lambda,\ w_e = \prod_{i=1}^d \lambda_{ie_i} \ \forall e \in E\Big\}, \tag{4.25}$$

where $\Lambda_i := \big\{\lambda_i \in \mathbb{R}^{n+1} \bigm| \sum_{j=0}^n \lambda_{ij} = 1,\ \lambda_i \ge 0\big\}$. It can be verified that a linear transformation that maps $\mathcal{M}^\Lambda_{\mathcal{I}}$ to $\mathcal{M}^Q$ is given by:

$$\begin{aligned} s_i &= \sum_{j=0}^n v_{ij}\lambda_{ij} \quad \text{for } i \in \{1,\ldots,d\}, \\ m_{(I,e)} &= \sum_{e' \in E}\Big(\prod_{i \in I} a_{ie_i \wedge e'_i}\Big) w_{e'} \quad \text{for } I \in \mathcal{I} \text{ and } e \in E, \end{aligned} \tag{4.26}$$

where $\{v_{10}, \ldots, v_{in}\}$ is the set of vertices of Q_i and $v_{ij} := (a_{i0}, \ldots, a_{ij-1}, a_{ij}, \ldots, a_{ij})$. As a result, in order to describe the convex hull of $\mathcal{M}^Q$, it suffices to derive the convex hull of $\mathcal{M}^\Lambda$.

Lemma 4.2.1 $\mathrm{RLT}_d(\mathcal{M}^\Lambda) = \mathrm{conv}(\mathcal{M}^\Lambda)$.

Proof Clearly, we have $\mathrm{conv}(\mathcal{M}^\Lambda) \subseteq \mathrm{RLT}_d(\mathcal{M}^\Lambda)$ because $\mathcal{M}^\Lambda \subseteq \mathrm{RLT}_d(\mathcal{M}^\Lambda)$ and the latter set is convex. To show the opposite containment, we consider a set

$$R := \left\{ (\lambda, w) \;\middle|\; \lambda \in \Lambda,\ w \geq 0,\ \lambda_{ij} = \sum_{e \in E : e_i = j} w_e \ \forall i\ \forall j \right\}. \tag{4.27}$$

It follows readily that $\mathrm{RLT}_d(\mathcal{M}^\Lambda) \subseteq R$. Therefore, the proof is complete if we show that $R \subseteq \mathrm{conv}(\mathcal{M}^\Lambda)$. Consider a point $(\lambda, w) \in R$. It follows readily that the point (λ, w) is expressible as follows:

$$(\lambda, w) = \sum_{e \in E} \big(\mu(e), \chi(e)\big) w_e,$$

where $\mu(e) \in \mathbb{R}^{d \times (n+1)}$ is a vector so that $\mu(e)_{ij} = 1$ if $e_i = j$ and $\mu(e)_{ij} = 0$ otherwise, and $\chi(e) \in \mathbb{R}^{|E|}$ is the indicator vector of e. The equality holds because, for all i and j, we have $\lambda_{ij} = \sum_{e \in E : e_i = j} w_e = \sum_{e \in E} \mu_{ij}(e) w_e$, and for all $e \in E$ we have $w_e = \sum_{e \in E} \chi(e) w_e$. Moreover, we have $w \geq 0$ and

$$\sum_{e \in E} w_e = \sum_{j=0}^{d} \Bigg(\sum_{e \in E : e_i = j} w_e \Bigg) = \sum_{j=0}^{d} \lambda_{ij} = 1.$$

Therefore, (λ, m) is expressible as convex combinations of points in $\mathcal{M}^\Lambda$. Hence, $R \subseteq \mathrm{conv}(\mathcal{M}^\Lambda)$. ■

Proposition 4.2.1 *If $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ is a multilinear function, i.e., $\phi(s_{1n}, \ldots, s_{dn}) = \sum_{I \in \mathcal{I}} c_I \prod_{i \in I} s_{in}$, where $\mathcal{I}$ is a collection of subsets of $\{1, \ldots, d\}$, the convex hull of Φ^Q is given by*

$$\left\{ (s, \phi) \;\middle|\; w \geq 0, \quad \phi = \sum_{I \in \mathcal{I}} \bigg(c_I \sum_{e \in E} \Big(\prod_{i \in I} a_{ie_i} \Big) w_e \bigg),\ s_i = \sum_{j=0}^{n} v_{ij} \sum_{e \in E : e_i = j} w_e \ \forall i \right\}. \tag{4.28}$$

Moreover, the convex hull of Φ^P is described by (4.28) and constraints $u \in P$, $u \leq s$ and $u_{\cdot n} = s_{\cdot n}$.

Proof The first statement follows from Lemma 4.2.1 and the linear transformation (4.26), and from the fact that convexification commutes with linear transformation. The second statement follows from Lemma 4.1.2. ∎

4.2.2 Improving RLT via recursive staircase expansion

In this subsection, we show how to recursively apply staircase expansion and termwise relaxations to improve RLT relaxations of the set X, which is a set of points in $\mathbb{R}^m$ satisfying polynomials inequalities $g_i(x) \geq 0$, $i \in \{1, \ldots, k\}$. We start with introducing the notion of recursive staircase expansion. Recall that a vector $u \in \mathbb{R}^{d \times (n+1)}$ is represented on the grid $\mathcal{G}$ given by $\{0, \ldots, n\}^d$, and that a subvector $(u_{1j_1}, \ldots, u_{dj_d})$ is denoted as $\bigl(u, (j_i)_{i=1}^d\bigr)$. For a given function $\phi : \mathbb{R}^d \mapsto \mathbb{R}$, and for a given $i \in \{1, \ldots, d\}$ and a point $p \in \{0, \ldots, n\}^d$, let $\partial_i(\phi)(\cdot; p) : \mathbb{R}^{d \times (n+1)} \mapsto \mathbb{R}$ be the difference function defined as follows:

$$\partial_i(\phi)(u; p) := \phi(u, p) - \phi(u, p - e_i),$$

where e_i denote the i^{th} principal vector in $\mathbb{R}^d$. More generally, for a subset $I := \{i_1, \ldots, i_t\} \subseteq \{1, \ldots, d\}$ let $\partial_I(\phi) : \mathbb{R}^{d \times (n+1)} \mapsto \mathbb{R}$ be the difference function defined as follows:

$$\partial_I(\phi)(u; p) := \partial_{i_1}\bigl(\cdots \partial_{i_t}(\phi)\bigr)(u; p).$$

Next, we expand the difference function $\partial_I(\phi)(u; p)$ according to a staircase move in a sub-grid of $\mathcal{G}$. We denote by $\mathcal{G}_{(I,p)}$ the sub-grid of $\mathcal{G}$ given by $\prod_{i=1}^d \bigl\{(0, p_I)_i, \ldots, p_i\bigr\}$, where $(0, p_I)$ is a point of $\{0, \ldots, n\}^d$ so that $(0, p_I)_i = p_i$ if $i \in I$ and $(0, p_I)_i = 0$ otherwise. For a staircase move $\pi = (\pi_1, \ldots, \pi_\tau)$ in the sub-grid $\mathcal{G}_{(I,p)}$, we can also track where the k^{th} move leaves us on the sub-grid using the transformation $\Pi_{(I,p)}$, which is defined as $\Pi_{(I,p)}(\pi, k) := (0, p_I) + \sum_{j=1}^{\tau} e_{\pi(j)}$. It follow readily that function $\partial_I(\phi)(u; p)$ can be expanded into the following way

$$\mathcal{D}^\pi\bigl(\partial_I(\phi)\bigr)(u; p) = \partial_I(\phi)\bigl(u; \Pi_{(I,p)}(\pi, 0)\bigr) + \sum_{k=1}^{m} \partial_{I \cup \pi(k)}\bigl(u; \Pi_{(I,p)}(\pi, k)\bigr).$$

As a result, for a given composite function $\phi \circ f$ and a given vector of functions $u : X \mapsto \mathbb{R}^{d\times(n+1)}$ so that $u_{\cdot n}(x) := f(x)$, we can expand $\phi \circ f$ into the sum of difference functions, that is

$$(\phi \circ f)(x) = \sum_{I \subseteq \mathcal{I}} \sum_{p \in J_I} \partial_I(\phi)\big(u(x); p\big), \tag{4.29}$$

where $\mathcal{I}$ is a collection of subsets of $\{1, \ldots, d\}$, and, for each $I \in \mathcal{I}$, J_I is set of points in $\{0, \ldots, n\}^d$.

Now, we are ready to present a relaxation procedure for the set X, which consists of the following four steps:

Composition: Construct a polynomial function $p : X \mapsto \mathbb{R}$ as follows: $p(x) = \phi\big(f_1(x), \ldots, f_d(x)\big)$, where $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ is a continuous function and f_i is a polynomial function;

Recursive staircase expansion: Expand the polynomial function $p(x)$ into sums of difference functions defined as in (4.29);

Termwise relaxation: Relax difference functions using special structures, *e.g.*, structures exploited in Section 4.1 and Section 4.2.1;

Linearization: Linearize nonlinear expressions by replacing difference functions and monomials with variables.

Our procedure can generate inequalities produced by RLT.

Example 4.2.3 (RLT) *We use our procedure to derive the RLT constraint obtained by linearizing monomials in the product constraint $g_1(x)g_2(x) \geq 0$. Clearly, the polynomial g_i can be represented as a weighted sum of monomials, that is $g_i(x) :=$*

$\sum_{j=0}^{n} a_{ij}x^{\alpha_{ij}}$, where $a_{ij} \in \mathbb{R}$ and $m_{ij}(x)$ is a monomial. Define $u_{ij}(x) := \sum_{j'=0}^{j} a_{ij}x^{\alpha_{ij}}$. Then, we obtain

$$\begin{aligned} g_1(x)g_2(x) = u_{10}(x)u_{20}(x) & \\ &+ \big(u_{11}(x)u_{20}(x) - u_{10}(x)u_{20}(x)\big) + \cdots + \big(u_{1n}(x)u_{20}(x) - u_{1(n-1)}(x)u_{20}(x)\big) \\ &+ \big(u_{1n}(x)u_{21}(x) - u_{1n}(x)u_{20}(x)\big) + \cdots + \big(u_{1n}(x)u_{2n}(x) - u_{1n}(x)u_{2(n-1)}(x)\big). \end{aligned}$$

Moreover, for $j \in \{1, \dots, n\}$ the difference function $u_{1n}(x)\big(u_{2j}(x) - u_{2j-1}(x)\big)$ can be further expanded as follows:

$$u_{10}(x)\big(u_{2j}(x)-u_{2j-1}(x)\big)+\sum_{j'=1}^{n} u_{ij'}(x)\big(u_{2j}(x)-u_{2j-1}(x)\big)-u_{ij'-1}(x)\big(u_{2j}(x)-u_{2j-1}(x)\big).$$

In other words, recursive staircase expansions yield $\sum_{j_1=0}^{n}\sum_{j_2=0}^{n} a_{1j_1}a_{2j_2}x^{\alpha_{1j_1}+\alpha_{2j_2}}$. On the other hand, the composite function $g_1(x)g_2(x)$ is lower bounded by 0 over X since $g_1(x) \geq 0$ and $g_2(x) \geq 0$ are valid for X. Therefore, for every $x \in X$ we have that $\sum_{j_1=0}^{n}\sum_{j_2=0}^{n} a_{1j_1}a_{2j_2}x^{\alpha_{1j_1}+\alpha_{2j_2}} = g_1(x)g_2(x) \geq 0$. Hence, after linearizing monomial $x^{\alpha_{1j_1}+\alpha_{2j_2}}$ by a new variable $y_{\alpha_{1j_1}+\alpha_{2j_2}}$, we obtain a valid inequality inequality for X, that is $\sum_{j_1=0}^{n}\sum_{j_2=0}^{n} a_{1j_1}a_{2j_2}y_{\alpha_{1j_1}+\alpha_{2j_2}} \geq 0$.

Example 4.2.4 (Lemma 4.2.1) *In this example, we apply our procedure to generate the convex hull of set $\mathcal{M}^{\Lambda}$ defined as in (4.25). First, we consider a function $p(\lambda) = \prod_{i=1}^{d}(\lambda_{i0} + \cdots + \lambda_{in})$, which can be treated as a composite function $\phi\big(f_1(\lambda_1), \dots, f_d(\lambda_d)\big)$ where $\phi(f_1, \dots, f_d) = \prod_{i=1}^{d} f_i$ and $f_i(\lambda_i) = \lambda_{i0}+\cdots+\lambda_{in}$. Then, by recursive staircase expansion the function $p(x)$ can be expanded into $\sum_{e\in E}\prod_{i=1}^{d}\lambda_{ie_i}$. Next, we observe that for every $e \in E$ the monomial $\prod_{i=1}^{d}\lambda_{ie_i}$ is bounded from below by 0 and that $p(\lambda) = 1$. After linearization, we obtain that $\sum_{e\in E} w_e = 1$ and, for every $e \in E$, $w_e \geq 0$. Last, the inequality $\lambda_{ij} = \sum_{e\in E: e_i=j} w_e$ can be obtained in a similar way.*

Our procedure generates valid linear inequalities which improve a certain level RLT relaxation using variables in a lower level RLT.

Example 4.2.5 *Consider the set $M_{(2,4)}$ defined by all monomials up to degree 4 in terms of 2 variables, that is, $M_{(2,4)} := \{m \mid x \in [0,1]^2,\ m_{(\gamma_1,\gamma_2)} = x_1^{\gamma_1} x_2^{\gamma_2},\ \gamma_1+\gamma_2 \leq 4\}$. In the following, we will derive linear inequalities for the set $M_{(2,4)}$ without using additional variables and show that those inequalities are not implied by the level-6 RLT.*

First, consider a polynomial p with degree of 4 defined as $p(x) = (1-2x_1+x_1^2)(1-2x_2+x_2^2)$. For $i = 1,2$ let $u_i(x)$ be a vector of underestimators of $1-2x_i+x_i^2$ defined as $u_{i0}(x) = 0$, $u_{i1}(x) = -x_i + 0.75$ and $u_{i2}(x) = (1-2x_1+x_1^2)$, and let $a_i := (0, 0.75, 1)$. Then, our procedure yields the following valid inequality for $M_{(2,4)}$:

$$p(x) \geq 0.25(-x_1+0.75)+0.75(1-2x_1+x_1^2)+0.25(-x_2+0.75)+0.75(1-2x_2+x_2^2)-0.935.$$

After replacing monomials by m variables we obtain a linear valid inequality $l_1(m) - l_2(m) \geq 0$ for $M_{(2,4)}$, where

$$l_1(z) := m_{2,2} - 2m_{2,1} - 2m_{1,2} + m_{2,0} + 4m_{1,1} + m_{0,2} - 2m_{1,0} - 2m_{0,1} + 1,$$
$$l_2(z) := 0.25(-m_{1,0} + 0.75) + 0.75(m_{2,0} - 2m_{1,0} + 1) + 0.25(-m_{0,1} + 0.75) + 0.75(m_{0,2} - 2m_{0,1} + 1) - 0.935.$$

Now, we conclude that the valid linear inequality $l_1(m)-l_2(m) \geq 0$ is not implied by the level-6 RLT of $M_{(2,4)}$ since $\min\{l_1(m) - l_2(m) \mid (m,w) \in \mathrm{RLT}_6(M_{(2,4)})\} = -0.021$.

Next, consider a polynomial with degree of 6 defined as $q(x) = (1-x_1)^3(1-x_2)^3$. One way to expand $q(x)$ is given as follows:

$$q(x) = 1+1(-3x_2)+(-3x_1)(-3x_2+1)+(-3x_1+1)(-x_2^3+3x_2^2)+(-x_1^3+3x_1^2)(1-x_2)^3, \tag{4.30}$$

where the last term can be relaxed to a degree 4 polynomial as follows:

$$(-x_1^3 + 3x_1^2)(1-x_2)^3 \leq -x_2^3 + 3x_2^2. \tag{4.31}$$

On the other hand, let $u_i(x)$ be a vector of underestimators of $(1-x_i)^3$ defined as $u_{i0}(x) = 0$, $u_{i1}(x) = -0.75x_1 + 0.5$, and $u_{i2}(x) = (1-x_i)^3$, and let $a_i := (0, 0.5, 1)$. Then, our procedure yields the following inequality for $M_{(2,4)}$:

$$q(x) \geq 0.5(-0.75x_1+0.5)+0.5(1-x_1)^3+0.5(-0.75x_2+0.5)+0.5(1-x_2)^3-0.75. \tag{4.32}$$

As a result, (4.30), (4.32) and (4.31) together imply a valid inequality for $M_{(2,4)}$ with degree of 4*, namely,*

$$1 + 1(-3x_2) + (-3x_1)(-3x_2+1) + (-3x_1+1)(-x_2^3+3x_2^2) + (-x_2^3+3x_2^2) \geq$$
$$0.5(-0.75x_1+0.5) + 0.5(1-x_1)^3 + 0.5(-0.75x_2+0.5) + 0.5(1-x_2)^3 - 0.75.$$

*Let $l(z) \geq 0$ denote the resulting linear inequality obtained by replacing monomials in the above polynomial inequality with corresponding m variables. It follows that that $l(z) \geq 0$ is not implied by the level-*6 *RLT relaxation of $M_{(2,4)}$ since* $\min\bigl\{l(m) \bigm| (m,w) \in \mathrm{RLT}_6(M_{(2,4)})\bigr\} = -0.125$.

Example 4.2.6 *Consider the product of simplices $Q := \prod_{i=1}^{3} Q_i$, where Q_i is the convex hull of* $\{(0,0),(3,3),(3,4)\}$*, and consider the epigraph Φ of $s_{12}s_{22}s_{32}$ over Q, that is $\Phi := \bigl\{(s,\phi) \bigm| \phi \geq s_{12}s_{22}s_{32},\ s \in Q\bigr\}$. Clearly, using the reformulation-linearization technique to derive relaxations for Φ, we need to utilize inequalities in the level-3 RLT of Q. In this example, we show that our procedure, using variables in the level-*2 *RLT of Q, can generate relaxations for Φ,*

$$\begin{aligned}\phi \geq s_{12}s_{22}s_{32} &= s_{11}s_{21}(s_{31}-0) + (s_{12}-s_{11})s_{21}s_{31} + s_{12}(s_{22}-s_{21})s_{31} + s_{12}s_{22}(s_{32}-s_{31}) \\ &\geq (3s_{11}+3s_{21}-9)s_{31} + (s_{12}-s_{11})(3s_{21}+s_{31}-9) + (3s_{12}+4s_{31}-12)(s_{22}-s_{21}) \\ &\quad + (3s_{12}+3s_{22}+s_{11}+s_{12}-15)(s_{32}-s_{21}),\end{aligned}$$

*where the first equality follows by staircase expansion, and the second inequality can be established by using Example and the fact that $s_{i2} \geq s_{i1}$. Last, linearizing bilinear monomials in the above inequality with corresponding variables in the level-*2 *RLT of Q, we obtain a valid linear inequality for Φ.*

4.3 Discrete relaxations for composite functions

In this section, we are interested in deriving *mixed-integer linear programming (MIP) relaxations* for the graph of a composite function $\phi \circ f : X \mapsto \mathbb{R}$, that is

$$\mathrm{gr}(\phi \circ f) = \bigl\{(x,\phi) \bigm| \phi = (\phi \circ f)(x),\ x \in X\bigr\},$$

where the inner-functions $f(\cdot)$ is assumed to be bounded over X. Here, for a set $y_1 \in D \subseteq \mathbb{R}^{d_1}$, we say that a set $E := L \cap \big(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^t \times \{0,1\}^\kappa\big)$, where L is a polytope in the space of variables (y_1, y_2, w, δ), is an MIP relaxation of Y if $D \subseteq \operatorname{proj}_{y_1}(E)$, and we say that L is the *linear programming (LP) relaxation* of E. If there exists E such that $D = \operatorname{proj}_{y_1}(E)$ then we say that the set D is *mixed-integer linear programming (MIP) representable* and E is a *mixed-integer linear programming (MIP) formulation* of D. Otherwise, we need to construct an MIP representable relaxation $\tilde{D} \subseteq \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ of D in an extended space of variables $y := (y_1, y_2)$, and then derive an MIP formulation E for $\tilde{D}$. Clearly, a tighter MIP representable relaxation $\tilde{D}$ is favorable since it leads to a better bound for the optimization problem over D. Moreover, since the most successful MIP solvers are currently based on the branch-and-bound (B&B) algorithm, we also interested in the strength of the LP relaxation L of the MIP formulation E of $\tilde{D}$. We say that a formulation E is *ideal* if $\operatorname{proj}_\delta(\operatorname{vert}(L)) \subseteq \{0,1\}^\kappa$. Ideal formulations are desirable since solving the LP relaxation yields an optimal solution that is integer and optimal for the original MIP problem.

4.3.1 Connecting incremental formulation and composite relaxation

Let $a := (a_1, \ldots, a_d)$ be a pre-specified vector in $\mathbb{R}^{d \times (n+1)}$ such that, for $i \in \{1, \ldots, d\}$, $f_i^L = a_{i0} < \cdots < a_{in} = f_i^U$, where for every $x \in X$ we have $f_i^L \leq f_i(x) \leq f_i^U$. For each $i \in \{1, \ldots, d\}$, we partition $[f_i^L, f_i^U]$ along points $a_{i\tau(i,0)}, \ldots, a_{i\tau(i,l_i)}$, that is, $\big[f_i^L, f_i^U\big] = \bigcup_{k=1}^{l_i} \big[a_{i\tau(i,k-1)}, a_{i\tau(i,k)}\big]$, where $0 = \tau(i,0) < \ldots < \tau(i,l_i) = n$. It follows readily that we obtain a subdivision $\mathcal{H}$ of the hypercube $[f^L, f^U]$, where

$$\mathcal{H} := \left\{ \prod_{i=1}^{d} [a_{i\tau(i,t_i-1)}, a_{i\tau(i,t_i)}] \;\middle|\; t_i \in \{1, \ldots, l_i\} \right\}.$$

Then, a typical MIP relaxation is constructed in two steps. First, we obtain a relaxation for the graph of composite function $\phi \circ f$ defined as follows:

$$\operatorname{gr}(\phi \circ f) \subseteq \left\{ (x, \phi) \;\middle|\; (f, \phi) \in \bigcup_{H \in \mathcal{H}} \operatorname{conv}(\phi^H),\ (x, f) \in W \right\}, \tag{4.33}$$

where ϕ^H denotes the graph of the outer-function $\phi(\cdot)$ when restricted to a subcube H, and W is a polyhderal outer-approximation of the inner-functions $f(\cdot)$. Second, we derive an MIP formulation for the right side of (4.33) by replacing the disjunctive constraint $\big\{\operatorname{conv}(\phi^H)\big\}_{H\in\mathcal{H}}$ by a valid MIP formulation. A standard formulation for $\big\{\operatorname{conv}(\phi^H)\big\}_{H\in\mathcal{H}}$ can be derived using results in [70,71], provided that for each $H\in\mathcal{H}$ the convex hull $\operatorname{conv}(\phi^H)$ is a polytope. Since it is constructed by treating each polytope separately and then using disjunctive programming, it requires a binary variable δ^H and a copy of (f^H,ϕ^H) for the convex hull $\operatorname{conv}(\phi^H)$, and thus $|\mathcal{H}|$ binary variables and $|\mathcal{H}|(d+1)$ continuous variables. However, convexifying the outer-function $\phi(\cdot)$ over a certain structure from incremental formulation [35,72], we obtain an ideal MIP formulation for $\big\{\operatorname{conv}(\phi^H)\big\}_{H\in\mathcal{H}}$, which requires $d(n+1)$ auxiliary continuous variables and $\sum_{i=1}^d(l_i-1)$ auxiliary binary variables. Moreover importantly, auxiliary variables in incremental formulation can be related to certain underestimators of inner-functions $f(\cdot)$, yielding an MIP relaxation for the graph of $\phi\circ f$, which is tighter than the relaxation defined as in (4.33) and the composite relaxation in Theorem 4.1.2. We also discuss a tightening strategy to further improve the strength of our MIP relaxation.

We start with an MIP formulation of selecting a subcube from $\mathcal{H}$:

$$z_i\in\Delta_i,\ \ \delta_i\in\{0,1\}^{l_i-1},\ \ z_{i\tau(i,t)}\geq\delta_{it}\geq z_{i\tau(i,t)+1}\quad\text{for } i\in\{1,\dots,d\}\ t\in\{1,\dots,l_i-1\}, \tag{4.34a}$$

$$f_i=a_{i0}z_{i0}+\sum_{j=1}^n(a_{ij}-a_{ij-1})z_{ij}=:F_i(z_i)\qquad\text{for } i\in\{1,\dots,d\}, \tag{4.34b}$$

where $\Delta_i:=\{z_i\in\mathbb{R}^{n+1}\mid 1=z_{i0}\geq z_{i1}\geq\cdots\geq z_{in}\geq 0\}$. Let $\Delta:=\prod_{i=1}^d\Delta$, and for $H\in\mathcal{H}$ let Δ_H be the face of Δ defined by $z_{i\tau(i,t_i(H)-1)}=1$ and $z_{i\tau(i,t_i(H))+1}=0$ for $i\in\{1,\dots,d\}$. Then, it can be verified that

$$(z,\delta)\in(4.34a)\iff z\in\bigcup_{H\in\mathcal{H}}\Delta_H. \tag{4.35}$$

Moreover, for $H\in\mathcal{H}$, the affine function $F(z):=\big(F_1(z_1),\dots,F_d(z_d)\big)$, where $F_i(\cdot)$ is defined in (4.34b), maps the face Δ_H to the subcube H. Therefore, (4.34) is a valid MIP formulation for selecting a subcube from $\mathcal{H}$. Notice that when $d = 1$ and $\{\tau(i,t)\}_{t=0}^{l_i} = \{0, \ldots, n\}$ the formulation (4.34) reduces to

$$f_1 = F_1(z_1),\ z_1 \in \Delta_1,\ \delta_1 \in \{0,1\}^{n-1},\ z_{1j} \geq \delta_{1j} \geq z_{1j+1} \quad \text{for } j \in \{1, \ldots, n-1\},$$

which is the classic *incremental formulation* of selecting an interval from $\big\{[a_{1j-1}, a_{1j}]\big\}_{j=1}^{n}$ in [35, 72]. In the following, we refer to (4.34) as the incremental formulation for selecting a cube from $\mathcal{H}$.

Let Ψ^{Δ} be the graph of function $\phi(F_1(z_1), \ldots, F_d(z_d))$ over Δ, that is

$$\Psi^{\Delta} := \Big\{(z,\phi)\ \Big|\ \phi = \phi\big(F(z)\big),\ z \in \Delta\Big\},$$

where $F(z) := \big(F_1(z_1), \ldots, F_d(z_d)\big)$ and $F_i(\cdot)$ is defined as in (4.34b). We will show that the convex hull of Ψ^{Δ}, together with the incremental formulation, yields an ideal MIP formulation for $\big\{\mathrm{conv}(\phi^H)\big\}_{H \in \mathcal{H}}$, provided that the convex hull is a polytope.

Proposition 4.3.1 *If* $\mathrm{conv}(\Psi^{\Delta})$ *is a polytope then an ideal MIP formulation for* $\big\{\mathrm{conv}(\phi^H)\big\}_{H \in \mathcal{H}}$ *is given by*

$$\Big\{(f,\phi,z,\delta)\ \Big|\ (z,\phi) \in \mathrm{conv}\big(\Psi^{\Delta}\big),\ (f,z,\delta) \in (4.34)\Big\}. \tag{4.36}$$

Proof Let E denote the formulation (4.37). Then, we observe that

$$\begin{aligned}
\mathrm{proj}_{(f,\phi)}(E) &= \mathrm{proj}_{(f,\phi)}\big(\mathrm{proj}_{(f,\phi,z)}(E)\big) \\
&= \mathrm{proj}_{(f,\phi)}\Big(\bigcup_{H \in \mathcal{H}} \big\{(f,\phi,z)\ \big|\ (z,\phi) \in \mathrm{conv}(\Psi^{\Delta}),\ z \in \Delta_H,\ f = F(z)\big\}\Big) \\
&= \mathrm{proj}_{(f,\phi)}\Big(\bigcup_{H \in \mathcal{H}} \mathrm{conv}\big\{(f,z,\phi)\ \big|\ (z,\phi) \in \Psi^{\Delta},\ z \in \Delta_H,\ f = F(z)\big\}\Big) \\
&= \bigcup_{H \in \mathcal{H}} \mathrm{conv}\Big(\mathrm{proj}_{(f,\phi)}\Big(\big\{(f,\phi,z)\ \big|\ (z,\phi) \in \Psi^{\Delta},\ z \in \Delta_H,\ f = F(z)\big\}\Big)\Big) \\
&= \bigcup_{H \in \mathcal{H}} \mathrm{conv}\big(\phi^H\big),
\end{aligned}$$

where the first equality holds because , the second equality holds due to (4.35), the third equality holds because $\Delta_H \cap \mathbb{R}$ is a face of $\mathrm{conv}(\Psi^{\Delta})$ and $F(\cdot)$ is an affine function, the fourth follows because projection commutes with convexification, and the last equality holds because $F(\Delta_H) = H$.

Next, we show that the formulation E is ideal. Consider an extreme point (f, ϕ, z, δ) of the LP relaxation R of E. We first argue that for all i and $t \in \{1, \dots, l_i - 1\}$ either $\delta_{it} = z_{i\tau(i,t)}$ or $\delta_{it} = z_{i\tau(i,t)+1}$, denoted as $\delta = L(z)$. Suppose that this is not the case, *i.e.*, $z_{i'\tau(i',t')} < \delta_{i't'} < z_{i'\tau(i',t'+1)}$ for some i' and t'. Then, since the convex hull of Ψ^Δ does not depend on variable δ, it follows readily that (f, ϕ, z, δ') and (f, ϕ, z, δ'') belong to R, where δ' and δ'' are two vectors so that $\delta'_{it} = \delta''_{it} = \delta_{it}$ for all i and $t \neq t'$, and $\delta'_{it} = z_{i'\tau(i',t')}$ and $\delta''_{it} = z_{i'\tau(i',t')+1}$ otherwise. However, the extreme point (f, ϕ, z, δ) can be expressed a convex combination of (f, ϕ, z, δ') and (f, ϕ, z, δ''), yielding a contradiction.

Now, we show $z \in \mathrm{vert}(\Delta)$ by contradiction. This implies that δ is binary, since $\mathrm{vert}(\Delta) \in \{0,1\}^{d\times(n+1)}$ and it has been shown that $\delta = L(z)$. Since $\mathrm{conv}(\Psi^\Delta)$ is a polytope and $(z, \phi) \in \mathrm{conv}(\Psi^\Delta)$, it follows readily that there exists $\bigl\{(\nu^k, \phi^k)\bigr\}_{k\in K} \subseteq \Psi^\Delta$ such that (z, ϕ) can be expressed as follows:

$$(z, \phi) = \sum_{k\in K} \lambda_k (\nu^k, \phi^k) \quad \lambda \geq 0 \quad \text{and} \quad \sum_{k \in K} \lambda_k = 1. \tag{4.37}$$

Observe that $(\nu^k, \phi^k) \in \Psi^\Delta$ implies that $\bigl(F(\nu^k), \phi^k, \nu^k, L(\nu^k)\bigr)$ belongs to R. Therefore, we obtain a contradiction by observing that the extreme point (f, ϕ, z, δ) can be expressed as follows:

$$(f, \phi, z, \delta) = \bigl(F(z), \phi, z, L(z)\bigr) = \sum_{k\in K} \lambda_k \bigl(F(\nu^k), \phi^k, \nu^k, L(\nu^k)\bigr),$$

where the first equality is established above and the second equality holds by (4.37) and the linearity of $F(\cdot)$ and $L(\cdot)$. ∎

The continuous auxiliary variables z in the incremental model can be incorporated into the convexification of the outer-function, yielding an MIP relaxation for the graph of $\phi \circ f$. On the other hand, we will show that continuous auxiliary variables z, via the invertible affine transformation Z^{-1} defined as in (4.18), can be related to certain underestimators $u(\cdot)$ of the inner-functions $f(\cdot)$, yielding a tighter MIP relaxation for the graph of $\phi \circ f$ than the typical one obtained from (4.33). Here, the affine transformation Z^{-1} maps a face Δ_H of Δ to a certain face of the polytope Q defined as in (4.16), that is

$$\begin{aligned} Z^{-1}(\Delta_H) &= Z^{-1}\Bigg(\operatorname{conv}\Bigg(\prod_{i=1}^{d}\big\{\zeta_{i\tau(i,t_i(H)-1)},\ldots,\zeta_{i\tau(i,t_i(H))}\big\}\Bigg)\Bigg) \\ &= \operatorname{conv}\Bigg(Z^{-1}\Bigg(\prod_{i=1}^{d}\big\{\zeta_{i\tau(i,t_i(H)-1)},\ldots,\zeta_{i\tau(i,t_i(H))}\big\}\Bigg)\Bigg) \\ &= \operatorname{conv}\Bigg(\prod_{i=1}^{d}\big\{v_{i\tau(i,t_i(H)-1)},\ldots,v_{i\tau(i,t_i(H))}\big\}\Bigg) \\ &=: Q_H, \end{aligned} \tag{4.38}$$

where the first equality holds by the definition of Δ_H, the second equality holds because convexification commutes with affine maps, and the third equality holds because Z_i^{-1} maps $\zeta_{ij} := \sum_{j'=0}^{j} e_{ij'}$ to $v_{ij} := (a_{i0},\ldots,a_{ij-1},a_{ij},\ldots,a_{ij})$, where $e_{ij'}$ is the j' principal vector in the space spanned by variables $(z_{i0},\ldots,z_{in})$. Conversely, Z maps Q_H to Δ_H. The polytope Q_H is indeed a face of the polytope $Q := \prod_{i=1}^{d} Q_i$ because it is defined as the Cartesian product of simplices, each a face of the simplex $Q_i := \operatorname{conv}(v_{i0},\ldots,v_{in})$. As a result, the incremental model (4.34) under the transformation Z model the union of faces $\bigcup_{H\in\mathcal{H}} Q_H$, that is,

$$\big(Z(s),\delta\big) \in (4.34a) \iff Z(s) \in \bigcup_{H\in\mathcal{H}} \Delta_H \iff s \in \bigcup_{H\in\mathcal{H}} Q_H. \tag{4.39}$$

Next, we define a vector of functions $u(\cdot)$ which can be related to faces $\{Q_H\}_{H\in\mathcal{H}}$. Let $u : W \mapsto \mathbb{R}^{d\times(n+1)}$ be a vector of functions defined as follows:

$$u_{i0}(x, s_{\cdot n}) = a_{i0},\ u_{in}(x, s_{\cdot n}) = s_{in},\ \text{and } a_{i0} \le u_{ij}(x, s_{\cdot n}) \le \min\big\{s_{in}, a_{ij}\big\}\ \forall i, \tag{4.40}$$

where W is an outer-approximation of the inner function $f(\cdot)$. We can assume without loss of generality that for all i and $j \neq n$ inequality $u_{ij}(x) \le s_{ij}$ is valid for W; otherwise the outer-approximation W can be strengthened using these inequalities.

Theorem 4.3.1 *Consider a composite function $\phi \circ f : X \mapsto \mathbb{R}$, where $\phi : \mathbb{R}^d \mapsto \mathbb{R}$ is continuous function and $f : \mathbb{R}^m \mapsto \mathbb{R}^d$ is a vector of bounded function over X. Let $\big(u(x, s_{\cdot n}), a\big)$ be a pair satisfying (4.40). If $\operatorname{conv}(\Phi^Q)$ is a polytope and $u(\cdot)$ is polyhedral then we obtain an MIP relaxation for the graph of $\phi \circ f$ defined as follows:*

$$R_{\mathcal{H}} := \left\{ (x, \phi, s, \delta) \;\middle|\; \begin{array}{l} (s, \phi) \in \operatorname{conv}(\Phi^Q),\ \big(Z(s), \delta\big) \in (4.34a) \\ u(x, s_{\cdot n}) \leq s,\ (x, s_{\cdot n}) \in W \end{array} \right\}. \tag{4.41}$$

Proof We first show that

$$\operatorname{proj}_{(x, s_{\cdot n})}(S) = W, \tag{4.42}$$

where $S := \big\{(x, s, \delta) \;\big|\; \big(Z(s), \delta\big) \in (4.34a),\ (x, s_{\cdot n}) \in W,\ s_{ij} = \min\{a_{ij}, s_{in}\}\ \forall i\ \forall j\big\}$. It follows clearly that $\operatorname{proj}_{(x,s_{\cdot n})}(S) \subseteq W$. To show $W \subseteq \operatorname{proj}_{(x,s_{\cdot n})}(S)$, we consider a point $(x, s_{\cdot n}) \in W$, and let H be a subcube from $\mathcal{H}$ so that $s_{\cdot n} \in H$. Then, we show that $s \in Q_H$. This, together with (4.39), implies that there exists a δ so that (s, δ) satisfies (4.34a), therefore showing that $(x, s, \delta) \in S$. Since $s_{in} \in [a_{i0}, \ldots, a_{in}]$, there exist a $j'_i \in \{1, \ldots, n\}$ and $t_i \in \{1, \ldots, l_i\}$ so that $a_{ij'_i-1} \leq f_i(x) \leq a_{ij'_i}$ and $\tau(i, t_i - 1) \leq j'_i - 1 \leq j'_i \leq \tau(i, t_i)$. By definition of s_{ij}, we have $s_{ij} = a_{ij}$ if $j < j'_i$ and $s_{ij} = s_{in}$ otherwise. Consider two adjacent extreme points of Q_i, *i.e.*, $v_{ij'_i-1} = (a_{i0}, \ldots, a_{ij'_i-1}, a_{j'_i-1}, \ldots, a_{ij'_i-1})$ and $v_{ij'_i} = (a_{i0}, \ldots, a_{ij'_i-1}, a_{ij'_i}, \ldots, a_{ij'_i})$. It follows by direct computation that $s_i = \lambda v_{ij'_i-1} + (1-\lambda) v_{ij'_i}$, where $\lambda = \frac{a_{ij'_i} - s_{in}}{a_{ij'_i} - a_{ij'_i-1}}$. Since $\tau(i, t_i - 1) \leq j'_i - 1 \leq j'_i \leq \tau(i, t_i)$, we can conclude $s \in Q_H$ where $H := \prod_{i=1}^d [a_{i\tau(i,t_i-1)}, a_{i\tau(i,t_i)}]$.

Now, the proof is complete by observing that

$$\operatorname{gr}(\phi \circ f) \subseteq \big\{(x, \phi) \;\big|\; (x, s, \delta) \in S,\ \phi = \phi(s_{1n}, \ldots, s_{dn})\big\} \subseteq \operatorname{proj}_{(x,\phi)}(R_{\mathcal{H}}),$$

where the first containment holds by (4.42) and $\operatorname{gr}(f) \subseteq W$, and the second containment holds because $\big\{(x, s, \delta, \phi) \in S \times \mathbb{R} \;\big|\; \phi = \phi(s_{1n, \ldots, s_{dn}})\big\} \subseteq R_{\mathcal{H}}$. ■

An natural following up question is to investigate whether the MIP relaxation for composite function $\phi \circ f$ obtained using Theorem 4.3.1 is tighter than the relaxation defined in (4.33) and the relaxation obtained using Theorem 4.1.2. To this end, we consider functions $\varphi_{\mathcal{H}}$, $\varphi_{\mathcal{H}-}$, and φ defined as follows:

$$\begin{aligned}
\varphi_{\mathcal{H}}(x, s_{\cdot n}) &:= \sup\Bigl\{\phi \Bigm| (x, \phi, s, \delta) \in (4.41)\Bigr\},\\
\varphi_{\mathcal{H}-}(x, s_{\cdot n}) &:= \sup\Bigl\{\phi \Bigm| (s, \phi) \in \mathrm{conv}(\Phi^Q),\ \bigl(Z(s), \delta\bigr) \in (4.34a),\ (x, s_{\cdot}) \in W\Bigr\},\\
\varphi(x, s_{\cdot n}) &:= \sup\Bigl\{\phi \Bigm| (s, \phi) \in \mathrm{conv}(\Phi^P),\ u(x, s_{\cdot n}) \le s,\ (x, s_{\cdot n}) \in W\Bigr\}.
\end{aligned}$$

It follows readily that $\varphi_{\mathcal{H}}(x, s_{\cdot n}) \le \varphi_{\mathcal{H}-}(x, s_{\cdot n})$ for every $(x, s_{\cdot n}) \in W$. In other words, the MIP relaxation in Theorem 4.3.1 is tighter than the standard one obtained from (4.33).

To show that $\varphi_{\mathcal{H}}(x, s_{\cdot n}) \le \varphi(x, s_{\cdot n})$ for every $(x, s_{\cdot n}) \in W$, we need a certain transformation which maps a point in the polytope P to a point in the polytope Q. With each point $u_i \in \mathbb{R}^{n+1}$, we associated a discrete univariate function $\xi(a; u_i) : [a_{i0}, a_{in}] \mapsto \mathbb{R}$ defined as follows:

$$\xi(a; u_i) = \begin{cases} u_{ij} & a = a_{ij} \quad \text{for } j \in \{0, \ldots, n\} \\ -\infty & \text{otherwise.} \end{cases} \tag{4.43}$$

Moreover, let $\hat{\xi} : [a_{i0}, a_{in}] \mapsto \mathbb{R}$ be the piecewise-linear interpolation of $\xi(a; u_i)$ such that $\hat{\xi}(a; u_i) = \xi(a; u_i)$ for $a \in \{a_{i0}, \ldots, a_{in}\}$ and, for all $j \in \{1, \ldots, n\}$, the restriction of $\hat{\xi}(a; u_i)$ to $[a_{ij-1}, a_{ij}]$ is linear. Although a point $u_i \in P_i$ may not lie in Q_i, it has shown by Proposition 2.3.1 that the point $s_i = \bigl(\mathrm{conc}(\xi)(a_{i0}; u_i), \ldots, \mathrm{conc}(\xi)(a_{in}; u_i)\bigr)$ lies in Q_i, where $\mathrm{conc}(\xi)(\cdot; u_i)$ is the concave envelope of $\xi(\cdot; u_i)$ over $[a_{i0}, a_{in}]$.

Proposition 4.3.2 *Assume the same setup as Theorem 4.3.1. Then,* $\varphi_{\mathcal{H}}(x, s_{\cdot n}) - \varphi_{\mathcal{H}-}(x, s_{\cdot n}) \le 0$ *and* $\varphi_{\mathcal{H}}(x, s_{\cdot n}) - \varphi(x, s_{\cdot n}) \le 0$ *for every* $(x, s_{\cdot n}) \in W$.

Proof We start with evaluating $\varphi(\cdot)$ at a point $(\bar{x}, \bar{s}_{\cdot n}) \in W$. Let $\bar{u} := (\bar{u}_1, \ldots, \bar{u}_d)$, where $\bar{u}_i := u_i(\bar{x}, \bar{s}_{\cdot n})$, and let $\bar{s} := (\bar{s}_1, \ldots, \bar{s}_d)$, where each subvector $\bar{s}_i$ is defined as $\bigl(\mathrm{conc}(\xi)(a_{i0}; \bar{u}_i), \ldots, \mathrm{conc}(\xi)(a_{ij}; \bar{u}_i)\bigr)$. Then, we obtain that

$$\varphi(\bar{x}, \bar{s}_{\cdot n}) = \sup\bigl\{\mathrm{conc}_P(\phi)(s) \bigm| \bar{u} \le s,\ s_{\cdot n} = \bar{s}_{\cdot n},\ s \in P\bigr\} = \mathrm{conc}_P(\phi)(\bar{u}) = \mathrm{conc}_Q(\phi)(\bar{s}),$$

where the first equality follows by definition, the second equality holds because it follows from Lemma 2.2.1 that $\mathrm{conc}_P(\phi)(\cdot)$ is non-increasing in s_{ij} for all i and $j \neq n$, and the last equality follows from Corollary 2.3.1.

Next, we evaluate $\varphi_{\mathcal{H}}(\cdot)$ at the point $(\bar{x}, \bar{s}_{\cdot n})$. It follows from (4.39) that $\varphi_{\mathcal{H}}(\bar{x}, \bar{s}_{\cdot n}) = \sup\bigl\{\mathrm{conc}_Q(\phi)(s) \bigm| s \in \mathcal{L}\bigr\}$, where

$$\mathcal{L} := \Bigl\{ s \Bigm| \bar{u} \leq s,\ s_{\cdot n} = \bar{s}_{\cdot n},\ s \in \bigcup_{H \in \mathcal{H}} Q_H \Bigr\}.$$

Let $\bar{H}$ be a subcube of $\mathcal{H}$ so that $\bar{s}_{\cdot n} \in \bar{H}$, and define $u^* := (u^*_1, \ldots, u^*_d)$ so that

$$u^*_{ij} = a_{ij} \text{ for } j \leq \tau(i, t_i(\bar{H}) - 1) \quad u^*_{ij} = \bar{s}_{in} \text{ for } j \geq \tau(i, t_i(\bar{H})) \text{ and } u^*_{ij} = \bar{u}_{ij} \text{ o.w..}$$

Then, we will show that $s^* := (s^*_1, \ldots, s^*_d)$ is the smallest element in $\mathcal{L}$, where

$$s^*_i := \bigl(\mathrm{conc}(\xi)(a_{i0}; u^*_i), \ldots, \mathrm{conc}(\xi)(a_{ij}; u^*_i)\bigr),$$

that is, $s^* \in \mathcal{L}$ and for every $s' \in \mathcal{L}$ we have $s^* \wedge s' = s^*$, where $\wedge$ denotes the component-wise minimum of two vectors. Since $u^* \in P$, it follows from Proposition 2.3.1 that $s^* \in Q$ and $s^* \geq u^*$, and therefore $s^* \in Q_{\bar{H}}$. As $u^* \geq \bar{u}$, we have $s^* \geq \bar{u}$. Moreover, for every $i \in \{1, \ldots, d\}$ we have $s^*_{in} = \mathrm{conc}(\xi)(a_{in}; u^*_i) = \xi(a_{in}; u^*_i) = u^*_{in} = \bar{s}_{in}$. Hence, $s^* \in \mathcal{L}$. Now, for every $s' \in \mathcal{L}$, we have $s^* \wedge s' = s^*$ by observing that for each $i \in \{1, \ldots, d\}$

$$(s'_i \wedge s^*_i)_j = \min\bigl\{\hat{\xi}(a_{ij}; s'_i), \mathrm{conc}(\xi)(a_{ij}; u^*_i)\bigr\} = \mathrm{conc}(\xi)(a_{ij}; u^*_i) = s^*_{ij} \quad \text{for } j \in \{0, \ldots, n\},$$

where the first and third equalities hold by the definition of $\hat{\xi}$ and s^*_i. To see the second equality, we show that $\mathrm{conc}(\xi)(a; u^*_i) \leq \hat{\xi}(a; s'_i)$ for every $a \in [a_{i0}, a_{in}]$. Observe that $s' \in Q_{\bar{H}}$ and $s'_{\cdot n} = \bar{s}_{\cdot n}$ imply that for every $i \in \{1, \ldots, d\}$

$$s'_{ij} = a_{ij} \quad \text{for } j \leq \tau(i, t_i(\bar{H}) - 1) \quad \text{and} \quad s'_{ij} = \bar{s}_{in} \quad \text{for } \tau(i, t_i(\bar{H}) - 1).$$

This, together with $\bar{u} \leq s$, implies that $u^* \leq s'$. Therefore, $\xi(a; u^*_i) \leq \xi(a; s'_i) \leq \hat{\xi}(a; s'_i)$ for every $a \in [a_{i0}, a_{in}]$. Moreover, by Lemma 2.3.4, $s'_i \in Q_i$ implies that $\hat{\xi}(a; s'_i)$ is a concave function. In other words, $\hat{\xi}(a; s'_i)$ is a concave overestimator of $\xi(a; u^*_i)$ over $[a_{i0}, a_{in}]$. Hence, $\mathrm{conc}(\xi)(a; u^*_i) \leq \hat{\xi}(a; s'_i)$.

Now, we can conclude that $\varphi(\bar{x}, \bar{s}_{\cdot n}) = \text{conc}_Q(\phi)(s^*)$ because

$$\varphi_{\mathcal{H}}(\bar{x}, \bar{s}_{\cdot n}) = \sup\{\text{conc}_Q(\phi)(s) \mid s \in \mathcal{L}\}$$

and

$$\text{conc}_Q(\phi)(s^*) \leq \sup\bigl\{\text{conc}_Q(\phi)(s) \bigm| s \in \mathcal{L}\bigr\} \leq \text{conc}_Q(\phi)(s^*),$$

where the first inequality is because $s^* \in \mathcal{L}$, and the second inequality holds because, by Lemma 2.4.3, $\text{conc}_Q(\phi)$ is non-increasing in s_{ij} for all i and $j \neq n$, and, for every point $s' \in \mathcal{L}$ we have $s' \geq s^*$ and $s'_{\cdot n} = s^*_{\cdot n}$.

The proof is complete by observing that $\varphi_{\mathcal{H}}(\bar{x}, \bar{s}_{\cdot n}) - \varphi(\bar{x}, \bar{s}_{\cdot n}) = \text{conc}_Q(\phi)(s^*) - \text{conc}_Q(\phi)(\bar{s}) \geq 0$, where the inequality holds because $s^* \geq \bar{s}$ and, by Lemma 2.4.3, $\text{conc}_Q(\phi)$ is non-increasing in s_{ij} for all i and $j \neq n$. ■

Example 4.3.1 *Consider $x_1^2 x_2^2$ over $[0,3] \times [0,2]$. Let $u : [0,3] \times [0,2] \mapsto \mathbb{R}^{4+2}$ be a vector of functions defined as follows:*

$$u_1(x) := \bigl(0, 2x_1 - 1, 4x_1 - 4, \max\{0, 2x_1 - 1, 4x_1 - 4, 6x_1 - 9\}\bigr)$$
$$u_2(x) := \bigl(0, \max\{0, 2x_2 - 1, 4x_2 - 4\}\bigr)$$

and let $a_1 = (0,5,8,9)$ and $a_2 = (0,4)$. Consider $\varpi(s) := \max\{0,\ 4s_{11} + 5s_{21} - 20,\ 4s_{12} + 8s_{21} - 32,\ 4s_{13} + 9s_{21} - 36\}$ obtained using Corollary 4.1.1, and consider two functions defined as follows:

$$\varphi_{\mathcal{H}}(x) := \min\left\{\phi \;\middle|\; \begin{aligned} & \phi \geq \varpi(s) \\ & 1 \geq \frac{s_{11} - 0}{5 - 0} \geq \delta \geq \frac{s_{12} - s_{11}}{8 - 5} \geq \frac{s_{13} - s_{12}}{9 - 8} \geq 0,\ \ \delta \in \{0,1\} \\ & s \geq u(x),\ \ x \in [0,3] \times [0,2] \end{aligned}\right\}$$

$$\varphi(x) := \min\bigl\{\phi \bigm| \phi \geq \varpi(s),\ \ s \geq u(x),\ \ x \in [0,3] \times [0,2]\bigr\}.$$

Then, we obtain that

$$\begin{aligned}\varphi_{\mathcal{H}}(2.5, 1.5) &= \min\left\{\phi \;\middle|\; \begin{aligned} & \phi \geq \varpi(s),\ \ s_1 \geq (4,6,6),\ \ s_{21} \geq 2,\ \ s_{11} = 5 \\ & 1 \geq \frac{s_{12} - s_{11}}{8 - 5} \geq \frac{s_{13} - s_{12}}{9 - 8} \geq 0 \end{aligned}\right\} \\ &= \varpi\bigl((5,6,6), 2\bigr) = 10. \end{aligned}$$

In constrast, $\varphi(2.5, 1.5) = \varpi\bigl((4,6,6), 2\bigr) = 6$.

Corollary 4.3.1 *Assume the same setup as Theorem 4.3.1. If the outer-function $\phi(\cdot)$ is multilinear, i.e., $\phi(s_{1n},\ldots,s_{dn}) = \sum_{I\in\mathcal{I}} c_I \prod_{i\in I} s_{in}$, where $\mathcal{I}$ is a collection of subsets of $\{1,\ldots,d\}$ so that for $I\in\mathcal{I}$ $|I|\geq 0$, then an explicit formulation for (4.41) is given as follows:*

$$\left\{ (s,\phi,\delta) \;\middle|\; \begin{aligned} & w\geq 0, \quad \phi = \sum_{I\in\mathcal{I}}\Bigl(c_I\sum_{e\in E}\Bigl(\prod_{i\in I} a_{ie_i}\Bigr)w_e\Bigr),\; u(x,s_{\cdot n})\leq s,\; (x,s_{\cdot n})\in W \\ & \bigl(Z(s),\delta\bigr)\in(4.34a), \quad s_i = \sum_{j=0}^{n} v_{ij}\sum_{e\in E: e_i = j} w_e \quad \text{for } i\in\{1,\ldots,d\} \end{aligned} \right\},$$

where v_{ij} is defined as in (4.16).

Proof This result follows from Theorem 4.3.1 and Proposition 4.2.1. ∎

Corollary 4.3.2 *If the outer-function $\phi(\cdot)$ is supermodular and concave-extendable from* $\mathrm{vert}(Q)$ *then*

$$\varphi_{\mathcal{H}}(x,s_{\cdot n}) = \sup\Bigl\{\min_{\pi\in\Omega}\hat{\mathcal{B}}^{\pi}(\psi)\bigl(Z(s)\bigr) \;\Bigm|\; \bigl(Z(s),\delta\bigr)\in(4.34a),\; u(x,s_{\cdot n})\leq s,\; (x,s_{\cdot n})\in W\Bigr\},$$

where $\psi(z) := \phi\bigl(Z^{-1}(z)\bigr)$.

Proof This result follows from Theorem 4.3.1 and Corollary 4.1.4. ∎

We end this subsection with an idea of tightening the MIP relaxation in Theorem 4.3.1. Let $u : W \mapsto \mathbb{R}$ be a vector of functions defined as in (4.40), and let $M := (M_1,\ldots,M_d)$ be a pair of matrices, where M_i is a matrix in $\mathbb{R}^{(n+1)\times l_i}$. For each $i\in\{1,\ldots,d\}$ and for each $t\in\{1,\ldots,l_i\}$, we assume that for every $(x,s_{\cdot})\in W\cap\bigl\{(x,s_{\cdot n}) \bigm| a_{i\tau(i,t-1)}\leq s_{in}\leq a_{i\tau(i,t)}\bigr\}$ entries in the t^{th} column of M_i satisfy the following inequalities:

$$m_{i0t}\leq\cdots\leq m_{int} \quad \text{and} \quad u_{ij}(x,s_{\cdot n})\leq m_{ijt} \quad \text{for all } j\in\{0,\ldots,n\}. \tag{4.44}$$

Without loss of generality, we can assume that for all i, j and t, $m_{ijt}\leq a_{ij}$. In particular, we assume that for all i and t

$$m_{ijt} = a_{ij} \text{ for } j\leq\tau(i,t-1) \quad \text{and} \quad m_{ijt}\leq a_{i\tau(i,t)} \text{ for } j\geq\tau(i,t). \tag{4.45}$$

Proposition 4.3.3 *Assume the same setup as Theorem 4.3.1. Let* $b : [0,1]^{\sum_{i=1}^{d}(l_i-1)} \mapsto \mathbb{R}$ *as a function so that*

$$b_{ij}(\delta) = m_{ij1} + \sum_{t=1}^{l_i-1}(m_{ijt+1} - m_{ijt})\delta_{it} \quad \textit{for all } i \in \{1,\dots,d\}\ j \in \{0,\dots,n\}.$$

If $\phi(\cdot)$ *is supermodular then* $\mathrm{hyp}(\phi\circ f) \subseteq \mathrm{proj}_{(x,\phi)}(R_{\mathcal{H}+})$, *where* $\mathrm{hyp}(\phi\circ f) := \big\{(x,\phi) \bigm| \phi \le (\phi\circ f)(x),\ x \in X\big\}$ *and* $R_{\mathcal{H}+}$ *is defined as follows:*

$$\phi \le \min_{\pi\in\Omega} \mathcal{B}^{\pi}(\phi)\big(s, b(\delta)\big),\ (x, s_{\cdot n}) \in W,\ u(x, s_{\cdot n}) \le s, \tag{4.46a}$$

$$(s_{ij+1} - s_{ij})\big(b_{ij}(\delta) - b_{ij-1}(\delta)\big) \le (s_{ij} - s_{ij-1})\big(b_{ij+1}(\delta) - b_{ij}(\delta)\big) \quad \forall i\ \forall j \ne 0,1,n, \tag{4.46b}$$

$$s_{i0} = b_{i0}(\delta),\ s_{i1} - s_{i0} \le b_{i1}(\delta) - b_{i0}(\delta),\ s_{in-1} \le s_{in} \quad \forall i \in \{1,\dots,d\}, \tag{4.46c}$$

$$\delta_i \in \{0,1\}^{l_i-1},\ \delta_{il_i-1} \le \cdots \le \delta_{i1} \quad \forall i \in \{1,\dots,d\}, \tag{4.46d}$$

where $\mathcal{B}^{\pi}(\phi)(\cdot,\cdot)$ *is defined as in (4.4).*

Proof Consider a point $(x,\phi) \in \mathrm{hyp}(\phi\circ f)$, we will show that there exists a pair (s,δ) such that the point (x,ϕ,s,δ) satisfies (4.46). Let $s_{\cdot n} := f(x)$. Since $s_{\cdot n} \in \prod_{i=1}^{d}[a_{i0}, a_{in}]$, there exists $H \in \mathcal{H}$ so that $s_{\cdot n} \in H$. In other words, $a_{i\tau(i,t_i(H)-1)} \le s_{in} \le a_{i\tau(i,t_i(H))}$. Let δ be a binary vector so that for $i \in \{1,\dots,d\}$

$$\delta_{it} = 1 \quad \text{for } t < t_i(H) \quad \text{and} \quad \delta_{it} = 0 \quad \text{otherwise.}$$

Now, consider a vector s defined as follows:

$$s_{ij} := \min\big\{s_{in}, b_{ij}(\delta)\big\} \qquad \text{for every } i \in \{1,\dots,d\}\ j \in \{0,\dots,n\}.$$

First, observe that $\phi \le \phi(s_{\cdot n}) \le \min_{\pi\in\Omega} \mathcal{B}^{\pi}\big(s, b(\delta)\big)$, where the first equality holds by $s_{\cdot n} = f(x)$, and the second equality follows from Theorem 4.1.1 since the pair $\big(s, b(\delta)\big)$ satisfies (4.3). Moreover, $(x, s_{\cdot n}) \in W$ as W is an outer-approximation of the graph of inner-functions $f(\cdot)$. In addition, for all i and j, we obtain

$$u_{ij}(x, s_{\cdot n}) \le \min\{s_{in}, m_{ijt_i(H)}\} = \min\{s_{in}, b_{ij}(\delta)\} = s_{ij},$$

where the inequality holds due to assumptions in (4.40) and (4.44), and the first equality follows from the assumption in (4.45) and the definition of $b_{ij}(\delta)$, and the last equality holds by the definition of s_{ij}.

Last, we verify that the point (s, δ) satisfies inequalities (4.46b). For every $i \in \{1, \ldots, d\}$, there exists j'_i so that $b_{ij'_i-1}(\delta) \leq s_{in} \leq b_{ij'_i}(\delta)$. Then, for each $i \in \{1, \ldots, d\}$, if $j \neq j'_i - 1$ then it can be verified that the inequality (4.46b) attains equality; otherwise, we obtain

$$\begin{aligned}(s_{ij+1} - s_{ij})\big(b_{ij}(\delta) - b_{ij-1}(\delta)\big) &= \big(s_{in} - b_{ij'_i-1}(\delta)\big)\big(b_{ij'_i-1}(\delta) - b_{ij'_i-2}(\delta)\big) \\ &\leq \big(b_{ij'_i}(\delta) - b_{ij'_i-1}(\delta)\big)\big(b_{ij'_i-1}(\delta) - b_{ij'_i-2}(\delta)\big) \\ &= (s_{ij} - s_{ij-1})\big(b_{ij+1}(\delta) - b_{ij}(\delta)\big),\end{aligned}$$

where the two equalities hold by the definition of s_{ij}, and the inequality holds since $b_{ij}(\delta) \geq b_{ij-1}(\delta)$ and $b_{ij'_i-1}(\delta) \leq s_{in} \leq b_{ij'_i}(\delta)$. Last, it is clearly that the point (s, δ) satisfies inequalities (4.46c). Therefore, we can conclude that $(x, s, \delta, \phi) \in R_{\mathcal{H}+}$. In other words, $\mathrm{hyp}(\phi \circ f) \subseteq \mathrm{proj}_{(x,\phi)}(R_{\mathcal{H}+})$. ■

Remark 4.3.1 *Notice that if $\phi(\cdot)$ is a multilinear function, $\mathcal{B}^\pi\big(\phi\big)(s, b(\delta)\big)$ is also a multilinear function, and each monomial of $\mathcal{B}^\pi\big(\phi\big)(s, b(\delta)\big)$ is in the form of $s_{ij}\prod_{i'\neq i}\delta_{i't_{i'}}$ for some $i \in \{1, \ldots, d\}$, $j \in \{0, \ldots, n\}$, and $(t_1, \ldots, t_{i-1}, t_i, \ldots, t_d) \in \prod_{i'\neq i}\{1, \ldots, l_{i'}\}$. Moreover, each monomial in (4.46b) is in the form of $s_{ij}\delta_{it}$ for some i, j, and t. Therefore, an MIP reformulation for (4.46) is obtained if $s_{ij}\prod_{i'\neq i}\delta_{i't_{i'}}$ and $s_{ij}\delta_{it}$ are linearized using*

$$\max\left\{0, s_{ij} + a_{ij}\Big(\sum_{i'\neq i}\delta_{i't_{i'}} - (d-1)\Big) - a_{ij}\right\} \leq w^{(i,j)}_{t_{-i}} \leq \min\{s_{ij}, a_{ij}\delta_{1t_1}, \cdots, a_{ij}\delta_{dt_d}\},$$

where $t_{-i} := (t_1, \ldots, t_{i-1}, t_i, \ldots, t_d)$, and $\max\{0, s_{ij}+a_{ij}-a_{ij}\} \leq \gamma_{ijt} \leq \min\{s_{ij}, a_{ij}\delta_{it}\}$, respectively.

Example 4.3.2 *Consider $x_1^2 x_2^2$ over $[0,3]\times[0,2]$, and consider underestimators $u : [0,3]\times[0,2] \mapsto \mathbb{R}^{4+2}$ defined as in Example 4.3.1. Let*

$$M_1 := \begin{bmatrix} 0 & 0 \\ 3.5 & 5 \\ 5 & 8 \\ 5 & 9 \end{bmatrix},$$

where, for $j \in \{0,\dots,3\}$, let m_{1j1} and m_{1j2} of M_1 are an upper bound of $u_{1j}(x)$ over $\{x \in [0,3]\times[0,2] \mid 0 \le u_{13}(x) \le 5\}$ and $\{x \in [0,3]\times[0,2] \mid 5 \le u_{13}(x) \le 9\}$, respectively. Define $b_1 : \{0,1\} \mapsto \mathbb{R}$ as follows:

$$b_{10}(\delta) := 0 \quad b_{11}(\delta) := 3.5 + 1.5\delta \quad b_{12}(\delta) := 5 + 3\delta \quad \textit{and} \quad b_{13}(\delta) := 5 + 4\delta,$$

and let $a_{20} = 0$ and $a_{21} = 4$. Consider a function $\varphi_{\mathcal{H}+} : [0,3]\times[0,2] \mapsto \mathbb{R}$ defined as follows:

$$\varphi_{\mathcal{H}+}(x) := \min\left\{ \phi \;\middle|\; \begin{aligned} & \phi \ge 0 \\ & \phi \ge 4s_{11} + (3.5+1.5\delta)s_{21} - 4(3.5+1.5 \\ & \phi \ge 4s_{12} + (5+3\delta)s_{21} - 4(5+3\delta) \\ & \phi \ge 4s_{13} + (5+4\delta)s_{21} - 4(5+4\delta) \\ & (3.5+1.5\delta) \ge s_{11},\; (1.5+1.5\delta)s_{11} \ge (3.5+1.5\delta)(s_{12}-s_{11}) \\ & \delta(s_{12}-s_{11}) \ge (1.5+1.5\delta)(s_{13}-s_{12}),\; s_{13} \ge s_{12} \\ & \delta \in \{0,1\},\; u(x) \le s,\; x \in [0,3]\times[0,2]. \end{aligned} \right\}$$

Then,

$$\begin{aligned} \varphi_{\mathcal{H}+}(2,1) &= \min\left\{ \phi \;\middle|\; \begin{array}{r} \phi \ge \max\{0, 4s_{11} + 3.5s_{21} - 14,\; 4s_{12} + 5s_{21} - 20\} \\ 3.5 \ge s_{11},\; 1.5s_{11} \ge 3.5s_{12} - 3.5s_{11},\; s_{13} \ge s_{12},\; s \ge \big((3,4,4),1\big) \end{array} \right\} \\ &= 1.5. \end{aligned}$$

In contrast, evaluating the function $\varphi_{\mathcal{H}}(\cdot)$ defined as in Example 4.3.1 at the point $(2,1)$, we obtain that $\varphi_{\mathcal{H}}(2,1) = \min\{\phi \mid \phi \ge \varpi(s),\; s_1 \ge (3,4,4),\; s_{21} \ge 1\} = 0$.

4.3.2 MIP representations for discrete composite functions

In this subsection, we consider a discrete set defined as as follows:

$$\mathcal{D} := \Bigl\{(x,\phi) \Bigm| \phi = \phi\bigl(f_1(x_1),\ldots,f_d(x_d)\bigr),\ x \in \prod_{i=1}^{d}\{a_{i0},\ldots,a_{in}\}\Bigr\}, \tag{4.47}$$

where $f_i(x_i)$ is a *univariate* function so that for every $j \in \{0,\ldots,n\}$ we have $f_i^L \leq f_i(a_{ij}) \leq f_i^U$ for some $f_i^L, f_i^U \in \mathbb{R}$. In Section 4.3.1, we lifted the subdivision $\mathcal{H}$ into the collections of faces $\{Q_H\}_{H\in\mathcal{H}}$ of the polytope Q. Here, we will lift the discrete domain $\prod_{i=1}^{d}\{a_{i0},\ldots,a_{in}\}$ into vertices of Q. This construction leads to an ideal MIP formulation for $\mathcal{D}$ which require dn additional continuous variables and $d\lceil\log_2(n+1)\rceil$ additional binary variables. In contrast, a standard ideal formulation from [71] for the discrete set $\mathcal{D}$ requires $d(n+1)$ binary variables and $(n+1)^d$ auxiliary continuous variables. When $\phi(\cdot)$ is a bilinear term, our result yields an ideal formulation for the graph of the product of two univariate discrete functions which requires $2n$ additional continuous variables and $2\lceil\log_2(n+1)\rceil$ additional binary variables. In contrast, the state-of-the-art formulation from [40] is not ideal (see Example 4.3.3), and requires a number of continuous variables linear in n and a number of binary variables logarithmic in n.

We start with lifting the discrete domain into the vertices of the polytope Q, which is defined as in (4.16). Let $\mathrm{id}(x) := \bigl(\mathrm{id}(x_1),\ldots,\mathrm{id}(x_d)\bigr)$, where $\mathrm{id}(x_i)$ denotes the identity function x_i. Consider a vector of special underestimators $s : \prod_{i=1}^{d}\{a_{i0},\ldots,a_{in}\} \mapsto \mathbb{R}^{d(n+1)}$ of inner-functions $\mathrm{id}(x)$ defined as

$$s_{ij}(x) = \min\{a_{ij}, x_i\} \quad \text{for } i \in \{1,\ldots,d\} \text{ and } j \in \{0,\ldots,n\}.$$

Let s denote $s(x)$. Then, it follows readily that for each $i \in \{1,\ldots,d\}$ and $j \in \{0,\ldots,n\}$

$$x_i = a_{ij} \iff s_i = v_{ij} := (a_{i0},\ldots,a_{ij},a_{ij},\ldots,a_{ij}).$$

Recall that $\text{vert}(Q) = \prod_{i=1}^{d}\{v_{i0}, \ldots, v_{in}\}$. Therefore, we obtain an extended formulation of the discrete set $\mathcal{D}$ defined as follows:

$$\bigcup_{s \in \text{vert}(Q)} \Big\{(v, \phi) \;\Big|\; \phi = \vartheta(s)\Big\}, \tag{4.48}$$

where $\vartheta : \text{vert}(Q) \mapsto \mathbb{R}$ so that $\vartheta(s) = \phi\big(f_1(s_{1n}), \ldots, f_d(s_{dn})\big)$. Therefore, it suffices to derive ideal logarithmic MIP formulations for (4.48).

To study MIP formulations for the vertex set $\text{vert}(Q_i)$, we show that $\text{vert}(Q_i)$ can be seen as an invertible affine transformation of $\text{vert}(\Lambda_i)$, where $\Lambda_i = \big\{\lambda_i \in \mathbb{R}^{n+1} \;\big|\; \lambda_i \geq 0,\; \sum_{j=0}^{n} \lambda_{ij} = 1\big\}$. The affine transformation that maps Λ_i to Q_i is $V_i : \mathbb{R}^{n+1} \mapsto \mathbb{R}^{n+1}$, and is defined as

$$s_i = \lambda_{i0} v_{i0} + \cdots + \lambda_{in} v_{in}. \tag{4.49}$$

After the transformation the vertex e_{ij} of Λ_i maps to the vertex v_{ij} of Q_i, where e_{ij} is the j principal vector the space spanned by variables $(\lambda_{i0}, \ldots, \lambda_{in})$. In addition, the inverse V_i^{-1} is defined as

$$\lambda_{ij} = z_{ij} - z_{ij+1} \quad \text{for } j = 0, \ldots, n-1 \qquad \text{and} \qquad \lambda_{in} = z_{in},$$

where $z_i = Z_i(s_i)$ defined in (4.17). Notice that the vertex set $\text{vert}(\Lambda_i) = \Lambda_i \cap \{0,1\}^{n+1}$ is exactly SOS1 constraints [73] over continuous variables $\delta_i \in [0,1]^{n+1}$, which allow at most one of continuous variables δ to be non-zero. MIP formulations for SOS1 constraints have been studied extensively. In particular, [74] introduces ideal MIP formulations for SOS1 constraints that have a number of binary variables logarithmic in the number of continuous variables λ_i. For any bijection $\eta_i : \big\{1, \ldots, 2^{\lceil \log_2(n+1) \rceil}\big\} \mapsto \mathcal{P}\big(\{1, \ldots, \lceil \log_2(n+1) \rceil\}\big)$, an MIP for SOS1 constraints from [74] is given as follows:

$$\begin{aligned}
&\lambda_i \in \Lambda_i,\; \delta_i \in \{0,1\}^{\lceil \log_2(n+1) \rceil},\\
&\sum_{j \notin T_{ik}} \lambda_{ij} \leq \delta_{ik},\; \sum_{j \notin \bar{T}_{ik}} \lambda_{ij} \leq 1 - \delta_{ik} \quad \text{for } k \in \big\{1, \ldots, \lceil \log_2(n+1) \rceil\big\},
\end{aligned}$$

where $T_{ik} := \big\{\eta_i^{-1}(I) \cap \{0, \ldots, n\} \;\big|\; I \subseteq \{1, \ldots, \lceil \log_2(n+1) \rceil\},\; k \notin I\big\}$ and $\bar{T}_{ik} = \{0, \ldots, n\} \setminus T_{ik}$. Now, observe that this formulation coincides with

$$\lambda_i \in \Lambda_i,\; \delta_i \in \{0,1\}^{\lceil \log_2(n+1) \rceil},\; \sum_{j \notin T_{ik}} \lambda_{ij} = \delta_{ik} \quad \text{for } k \in \{1, \ldots, \lceil \log_2(n+1) \rceil\}. \tag{4.50}$$

To see this, we subtract $\sum_{j=0}^{n} \lambda_{ij} = 1$ from each constraint $\sum_{j \notin \bar{T}_{ik}} \lambda_{ik} \leq 1 - \delta_{ik}$ and obtain $\sum_{j \notin T_{ik}} \lambda_{ik} \geq \delta_{ik}$.

Theorem 4.3.2 *An ideal MIP formulation for (4.48) is given by*

$$\Bigl\{(s,\phi,\delta) \Bigm| \mathrm{conv}_Q(\vartheta)(s) \leq \phi \leq \mathrm{conc}_Q(\vartheta)(s),\ \bigl(V_i^{-1}(s_i),\delta_i\bigr) \in (4.50) \text{ for } i \in \{1,\dots,d\}\Bigr\}, \tag{4.51}$$

where $\vartheta : \mathrm{vert}(Q) \mapsto \mathbb{R}$ *so that* $\vartheta(s) = \phi\bigl(f_1(s_{1n}),\dots,f_d(s_{dn})\bigr)$.

Proof Let E denotes the formulation (4.51). To establish that E is a valid formulation for (4.47), it suffices to show that its validity for (4.48). This follows by observing that

$$\mathrm{proj}_{(s,\phi)}(E) = \bigcup_{s \in \mathrm{vert}(Q)} \bigl\{(s,\phi) \bigm| \mathrm{conc}_Q(\vartheta)(s) \leq \phi \leq \mathrm{conc}_Q(\vartheta)(s)\bigr\} = \bigcup_{s \in \mathrm{vert}(Q)} \Bigl\{\bigl(s,\vartheta(s)\bigr)\Bigr\},$$

where the first equality holds because (4.50) is a valid formulation for $\mathrm{vert}(\Lambda_i)$ and V_i is an invertible mapping from $\mathrm{vert}(\Lambda_i)$ to $\mathrm{vert}(Q_i)$, and the second equality holds because for every $s \in \mathrm{vert}(Q)$ we have $\mathrm{conv}_Q(\vartheta)(s) = \vartheta(s) = \mathrm{conc}_Q(\vartheta)(s)$.

Next, we prove that the formulation E is ideal by contradiction. Consider an extreme point (s,ϕ,δ) of the LP relaxation R of E, and assume that δ is not binary. Then, it follows readily that $V_i^{-1}(s_i)$ is not binary, and thus $s \notin \mathrm{vert}(Q)$. Moreover, (s,ϕ) belongs to the convex hull of the graph of $\vartheta(\cdot)$ and $\delta = A(s)$ for some affine map A. Therefore, there exists a convex multiplier $\{\gamma_v\}_{v \in \mathrm{vert}(Q)}$ so that (s,ϕ,δ) is expressible as follows:

$$(s,\phi,\delta) = \sum_{v \in \mathrm{vert}(Q)} \gamma_v \bigl(v, \vartheta(v), A(v)\bigr),$$

Now, observe that each point $\bigl(v,\vartheta(v),A(v)\bigr)$ belongs to R, a contradiction. ∎

It has been shown in Corollary 4.1.4 the concave envelope $\mathrm{conc}_Q(\vartheta)$ can be derived using the technique in Theorem 4.1.1 provided that $\vartheta(\cdot)$ is supermodular when restricted to the vertex set $\mathrm{vert}(Q)$. It is clear that the supermodularity of $\phi(\cdot)$ does not guarantee that the supermodularity of $\vartheta(\cdot)$ over $\mathrm{vert}(Q)$. However, it can be shown that $\vartheta(\cdot)$ is supermodular by switching with a certain σ, that is, $\vartheta(\sigma)(\cdot)$ is supermodular for a vector of permutation $\sigma = (\sigma_1, \dots, \sigma_d)$, where $\vartheta(\sigma) : \mathrm{vert}(Q) \mapsto \mathbb{R}$ so that $\vartheta(\sigma)(s) := \vartheta\big(A^\sigma(s)\big)$ for every $s \in \mathrm{vert}(Q)$, and A^σ is defined as in (4.20).

Lemma 4.3.1 *Let* $\vartheta : \mathrm{vert}(Q) \mapsto \mathbb{R}$ *be a function so that* $\vartheta(s) = \phi\big(f_1(s_{1n}), \dots, f_d(s_{dn})\big)$ *for every* $s \in \mathrm{vert}(Q)$*. For* $i \in \{1, \dots, d\}$*, let* σ_i *be a permutation of* $\{0, \dots, n\}$ *so that*

$$f_i\big(a_{i\sigma_i(0)}\big) \leq \dots \leq f_i\big(a_{i\sigma_i(n)}\big). \tag{4.52}$$

Then, if $\phi(\cdot)$ *is supermodular over* $[f^L, f^U]$ *then function* $\vartheta(\sigma)(s)$ *is supermodular over* $\mathrm{vert}(Q)$*.*

Proof Let $(j_i)_{i=1}^d$ denotes the vertex $(v_{1j_1}, \dots, v_{dj_d})$ of Q, where recall that $v_{ij_i} = (a_{i0}, \dots, a_{ij_i}, a_{ij_i}, \dots, a_{in})$. Consider vertices v' and v'' of Q corresponding to $(j'_i)_{i=1}^d$ and $(j''_i)_{i=1}^d$, respectively. Then, observe that

$$f_i\big(A^\sigma(v' \wedge v'')_{in}\big) = f_i\big(a_{i\sigma_i(j'_i \wedge j''_i)}\big) = f_i\big(a_{i\sigma_i(j'_i)}\big) \wedge f_i\big(a_{i\sigma_i(j''_i)}\big) = f_i\big(A^\sigma(v')_{in}\big) \wedge f_i\big(A^\sigma(v'')_{in}\big), \tag{4.53}$$

where the first and the last equalities hold by the definition of A^σ, and the second equality follows because σ_i is defined such that inequalities in (4.52) are satisfied. Similarly, we obtain

$$f_i\big(A^\sigma(v' \vee v'')_{in}\big) = f_i\big(a_{i\sigma_i(j'_i \vee j''_i)}\big) = f_i\big(a_{i\sigma_i(j'_i)}\big) \vee f_i\big(a_{i\sigma_i(j''_i)}\big) = f_i\big(A^\sigma(v')_{in}\big) \vee f_i\big(A^\sigma(v'')_{in}\big). \tag{4.54}$$

The proof is complete by observing that

$$\begin{aligned}
&\vartheta(\sigma)(v' \vee v'') + \vartheta(\sigma)(v' \wedge v'') \\
&= \phi\Big(f_1\big(A^\sigma(v' \vee v'')_{1n}\big), \dots, f_d\big(A^\sigma(v' \vee v'')_{dn}\big)\Big) + \\
&\qquad \phi\Big(f_1\big(A^\sigma(v' \wedge v'')_{1n}\big), \dots, f_d\big(A^\sigma(v' \wedge v'')_{dn}\big)\Big) \\
&= \phi\Big(f_1\big(A^\sigma(v')_{1n}\big) \vee f_1\big(A^\sigma(v'')_{1n}\big), \dots, f_d\big(A^\sigma(v')_{dn}\big) \vee f_d\big(A^\sigma(v'')_{dn}\big)\Big) + \\
&\qquad \phi\Big(f_1\big(A^\sigma(v')_{1n}\big) \wedge f_1\big(A^\sigma(v'')_{1n}\big), \dots, f_d\big(A^\sigma(v')_{dn}\big) \wedge f_d\big(A^\sigma(v'')_{dn}\big)\Big) \\
&\geq \phi\Big(f_1\big(A^\sigma(v')_{1n}\big), \dots, f_d\big(A^\sigma(v')_{dn}\big)\Big) + \phi\Big(f_1\big(A^\sigma(v'')_{1n}\big), \dots, f_d\big(A^\sigma(v'')_{dn}\big)\Big) \\
&= \vartheta(\sigma)(v') + \vartheta(\sigma)(v''),
\end{aligned}$$

where the first and last equalities follows from the definition of $\vartheta(\sigma)(\cdot)$, the second equality has been established in (4.53) and (4.54), and the inequality holds by the supermodularity of $\phi(\cdot)$ over $[f^L, f^U]$. ■

This lemma, together with Theorem 4.3.2, yields an ideal MIP formulation for the hypograph $\mathrm{hyp}(\vartheta) := \big\{(v,\phi) \bigm| \phi \leq \vartheta(s), s \in \mathrm{vert}(Q)\big\}$, where recall that $\vartheta : \mathrm{vert}(Q) \mapsto \mathbb{R}$ so that $\vartheta(s) = \phi\big(f_1(s_{1n}), \dots, f_d(s_{dn})\big)$ for every $s \in \mathrm{vert}(Q)$.

Proposition 4.3.4 *If $\phi(\cdot)$ is supermodular over $[f^L, f^U]$ then an ideal MIP formulation for the hypograph $\mathrm{hyp}(\vartheta)$ is given by,*

$$\left\{(s,\delta,\phi) \;\middle|\; \phi \leq \min_{\pi \in \Omega} \hat{\mathcal{B}}^\pi\big(\tilde{\vartheta}(\sigma)\big)\Big(Z\big((A^\sigma)^{-1}(s)\big)\Big),\ \big(V_i^{-1}(s_i), \delta_i\big) \in (4.50)\ \forall i\right\},$$

where $\tilde{\vartheta}(\sigma) : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ defined as $\tilde{\vartheta}(\sigma)(z) = \vartheta(\sigma)\big(Z^{-1}(z)\big)$, and for $\psi : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ function $\hat{\mathcal{B}}^\pi(\psi)$ is defined as in (4.14).

Proof By Theorem 4.3.2, it suffices to characterize the envelope $\mathrm{conc}_Q(\vartheta)(s)$. It follows from Lemma 4.3.1 that $\vartheta(\sigma)(\cdot)$ is supermodular over $\mathrm{vert}(Q)$, and therefore, by Corollary 4.1.3, $\min_{\pi\in\Omega} \hat{\mathcal{B}}^\pi\big(\tilde{\vartheta}(\sigma)\big)\big(Z(s)\big)$ describes the envelope $\mathrm{conc}_Q\big(\vartheta(\sigma)\big)(s)$. Hence, $\min_{\pi\in\Omega} \hat{\mathcal{B}}^\pi\big(\tilde{\vartheta}(\sigma)\big)\Big(Z\big((A^\sigma)^{-1}(s)\big)\Big)$ describe envelopes $\mathrm{conc}_Q(\vartheta)(s)$. ■

Now, we specify our study to the case where the function $\phi(\cdot)$ is a bilinear term. Then, our results yield an explicit ideal logarithmic formulation for the following discrete bilinear set

$$D := \big\{(x,\phi) \bigm| \phi = f_1(x_1)f_2(x_2),\ x_i \in \{a_{i0},\ldots,a_{in}\}\ i=1,2\big\},$$

where $f_i(\cdot)$ is a univariate function.

Corollary 4.3.3 *Let $\check{\sigma} := (\check{\sigma}_1, \check{\sigma}_2)$ be a pair of permutations of $\{0,\ldots,n\}$ such that*

$$f_1\big(a_{1\check{\sigma}_1(0)}\big) \leq \cdots \leq f_1\big(a_{1\check{\sigma}_1(n)}\big) \quad \textit{and} \quad f_2\big(a_{2\check{\sigma}_2(0)}\big) \geq \cdots \geq f_2\big(a_{2\check{\sigma}_2(n)}\big),$$

and let $\hat{\sigma} := (\hat{\sigma}_1, \hat{\sigma}_2)$ be a pair of permutations of $\{0,\ldots,n\}$ such that

$$f_i\big(a_{i\hat{\sigma}_i(0)}\big) \leq \cdots \leq f_i\big(a_{i\hat{\sigma}_i(n)}\big).$$

Then, an ideal MIP formulation for D is given by

$$\left\{ (s,\phi,\delta) \;\middle|\; \begin{array}{l} \hat{\mathcal{B}}^\pi\big(\tilde{\vartheta}(\check{\sigma})\big)\Big(Z\big((A^{\check{\sigma}})^{-1}(s)\big)\Big) \leq \phi \leq \hat{\mathcal{B}}^\pi\big(\tilde{\vartheta}(\hat{\sigma})\big)\Big(Z\big((A^{\hat{\sigma}})^{-1}(s)\big)\Big)\ \forall \pi \in \Omega \\ \big(V_i^{-1}(s_1), \delta_i\big) \in (4.50)\ i=1,2 \end{array} \right\},$$

where $\vartheta(s) : \mathrm{vert}(Q) \mapsto \mathbb{R}$ so that $\vartheta(s) = f_1(s_{1n})f_2(s_{2n})$ for every $s \in \mathrm{vert}(Q)$, for any pair of permutations σ, $\tilde{\vartheta}(\sigma) : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ defined as $\tilde{\vartheta}(\sigma)(z) = \vartheta(\sigma)\big(Z^{-1}(z)\big)$, and for $\psi : \mathrm{vert}(\Delta) \mapsto \mathbb{R}$ function $\hat{\mathcal{B}}^\pi(\psi)$ is defined as in (4.14).

Proof By Theorem 4.3.2, it suffices to characterize the convex envelope $\mathrm{conv}_Q(\vartheta)(s)$ and the concave envelope $\mathrm{conc}_Q(\vartheta)(s)$. It follows readily from Proposition 4.3.4 that $\min_{\pi\in\Omega} \hat{\mathcal{B}}^\pi\big(\tilde{\vartheta}(\hat{\sigma})\big)\Big(Z\big((A^{\hat{\sigma}})^{-1}(s)\big)\Big)$ coincides with the concave envelope $\mathrm{conc}_Q(\vartheta)(s)$ since the bilinear term is supermodular over $\mathbb{R}^2$. The convex envelope $\mathrm{conv}_Q(\vartheta)(s)$

can be obtained in a similar way since the submodularity of $\vartheta(\check{\sigma})(\cdot)$ can be established by observing that for vertices v' and v'' of Q corresponding to $(j'_i)_{i=1}^d$ and $(j''_i)_{i=1}^d$,

$$\begin{aligned}
&\vartheta(\check{\sigma})(v' \vee v'') + \vartheta(\check{\sigma})(v' \wedge v'') \\
&= f_1\big(A^{\check{\sigma}}(v' \vee v'')_{1n}\big) \cdot f_2\big(A^{\check{\sigma}}(v' \vee v'')_{2n}\big) + f_1\big(A^{\check{\sigma}}(v' \wedge v'')_{1n}\big) \cdot f_2\big(A^{\check{\sigma}}(v' \wedge v'')_{2n}\big) \\
&= \Big(f_1\big(A^{\check{\sigma}}(v')_{1n}\big) \vee f_1\big(A^{\check{\sigma}}(v'')_{1n}\big)\Big) \cdot \Big(f_2\big(A^{\check{\sigma}}(v')_{2n}\big) \wedge f_2\big(A^{\check{\sigma}}(v'')_{2n}\big)\Big) + \\
&\qquad \Big(f_1\big(A^{\check{\sigma}}(v')_{1n}\big) \wedge f_1\big(A^{\check{\sigma}}(v'')_{1n}\big)\Big) \cdot \Big(f_2\big(A^{\check{\sigma}}(v')_{2n}\big) \vee f_2\big(A^{\check{\sigma}}(v'')_{2n}\big)\Big) \\
&\le f_1\big(A^{\check{\sigma}}(v')_{1n}\big) \cdot f_2\big(A^{\check{\sigma}}(v')_{2n}\big) + f_1\big(A^{\check{\sigma}}(v'')_{1n}\big) \cdot f_2\big(A^{\check{\sigma}}(v'')_{2n}\big) \\
&= \vartheta(\check{\sigma})(v') + \vartheta(\check{\sigma})(v''),
\end{aligned}$$

where the first and the last equalities hold by definition, the second equality can be shown using arguments similar to those in (4.53) and (4.54) and using the definition of $\check{\sigma}$, and the inequality holds because for any two vectors $c', c'' \in \mathbb{R}^2$ we have $\min\{c'_1, c''_1\} \max\{c'_2, c''_2\} + \max\{c'_1, c''_1\} \min\{c'_2, c''_2\} \le c'_1 c'_2 + c''_1 c''_2$. ■

We provide an ideal formulation for the product of two discrete univariate functions, which requires $2n$ continuous variables and $2\lceil \log_2(n+1) \rceil$ binary variables. In contrast, the formulation for the product of two discrete univariate functions from [40] is not ideal and requires $2 + 3(n+1) + \lceil \log_2(n+1) \rceil$ additional continuous variable and $2\lceil \log_2(n+1) \rceil$ binary variables.

Example 4.3.3 *Consider a discrete set* $D := \big\{(x, \phi) \bigm| \phi = x_1^2 x_2^2,\ x_i \in \{0,1,2\}\ i = 1, 2\big\}$*, whose convex hull is given as follows:*

$$\operatorname{conv}(D) = \left\{ (x, \phi) \left| \begin{array}{l} \max\{0, 4x_1 + 4x_2 - 8, 12x_1 + 12x_2 - 32\} \le \phi \le \min\{8x_1, 8x_2\} \\ x \in [0,2]^2 \end{array} \right. \right\}.$$

Formulas (13)-(18) from [40] yield an MIP formulation for D whose projection onto the space of (x, ϕ) variables is given as follows:

$$\begin{aligned}
&\phi \ge \max\{12x_1 + 16x_2 - 40, 4x_1 + 16x_2 - 32, 0, 15x_1 + 12x_2 - 38, 9x_1 + 4x_2 - 18\}, \\
&\phi \le \min\{8x_1, 8x_2\}, \\
&x \in [0,2]^2,
\end{aligned}$$

which differs from the convex hull of D. In other words, the formulation from [40] for D is not sharp, and thus is not ideal.

For $i = 1, 2$, let s_{i2} denotes x_i and consider $Q_i = \operatorname{conv}\big(\{v_{i0}, v_{i1}, v_{i2}\}\big)$, where $v_{i0} = (0,0,0)$, $v_{i1} = (0,1,1)$ and $v_{i2} = (0,1,2)$. Then, the set G is expressible as $\big\{(s,\phi) \bigm| \phi = \vartheta(s),\ s_i \in \operatorname{vert}(Q_i)\ i = 1,2\big\}$, where $\vartheta : \operatorname{vert}(Q_1)\times\operatorname{vert}(Q_2) \mapsto \mathbb{R}$ so that $\vartheta(s) = s_{12}^2 s_{22}^2$. It follows readily that the vertex set $\operatorname{vert}(Q_i)$ as follows:

$$s_{i0} = 0, \quad 1 \geq s_{i1} \geq s_{i2} - s_{i1} \geq 0, \quad 1 - s_{i1} = \delta_{i1}, \quad s_{i1} - (s_{i2} - s_{i1}) = \delta_{i2}, \quad \delta_i \in \{0,1\}^2.$$

Since $\vartheta(\cdot)$ is supermodular over $\operatorname{vert}(Q)$, for a movement vector $\pi = (1,2,1,2)$, formula in (4.14) yields

$$\begin{aligned}
\vartheta(s) &\leq \vartheta(v_{10}, v_{20}) + \big(\vartheta(v_{11}, v_{20}) - \vartheta(v_{10}, v_{20})\big)\frac{s_{11} - s_{10}}{a_{11} - a_{10}} \\
&\qquad + \big(\vartheta(v_{11}, v_{21}) - \vartheta(v_{11}, v_{20})\big)\frac{s_{21} - s_{20}}{a_{21} - a_{20}} + \big(\vartheta(v_{12}, v_{21}) - \vartheta(v_{11}, v_{21})\big)\frac{s_{12} - s_{11}}{a_{12} - a_{11}} \\
&\qquad + \big(\vartheta(v_{12}, v_{22}) - \vartheta(v_{12}, v_{21})\big)\frac{s_{22} - s_{21}}{a_{22} - a_{21}} \\
&= -3s_{11} + 3s_{12} - 11s_{21} + 12s_{22}.
\end{aligned}$$

It turns out that $\operatorname{conc}_Q(\vartheta)(s)$ can be explicitly derived as follows:

$$\min\left\{\begin{aligned} &-11s_{11} + 12s_{12} - 3s_{21} + 3s_{22}, -3s_{11} + 3s_{12} - 11s_{21} + 12s_{22}, -8s_{11} + 12s_{12} \\ &-8s_{21} + 12s_{22}, -2s_{11} + 3s_{12} - 12s_{21} + 12s_{22}, -12s_{11} + 12s_{12} - 2s_{21} + 3s_{22}\end{aligned}\right\}.$$

To derive the convex envelope of ϑ over Q, we consider $\breve{\sigma} = (\breve{\sigma}_1, \breve{\sigma}_2)$, where $\breve{\sigma}_1 = (0,1,2)$ and $\breve{\sigma}_2 = (2,1,0)$. It follows readily that $\vartheta(\breve{\sigma})(v_{1j_1}, v_{2j_2}) = \vartheta(v_{1j_1}, v_{2(2-j_2)})$ for

$(j_1, j_2) \in \{0,1,2\}^2$, *which is submodular over* $\mathrm{vert}(Q_1) \times \mathrm{vert}(Q_2)$. *Let* $\tilde{s} = A^\sigma(s)$, *namely,* $\tilde{s}_1 = s_1$, *and* $\tilde{s}_{20} = 0$, $\tilde{s}_{21} = 1 + s_{21} - s_{22}$ *and* $\tilde{s}_{22} = 2 - s_{22}$.

$$
\begin{aligned}
\vartheta(s) \geq \vartheta(\check{\sigma})(v_{10}, v_{20}) & \\
&+ \big(\vartheta(\check{\sigma})(v_{11}, v_{20}) - \vartheta(\check{\sigma})(v_{10}, v_{20})\big)\frac{s_{11} - s_{10}}{a_{11} - a_{10}} \\
&+ \big(\vartheta(\check{\sigma})(v_{11}, v_{21}) - \vartheta(\check{\sigma})(v_{11}, v_{20})\big)\frac{\tilde{s}_{21} - \tilde{s}_{20}}{a_{21} - a_{20}} \\
&+ \big(\vartheta(\check{\sigma})(v_{12}, v_{21}) - \vartheta(\check{\sigma})(v_{11}, v_{21})\big)\frac{s_{12} - s_{11}}{a_{12} - a_{11}} \\
&+ \big(\vartheta(\check{\sigma})(v_{12}, v_{22}) - \vartheta(\check{\sigma})(v_{12}, v_{21})\big)\frac{\tilde{s}_{22} - \tilde{s}_{21}}{a_{22} - a_{21}} \\
= s_{11} + 3s_{12} + s_{21} + 3s_{22} - 7 &
\end{aligned}
$$

It turns out that $\mathrm{conv}_Q(\vartheta)(s)$ *can be explicitly derived as follows:*

$$
\max\left\{\begin{aligned} 4s_{11} - 2s_{21} + 3s_{22} - 4, s_{11} + 3s_{12} + s_{21} + 3s_{22} - 7, s_{11} + s_{21} - 1 \\ -2s_{11} + 3s_{12} + 4s_{21} - 4, -8s_{11} + 12s_{12} - 8s_{11} + 12s_{22} - 16, 0 \end{aligned}\right\}.
$$

4.3.3 Approximating the concave envelope of special composite functions

In this subsection, we will restrict our attention to composite function $\phi \circ f : [x^L, x^U] \mapsto \mathbb{R}$ such that $(\phi \circ f)(x) = \phi\big(f_1(x_1), \ldots, f_d(x_d)\big)$, where $f_i(x_i)$ is a *univariate* function. We assume that without loss of generality that $x^L = (0, \ldots, 0)$ and $x^U = (1, \ldots, 1)$ since otherwise we treat $f_i'(x_i) := f_i\big((x_i - x_i^L)/(x_i^U - x_i^L)\big)$ as an inner function. We will show that, combined with staircase expansion, Theorem 4.3.2, an idea for constructing ideal formulations for discrete set, is also useful to generate a sequence of polyhedral relaxations of $\phi \circ f$ that converges uniformly to its concave envelope on $[0,1]^d$. We start with the case where each inner-function $f_i(\cdot)$ is a piecewise-linear function.

Corollary 4.3.4 *Let* $\phi \circ f : [0,1]^d \mapsto \mathbb{R}$ *be a composite function so that* $(\phi \circ f)(x) = \phi\big(f_1(x_1), \ldots, f_d(x_d)\big)$, *where* $f_i(\cdot)$ *is a univariate function. If* $f_i(\cdot)$ *is a piecewise linear function characterized by breakpoints* $0 = a_{i0} < \ldots\ldots < a_{in} = 1$, *and the* *outer-function $\phi(\cdot)$ is a multilinear function, an extended formulation for the convex hull of the hypograph of $\phi \circ f$ is given by*

$$\big\{(s,\phi) \;\big|\; \mathrm{conv}_Q(\vartheta)(s) \leq \phi \leq \mathrm{conc}_Q(\vartheta)(s),\;\; s \in Q\big\}, \tag{4.55}$$

where $\vartheta : \mathrm{vert}(Q) \mapsto \mathbb{R}$ so that $\vartheta(s) = \phi\big(f_1(s_{1n}),\ldots,f_d(s_{dn})\big)$.

Proof We first show that $\mathrm{conv}\big(\mathrm{gr}(\phi\circ f)\big) = \mathrm{conv}(\mathcal{D})$, where $\mathcal{D}$ is defined as in (4.47). Clearly, we have $\mathcal{D} \subseteq \mathrm{gr}(\phi \circ f)$, thus implying that $\mathrm{conv}(\mathcal{D}) \subseteq \mathrm{conv}\big(\mathrm{gr}(\phi \circ f)\big)$. The opposite inclusion holds because for every $i \in \{1,\ldots,d\}$ and $j_i \in \{1,\ldots,n\}$ the restriction of $\phi \circ f$ to $\prod_{i=1}^{d}[a_{ij_i-1}, a_{ij_i}]$ is a multilinear function.

Now, let R be the set defined by (4.55). Then, the result holds by observing that

$$\mathrm{conv}(\mathcal{D}) = \mathrm{conv}\Big(\mathrm{proj}_{(s_{\cdot n},\phi)}\big(\mathrm{gr}(\vartheta)\big)\Big) = \mathrm{proj}_{(s_{\cdot n},\phi)}\Big(\mathrm{conv}\big(\mathrm{gr}(\vartheta)\big)\Big) = \mathrm{proj}_{(s_{\cdot n},\phi)}(R),$$

where the first equality has been established in (4.48), the second equality holds because projection commutes with convexification, and the last equality holds because the projection of the LP relaxation of (4.51) onto the space of (s,ϕ) variables coincides with the set R, and, by Theorem 4.3.2, (4.51) is an ideal formulation for $\mathrm{gr}(\vartheta)$. ■

Next, we study the case where the inner-function $f_i(\cdot)$ is not a piecewise-linear function. To construct a sequence of polyhedral relaxations for composite function $\phi \circ f$, we assume that for each $i \in \{1,\ldots,d\}$ there exist a sequence $\big(l_i^{(n)}\big)_{n\in\mathbb{N}}$ of piecewise-linear function mapping from $[0,1]$ to $\mathbb{R}$ and a sequence $\big(\beta_i^{(n)}\big)_{n\in\mathbb{N}}$ of constants such that for every $x_i \in [0,1]$ we have

$$\big|f_i(x_i) - l_i^{(n)}(x_i)\big| \leq \beta_i^{(n)} \quad \text{for } n \in \mathbb{N} \qquad \text{and} \qquad \beta_i^{(n)} \to 0 \quad \text{as } n \to \infty. \tag{4.56}$$

We will assume that the outer-function $\phi(\cdot)$ is Lipschitz continuous. Recall that a function $g : Y \subseteq \mathbb{R}^m \mapsto \mathbb{R}$ is called *Lipschitz continuous* if there exists a real constant $K \geq 0$ such that, for all y_1 and y_2 in Y, $|g(y_1) - g(y_2)| \leq K\|x_1 - x_2\|_2$, and such K is referred to as a *Lipschitz constant* for the function g.

Theorem 4.3.3 *Let $\phi \circ f : [0,1]^d \mapsto \mathbb{R}$ be a composite function so that $(\phi \circ f)(x) = \phi\big(f_1(x_1), \dots, f_d(x_d)\big)$, where $f_i(x)$ is a continuous univariate function. For $i \in \{1, \dots, d\}$, let $\big(l_i^{(n)}, \beta_i^{(n)}\big)_{n \in \mathbb{N}}$ be a sequence of pairs satisfying (4.56). If $\phi(\cdot)$ is Lipschitz continuous with a Lipschitz constant K then the sequence $\big(\varphi_{pw}^{(n)}\big)_{n \in \mathbb{N}}$ of functions converges uniformly to $\mathrm{conc}_{[0,1]^d}(\phi \circ f)$ on $[0,1]^d$, where*

$$\varphi_{pw}^{(n)}(x) := \mathrm{conc}_{[0,1]^d}\big(\phi \circ l^{(n)}\big)(x) + \sum_{i=1}^{d} K\beta_i^{(n)},$$

and $l^{(n)}(x)$ denotes $\big(l_1^{(n)}(x_1), \dots, l_d^{(n)}(x_d)\big)$. Moreover, for each $n \in \mathbb{N}$, the function $\varphi_{pw}^{(n)}(\cdot)$ is a concave overestimator of $(\phi \circ f)(\cdot)$.

Proof It follows clearly that $\varphi_{\mathrm{pw}}^{(n)}(\cdot)$ is a concave function, and that, for every $x \in [0,1]^d$,

$$\begin{aligned}
(\phi \circ f)(x) &= \phi\big(l^{(n)}(x)\big) + \Big(\phi\big(f_1(x_1), l_2^{(n)}(x_2) \dots, l_d^{(n)}(x_d)\big) - \phi\big(l_1^{(n)}(x_1), \dots, l_d^{(n)}(x_d)\big)\Big) \\
&\qquad + \cdots + \Big(\phi\big(f_1(x_1), \dots, f_d(x_d)\big) - \phi\big(f_1(x_1), \dots, f_{d-1}(x_{d-1}), l_d^{(n)}(x_d)\big)\Big) \\
&\le \mathrm{conc}_{[0,1]^d}\big(\phi \circ l^{(n)}\big)(x) + K\left\|\big(f_1(x_1), l_2^{(n)}(x_2), \dots, l_d^{(n)}(x_d)\big) - \big(l_1^{(n)}(x_1), \dots, l_d^{(n)}(x_d)\big)\right\|_2 \\
&\qquad + \cdots + K\left\|\big(f_1(x_1), \dots, f_d(x_d)\big) - \big(f_1(x_1), \dots, f_{d-1}(x_{d-1}), l_d^{(n)}(x_d)\big)\right\|_2 \\
&\le \varphi_{\mathrm{pw}}^{(n)}(x),
\end{aligned}$$

where the first equality holds by staircase expansion, the inequality holds by the Lipschitz continuity of $\phi(\cdot)$, and the last inequality holds by the first requirement in (4.56).

Next, we show that $\big(\varphi_{\mathrm{pw}}^{(n)}(x)\big)_{n \in \mathbb{N}}$ converges uniformly to $\mathrm{conc}_{[0,1]^d}(\phi \circ f)(x)$ on $[0,1]^d$. Let $x \in [0,1]^d$, and let λ be a convex multiply and $(x^t)_{t=1}^m$ be points in $[0,1]^d$ such that $x = \sum_{t=1}^m \lambda_t x^t$. Then, it follows readily that

$$\begin{aligned}
\sum_{t=1}^{m} \lambda_t\big(\phi \circ l^{(n)}\big)(x^t) + \sum_{i=1}^{d} K\beta_i^{(n)} &\le \sum_{t=1}^{m} \lambda_t(\phi \circ f)(x^t) + \sum_{i=1}^{d} K\beta_i^{(n)} + K\big\|f(x) - l^{(n)}(x)\big\|_2 \\
&\le \mathrm{conc}_{[0,1]^d}(\phi \circ f)(x) + \sum_{i=1}^{d} K\beta_i^{(n)} + K\big\|\beta_1^{(n)}, \dots, \beta_d^{(n)}\big\|_2,
\end{aligned}$$

where the first inequality holds by the Lipschitz continuity of $\phi(\cdot)$, and the second inequality holds by the concavity of $\mathrm{conc}_{[0,1]^d}\big(\phi\circ f\big)$ and the first requirement in (4.56). By Proposition 2.31 in [75], for any function $g : Y \mapsto \mathbb{R}$, where Y is a convex subset of $\mathbb{R}^m$, we obtain that $\mathrm{conc}_Y(g)(y) = \sup\big\{\sum_{j=1}^k \alpha_j g(y_j) \;\big|\; y = \sum_{j=1}^k \alpha_j y_j,\ \alpha \geq 0,\ \sum_{j=1}^k \alpha_j = 1,\ y_j \in Y\big\}$. Therefore, we obtain that $\varphi^{(n)}_{\mathrm{pw}}(x) \leq \mathrm{conc}_{[0,1]^d}(\phi\circ f)(x) + \sum_{i=1}^d K\beta_i^{(n)} + K\big\|\beta_1^{(n)},\dots,\beta_d^{(n)}\big\|_2$. Hence, the proof is complete by observing that for every $i \in \{1,\dots,d\}$ we assume in (4.56) that $\beta_i^{(n)}$ is independent of x variables and $\beta_i^{(n)} \to 0$. ∎

4.4 Ordered outer-approximations of inner functions

Consider the graph of the vector of bounded inner-functions, *i.e.*, $\mathrm{gr}(f) := \{(x,u_{\cdot n}) \in \mathbb{R}^{m+d} \mid u_{\cdot n} = f(x),\ x \in X\}$ where $u_{\cdot n}$ denotes $(u_{1n},\dots,u_{dn})$. Let $\kappa := m + d + m'$ for some nonnegative integer m', and let F as a bounded convex set in $\mathbb{R}^\kappa$ so that $\mathrm{gr}(f) \subseteq \mathrm{proj}_{(x,u_{\cdot n})}(F)$ and $\mathrm{int}(F) \neq \emptyset$. For convenience, define

$$K := \left\{\lambda \begin{pmatrix} 1 \\ y \end{pmatrix} \;\middle|\; y \in F,\ \lambda \geq 0\right\},$$

the homogenization of F; K is a convex cone in $\mathbb{R}^{\kappa+1}$ (the additional coordinate is indexed by 0) and $F = \{y \in \mathbb{R}^\kappa \mid \begin{pmatrix} 1 \\ y \end{pmatrix} \in K\}$. Let K^* be the dual cone of K, that is, $K^* = \big\{\alpha \in \mathbb{R}^{\kappa+1} \;\big|\; \langle \alpha, z\rangle \geq 0 \text{ for all } z \in K\big\}$. Since it is assumed that $\mathrm{int}(F) \neq \emptyset$, it follows that $\mathrm{int}(K) \neq \emptyset$, and, thus, K^* is a pointed convex cone, that is, if $\alpha \in K^*$ and $-\alpha \in K^*$ then $\alpha = 0$. The pointed convex cone K^* defines a partial ordering on $\mathbb{R}^{\kappa+1}$ as follows: we say that $\alpha^1 \preceq_{K^*} \alpha^2$ provided that $\alpha^2 - \alpha^1 \in K^*$. In the following, we derive a pair $\big(u(\cdot), a(\cdot)\big)$ satisfying the requirements in (4.3) by solving optimization problems over the pointed convex cone K^*.

Proposition 4.4.1 *Let F be a bounded convex set in $\mathbb{R}^\kappa$ such that* $\mathrm{int}(F) \neq \emptyset$ *and* $\mathrm{gr}(f) \subseteq \mathrm{proj}_{(x,f)}(F)$. *For $i \in \{1,\dots,d\}$, let $\big(\alpha^{i,j}\big)_{j=0}^n$ be a chain in $(\mathbb{R}^{\kappa+1}, \preceq_{K^*})$ so that $\alpha^{i,0} \preceq_{K^*} e_{u_{in}} \preceq_{K^*} \alpha^{i,n}$, where $e_{u_{in}}$ is the principle vector in $\mathbb{R}^{\kappa+1}$ corresponding* *to variable* u_{in}*. For any given* $\bar{y} \in F$ *and for* $i \in \{1, \ldots, d\}$*, define* $\left(\beta^{i,j}\right)_{j=0}^{n}$ *so that* $\beta^{i,n} = u_{in}$ *and, for* $j \in \{0, \ldots, n-1\}$

$$\beta^{i,j} \in \arg\max\Big\{\big\langle \beta, (1, \bar{y})\big\rangle \;\Big|\; \beta \preceq_{K^*} e_{u_{in}},\ \beta \preceq_{K^*} \alpha^{i,j}\Big\}. \tag{4.57}$$

Let $a_{ij}(y) = \big\langle \alpha^{i,j}, (1, y)\big\rangle$ *and* $u_{ij}(y) = \big\langle \beta^{i,j}, (1, y)\big\rangle$*. Then, the pair* $\big(u(y), a(y)\big)$ *satisfies requirement in (4.3).*

REFERENCES

REFERENCES

[1] M. Tawarmalani and N. V. Sahinidis, "A polyhedral branch-and-cut approach to global optimization," *Mathematical Programming*, vol. 103, no. 2, pp. 225–249, 2005.

[2] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wächter, "Branching and bounds tighteningtechniques for non-convex minlp," *Optimization Methods & Software*, vol. 24, no. 4-5, pp. 597–634, 2009.

[3] R. Misener and C. A. Floudas, "Antigone: Algorithms for continuous / integer global optimization of nonlinear equations," *J. Global Optimization*, vol. 59, no. 2-3, pp. 503–526, 2014.

[4] A. Gleixner, L. Eifler, T. Gally, G. Gamrath, P. Gemander, R. L. Gottwald, G. Hendel, C. Hojny, T. Koch, M. Miltenberger, B. Müller, M. E. Pfetsch, C. Puchert, D. Rehfeldt, F. Schlösser, F. Serrano, Y. Shinano, J. M. Viernickel, S. Vigerske, D. Weninger, J. T. Witt, and J. Witzig, "The scip optimization suite 5.0," ZIB, Takustr.7, 14195 Berlin, Tech. Rep. 17-61, 2017.

[5] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part i convex underestimating problems," *Mathematical programming*, vol. 10, no. 1, pp. 147–175, 1976.

[6] R. Horst and H. Tuy, *Global optimization - deterministic approaches (3. Aufl.)*. Springer, 1996.

[7] J. Luedtke, M. Namazifar, and J. Linderoth, "Some results on the strength of relaxations of multilinear functions," *Mathematical programming*, vol. 136, no. 2, pp. 325–351, 2012.

[8] N. Boland, S. S. Dey, T. Kalinowski, M. Molinaro, and F. Rigterink, "Bounding the gap between the mccormick relaxation and the convex hull for bilinear functions," *Mathematical Programming*, pp. 1–13, 2015.

[9] A. D. Rikun, "A convex envelope formula for multilinear functions," *Journal of Global Optimization*, vol. 10, no. 4, pp. 425–437, 1997.

[10] H. D. Sherali, "Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets," *Acta mathematica vietnamica*, vol. 22, no. 1, pp. 245–270, 1997.

[11] H. P. Benson, "Concave envelopes of monomial functions over rectangles," *Naval Research Logistics (NRL)*, vol. 51, no. 4, pp. 467–476, 2004.

[12] C. A. Meyer and C. A. Floudas, "Trilinear monomials with mixed sign domains: Facets of the convex and concave envelopes," *Journal of Global Optimization*, vol. 29, no. 2, pp. 125–155, 2004.

[13] S. Cafieri, J. Lee, and L. Liberti, "On convex relaxations of quadrilinear terms," *Journal of Global Optimization*, vol. 47, no. 4, pp. 661–685, 2010.

[14] M. Tawarmalani and N. V. Sahinidis, "Semidefinite relaxations of fractional programs via novel convexification techniques," *Journal of Global Optimization*, vol. 20, no. 2, pp. 133–154, 2001.

[15] M. Jach, D. Michaels, and R. Weismantel, "The convex envelope of (n–1)-convex functions," *SIAM Journal on Optimization*, vol. 19, no. 3, pp. 1451–1466, 2008.

[16] M. Tawarmalani, J.-P. P. Richard, and C. Xiong, "Explicit convex and concave envelopes through polyhedral subdivisions," *Mathematical Programming*, pp. 1–47, 2013.

[17] A. Khajavirad and N. V. Sahinidis, "Convex envelopes of products of convex and component-wise concave functions," *Journal of global optimization*, vol. 52, no. 3, pp. 391–409, 2012.

[18] ——, "Convex envelopes generated from finitely many compact convex sets," *Mathematical Programming*, vol. 137, no. 1-2, pp. 371–408, 2013.

[19] M. Padberg, "The boolean quadric polytope: some characteristics, facets and relatives," *Mathematical programming*, vol. 45, no. 1, pp. 139–172, 1989.

[20] X. Bao, A. Khajavirad, N. V. Sahinidis, and M. Tawarmalani, "Global optimization of nonconvex problems with multilinear intermediates," *Mathematical Programming Computation*, vol. 7, no. 1, pp. 1–37, 2015.

[21] R. Misener, J. B. Smadbeck, and C. A. Floudas, "Dynamically generated cutting planes for mixed-integer quadratically constrained quadratic programs and their incorporation into glomiqo 2," *Optimization Methods and Software*, vol. 30, no. 1, pp. 215–249, 2015.

[22] A. Del Pia and A. Khajavirad, "A polyhedral study of binary polynomial programs," *Mathematics of Operations Research*, vol. 42, no. 2, pp. 389–410, 2016.

[23] Y. Crama and E. Rodríguez-Heck, "A class of valid inequalities for multilinear 0–1 optimization problems," *Discrete Optimization*, vol. 25, pp. 28–47, 2017.

[24] A. Del Pia and A. Khajavirad, "The multilinear polytope for acyclic hypergraphs," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1049–1076, 2018.

[25] M. Tawarmalani, J.-P. P. Richard, and K. Chung, "Strong valid inequalities for orthogonal disjunctions and bilinear covering sets," *Mathematical Programming*, vol. 124, no. 1-2, pp. 481–512, 2010.

[26] O. Günlük and J. Linderoth, "Perspective reformulations of mixed integer nonlinear programs with indicator variables," *Mathematical programming*, vol. 124, no. 1-2, pp. 183–205, 2010.

[27] K. M. Anstreicher and S. Burer, "Computable representations for convex hulls of low-dimensional quadratic forms," *Mathematical programming*, vol. 124, no. 1-2, pp. 33–43, 2010.

[28] F. Kılınç-Karzan and S. Yıldız, "Two-term disjunctions on the second-order cone," *Mathematical Programming*, vol. 154, no. 1-2, pp. 463–491, 2015.

[29] S. Modaresi and J. P. Vielma, "Convex hull of two quadratic or a conic quadratic and a quadratic inequality," *Mathematical Programming*, vol. 164, no. 1-2, pp. 383–409, 2017.

[30] S. Burer and F. Kılınç-Karzan, "How to convexify the intersection of a second order cone and a nonconvex quadratic," *Mathematical Programming*, vol. 162, no. 1-2, pp. 393–429, 2017.

[31] T. T. Nguyen, J.-P. P. Richard, and M. Tawarmalani, "Deriving convex hulls through lifting and projection," *Mathematical Programming*, pp. 1–39, 2017.

[32] H. D. Sherali and W. P. Adams, "A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems," *SIAM Journal on Discrete Mathematics*, vol. 3, no. 3, pp. 411–430, 1990.

[33] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Mathematical Programming*, vol. 96, no. 2, pp. 293–320, 2003.

[34] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, 2001.

[35] G. B. Dantzig, "On the significance of solving linear programming problems with some integer variables," *Econometrica, Journal of the Econometric Society*, pp. 30–44, 1960.

[36] R. Misener and C. A. Floudas, "Global optimization of mixed-integer quadratically-constrained quadratic programs (miqcqp) through piecewise-linear and edge-concave relaxations," *Mathematical Programming*, vol. 136, no. 1, pp. 155–182, 2012.

[37] J. Huchette and J. P. Vielma, "A combinatorial approach for small and strong formulations of disjunctive constraints," *arXiv preprint arXiv:1607.04803*, 2016.

[38] C. Villani, *Optimal transport, old and new.* Berlin: Springer, 2008.

[39] H. D. Sherali and W. P. Adams, *A reformulation-linearization technique for solving discrete and continuous nonconvex problems.* Springer Science & Business Media, 2013, vol. 31.

[40] W. P. Adams and S. M. Henry, "Base-2 expansions for linearizing products of functions of discrete variables," *Operations research*, vol. 60, no. 6, pp. 1477–1490, 2012.

[41] M. Tawarmalani and N. V. Sahinidis, "Global optimization of mixed-integer nonlinear programs: A theoretical and computational study," *Mathematical programming*, vol. 99, no. 3, pp. 563–591, 2004.

[42] T. He and M. Tawarmalani, "Tractable relaxations of composite functions with supermodular outer function," 2018, working paper.

[43] J. Kim, M. Tawarmalani, and J.-P. P. Richard, "Convexification of permutation-invariant sets," 2018, working paper.

[44] A. M. Gleixner, T. Berthold, B. Müller, and S. Weltge, "Three enhancements for optimization-based bound tightening," *Journal of Global Optimization*, vol. 67, pp. 731–757, 2017.

[45] J. Najman and A. Mitsos, "Tighter mccormick relaxations through subgradient propagation," *arXiv preprint arXiv:1710.09188*, 2017.

[46] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of convex analysis.* Springer Science & Business Media, 2012.

[47] M. Tawarmalani and N. V. Sahinidis, "Convex extensions and envelopes of lower semi-continuous functions," *Mathematical Programming*, vol. 93, no. 2, pp. 247–263, 2002.

[48] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization.* Springer Science & Business Media, 2012, vol. 2.

[49] T. M. Chan, "Optimal output-sensitive convex hull algorithms in two and three dimensions," *Discrete & Computational Geometry*, vol. 16, no. 4, pp. 361–368, 1996.

[50] R. L. Graham, "An efficient algorithm for determining the convex hull of a finite planar set," *Information processing letters*, vol. 1, no. 4, pp. 132–133, 1972.

[51] M. Tawarmalani, "Inclusion certificates and simultaneous convexification of functions," *Mathematical Programming*, 2010.

[52] J. De Loera, J. Rambau, and F. Santos, *Triangulations: Structures for Algorithms and Applications*, ser. Algorithms and Computation in Mathematics. Springer Berlin Heidelberg, 2010. [Online]. Available: https://books.google.com/books?id=SxY1Xrr12DwC

[53] L. Lovász, "Submodular functions and convexity," in *Mathematical Programming The State of the Art.* Springer, 1983, pp. 235–257.

[54] D. M. Topkis, *Supermodularity and complementarity.* Princeton university press, 2011.

[55] Y. Crama, "Recognition problems for special classes of polynomials in 0–1 variables," *Mathematical Programming*, vol. 44, no. 1, pp. 139–155, 1989.

[56] T. He and M. Tawarmalani, "On relaxing composite functions exploiting inner function structure: Polynomially equivalent formulations," 2018, working paper.

[57] D. E. Knuth, *The art of computer programming.* Pearson Education, 1997, vol. 3.

[58] F. Harary *et al.*, "On the notion of balance of a signed graph." *The Michigan Mathematical Journal*, vol. 2, no. 2, pp. 143–146, 1953.

[59] S. T. Rachev and L. Rüschendorf, *Mass Transportation Problems: Volume I: Theory.* Springer Science & Business Media, 1998, vol. 1.

[60] A. H. Tchen, "Inequalities for distributions with given marginals," *The Annals of Probability*, pp. 814–827, 1980.

[61] G. Lorentz, "An inequality for rearrangements," *The American Mathematical Monthly*, vol. 60, no. 3, pp. 176–179, 1953.

[62] P. Billingsley, *Probability and measure.* John Wiley & Sons, 2008.

[63] R. T. Rockafellar and J. O. Royset, "Random variables, monotone relations, and convex analysis," *Mathematical Programming*, vol. 148, no. 1-2, pp. 297–331, 2014.

[64] R. T. Rockafellar, *Convex analysis.* Princeton university press, 2015.

[65] W. Höffding, "Masstabinvariante korrelationstheorie," *Schriften des Mathematischen Instituts und Instituts fur Angewandte Mathematik der Universitat Berlin*, vol. 5, pp. 181–233, 1940.

[66] M. Fréchet, "Sur les tableaux de corrélation dont les marges sont données," *Ann. Univ. Lyon, 3ˆ e serie, Sciences, Sect. A*, vol. 14, pp. 53–77, 1951.

[67] M. Shaked and J. G. Shanthikumar, *Stochastic orders.* Springer Science & Business Media, 2007.

[68] K. Fan and G. Lorentz, "An integral inequality," *The American Mathematical Monthly*, vol. 61, no. 9, pp. 626–631, 1954.

[69] H. D. Sherali and C. H. Tuncbilek, "A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique," *Journal of Global Optimization*, vol. 2, no. 1, pp. 101–112, 1992.

[70] E. Balas, "Disjunctive programming: Properties of the convex hull of feasible points," *Discrete Applied Mathematics*, vol. 89, no. 1-3, pp. 3–44, 1998.

[71] R. G. Jeroslow and J. K. Lowe, "Modelling with integer variables," in *Mathematical Programming at Oberwolfach II.* Springer, 1984, pp. 167–184.

[72] H. M. Markowitz and A. S. Manne, "On the solution of discrete programming problems," *Econometrica: journal of the Econometric Society*, pp. 84–110, 1957.

[73] E. M. L. Beale and J. A. Tomlin, "Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables," *OR*, vol. 69, no. 447-454, p. 99, 1970.

[74] J. P. Vielma and G. L. Nemhauser, "Modeling disjunctive constraints with a logarithmic number of binary variables and constraints," *Mathematical Programming*, vol. 128, no. 1-2, pp. 49–72, 2011.

[75] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis.* Springer Science & Business Media, 2009, vol. 317.

VITA

VITA

Taotao He graduated from Jinan University, China in May 2008, with a B.B.A in Management Science. He received his MBA from Purdue University in May 2013, and then pursued his Ph.D. in Management. His research interests are: (1) Building strong and tractable relaxations for mixed-integer nonlinear programs; (2) Developing theoretical and computational tools to solve practical mixed-integer nonlinear programs.