Joint Chance-Constrained Programs and the Intersection of Mixing Sets through a Submodularity Lens

Fatma Kılınç-Karzan* Simge Küçükyavuz† Dabeen Lee‡

October 3, 2019

Abstract

A particularly important substructure in modeling joint linear chance-constrained programs with random right-hand sides and finite sample space is the intersection of mixing sets with common binary variables (and possibly a knapsack constraint). In this paper, we first revisit basic mixing sets by establishing a strong and previously unrecognized connection to submodularity. In particular, we show that mixing inequalities with binary variables are nothing but the polymatroid inequalities associated with a specific submodular function. This submodularity viewpoint enables us to unify and extend existing results on valid inequalities and convex hulls of the intersection of multiple mixing sets with common binary variables. Then, we study such intersections under an additional linking constraint lower bounding a linear function of the continuous variables. This is motivated from the desire to exploit the information encoded in the knapsack constraint arising in joint linear CCPs via the quantile cuts. We propose a new class of valid inequalities and characterize when this new class along with the mixing inequalities are sufficient to describe the convex hull.

1 Introduction

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a joint linear chance-constrained program (CCP) with right-hand side uncertainty is an optimization problem of the following form:

$$\begin{aligned} \min \quad & h^\top x && \text{(1a)}\\ \text{s.t.} \quad & \mathbb{P}\left[Ax \geq b(\omega)\right] \geq 1-\epsilon && \text{(1b)}\\ & x \in \mathcal{X} \subseteq \mathbb{R}^m, && \text{(1c)} \end{aligned}$$

where $\mathcal{X} \subseteq \mathbb{R}^m$ is a domain for the decision variables x, $\epsilon \in (0,1)$ is a risk level, $b(\omega) \in \mathbb{R}^k$ is the random right-hand side vector that depends on the random variable $\omega \in \Omega$, and A, h are matrices of appropriate dimension.

*Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA, fkilinc@andrew.cmu.edu

†Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208, USA, simge@northwestern.edu

‡Discrete Mathematics Group, Institute for Basic Science (IBS), Daejeon 34126, Republic of Korea, dabeenl@ibs.re.kr

For $k = 1$ (resp., $k > 1$), inequality (1b) is referred to as an individual (resp., joint) chance constraint. Here, we seek to find a solution $x \in \mathcal{X}$ satisfying the chance constraint (1b), enforcing that $Ax \geq b(\omega)$ holds with probability at least the given confidence level $1 - \epsilon$, while minimizing the objective (1a). In the case of continuous distributions governing the uncertainty, i.e., when Ω is continuous, a classical technique is to use the Sample Average Approximation (SAA) to approximate Ω via a set of sample scenarios $\omega^1, \ldots, \omega^n$ and reduce the problem to the case with a finite-sample distribution; we refer the interested reader to [7, 8, 22] for further details of SAA for CCPs.

Joint chance constraints are used to model risk-averse decision-making problems in various applications, such as supply chain logistics [16, 18, 25, 37], chemical processes [13, 14], water quality management [31], and energy [32]. Problems with joint chance constraints have been extensively studied (see [27] for background and an extensive list of references) and they are known to be notoriously challenging because the resulting feasible region is nonconvex even if all other constraints $x \in \mathcal{X}$ and the restrictions inside the chance constraints are convex. Consequently, the classical techniques to model CCPs with discrete distributions rely on converting them into equivalent mixed-integer programs with binary variables and big-M constraints.

In this paper, we consider joint linear CCPs with random right-hand sides under the finite sample space assumption. In particular, we assume that $\Omega = \{\omega^1, \ldots, \omega^n\}$ for some integer $n \geq 1$ and that $\mathbb{P}\left[\omega = \omega^i\right] = p_i$ for $i \in [n]$ for some $p_1, \ldots, p_n \geq 0$ with $\sum_{i \in [n]} p_i = 1$, where for any positive integer n, we define $[n]$ to be the set $\{1, \ldots, n\}$. In this setting, Luedtke et al. [23], Ruszczyński [29] observed that the joint linear CCP, defined by (1), can be reformulated as a mixed-integer linear program as follows:

$$
\begin{aligned}
\min \quad & h^\top x && \text{(2a)}\\
\text{s.t.} \quad & x \in \mathcal{X} \subseteq \mathbb{R}^m, \qquad Ax = b + y, && \text{(2b)}\\
& y_j \geq w_{ij}(1 - z_i), \quad \forall i \in [n],\ \forall j \in [k], && \text{(2c)}\\
& \sum_{i \in [n]} p_i z_i \leq \epsilon, && \text{(2d)}\\
& y \in \mathbb{R}^k_+,\ z \in \{0,1\}^n, && \text{(2e)}
\end{aligned}
$$

where $b \in \mathbb{R}^k$ is some vector satisfying $b(\omega^i) \geq b$ for all i and $w_i = (w_{i1}, \ldots, w_{ik})^\top$ denotes $b(\omega^i) - b$. Note that by definition of b, it follows that the data vector w_i is nonnegative for all $i \in [n]$. Observe that $Ax \geq b$ are implicit inequalities, due to the chance constraint (1b) with $1 - \epsilon > 0$. Here, z_i is introduced as an indicator variable to model the event $Ax \geq b(\omega^i)$. More precisely, when $z_i = 0$, the constraints (2c) enforce that $y \geq w_i$ holds and thus $Ax \geq b(\omega^i)$ is satisfied. On the other hand, when $z_i = 1$, it follows that $y_j \geq 0$ and $Ax \geq b$, which is satisfied by default. Therefore, constraints (2c) are referred to as big-M constraints. Finally, (2d) enforces that the probability of $Ax \geq b(\omega^i)$ being violated is at most ϵ.

The size of the deterministic equivalent formulation of the joint CCP given by (2) grows linearly with the number of scenarios. Unfortunately, such a reformulation based on big-M constraints comes with the disadvantage that the corresponding relaxations obtained by relaxing the binary variables into continuous are weak. Thus, in order to achieve effectiveness in practical implementation, these reformulations must be strengthened with additional valid inequalities.

A particularly important and widely applicable class of valid inequalities that strengthen the big-M reformulations of CCPs rely on a critical specific substructure in the formulation (2), called a *mixing set with binary variables*; see e.g., Luedtke et al. [23] and Küçükyavuz [15]. Formally, given a vector $w = \{w_i\} \in \mathbb{R}^n_+$, a mixing set with binary variables is defined as follows:

$$\mathrm{MIX}_j := \{(y_j, z) \in \mathbb{R}_+ \times \{0,1\}^n : \ y_j + w_i z_i \geq w_i \ \forall i \in [n]\} ;$$

hence the set defined by (2c) and (2e), i.e.,

$$\left\{(y, z) \in \mathbb{R}^k_+ \times \{0,1\}^n : \ y_j + w_i z_i \geq w_i \ \forall i \in [n], \ \forall j \in [k]\right\} ,$$

is nothing but a joint mixing set that shares common binary variables z, but independent continuous variables $y_j, j \in [k]$. Also, it is worthwhile to note that the constraint (2d) is a knapsack constraint. Therefore, the formulation (2) can be strengthened by the inclusion of valid inequalities originating from the set defined by (2c)–(2e).

The term *mixing set* is originally coined by Günlük and Pochet [12] for the sets of the form

$$\mathrm{GMIX} := \{(y, z) \in \mathbb{R}_+ \times \mathbb{Z}^n : \ y + u z_i \geq q_i \ \forall i \in [n]\} ,$$

where the parameters are $u \in \mathbb{R}_+$ and $q = (q_1, \ldots, q_n)^\top \in \mathbb{R}^n$. Such sets GMIX with general integer variables have applications in lot sizing and capacitated facility location problems; see e.g., [9, 10, 12, 24, 39] (see also [33] for a survey of the area). For mixing sets with general integer variables such as GMIX defined above, Günlük and Pochet [12] introduced the so-called *mixing inequalities*—an exponential family of linear inequalities that admits an efficient separation oracle—and showed that this class of inequalities are sufficient to describe the associated convex hull of the sets GMIX. In fact, prior to [12], in the context of lot-sizing problems, Pochet and Wolsey [26, Theorem 18] obtained the same result, albeit without using the naming convention of mixing sets/inequalities. Furthermore, the equivalence of MIX_j and GMIX under the additional domain restrictions $z \in \{0,1\}^n$ and the assumption $u \geq \max_i q_i$ is immediate. The appearance of mixing sets with binary variables dates back to the work of Atamtürk et al. [5] on vertex covering. Essentially, it was shown in [5] that the intersection of several sets of the form MIX_j with common binary variables z but separate continuous variables $y_j, j \in [k]$ can be characterized by the intersection of the corresponding *star inequalities*; see [5, Theorem 3]. Furthermore, it is well-known [23] that mixing inequalities for MIX_j are equivalent to the star inequalities introduced in [5]. We will give a formal definition of mixing (star) inequalities for mixing sets with binary variables in Section 3.

Due to the importance of their use in joint CCPs, the mixing (with knapsack) substructure (2c)–(2e) present in the reformulations of joint CCPs has received a lot of attention in the more recent literature.

- For general k, i.e., when the number of linear constraints inside the chance constraint is more than one, Atamtürk et al. [5] proved that the convex hull of a joint mixing set of the form (2c) and (2e) can be described by applying the mixing inequalities.

- For $k = 1$, Luedtke et al. [23], Küçükyavuz [15], and Abdi and Fukasawa [1] suggested valid inequalities for a single mixing set subject to the knapsack constraint (2d).
- For general k, Küçükyavuz [15] and Zhao et al. [40] proposed valid inequalities for a joint mixing set with a knapsack constraint.

Luedtke et al. [23] showed that the problem is NP-hard for $k > 1$ even when the restrictions inside the chance constraints are linear and each scenario has equal probability, in which case the knapsack constraint (2d) becomes a cardinality constraint. However, Küçükyavuz [15] argued that the problem for $k = 1$ under equiprobable scenarios is polynomial-time solvable and gave a compact and tight extended formulation based on disjunctive programming. Note that while not explicitly stated in [15], when $k = 1$ the polynomial-time solvability argument extends for the unequal probability case.

Many of these prior works aim to convexify a (joint) mixing set with a knapsack constraint directly. In contrast, in our paper we exploit the knapsack structure through an indirect approach based on *quantile inequalities*. Given $c \in \mathbb{R}^k_+$ and $\delta > 0$, the $(1-\delta)$-*quantile for* $c^\top y$ is defined as

$$q_{c,\delta} := \min \left\{ c^\top y : \sum_{i \in [n]} p_i z_i \leq \delta,\ (y, z) \text{ satisfies (2c), (2e)} \right\},$$

and the inequality $c^\top y \geq q_{c,\delta}$ is called a $(1-\delta)$-*quantile cut*. By definition, a $(1-\epsilon)$-quantile cut is valid for the solutions satisfying (2c)–(2e). The quantile cuts have been studied in [2, 17, 21, 28, 30, 35], and their computational effectiveness has been observed in practice. As opposed to mixing sets and associated mixing inequalities, the quantile cuts link many continuous variables together; it is plausible to conjecture that this linking of the continuous variables is the one of the main sources of their effectiveness in practice.

In this paper we study a generalization of the mixing sets as follows: given integers $n, k \geq 1$, a matrix $W = \{w_{ij}\} \in \mathbb{R}^{n \times k}_+$, a vector $\ell \in \mathbb{R}^k_+$ and a nonnegative number $\varepsilon \geq 0$, we consider the set defined by

$$\begin{aligned} & y_j + w_{ij} z_i \geq w_{ij}, && \forall i \in [n],\ \forall j \in [k], && \text{(3a)} \\ & y_j \geq \ell_j, && \forall j \in [k], && \text{(3b)} \\ & y_1 + \cdots + y_k \geq \varepsilon + \sum_{j \in [k]} \ell_j, && && \text{(3c)} \\ & y \in \mathbb{R}^k,\ z \in \{0,1\}^n. && && \text{(3d)} \end{aligned}$$

We denote this set by $\mathcal{M}(W, \ell, \varepsilon)$. When $W \in \mathbb{R}^{n \times k}_+$, constraints (3a) are often called big-M constraints, and constraints (3b) impose lower bounds on the continuous variables y. Notice that (3c) is a constraint linking all continuous variables, but it is non-redundant only if ε is strictly positive. We will refer to (3c) as the *linking constraint*. When $k = 1$, $\ell = \mathbf{0}$, and $\varepsilon = 0$, the set $\mathcal{M}(W, \ell, \varepsilon)$ is nothing but MIX_1, i.e., the mixing set with binary variables, studied in the literature [1, 15, 19, 23, 40]. Sets of the form $\mathcal{M}(W, \mathbf{0}, 0)$ for general $k > 1$ were first considered by Atamtürk et al. [5]; we will call the set $\mathcal{M}(W, \mathbf{0}, 0)$ a *joint mixing set* in order to emphasize that k can be taken to be strictly greater than 1. We will refer to a set of the form $\mathcal{M}(W,\ell,\varepsilon)$ for general ℓ,ε as a *joint mixing set with lower bounds*.

The structure of a joint mixing set with lower bounds $\mathcal{M}(W,\ell,\varepsilon)$ is flexible enough to simultaneously work with quantile cuts. For $j\in[k]$, let ℓ_j denote the $(1-\epsilon)$-quantile for $c^\top y = y_j$. Then, for any $j\in[k]$, we have

$$\ell_j = \min\left\{\max_{i\in[n]}\{w_{ij}(1-z_i)\}:\ z \text{ satisfies (2d), (2e)}\right\}.$$

Note that ℓ_j can be computed in $O(n\log n)$ time, because without loss of generality we can assume $w_{1j}\geq\cdots\geq w_{nj}$ after possible reordering of $[n]$, and the optimum value of the above optimization problem is precisely w_{tj} where t is the index such that $\sum_{i\leq t-1}p_i\leq\epsilon$ and $\sum_{i\leq t}p_i>\epsilon$. Although the $(1-\epsilon)$-quantile for $\sum_{j\in[k]}y_j$ seems harder to compute, at least we know that the value is greater than or equal to $\sum_{j\in[k]}\ell_j$. Therefore, we have quantile cuts $y_j\geq\ell_j$ for $j\in[k]$ and $\sum_{j\in[k]}y_j\geq\varepsilon+\sum_{j\in[k]}\ell_j$ for some $\varepsilon\geq 0$, and the set defined by these quantile cuts and the constraints (2c), (2e) is precisely a set of the form $\mathcal{M}(W,\ell,\varepsilon)$. Similarly, it is straightforward to capture the quantile cut $c^\top y\geq\varepsilon+\sum_{j\in[k]}c_j\ell_j$ for general $c\in\mathbb{R}^k_+$, because we can rewrite $y_j\geq\ell_j$ for $j\in[k]$, (2c) and (2e) in terms of $c_1y_1,\ldots,c_jy_j$, and thus the resulting system is equivalent to a joint mixing set with lower bounds.

Next, we summarize our contributions and provide an outline of the paper.

1.1 Contributions and outline

In this paper, we study the polyhedral structure of $\mathcal{M}(W,\ell,\varepsilon)$, i.e., joint mixing sets with lower bounds, mainly in the context of joint linear CCPs with random right-hand sides and a discrete probability distribution. Our approach is based on a connection between mixing sets and submodularity that has been overlooked in the literature. Therefore, in Section 2.1, we first discuss basics of submodular functions and polymatroid inequalities as they relate to our work. In addition, we devote Section 2.2 to establish new tools on a particular joint submodular structure; these new tools play a critical role in our analysis of the joint mixing sets.

Our contributions are as follows:

(i) We first establish a strong and somewhat surprising connection between polymatroids and the basic mixing sets with binary variables (Section 3). It is well-known that submodularity imposes favorable characteristics in terms of explicit convex hull descriptions via known classes of inequalities and their efficient separation. In particular, the idea of utilizing polymatroid inequalities from submodular functions has appeared in various papers in other contexts for specific binary integer programs [3, 4, 6, 34, 36, 38]. Notably, mixing sets have been known to be examples of simple structured sets whose convex hull descriptions possess similar favorable characteristics. However, to the best of our knowledge, the connection between submodularity and mixing sets has not been recognized before. Establishing this connection enables us to unify and generalize various existing results on mixing sets with binary variables.

(ii) In Section 4, we propose a new class of valid inequalities, referred to as the *aggregated mixing inequalities*, for the set $\mathcal{M}(W, \ell, \varepsilon)$. One important feature of the class of aggregated mixing inequalities as opposed to the standard mixing inequalities is that it is specifically designed to simultaneously exploit the information encoded in multiple mixing sets with common binary variables.

(iii) In Section 5, we establish conditions under which the convex hull of the set $\mathcal{M}(W, \ell, \varepsilon)$ can be characterized through a submodularity lens. We show that the new class of aggregated mixing inequalities, in addition to the classical mixing inequalities, are sufficient under appropriate conditions.

(iv) In Section 6, we revisit the results from a recent paper by Liu et al. [19] on modeling two-sided CCPs. We show that mixing sets of the particular structure considered in [19] is nothing but a joint mixing set with lower bound structure with $k = 2$ and two additional constraints involving only the continuous variables y. Thus, our results on aggregated mixing inequalities are immediately applicable to two-sided CCPs. In addition, we show that, due to the simplicity of the additional constraints on the variables y in two-sided CCPs, our general convex hull results on $\mathcal{M}(W, \ell, \varepsilon)$ can be extended easily to accommodate the additional constraints on y and recover the convex hull results from [19].

Finally, we would like to highlight that although our results are motivated by joint CCPs, they are broadly applicable to other settings where the intersection of mixing sets with common binary variables is present. In addition, applicability of our results from Section 2.2 extend to other cases where epigraphs of general submodular functions appear in a similar structure.

1.2 Notation

Given a positive integer n, we let $[n] := \{1, \dots, n\}$. We let $\mathbf{0}$ denote the vector of all zeros whose dimension varies depending on the context, and similarly, $\mathbf{1}$ denotes the vector of all ones. e^j denotes the unit vector whose j^{th} coordinate is 1, and its dimension depends on the context. For $V \subseteq [n]$, $\mathbf{1}_V \in \{0,1\}^n$ denotes the characteristic vector, or the incidence vector, of V. For a set Q, we denote its convex hull and the extreme points of its convex hull by $\operatorname{conv}(Q)$ and $\operatorname{ext}(Q)$ respectively. For $\alpha \in \mathbb{R}$, $(\alpha)_+$ denotes $\max\{0, \alpha\}$. Given a vector $\pi \in \mathbb{R}^n$, and a set $V \subseteq [n]$, we define $\pi(V) = \sum_{i \in V} \pi_i$. For notational purposes, when $S = \emptyset$, we define $\max_{i \in S} s_i = 0$ and $\sum_{i \in S} s_i = 0$.

2 Submodular functions and polymatroid inequalities

In this section, we start with a brief review of submodular functions and polymatroid inequalities, and then in Section 2.2 we establish tools on joint submodular constraints that are useful for our analysis of $\mathcal{M}(W, \ell, \varepsilon)$.

2.1 Preliminaries

Consider an integer $n \geq 1$ and a set function $f : 2^{[n]} \to \mathbb{R}$. Recall that f is *submodular* if

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B), \quad \forall A, B \subseteq [n].$$

Given a submodular set function f, Edmonds [11] introduced the notion of *extended polymatroid of* f, which is a polyhedron associated with f defined as follows:

$$EP_f := \{\pi \in \mathbb{R}^n : \ \pi(V) \leq f(V), \ \forall V \subseteq [n]\}. \tag{4}$$

Observe that EP_f is nonempty if and only if $f(\emptyset) \geq 0$. In general, a submodular function f need not satisfy $f(\emptyset) \geq 0$. Nevertheless, it is straightforward to see that the function $f - f(\emptyset)$ is submodular whenever f is submodular, and that $(f - f(\emptyset))(\emptyset) = 0$. Hence, $EP_{f-f(\emptyset)}$ is always nonempty. Hereinafter, we use notation $\tilde{f}$ to denote $f - f(\emptyset)$ for any set function f.

A function on $\{0,1\}^n$ can be interpreted as a set function over the subsets of $[n]$, and thus, the definitions of submodular functions and extended polymatroids extend to functions over $\{0,1\}^n$. To see this, consider any integer $n \geq 1$ and any function $f : \{0,1\}^n \to \mathbb{R}$. With a slight abuse of notation, define $f(V) := f(\mathbf{1}_V)$ for $V \subseteq [n]$ where $\mathbf{1}_V$ denotes the characteristic vector of V. We say that $f : \{0,1\}^n \to \mathbb{R}$ is a submodular function if the corresponding set function over $[n]$ is submodular. We can also define the extended polymatroid of $f : \{0,1\}^n \to \mathbb{R}$ as in (4). Throughout this paper, given a function $f : \{0,1\}^n \to \mathbb{R}$, we will switch between its set function interpretation and its original form, depending on the context.

Given a submodular function $f : \{0,1\}^n \to \mathbb{R}$, its *epigraph* is the mixed-integer set given by

$$Q_f = \{(y, z) \in \mathbb{R} \times \{0,1\}^n : \ y \geq f(z)\}.$$

It is well-known that when f is submodular, one can characterize the convex hull of Q_f through the extended polymatroid of $\tilde{f}$.

Theorem 2.1 (Lovász [20], Atamtürk and Narayanan [4, Proposition 1]). *Let $f : \{0,1\}^n \to \mathbb{R}$ be a submodular function. Then*

$$\operatorname{conv}(Q_f) = \left\{(y, z) \in \mathbb{R} \times [0,1]^n : \ y \geq \pi^\top z + f(\emptyset), \ \forall \pi \in EP_{\tilde{f}}\right\}.$$

The inequalities $y \geq \pi^\top z + f(\emptyset)$ for $\pi \in EP_{\tilde{f}}$ are called the *polymatroid inequalities of* f. Although there are infinitely many polymatroid inequalities of f, for the description of $\operatorname{conv}(Q_f)$, it is sufficient to consider only the ones corresponding to the extreme points of $EP_{\tilde{f}}$. We refer to the polymatroid inequalities defined by the extreme points of $EP_{\tilde{f}}$ as the *extremal polymatroid inequalities of* f. Moreover, Edmonds [11] provided the following explicit characterization of the extreme points of $EP_{\tilde{f}}$.

Theorem 2.2 (Edmonds [11]). *Let $f : \{0,1\}^n \to \mathbb{R}$ be a submodular function. Then $\pi \in \mathbb{R}^n$ is an extreme point of $EP_{\tilde{f}}$ if and only if there exists a permutation σ of $[n]$ such that $\pi_{\sigma(t)} = f(V_t) - f(V_{t-1})$, where $V_t = \{\sigma(1), \ldots, \sigma(t)\}$ for $t \in [n]$ and $V_0 = \emptyset$.*

The algorithmic proof of Theorem 2.2 from Edmonds [11] is of interest. Suppose that we are given a linear objective $\bar{z} \in \mathbb{R}^n$; then $\max_\pi \left\{\bar{z}^\top \pi : \pi \in EP_{\tilde{f}}\right\}$ can be solved by the following "greedy" algorithm: given $\bar{z} \in \mathbb{R}^n$, first find an ordering σ such that $\bar{z}_{\sigma(1)} \geq \cdots \geq \bar{z}_{\sigma(n)}$, and let $V_t := \{\sigma(1), \dots, \sigma(t)\}$ for $t \in [n]$ and $V_0 = \emptyset$. Then, $\pi \in \mathbb{R}^n$ where $\pi_{\sigma(t)} = f(V_t) - f(V_{t-1})$ for $t \in [n]$ is an optimal solution to $\max_\pi \left\{\bar{z}^\top \pi : \pi \in EP_{\tilde{f}}\right\}$. Note that the implementation of this algorithm basically requires a sorting algorithm to compute the desired ordering σ, and this can be done in $O(n \log n)$ time. Thus, the overall complexity of this algorithm is $O(n \log n)$.

Consequently, given a point $(\bar{y}, \bar{z}) \in \mathbb{R} \times \mathbb{R}^n$, separating a violated polymatroid inequality amounts to solving the optimization problem $\max_\pi \left\{\bar{z}^\top \pi : \pi \in EP_{\tilde{f}}\right\}$, and thus we arrive at the following result.

Corollary 1 (Atamtürk and Narayanan [4, Section 2]). *Let $f : \{0,1\}^n \to \mathbb{R}$ be a submodular function. Then the separation problem for polymatroid inequalities can be solved in $O(n \log n)$ time.*

2.2 Joint submodular constraints

In this section, we establish tools that will be useful throughout this paper. Recall that when f is submodular, the convex hull of its epigraph Q_f is described by the extremal polymatroid inequalities of f. Henceforth, we use the restriction $(y, z) \in \operatorname{conv}(Q_f)$ as a constraint to indicate the inclusion of the corresponding extremal polymatroid inequalities of f in the constraint set.

Let $f_1, \dots, f_k : \{0,1\}^n \to \mathbb{R}$ be k submodular functions. Let us examine the convex hull of the following mixed-integer set:

$$Q_{f_1,\dots,f_k} := \left\{(y, z) \in \mathbb{R}^k \times \{0,1\}^n : y_1 \geq f_1(z), \dots, y_k \geq f_k(z)\right\}.$$

When $k = 1$, the set Q_{f_1} is just the epigraph of the submodular function f_1 on $\{0,1\}^n$. For general k, $Q_{f_1,\dots,f_k}$ is described by k submodular functions that share the same set of binary variables. For $(y, z) \in Q_{f_1,\dots,f_k}$, constraint $y_j \geq f_j(z)$ can be replaced with $(y_j, z) \in Q_{f_j}$ for $j \in [k]$. Therefore, the polymatroid inequalities of f_j with left-hand side y_j, of the form $y_j \geq \pi^\top z + f_j(\emptyset)$ with $\pi \in EP_{\tilde{f}_j}$, are valid for $Q_{f_1,\dots,f_k}$. In fact, these inequalities are sufficient to describe $\operatorname{conv}(Q_{f_1,\dots,f_k})$ as well.

Proposition 1 (Baumann et al. [6, Theorem 2]). *Let the functions $f_1, \dots, f_k : \{0,1\}^n \to \mathbb{R}$ be submodular. Then,*

$$\operatorname{conv}\left(Q_{f_1,\dots,f_k}\right) = \left\{(y, z) \in \mathbb{R}^k \times [0,1]^n : (y_j, z) \in \operatorname{conv}(Q_{f_j}),\ \forall j \in [k]\right\}.$$

By Proposition 1, when $f_1, \dots, f_k$ are submodular, $\operatorname{conv}\left(Q_{f_1,\dots,f_k}\right)$ can be described by the polymatroid inequalities of f_j with left-hand side y_j for $j \in [k]$. The submodularity requirement on all of the functions f_j in Proposition 1 is indeed critical. We demonstrate in the next example that even when $k = 2$, and only one of the functions f_i is not submodular, we can no longer describe the corresponding convex hull using the polymatroid inequalities for f_j.

Example 1. Let $f_1, f_2 : \{0,1\}^2 \to \mathbb{R}$ be defined by

$$f_1(0,0) = f_1(1,1) = 0,\ f_1(0,1) = f_1(1,0) = 1 \quad \text{and} \quad f_2(0,0) = f_2(1,1) = 1,\ f_2(0,1) = f_2(1,0) = 0.$$

While f_1 is submodular, f_2 is not. Since $f_1(0,0) = f_1(1,1) = 0$, we deduce that $(0, 1/2, 1/2) \in \mathrm{conv}(Q_{f_1})$. Similarly, as $f_2(0,1) = f_2(1,0) = 0$, it follows that $(0, 1/2, 1/2) \in \mathrm{conv}(Q_{f_2})$. This implies that

$$(0,0,1/2,1/2) \in \left\{(y,z) \in \mathbb{R}^2 \times [0,1]^n :\ (y_1, z) \in \mathrm{conv}(Q_{f_1}),\ (y_2, z) \in \mathrm{conv}(Q_{f_2})\right\}.$$

Notice that, by definition of f_1, f_2, we have $f_1(z) + f_2(z) = 1$ for each $z \in \{0,1\}^n$, implying in turn that $y_1 + y_2 \geq 1$ is valid for $\mathrm{conv}\left(Q_{f_1,f_2}\right)$. Therefore, the point $(0,0,1/2,1/2)$ cannot be in $\mathrm{conv}\left(Q_{f_1,f_2}\right)$. So, it follows that $\mathrm{conv}\left(Q_{f_1,f_2}\right) \neq \left\{(y,z) \in \mathbb{R}^k \times [0,1]^n :\ (y_j, z) \in \mathrm{conv}(Q_{f_j}),\ \forall j \in [2]\right\}$. □

In Section 3, we will discuss how Proposition 1 can be used to provide the convex hull description of a joint mixing set $\mathcal{M}(W, \mathbf{0}, 0)$.

We next highlight a slight generalization of Proposition 1 that is of interest for studying $\mathcal{M}(W, \ell, \varepsilon)$. Observe that $Q_{f_1,\ldots,f_k}$ is defined by multiple submodular constraints with independent continuous variables y_j. We can replace this independence condition by a certain type of dependence. Consider the following mixed-integer set:

$$\mathcal{P} = \left\{(y,z) \in \mathbb{R}^k \times \{0,1\}^n :\ a_1^\top y \geq f_1(z), \ldots, a_m^\top y \geq f_m(z)\right\} \tag{5}$$

where $a_1, \ldots, a_m \in \mathbb{R}_+^k \setminus \{\mathbf{0}\}$ and $f_1, \ldots, f_m : \{0,1\}^n \to \mathbb{R}$ are submodular functions. Here, m can be larger than k, so $a_1, \ldots, a_m$ need not be linearly independent. Now consider $\alpha = \sum_{j \in [m]} c_j a_j$ for some $c \in \mathbb{R}_+^m$. Notice that $f_\alpha \geq \sum_{j \in [m]} c_j f_j$ where $f_\alpha : \{0,1\}^n \to \mathbb{R}$ is defined as

$$f_\alpha(z) := \min\left\{\alpha^\top y :\ (y,z) \in \mathcal{P}\right\}, \quad \forall z \in \{0,1\}^n. \tag{6}$$

Definition 1. We say that $a_1 y, \ldots, a_m y$ are *weakly independent with respect to* $f_1, \ldots, f_m$ if for any $\alpha = \sum_{j \in [m]} c_j a_j$ with $c \in \mathbb{R}_+^m$, we have $f_\alpha = \sum_{j \in [m]} c_j f_j$. □

It is straightforward to see that if $a_1, \ldots, a_m$ are distinct unit vectors, i.e., $m = k$ and $a_j^\top y = y_j$ for $j \in [k]$, then $a_1^\top y, \ldots, a_m^\top y$ are weakly independent. It is also easy to see that if $a_1, \ldots, a_m$ are linearly independent, then $a_1^\top y, \ldots, a_m^\top y$ are weakly independent. Based on this definition, we have the following slight extension of Proposition 1.

Proposition 2. *Let $\mathcal{P}$ be defined as in (5). If $a_1^\top y, \ldots, a_m^\top y$ are weakly independent with respect to $f_1, \ldots, f_m$, then*

$$\mathrm{conv}\left(\mathcal{P}\right) = \left\{(y,z) \in \mathbb{R}^k \times [0,1]^n :\ (a_j^\top y, z) \in \mathrm{conv}(Q_{f_j}),\ \forall j \in [m]\right\}.$$

Proof. Define $\mathcal{R} := \left\{(y,z) \in \mathbb{R}^k \times [0,1]^n :\ (a_j^\top y, z) \in \mathrm{conv}(Q_{f_j}),\ \forall j \in [m]\right\}$. It is clear that $\mathrm{conv}\left(\mathcal{P}\right) \subseteq \mathcal{R}$. For the direction $\mathrm{conv}\left(\mathcal{P}\right) \supseteq \mathcal{R}$, we need to show that any inequality $\alpha^\top y + \beta^\top z \geq \gamma$ valid for $\mathrm{conv}\left(\mathcal{P}\right)$ is also valid for $\mathcal{R}$. To that end, take an inequality $\alpha^\top y + \beta^\top z \geq \gamma$ valid for $\operatorname{conv}(\mathcal{P})$. Then, since $\alpha^\top r \geq 0$ for every recessive direction $(r, \mathbf{0})$ of $\operatorname{conv}(\mathcal{P})$, we deduce by Farkas' lemma that $\alpha = \sum_{j\in[m]} c_j a_j$ for some $c \in \mathbb{R}^m_+$. Moreover, $\alpha^\top y + \beta^\top z \geq \gamma$ is valid for

$$\mathcal{Q} := \left\{(y,z) \in \mathbb{R}^k \times \{0,1\}^n : \ \alpha^\top y \geq f_\alpha(z)\right\},$$

where f_α is defined as in (6). Since $a_1^\top y, \ldots, a_m^\top y$ are weakly independent with respect to $f_1, \ldots, f_m$, it follows that $f_\alpha = \sum_{j\in[m]} c_j f_j$, and therefore, f_α is submodular. Then it is not difficult to see that

$$\operatorname{conv}(\mathcal{Q}) = \left\{(y,z) \in \mathbb{R}^k \times [0,1]^n : \ (\alpha^\top y, z) \in \operatorname{conv}(Q_{f_\alpha})\right\}.$$

Therefore, to show that $\alpha^\top y + \beta^\top z \geq \gamma$ is valid for $\mathcal{R}$, it suffices to argue that $\mathcal{R} \subseteq \operatorname{conv}(\mathcal{Q})$. Let $(\bar{y}, \bar{z}) \in \mathcal{R}$. Then, by Theorem 2.1, it suffices to show that $\alpha^\top \bar{y} \geq \pi^\top \bar{z} + f_\alpha(\emptyset)$ holds for every extreme point π of $EP_{\tilde{f}_\alpha}$. To this end, take an extreme point π of $EP_{\tilde{f}_\alpha}$. By Theorem 2.2, there exists a permutation σ of $[n]$ such that $\pi_{\sigma(t)} = f_\alpha(V_t) - f_\alpha(V_{t-1})$ where $V_t = \{\sigma(1), \ldots, \sigma(t)\}$ for $t \in [n]$ and $V_0 = \emptyset$. Now, for $j \in [m]$, let $\pi^j \in \mathbb{R}^n$ be the vector such that $\pi^j_{\sigma(t)} = f_j(V_t) - f_j(V_{t-1})$ for $t \in [n]$. Then, we have $\pi = \sum_{j\in[m]} c_j \pi^j$ because $f_\alpha = \sum_{j\in[m]} c_j f_j$. Moreover, by Theorem 2.2, π^j is an extreme point of $EP_{\tilde{f}_j}$. Hence, due to our assumption that $(a_j^\top \bar{y}, \bar{z}) \in \operatorname{conv}(Q_{f_j})$, Theorem 2.1 implies $a_j^\top \bar{y} \geq (\pi^j)^\top \bar{z} + \pi_j(\emptyset)$ is valid for all $j \in [m]$. Since $\alpha^\top \bar{y} \geq \pi^\top \bar{z} + f_\alpha(\emptyset)$ is obtained by adding up $a_j^\top \bar{y} \geq (\pi^j)^\top \bar{z} + \pi_j(\emptyset)$ for $j \in [m]$, it follows that $\alpha^\top \bar{y} \geq \pi^\top \bar{z} + f_\alpha(\emptyset)$ is valid, as required. We just have shown that $\mathcal{R} \subseteq \operatorname{conv}(\mathcal{Q})$, thereby completing the proof. $\square$

In Section 5, we will use Proposition 2 to study the convex hull of $\mathcal{M}(W, \ell, \varepsilon)$, i.e., a joint mixing set with lower bounds. Again, the submodularity assumption on $f_1, \ldots, f_m$ is important in Proposition 2. Recall that Example 1 demonstrates that in Proposition 2 even when m is taken to be equal to k and the vectors $a_j \in \mathbb{R}^k$, $j \in [m] = [k]$, are taken to be the unit vectors in $\mathbb{R}^k$, the statement does not hold if one of the functions f_j is not submodular.

3 Mixing inequalities and joint mixing sets

In this section, we establish that mixing sets with binary variables are indeed nothing but the epigraphs of certain submodular functions. In addition, through this submodularity lens, we prove that the well-known mixing (or star) inequalities for mixing sets are nothing but the extremal polymatroid inequalities.

Recall that a joint mixing set with lower bounds $\mathcal{M}(W, \ell, \varepsilon)$, where $W \in \mathbb{R}^{n\times k}_+$, $\ell \in \mathbb{R}^k_+$ and $\varepsilon \geq 0$, is defined by (3). In this section, we study the case when $\varepsilon = 0$, and characterize the convex hull of $\mathcal{M}(W, \ell, 0)$ for any $W \in \mathbb{R}^{n\times k}_+$ and $\ell \in \mathbb{R}^k_+$. As corollaries, we prove that the famous star/mixing inequalities are in fact polymatroid inequalities, and we recover the result of Atamtürk et al. [5, Theorem 3] on joint mixing sets $\mathcal{M}(W, \mathbf{0}, 0)$.

Given a matrix $W = \{w_{ij}\} \in \mathbb{R}^{n\times k}_+$ and a vector $\ell \in \mathbb{R}^k_+$, we define the following mixed-integer set:

$$\mathcal{P}(W, \ell, \varepsilon) = \left\{(y,z) \in \mathbb{R}^k \times \{0,1\}^n : \ (8) - (10)\right\} \tag{7}$$

where

$$
\begin{aligned}
& y_j \geq w_{ij} z_i, && \forall i \in [n],\ j \in [k], && (8)\\
& y_j \geq \ell_j, && \forall j \in [k], && (9)\\
& \sum_{j \in [k]} y_j \geq \varepsilon + \sum_{j \in [k]} \ell_j. && && (10)
\end{aligned}
$$

Remark 1. By definition, $(y, z) \in \mathcal{M}(W, \ell, \varepsilon)$ if and only if $(y, \mathbf{1} - z) \in \mathcal{P}(W, \ell, \varepsilon)$. Thus, the convex hull of $\mathcal{M}(W, \ell, \varepsilon)$ can be obtained after taking the convex hull of $\mathcal{P}(W, \ell, \varepsilon)$ and complementing the z variables. $\square$

For $j \in [k]$, we define

$$
f_j(z) := \max \left\{ \ell_j,\ \max_{i \in [n]} \{w_{ij} z_i\} \right\}, \quad \forall z \in \{0,1\}^n. \tag{11}
$$

Then, the set $\mathcal{P}(W, \ell, 0)$ admits a representation as the intersection of epigraphs of the functions $f_j(z)$:

$$
\mathcal{P}(W, \ell, 0) = \left\{ (y, z) \in \mathbb{R}^k \times \{0,1\}^n :\ y_j \geq f_j(z),\ \forall j \in [k] \right\}.
$$

We next establish that the functions $f_j(z), j \in [k]$ are indeed submodular.

Lemma 1. *Let $\ell \in \mathbb{R}^k_+$. For each $j \in [k]$, the function f_j defined as in (11) satisfy $f_j(\emptyset) = \ell_j$ and it is submodular.*

Proof. Let $j \in [k]$. Notice that $f_j(\emptyset) = f_j(\mathbf{0}) = \max\{\ell_j, 0\} = \ell_j$. In order to establish the submodularity of f_j, for ease of notation, we drop the index j and use f to denote f_j. As before, for each $V \subseteq [n]$, let $f(V)$ be defined as $f(\mathbf{1}_V)$ where $\mathbf{1}_V \in \{0,1\}^n$ denotes the characteristic vector of V. Consider two sets $U, V \subseteq [n]$. By definition of f, we have $\max\{f(U), f(V)\} = f(U \cup V)$, and $\min\{f(U), f(V)\} \geq f(U \cap V)$. Then we immediately get

$$
f(U) + f(V) = \max\{f(U), f(V)\} + \min\{f(U), f(V)\} \geq f(U \cup V) + f(U \cap V),
$$

thereby proving that f_j is submodular, as required. $\square$

Corollary 2. *Let $\ell \in \mathbb{R}^k_+$ and f_j be as defined in (11). Then,*

$$
\operatorname{conv}(\mathcal{M}(W, \ell, 0)) = \left\{ (y, z) \in \mathbb{R}^k \times [0,1]^n :\ (y_j, \mathbf{1} - z) \in \operatorname{conv}(Q_{f_j}),\ \forall j \in [k] \right\},
$$

i.e., the convex hull of $\mathcal{M}(W, \ell, 0)$ is given by the extremal polymatroid inequalities of particular submodular functions.

Proof. We deduce from Proposition 1 that

$$
\operatorname{conv}(\mathcal{P}(W, \ell, 0)) = \left\{ (y, z) \in \mathbb{R}^k \times [0,1]^n :\ (y_j, z) \in \operatorname{conv}(Q_{f_j}),\ \forall j \in [k] \right\},
$$

which immediately implies the desired relation via Remark 1 and Theorem 2.1 since the constraint $(y_j, \mathbf{1} - z) \in \operatorname{conv}(Q_{f_j})$ is equivalent to the set of the corresponding extremal polymatroid inequalities. $\square$

Corollary 2 establishes a strong connection between the mixing sets with binary variables and the epigraphs of submodular functions, and implies that the convex hull of joint mixing sets are given by the extremal polymatroid inequalities. To the best of our knowledge this connection between mixing sets with binary variables and submodularity has not been identified in the literature before.

An explicit characterization of the convex hull of a mixing set with binary variables in the original space has been studied extensively in the literature. Specifically, Atamtürk et al. [5] gave the explicit characterization of $\mathrm{conv}(\mathcal{M}(W,\mathbf{0},0))$ in terms of the so called *mixing (star) inequalities*. Let us state the definition of these inequalities here.

Definition 2. We call a sequence $\{j_1 \to \cdots \to j_\tau\}$ of indices in $[n]$ a *j-mixing-sequence* if $w_{j_1 j} \geq w_{j_2 j} \geq \cdots \geq w_{j_\tau j} \geq \ell_j$. ☐

For $W = \{w_{ij}\} \in \mathbb{R}_+^{n\times k}$ and $\ell \in \mathbb{R}_+^k$, the *mixing inequality derived from a j-mixing-sequence* $\{j_1 \to \cdots \to j_\tau\}$ is defined as the following (see [12, Section 2]):

$$y_j + \sum_{s\in[\tau]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s} \geq w_{j_1 j}, \tag{$\mathrm{Mix}_{W,\ell}$}$$

where $w_{j_{\tau+1} j} := \ell_j$ for convention. Atamtürk et al. [5, Proposition 3] showed that the inequality ($\mathrm{Mix}_{W,\ell}$) for any j-mixing-sequence $\{j_1 \to \cdots \to j_\tau\}$ is valid for $\mathcal{M}(W,\ell,0)$ when $\ell = \mathbf{0}$. Luedtke [21, Theorem 2] later observed that the inequality ($\mathrm{Mix}_{W,\ell}$) for any j-mixing-sequence $\{j_1 \to \cdots \to j_\tau\}$ is valid for $\mathcal{M}(W,\ell,0)$ for any $\ell \in \mathbb{R}_+^k$.

Given these results from the literature on the convex hull characterizations of mixing sets and Corollary 2, it is plausible to think that there must be a strong connection between the extremal polymatroid inequalities and the mixing (star) inequalities. We next argue that the extremal polymatroid inequalities given by the constraint $(y_j, \mathbf{1} - z) \in \mathrm{conv}(Q_{f_j})$ are precisely the mixing (star) inequalities.

Proposition 3. *Let $W = \{w_{ij}\} \in \mathbb{R}_+^{n\times k}$ and $\ell \in \mathbb{R}_+^k$. Consider any $j \in [k]$. Then, for every extreme point π of $EP_{\tilde f_j}$, there exists a j-mixing-sequence $\{j_1 \to \cdots \to j_\tau\}$ in $[n]$ that satisfies the following:*

(1) $w_{j_1 j} = \max\{w_{ij} : i \in [n]\}$,

(2) the corresponding polymatroid inequality $y_j + \sum_{i\in[n]} \pi_i z_i \geq \ell_j + \sum_{i\in[n]} \pi_i$ is equivalent to the mixing inequality ($\mathrm{Mix}_{W,\ell}$) derived from the sequence $\{j_1 \to \cdots \to j_\tau\}$.

In particular, for any $j \in [k]$, the extremal polymatroid inequality is of the form

$$y_j + \sum_{s\in[\tau]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s} \geq \max\{w_{ij} : i \in [n]\}, \tag{$\mathrm{Mix}^*_{W,\ell}$}$$

where $w_{j_1 j} = \max\{w_{ij} : i \in [n]\}$ and $w_{j_{\tau+1} j} := \ell_j$.

Proof. By Theorem 2.2, there exists a permutation σ of $[n]$ such that $\pi_{\sigma(t)} = f_j(V_t) - f_j(V_{t-1})$ where $V_t = \{\sigma(1), \ldots, \sigma(t)\}$ for $t \in [n]$ and $V_0 = \emptyset$. By definition of f_j in (11), we have $\ell_j = f_j(V_0) \leq f_j(V_1) \leq$

$\cdots \leq f_j(V_n)$, because $\emptyset = V_0 \subset V_1 \subset \cdots \subset V_n$. Let $\{t_1, \ldots, t_\tau\}$ be the collection of all indices t satisfying $f_j(V_{t-1}) < f_j(V_t)$. Without loss of generality, we may assume that $w_{\sigma(t_1)j} \geq \cdots \geq w_{\sigma(t_\tau)j}$. Notice that $w_{\sigma(t_\tau)j} > \ell_j$, because $f_j(V_{t_\tau}) > f_j(V_{t_\tau - 1}) \geq \ell_j$. Then, after setting $j_s = \sigma(t_s)$ for $s \in [\tau]$, it follows that $\{j_1 \to \cdots \to j_\tau\}$ is a j-mixing-sequence. Moreover, we have $w_{j_1 j} = f_j(V_{t_1}) = f_j([n]) = \max\{w_{ij} : i \in [n]\}$. Therefore, we deduce that $\pi_i = w_{j_s j} - w_{j_{s+1} j}$ if $i = \sigma(t_s) = j_s$ for some $s \in [\tau]$ and $\pi_i = 0$ otherwise. $\square$

As the name "mixing" inequalities is more commonly used in the literature than "star" inequalities, we will stick to the term "mixing" hereinafter to denote the inequalities of the form ($\text{Mix}_{W,\ell}$) or ($\text{Mix}^*_{W,\ell}$).

Proposition 1 and consequently Corollary 2 imply that for any facet defining inequality of the set $\text{conv}(\mathcal{M}(W, \ell, 0))$, there is a corresponding extremal polymatroid inequality. Proposition 3 implies that mixing inequalities are nothing but the extremal polymatroid inequalities. Therefore, an immediate consequence of Corollary 2 and Proposition 3 is the following result.

Theorem 3.1. *Given $W = \{w_{ij}\} \in \mathbb{R}^{n \times k}_+$ and any $\ell \in \mathbb{R}^k_+$, the convex hull of $\mathcal{M}(W, \ell, 0)$ is described by the mixing inequalities of the form ($\text{Mix}^*_{W,\ell}$) for $j \in [k]$ and the bounds $\mathbf{0} \leq z \leq \mathbf{1}$.*

A few remarks are in order.

Remark 2. First, note that Luedtke et al. [23, Theorem 2] showed the validity of inequality ($\text{Mix}^*_{W,\ell}$) and its facet condition for a particular choice of $\ell \in \mathbb{R}^k_+$ in the case of $k = 1$. Also, recall that $\mathcal{M}(W, \mathbf{0}, 0)$ is called a joint mixing set, and Atamtürk et al. [5, Theorem 3] proved that $\text{conv}(\mathcal{M}(W, \mathbf{0}, 0))$ is described by the mixing inequalities and the bound constraints $y \geq 0$ and $z \in [0, 1]^n$. Since Theorem 3.1 applies to $\mathcal{M}(W, \ell, 0)$ for arbitrary ℓ, it immediately extends [5, Theorem 3] and further extends the validity inequality component of [23, Theorem 2]. $\square$

Remark 3. Our final remark is that, since the mixing inequalities ($\text{Mix}^*_{W,\ell}$) for $j \in [k]$ are polymatroid inequalities, they can be separated in $O(k\, n \log n)$ time by a simple greedy algorithm, thanks to Corollary 1. $\square$

4 Aggregated mixing inequalities

As discussed in Section 1, in order to make use of the knapsack constraint in the MIP formulation of joint CCPs via quantile cuts, we need to study the set $\mathcal{M}(W, \ell, \varepsilon)$ for general $\varepsilon \geq 0$. Unfortunately, in contrast to our results in Section 3 for the convex hull of $\mathcal{M}(W, \ell, 0)$, the convex hull of $\mathcal{M}(W, \ell, \varepsilon)$ for general $\varepsilon \geq 0$ may be complicated; we will soon see this in Example 2. In this section, we introduce a new class of valid inequalities for $\mathcal{M}(W, \ell, \varepsilon)$ for arbitrary $\varepsilon \geq 0$. In Sections 5.2 and 5.3, we identify conditions under which these new inequalities along with the original mixing inequalities are sufficient to give the complete convex hull characterization.

For general $\varepsilon \geq 0$, $\mathcal{M}(W, \ell, \varepsilon)$, given by (3), is a subset of $\mathcal{M}(W, \ell, 0)$, which means that any inequality valid for $\mathcal{M}(W, \ell, 0)$ is also valid for $\mathcal{M}(W, \ell, \varepsilon)$. In particular, Theorem 3.1 implies that the mixing inequalities of the form ($\text{Mix}_{W,\ell}$) are valid for $\mathcal{M}(W,\ell,\varepsilon)$. However, unlike the $\varepsilon = 0$ case, we will see that the mixing inequalities are not sufficient to describe the convex hull of $\mathcal{M}(W,\ell,\varepsilon)$ if $\varepsilon > 0$.

We first present a simplification of $\mathcal{M}(W,\ell,\varepsilon)$. Although it is possible that $w_{ij} < \ell_j$ for some i, j when W, ℓ are arbitrary, we can reduce $\mathcal{M}(W,\ell,\varepsilon)$ to a set of the form $\mathcal{M}(W^\ell, \mathbf{0}, \varepsilon)$ for some $W^\ell = \{w_{ij}^\ell\} \in \mathbb{R}_+^{n\times k}$.

Lemma 2. *Let $\ell \in \mathbb{R}_+^k$. Then $\mathcal{M}(W,\ell,\varepsilon) = \{(y,z) \in \mathbb{R}^k \times \mathbb{R}^n : \ (y-\ell, z) \in \mathcal{M}(W^\ell, \mathbf{0}, \varepsilon)\}$, where $W^\ell = \{w_{ij}^\ell\} \in \mathbb{R}_+^{n\times k}$ is the matrix whose entries are given by*

$$w_{ij}^\ell = (w_{ij} - \ell_j)_+ \quad \forall i \in [n],\ j \in [k].$$

Proof. By definition, $(y-\ell, z) \in \mathcal{M}(W^\ell, \mathbf{0}, \varepsilon)$ if and only if

$$y_j + (w_{ij} - \ell_j)_+ z_i \geq \ell_j + (w_{ij} - \ell_j)_+, \quad \forall i \in [n],\ j \in [k], \tag{12}$$

and (y,z) satisfies (3b)–(3d). Consider any $j \in [k]$. If $\ell_j > w_{ij}$, then the constraint (12) becomes $y_j \geq \ell_j$ and the inequality $y_j + w_{ij} z_i \geq w_{ij}$ is a consequence of $y_j \geq \ell_j$. On the other hand, if $\ell_j \leq w_{ij}$, then (12) is equivalent to $y_j + (w_{ij} - \ell_j) z_i \geq w_{ij}$, and therefore we have $y_j \geq w_{ij}$ when $z_i = 0$ and have $y_j \geq \ell_j$ when $z_i = 1$. Then, in both cases, it is clear that

$$\{(y_j, z_i) \in \mathbb{R} \times \{0,1\} : \ y_j \geq \ell_j,\ y_j + (w_{ij} - \ell_j)_+ z_i \geq \ell_j + (w_{ij} - \ell_j)_+\}$$

is equal to

$$\{(y_j, z_i) \in \mathbb{R} \times \{0,1\} : \ y_j \geq \ell_j,\ y_j + w_{ij} z_i \geq w_{ij}\},$$

because $\ell_j \geq 0$. Hence, we have $(y-\ell, z) \in \mathcal{M}(W^\ell, \mathbf{0}, \varepsilon)$ if and only if $(y,z) \in \mathcal{M}(W,\ell,\varepsilon)$, as required. $\square$

We deduce from Lemma 2 that

$$\operatorname{conv}(\mathcal{M}(W,\ell,\varepsilon)) = \{(y,z) \in \mathbb{R}^k \times \mathbb{R}^n : \ (y-\ell, z) \in \operatorname{conv}(\mathcal{M}(W^\ell, \mathbf{0}, \varepsilon))\},$$

and thus the convex hull description of $\mathcal{M}(W,\ell,\varepsilon)$ can be obtained by taking the convex hull of $\mathcal{M}(W^\ell, \mathbf{0}, \varepsilon)$. Moreover, any inequality $\alpha^\top y + \beta^\top z \geq \gamma$ is valid for $\mathcal{M}(W^\ell, \mathbf{0}, \varepsilon)$ if and only if $\alpha^\top (y - \ell) + \beta^\top z \geq \gamma$ is valid for $\mathcal{M}(W,\ell,\varepsilon)$.

So, from now on, we assume that $\ell = \mathbf{0}$, and we work over $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ with $W \in \mathbb{R}_+^{n\times k}$ and $\varepsilon \geq 0$. Recall that $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ is the mixed-integer set defined by

$$y_j + w_{ij} z_i \geq w_{ij}, \qquad \forall i \in [n],\ j \in [k], \tag{13a}$$

$$y_j \geq 0, \qquad \forall j \in [k], \tag{13b}$$

$$y_1 + \cdots + y_k \geq \varepsilon, \tag{13c}$$

$$y \in \mathbb{R}^k,\ z \in \{0,1\}^n. \tag{13d}$$

Let us begin with an example.

Example 2. Consider the following mixing set with lower bounds, i.e., $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ with $\varepsilon = 7 > 0$.

$$\left\{ (y,z) \in \mathbb{R}^2_+ \times \{0,1\}^5 : \begin{array}{c} y_1 + 8z_1 \geq 8 \\ y_1 + 6z_2 \geq 6 \\ y_1 + 13z_3 \geq 13 \\ y_1 + z_4 \geq 1 \\ y_1 + 4z_5 \geq 4 \end{array}, \begin{array}{c} y_2 + 3z_1 \geq 3 \\ y_2 + 4z_2 \geq 4 \\ y_2 + 2z_3 \geq 2 \\ y_2 + 2z_4 \geq 2 \\ y_2 + z_5 \geq 1 \end{array}, \ y_1 + y_2 \geq 7 \right\}. \tag{14}$$

The convex hull of this set is given by

$$\left\{ (y,z) \in \mathbb{R}^2_+ \times [0,1]^5 : \begin{array}{ll} y_1 + y_2 & +z_1 + z_2 + 8z_3 \geq 17 \\ y_1 + y_2 & +2z_2 + 8z_3 \geq 17 \\ y_1 + y_2 & +3z_2 + 7z_3 \geq 17 \\ y_1 + y_2 & +2z_1 + 3z_2 + 5z_3 \geq 17 \\ y_1 + y_2 & +4z_1 + z_2 + 5z_3 \geq 17 \end{array}, \begin{array}{c} \text{the mixing inequalities} \\ \text{with } y_j \text{ for } j = 1, 2 \end{array} \right\}.$$

In this example, the inequalities $y_1 + 2z_1 + 2z_2 + 5z_3 + z_4 + 3z_5 \geq 13$ and $y_2 + 2z_2 + z_4 + z_5 \geq 4$ are examples of mixing inequalities that are facet-defining. Note that the five inequalities with $y_1 + y_2$ are not of the form ($\text{Mix}_{W,\ell}$). Moreover, these non-mixing inequalities cannot be obtained by simply adding one mixing inequality involving y_1 and another mixing inequality involving y_2. The developments we present next on a new class of inequalities will demonstrate this point, and we will revisit this example again in Example 3. □

The five inequalities with $y_1 + y_2$ in Example 2 admit a common interpretation. To explain them, take an integer $\theta \in [n]$ and a sequence Θ of θ indices in $[n]$ given by $\{i_1 \to i_2 \to \cdots \to i_\theta\}$. Given two indices in the sequence i_p, i_q, we say that i_p *precedes* i_q *in* Θ if $p < q$. Our description is based on the following definition.

Definition 3. Given a sequence Θ, a *j-mixing-subsequence of* Θ is the subsequence $\{j_1 \to \cdots \to j_{\tau_j}\}$ of Θ that satisfies the following property:

$$\{j_1, \ldots, j_{\tau_j}\} \text{ is the collection of all indices } i^* \in \Theta \text{ satisfying } w_{i^*j} \geq \max\{w_{ij} : \ i^* \text{ precedes } i \text{ in } \Theta\},$$

where we define $\max\{w_{ij} : \ i_\theta \text{ precedes } i \text{ in } \Theta\} = 0$ for convention (i_θ is the last element, so it precedes no element in Θ). □

Based on Definition 3, we deduce that the j-mixing-subsequence of Θ is unique for each $j \in [k]$ and admits a few nice structural properties as identified below.

Lemma 3. *If $\{j_1 \to \cdots \to j_{\tau_j}\}$ is the j-mixing-subsequence of Θ, then j_{τ_j} is always the last element i_θ of Θ and $w_{j_1 j} \geq \cdots \geq w_{j_\tau j} \geq 0$.*

Proof. When $p < q$, because j_p precedes j_q in Θ, it follows that $w_{j_1 j} \geq \cdots \geq w_{j_{\tau_j} j} \geq 0$. The last element i_θ always satisfies $w_{i_\theta j} \geq \max\{w_{ij} : \ i_\theta \text{ precedes } i \text{ in } \Theta\} = 0$. Therefore, i_θ is part of the j-mixing-subsequence as its last element. □

Given $\Theta = \{i_1 \to i_2 \to \cdots \to i_\theta\}$, for any $j \in [n]$, we denote by $\Theta_j = \{j_1 \to \cdots \to j_{\tau_j}\}$ the j-mixing-subsequence of Θ. By Definition 2 and Lemma 3, we deduce that $\{j_1 \to \cdots \to j_{\tau_j}\}$ is a j-mixing-sequence.

Recall that for any j-mixing-sequence $\{j_1 \to \cdots \to j_{\tau_j}\}$, the corresponding mixing inequality (Mix$_{W,\ell}$) is of the following form:

$$y_j + \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s} \geq w_{j_1 j}, \tag{Mix}$$

where $w_{j_{\tau_j+1} j} := 0$, and it is valid for $\mathcal{M}(W, \mathbf{0}, \varepsilon)$. In particular, when $w_{j_1 j} = \max\{w_{ij} : i \in [n]\}$, (Mix) is

$$y_j + \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s} \geq \max\{w_{ij} : i \in [n]\}. \tag{Mix*}$$

Also, for $t \in [\theta]$,

$$\left(w_{i_t j} - \max\left\{w_{ij} : i_t \text{ precedes } i \text{ in } \Theta\right\}\right)_+ = \begin{cases} w_{j_s j} - w_{j_{s+1} j} & \text{if } i_t = j_s \text{ for some } s \in [\tau_j], \\ 0 & \text{if } i_t \text{ is not on } \Theta_j. \end{cases} \tag{15}$$

Then (Mix) can be rewritten as

$$y_j + \sum_{t\in[\theta]} \left(w_{i_t j} - \max\left\{w_{ij} : i_t \text{ precedes } i \text{ in } \Theta\right\}\right)_+ z_{i_t} \geq w_{j_1 j}.$$

In order to introduce our new class of inequalities, we define a constant $L_{W,\Theta}$ that depends on W and Θ as follows:

$$\begin{aligned} L_{W,\Theta} &:= \min\left\{\sum_{j\in[k]} \left(w_{i_t j} - \left(w_{i_t j} - \max\left\{w_{ij} : i_t \text{ precedes } i \text{ in } \Theta\right\}\right)_+\right) : t \in [\theta]\right\} \\ &= \min\left\{\sum_{j\in[k]} \min\left\{w_{i_t j},\ \max\left\{w_{ij} : i_t \text{ precedes } i \text{ in } \Theta\right\}\right\} : t \in [\theta]\right\} \end{aligned} \tag{16}$$

Now we are ready to introduce our new class of inequalities.

Definition 4. Given a sequence $\Theta = \{i_1 \to i_2 \to \cdots \to i_\theta\}$, let $L_{W,\Theta}$ be defined as in (16). Then, the *aggregated mixing inequality derived from* Θ is defined as the following:

$$\sum_{j\in[k]} \left(y_j + \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s}\right) - \min\left\{\varepsilon,\ L_{W,\Theta}\right\} z_{i_\theta} \geq \sum_{j\in[k]} \max\left\{w_{ij} : i \in \Theta\right\}. \tag{A-Mix}$$

□

Remark 4. Since $\min\{\varepsilon,\ L_{W,\Theta}\} \geq 0$, the aggregated mixing inequality (A-Mix) dominates what is obtained after adding up the mixing inequalities (Mix) for $j \in [k]$. □

Before proving validity of (A-Mix), we present an example illustrating how the aggregated mixing inequalities are obtained.

Example 3. We revisit the mixed-integer set in Example 2. Now take a sequence $\Theta = \{2 \to 1 \to 3\}$. Then $\{3\}$ and $\{2 \to 1 \to 3\}$ are the 1-mixing-subsequence and 2-mixing-subsequence of Θ, respectively. Moreover,

$$\begin{aligned} L_{W,\Theta} &= \min\left\{(6-(6-13)_+) + (4-(4-3)_+),\ \ (8-(8-13)_+) + (3-(3-2)_+),\ \ 13+2\right\} \\ &= \min\left\{6+3, 8+2, 13+2\right\} = 9. \end{aligned}$$

In (14), we have $\varepsilon = 7$. Since $\varepsilon \le L_{W,\Theta}$, the corresponding (A-Mix) is

$$(y_1 + 13z_3) + (y_2 + (4-3)z_2 + (3-2)z_1 + 2z_3) - 7z_3 \ge 13 + 4,$$

that is $y_1 + y_2 + z_1 + z_2 + 8z_3 \ge 17$. In Example 2, the other four inequalities with $y_1 + y_2$ are also of the form (A-Mix), and they are derived from the sequences $\{2 \to 3\}$, $\{3 \to 2\}$, $\{3 \to 1 \to 2\}$, and $\{3 \to 2 \to 1\}$. So, in this example, the convex hull of (14) is obtained after applying the mixing inequalities (Mix) and the aggregated mixing inequalities (A-Mix). $\square$

We will next present the proof of validity of (A-Mix). To this end, the following lemma is useful. As the proof of this lemma is technical, we defer its proof to the appendix. Lemma 4 will be used again in Section 5.3.

Lemma 4. *Let $(\bar{y}, \bar{z}) \in \mathbb{R}^k_+ \times [0,1]^n$ be a point satisfying (13a)–(13c). If $(\bar{y}, \bar{z})$ satisfies (A-Mix) for all sequences contained in $\{i \in [n] :\ \bar{z}_i < 1\}$, then $(\bar{y}, \bar{z})$ satisfies (A-Mix) for all the other sequences as well.*

Now we are ready to prove the following theorem:

Theorem 4.1. *The aggregated mixing inequalities defined as in (A-Mix) are valid for $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ where $W \in \mathbb{R}^{n\times k}_+$.*

Proof. We will argue that every point in $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ with $W \in \mathbb{R}^{n\times k}_+$ satisfies (A-Mix) for all sequences in $[n]$. To this end, take a point $(\bar{y}, \bar{z}) \in \mathcal{M}(W, \mathbf{0}, \varepsilon)$. Then, $\bar{z}_i \in \{0,1\}^n$ holds by definition of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$. If $\bar{z} = \mathbf{1}$, then $(\bar{y}, \bar{z})$ satisfies (A-Mix) if and only if $\sum_{j\in[k]} \bar{y}_j \ge \min\{\varepsilon,\ L_{W,\Theta}\}$. Since $\sum_{j\in[k]} \bar{y}_j \ge \varepsilon$, it follows that $(\bar{y}, \bar{z})$ satisfies (A-Mix). Thus, we may assume that $\{i \in [n] :\ \bar{z}_i < 1\} = \{i \in [n] :\ \bar{z}_i = 0\}$ is nonempty. By Lemma 4, it is sufficient to show that $(\bar{y}, \bar{z})$ satisfies (A-Mix) for every sequence contained in the nonempty set $\{i \in [n] :\ \bar{z}_i < 1\}$. Take a nonempty sequence $\Theta = \{i_1 \to \cdots \to i_\theta\}$ in $\{i \in [n] :\ \bar{z}_i = 0\}$. By our choice of Θ, we have $\bar{z}_{i_\theta} = 0$, so $(\bar{y}, \bar{z})$ satisfies (A-Mix) if and only if

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) \bar{z}_{j_s} \right) \ge \sum_{j\in[k]} w_{j_1 j}.$$

This inequality is precisely what is obtained by adding up the mixing inequalities (Mix) for $j \in [k]$, and therefore, $(\bar{y}, \bar{z})$ satisfies it, as required. $\square$

In Example 3, $\varepsilon = 7$ and $L_{W,\{2\to1\to3\}} = 9$. It can also be readily checked that $L_{W,\{2\to3\}} = L_{W,\{3\to2\}} = 8$ and $L_{W,\{3\to1\to2\}} = L_{W,\{3\to2\to1\}} = 9$, which means $\min\{\varepsilon, L_{W,\Theta}\} = \varepsilon$ for the sequences corresponding to the five aggregated mixing inequalities in the convex hull description of (14). In general, the following holds:

Proposition 4. *If $\varepsilon \le L_{W,\Theta}$, then the aggregated mixing inequality (A-Mix) obtained from Θ dominates the linking constraint $y_1 + \cdots + y_k \ge \varepsilon$.*

Proof. The inequality (A-Mix) is equivalent to

$$\sum_{j\in[k]} y_j \ge \varepsilon z_{i_\theta} + \sum_{j\in[k]} \left(w_{j_1 j} - \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s} \right).$$

Since $\sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) = w_{j_1 j}$, by Lemma 3, we have for all $j \in [k]$

$$\begin{aligned} w_{j_1 j} - \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s} &= \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j})(1 - z_{j_s}) \\ &\ge (w_{j_{\tau_j} j} - w_{j_{\tau_j+1} j})(1 - z_{j_{\tau_j}}) = w_{i_\theta j}(1 - z_{i_\theta}), \end{aligned}$$

where the inequality follows from the facts that $w_{j_s j} - w_{j_{s+1} j} \ge 0$ for all $j_s \in [\tau_j]$ and thus each summand is nonnegative, and the last equation follows from $j_{\tau_j} = i_\theta$ and by our convention that $w_{j_{\tau_j+1}} = 0$. Therefore, the following inequality is a consequence of (A-Mix):

$$\sum_{j\in[k]} y_j \ge \sum_{j\in[k]} w_{i_\theta j} + \left(\varepsilon - \sum_{j\in[k]} w_{i_\theta j} \right) z_{i_\theta}.$$

Since $0 \le z_{i_\theta} \le 1$, its right-hand side is always greater than or equal to $\min \left\{ \sum_{j\in[k]} w_{i_\theta j},\ \varepsilon \right\}$. Since $\max \{ w_{ij} :\ i_\theta \text{ precedes } i \text{ in } \Theta \} = 0$, it follows from the definition of $L_{W,\Theta}$ in (16) that $\sum_{j\in[k]} w_{i_\theta j} \ge L_{W,\theta}$. Then, by our assumption that $L_{W,\Theta} \ge \varepsilon$, we have $\min \left\{ \sum_{j\in[k]} w_{i_\theta j},\ \varepsilon \right\} = \varepsilon$, implying in turn that $y_1 + \cdots + y_k \ge \varepsilon$ is implied by (A-Mix), as required. ☐

We next demonstrate that when ε is large, applying the aggregated mixing inequalities is not always enough to describe the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ via an example.

Example 4. The following set is the same as (14) in Examples 2 and 3 except that $\varepsilon = 9$.

$$\left\{ (y,z) \in \mathbb{R}^2_+ \times \{0,1\}^5 : \begin{array}{c} y_1 + 8z_1 \ge 8 \\ y_1 + 6z_2 \ge 6 \\ y_1 + 13z_3 \ge 13 \\ y_1 + z_4 \ge 1 \\ y_1 + 4z_5 \ge 4 \end{array},\ \begin{array}{c} y_2 + 3z_1 \ge 3 \\ y_2 + 4z_2 \ge 4 \\ y_2 + 2z_3 \ge 2 \\ y_2 + 2z_4 \ge 2 \\ y_2 + z_5 \ge 1 \end{array},\ \mathbf{y_1 + y_2 \ge 9} \right\}. \tag{17}$$

Recall that $L_{W,\{2\to 3\}} = 8$, so $\varepsilon > L_{W,\{2\to 3\}}$ in this case. The convex hull of (17) is given by

$$\left\{ (y,z) \in \mathbb{R}^2_+ \times [0,1]^5 : \begin{array}{ll} 7y_1 + 6y_2 & +12z_2 + 49z_3 \ge 115 \\ 6y_1 + 5y_2 & +10z_2 + 42z_3 + z_4 \ge 98 \\ 3y_1 + 2y_2 & +4z_2 + 21z_3 + z_4 + 3z_5 \ge 47 \\ 3y_1 + 2y_2 & +4z_2 + 21z_3 + 4z_5 \ge 47 \\ 2y_1 + 3y_2 & +6z_2 + 14z_3 \ge 38 \\ y_1 + 2y_2 & +4z_2 + 7z_3 + z_5 \ge 21 \\ y_1 + y_2 & +z_1 + z_2 + 6z_3 \ge 17 \\ y_1 + y_2 & +2z_1 + z_2 + 5z_3 \ge 17 \end{array},\ \begin{array}{c} \text{the mixing inequalities (Mix)} \\ \text{for } j = 1, 2 \end{array} \right\}.$$

In this convex hull description, there are still two inequalities with y_1+y_2, and it turns out that these are aggregated mixing inequalities. To illustrate, take a sequence $\Theta = \{2 \to 1 \to 3\}$. We observed in Example 3 that $\{3\}$ and $\{2 \to 1 \to 3\}$ are the 1-mixing subsequence and the 2-mixing subsequence of Θ and that $L_{W,\Theta} = 9$. So, the corresponding aggregated mixing inequality (A-Mix) is $y_1 + y_2 + z_1 + z_2 + 6z_3 \geq 17$. Similarly, we obtain $y_1 + y_2 + 2z_1 + z_2 + 5z_3 \geq 17$ from $\{3 \to 2 \to 1\}$. However, unlike the system (14) in Example 2, there are facet-defining inequalities for the convex hull of this set other than the aggregated mixing inequalities, i.e., the first 6 inequalities in the above description of the convex hull have different coefficient structures on the y variables. $\square$

So, a natural question is: When are the mixing inequalities and the aggregated mixing inequalities sufficient to describe the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$? Examples 2–4 suggest that whether or not the mixing and the aggregated mixing inequalities are sufficient depends on the value of ε. In the next section, we find a necessary and sufficient condition for the sufficiency of the mixing and the aggregated mixing inequalities.

5 Joint mixing sets with lower bounds

In this section, we study the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$, where $W = \{w_{ij}\} \in \mathbb{R}_+^{n\times k}$ and $\varepsilon \in \mathbb{R}_+$. More specifically, we focus on the question of when the convex hull of this set is obtained after applying the mixing inequalities and the aggregated mixing inequalities. By Remark 1, we have $(y, z) \in \mathcal{M}(W, \mathbf{0}, \varepsilon)$ if and only if $(y, \mathbf{1} - z) \in \mathcal{P}(W, \mathbf{0}, \varepsilon)$. In Section 3, we identified that $\mathcal{P}(W, \ell, 0)$ defined as in (7) has an underlying submodularity structure (due to Lemma 1 and Proposition 1). In this section, we will first establish that $\mathcal{P}(W, \mathbf{0}, \varepsilon)$ has a similar submodularity structure for particular values of ε. In fact, for those favorable values of ε, we show that the mixing and the aggregated mixing inequalities are sufficient to describe the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ if and only if $\mathcal{P}(W, \mathbf{0}, \varepsilon)$ has the desired submodularity structure; this is the main result of this section.

5.1 Submodularity in joint mixing sets with lower bounds

In order to make a connection with submodularity, we first define the following functions $f_1, \dots, f_k, g : \{0,1\}^n \to \mathbb{R}$: for $z \in \{0,1\}^n$,

$$f_j(z) := \max_{i\in[n]} \{w_{ij} z_i\} \quad \text{for } j \in [k] \quad \text{and} \quad g(z) := \max\left\{\varepsilon, \sum_{j\in[k]} f_j(z)\right\}. \tag{18}$$

Then, we immediately arrive at the following representation of $\mathcal{P}(W, \mathbf{0}, \varepsilon)$.

Lemma 5. *Let $f_1, \dots, f_k, g : \{0,1\}^n \to \mathbb{R}$ be as defined in (18). Then,*

$$\mathcal{P}(W, \mathbf{0}, \varepsilon) = \left\{(y, z) \in \mathbb{R}^k \times \{0,1\}^n : \; y_j \geq f_j(z),\; \forall j \in [k],\;\; y_1 + \cdots + y_k \geq g(z)\right\}. \tag{19}$$

Proof. We deduce the equivalence of the relations $y_j \geq f_j(z)$ for $j \in [k]$ to the first set of constraints in

$\mathcal{P}(W,\mathbf{0},\varepsilon)$ from the corresponding definition of this set in (7). Also, we immediately have $\sum_{j\in[k]} y_j \geq \max\left\{\varepsilon,\ \sum_{j\in[k]} f_j(z)\right\}$. The result then follows from the definition of the function g. $\square$

We would like to understand the convex hull of $\mathcal{P}(W,\mathbf{0},\varepsilon)$ for $W \in \mathbb{R}_+^{n\times k}$ and $\varepsilon \in \mathbb{R}_+$ using Lemma 5. Observe that $f_1,\dots,f_k$ defined in (18) coincide with the functions $f_1,\dots,f_k$ defined in (11) for the $\ell = \mathbf{0}$ case. So, the following is a direct corollary of Lemma 1.

Corollary 3. *For any $j \in [k]$, the function f_j defined as in (18) is submodular and satisfies $f_j(\emptyset) \geq 0$.*

In contrast to the functions $f_1,\dots,f_k$, the function g is not always submodular. However, we can characterize exactly when g is submodular in terms of ε. For this characterization, we need to define several parameters based on W and ε. For a given ε, let $\bar{I}(\varepsilon)$ be the following subset of $[n]$:

$$\bar{I}(\varepsilon) := \left\{ i \in [n] :\ \sum_{j\in[k]} w_{ij} \leq \varepsilon \right\}. \tag{20}$$

With $\bar{I}(\varepsilon)$, we define another parameter $L_W(\varepsilon) \in \mathbb{R}_+$ as follows:

$$L_W(\varepsilon) := \begin{cases} \min\limits_{p,q\in[n]\setminus \bar{I}(\varepsilon)} \left\{ \sum\limits_{j\in[k]} \min\left\{w_{pj},\ w_{qj}\right\} \right\}, & \text{if } \bar{I}(\varepsilon) \neq [n], \\ +\infty, & \text{if } \bar{I}(\varepsilon) = [n]. \end{cases} \tag{21}$$

Note that $\bar{I}(\varepsilon)$ can be found in $O(kn)$ time and that $L_W(\varepsilon)$ can be computed in $O(kn^2)$ time.

Example 5. In Example 2, we have $\bar{I}(\varepsilon) = \{4,5\}$ and $L_W(\varepsilon) = w_{21} + w_{32} = 8$. Moreover, $\bar{I}(\varepsilon) = \{4,5\}$ in Example 4 as well, so we still have $L_W(\varepsilon) = 8$ in Example 4. $\square$

In Section 4, we introduced the parameter $L_{W,\Theta}$ that depends on W and a sequence Θ of indices in $[n]$ to define the aggregated mixing inequality (A-Mix) derived from Θ. The following lemma illustrates a relationship between $L_W(\varepsilon)$ and $L_{W,\Theta}$:

Lemma 6. *If $\bar{I}(\varepsilon) \neq [n]$, then $L_W(\varepsilon) = \min\left\{L_{W,\Theta} :\ \Theta \text{ is a nonempty sequence in } [n]\setminus\bar{I}(\varepsilon)\right\}$.*

Proof. Take a nonempty sequence Θ in $[n]\setminus\bar{I}(\varepsilon)$. When $\Theta = \{r\}$ for some $r \in [n]\setminus\bar{I}(\varepsilon)$, $L_{W,\Theta} = \sum_{j\in[k]} w_{rj}$, so $L_W(\varepsilon) \leq L_{W,\Theta}$ in this case. When $\Theta = \{i_1 \to \cdots \to i_\theta\}$ with $\theta \geq 2$, for any $s \in [\theta]$ we have

$$\min\left\{w_{i_s j},\ \max\left\{w_{ij} :\ i_s \text{ precedes } i \text{ in } \Theta\right\}\right\} \geq \min\left\{w_{i_s j},\ w_{i_{s+1} j}\right\}$$

where $w_{i_{\theta+1} j}$ is set to 0 for convention. Then it follows from the definition of $L_{W,\Theta}$ in (16) that $L_{W,\Theta} \geq \min\left\{\sum_{j\in[k]} \min\left\{w_{i_s j},\ w_{i_{s+1} j}\right\} :\ s \in [\theta]\right\}$. Consequently, from the definition of $L_W(\varepsilon)$, we deduce that $L_W(\varepsilon) \leq L_{W,\Theta}$. In both cases, we get $L_W(\varepsilon) \leq L_{W,\Theta}$.

Now it remains to show $L_W(\varepsilon) \geq \min\left\{L_{W,\Theta} :\ \Theta \text{ is a nonempty sequence in } [n]\setminus\bar{I}(\varepsilon)\right\}$. Since $\bar{I}(\varepsilon) \neq [n]$, either there exist distinct $p,q \in [n]\setminus\bar{I}(\varepsilon)$ such that $L_W(\varepsilon) = \sum_{j\in[k]} \min\left\{w_{pj},\ w_{qj}\right\} = L_{W,\{p\to q\}}$ or there exists $r \in [n] \setminus \bar{I}(\varepsilon)$ such that $L_W(\varepsilon) = \sum_{j\in[k]} w_{rj} = L_{W,\{r\}}$, implying in turn that $L_W(\varepsilon) \geq L_{W,\Theta}$ for some nonempty sequence Θ in $[n] \setminus \bar{I}(\varepsilon)$, as required. $\square$

Our last concept for understanding submodularity of g is the notion of ε-*negligibility*.

Definition 5. We say that $\bar{I}(\varepsilon)$ is ε-*negligible* if either

- $\bar{I}(\varepsilon) = \emptyset$ or
- $\bar{I}(\varepsilon) \neq \emptyset$ and $\bar{I}(\varepsilon)$ satisfies both of the following two conditions:

$$\max_{i\in\bar{I}(\varepsilon)} \{w_{ij}\} \leq w_{ij} \text{ for every } i \in [n] \setminus \bar{I}(\varepsilon) \text{ and } j \in [k], \tag{C1}$$

$$\sum_{j\in[k]} \max_{i\in\bar{I}(\varepsilon)} \{w_{ij}\} \leq \varepsilon. \tag{C2}$$

$\square$

Example 6. In both Examples 2 and 4, we have seen that $\bar{I}(\varepsilon) = \{4, 5\}$. Moreover, it can be readily checked that $\bar{I}(\varepsilon)$ is ε-negligible in both examples. $\square$

Let us consider examples where $\bar{I}(\varepsilon)$ is not ε-negligible.

Example 7. Let us consider Examples 2 with a slight modification. The following set is the same as (14) except that w_{42} is now 3.

$$\left\{ (y,z) \in \mathbb{R}_+^2 \times \{0,1\}^5 : \begin{array}{c} y_1 + 8z_1 \geq 8 \\ y_1 + 6z_2 \geq 6 \\ y_1 + 13z_3 \geq 13 \\ y_1 + z_4 \geq 1 \\ y_1 + 4z_5 \geq 4 \end{array} , \begin{array}{c} y_2 + 3z_1 \geq 3 \\ y_2 + 4z_2 \geq 4 \\ y_2 + 2z_3 \geq 2 \\ \mathbf{y_2 + 3z_4 \geq 3} \\ y_2 + z_5 \geq 1 \end{array} , \; y_1 + y_2 \geq 7 \right\}. \tag{22}$$

In this example, $\bar{I}(\varepsilon)$ is still $\{4, 5\}$. But, $\bar{I}(\varepsilon)$ is no longer ε-negligible because $3 \notin \bar{I}(\varepsilon)$ yet $w_{42} > w_{32}$ implying that condition (C1) is violated. The following set is the same as (14) except that w_{51} is now 6.

$$\left\{ (y,z) \in \mathbb{R}_+^2 \times \{0,1\}^5 : \begin{array}{c} y_1 + 8z_1 \geq 8 \\ y_1 + 6z_2 \geq 6 \\ y_1 + 13z_3 \geq 13 \\ y_1 + z_4 \geq 1 \\ \mathbf{y_1 + 6z_5 \geq 6} \end{array} , \begin{array}{c} y_2 + 3z_1 \geq 3 \\ y_2 + 4z_2 \geq 4 \\ y_2 + 2z_3 \geq 2 \\ y_2 + 2z_4 \geq 2 \\ y_2 + z_5 \geq 1 \end{array} , \; y_1 + y_2 \geq 7 \right\}. \tag{23}$$

Again, $\bar{I}(\varepsilon)$ is $\{4, 5\}$. However, $\bar{I}(\varepsilon)$ is not ε-negligible because $\sum_{j\in[k]} \max_{i\in\bar{I}(\varepsilon)} \{w_{ij}\} = 6 + 2 > \varepsilon$ implying that condition (C2) is violated. One can check that there are facet-defining inequalities other than the mixing and the aggregated mixing inequalities in both of these examples. For instance, $2y_1 + 3y_2 + 3z_2 + 18z_3 + 3z_4 \geq 38$ is facet-defining for the convex hull of (22) and $2y_1 + y_2 + z_1 + z_2 + 14z_3 + z_4 + 6z_5 \geq 30$ is facet-defining for the convex hull of (23). $\square$

The ε-negligibility property of a set $\bar{I}(\varepsilon)$ is closely connected to a favorable property of the g function defined in (18).

Lemma 7. *Let g be as defined in (18). If $\bar{I}(\varepsilon)$ is ε-negligible, then $g(U) = g(U \setminus \bar{I}(\varepsilon))$ for every $U \subseteq [n]$.*

Proof. Suppose $\bar{I}(\varepsilon)$ is nonempty and satisfies conditions (C1) and (C2). Take a subset U of $[n]$. If $U \subseteq \bar{I}(\varepsilon)$, then $g(U) \leq g(\bar{I}(\varepsilon))$ because g is a monotone function. Since $\sum_{j\in[k]} \max_{i\in\bar{I}(\varepsilon)} \{w_{ij}\} \leq \varepsilon$, we obtain $g(\bar{I}(\varepsilon)) = \varepsilon$ by definition of g in (18). So, $g(U) = g(\emptyset) = \varepsilon$ in this case. If $U \setminus \bar{I}(\varepsilon) \neq \emptyset$, then $\sum_{j\in[k]} w_{pj} > \varepsilon$ for some $p \in U$, implying in turn that $\sum_{j\in[k]} \max_{i\in U} \{w_{ij}\} > \varepsilon$. Moreover, as $\bar{I}(\varepsilon)$ satisfies (C1), $\sum_{j\in[k]} \max_{i\in U} \{w_{ij}\} = \sum_{j\in[k]} \max_{i\in U\setminus\bar{I}(\varepsilon)} \{w_{ij}\}$, and therefore, $g(U) = g(U \setminus \bar{I}(\varepsilon))$, as required. $\square$

We will next establish that whether the function g is submodular or not is determined entirely by $\bar{I}(\varepsilon)$ and $L_W(\varepsilon)$ defined as in (20) and (21).

Lemma 8. *The function g defined as in (18) is submodular if and only if $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$.*

Proof. ($\Rightarrow$): Assume that g is submodular. Suppose for a contradiction that $\bar{I}(\varepsilon)$ is not ε-negligible. Then $\bar{I}(\varepsilon)$ is nonempty, and (C1) or (C2) is violated. Assume that $\bar{I}(\varepsilon)$ does not satisfy (C1). Then $w_{qj} > w_{pj}$ for some $j \in [k]$, $p \in [n] \setminus \bar{I}(\varepsilon)$ and $q \in \bar{I}(\varepsilon)$. By our choice of q, we have $g(\{q\}) = \varepsilon$. Moreover, $w_{qj} > w_{pj}$ implies that $g(\{p,q\}) = \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\} > \sum_{j\in[k]} w_{pj} = g(\{p\})$. Since $g(\emptyset) = \varepsilon$, it follows that $g(\{p\}) + g(\{q\}) < g(\emptyset) + g(\{p,q\})$, a contradiction to the submodularity of g. Thus, we may assume that $\bar{I}(\varepsilon)$ does not satisfy (C2). Then $\sum_{j\in[k]} \max_{i\in\bar{I}(\varepsilon)} \{w_{ij}\} > \varepsilon$, so $g(\bar{I}(\varepsilon)) = \sum_{j\in[k]} \max_{i\in\bar{I}(\varepsilon)} \{w_{ij}\}$. Now take a minimal subset I of $\bar{I}(\varepsilon)$ with $g(I) > \varepsilon$. Since $I \subseteq \bar{I}(\varepsilon)$ and $g(I) > \varepsilon$, we know that $|I| \geq 2$. That means that one can find two nonempty subsets U, V of I partitioning I. By our minimal choice of I, we have $g(U) = g(V) = \varepsilon$, but this indicates that $g(U) + g(V) < g(\emptyset) + g(I) = g(U \cap V) + g(U \cup V)$, a contradiction to the submodularity of g. Therefore, $\bar{I}(\varepsilon)$ is ε-negligible.

Lastly, suppose for a contradiction that $\varepsilon > L_W(\varepsilon)$. Then, $L_W(\varepsilon) \neq \infty$, implying $\bar{I}(\varepsilon) \neq [n]$ and $\varepsilon > \sum_{j\in[k]} \min\{w_{pj},\, w_{qj}\}$ for some $p, q \in [n] \setminus \bar{I}(\varepsilon)$. Moreover, because both $\sum_{j\in[k]} w_{pj}$ and $\sum_{j\in[k]} w_{qj}$ are greater than ε, we deduce that p and q are distinct. Then,

$$\begin{aligned} g(\{p\}) + g(\{q\}) = \sum_{j\in[k]} w_{pj} + \sum_{j\in[k]} w_{qj} &= \sum_{j\in[k]} \max\{w_{pj},\, w_{qj}\} + \sum_{j\in[k]} \min\{w_{pj},\, w_{qj}\} \\ &= g(\{p,q\}) + \sum_{j\in[k]} \min\{w_{pj},\, w_{qj}\} < g(\{p,q\}) + g(\emptyset)\,, \end{aligned}$$

where the strict inequality follows from $g(\emptyset) = \varepsilon$. This is a contradiction to the assumption that g is submodular. Hence, $\varepsilon \leq L_W(\varepsilon)$, as required.

($\Leftarrow$): Assume that $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$. We will show that $g(U) + g(V) \geq g(U\cup V) + g(U\cap V)$ for every two sets $U, V \subseteq [n]$. If $\bar{I}(\varepsilon) = [n]$, then we have $g(U) = \varepsilon$ for every subset U of $[n]$ due to (C2). Thus, we may assume that $\bar{I}(\varepsilon) \neq [n]$. By Lemma 7, for every two subsets $U, V \subseteq [n]$, $g(U)+g(V) \geq g(U\cup V)+g(U\cap V)$ holds if and only if $g(U') + g(V') \geq g(U' \cup V') + g(U' \cap V')$, where $U' := U \setminus \bar{I}(\varepsilon)$ and $V' := V \setminus \bar{I}(\varepsilon)$, holds. This means that it is sufficient to consider subsets of $[n] \setminus \bar{I}(\varepsilon)$. Consider two sets $U, V \subseteq [n] \setminus \bar{I}(\varepsilon)$. If $U = \emptyset$ or $V = \emptyset$, the inequality trivially holds due to the monotonicity of g. So, we may assume that $U, V \neq \emptyset$. First, suppose that $U \cap V \neq \emptyset$. Because $U, V \subseteq [n] \setminus \bar{I}(\varepsilon)$, we deduce that $g(X) = \sum_{j\in[k]} f_j(X)$ for any $X \in \{U, V, U\cup V, U\cap V\}$. Then, Corollary 3 implies that $g(U)+g(V) \geq g(U\cup V)+g(U\cap V)$. Now, consider the case of $U \cap V = \emptyset$. Note that for each $j \in [k]$, the definition of $f_j(V) = \max_{i\in V}\{w_{ij}\}$ implies that

$$f_j(U)+f_j(V)-f_j(U\cup V) = \max\{f_j(U), f_j(V)\}+\min\{f_j(U), f_j(V)\}-f_j(U\cup V) = \min\{f_j(U), f_j(V)\}.$$

Hence, we have

$$g(U) + g(V) - g(U \cup V) = \sum_{j\in[k]} (f_j(U) + f_j(V) - f_j(U \cup V)) = \sum_{j\in[k]} \min\{f_j(U), f_j(V)\}.$$

So, it suffices to argue that $\sum_{j\in[k]} \min\{f_j(U), f_j(V)\} \geq g(\emptyset) = \varepsilon$. Since $U, V \neq \emptyset$ and $U \cap V = \emptyset$, there exist distinct $p, q \in [n]\setminus\bar{I}(\varepsilon)$ such that $p \in U$ and $q \in V$. Then $f_j(U) \geq f_j(\{p\}) = w_{pj}$ and $f_j(V) \geq f_j(\{q\}) = w_{qj}$, implying in turn that

$$\sum_{j\in[k]} \min\{f_j(U), f_j(V)\} \geq \sum_{j\in[k]} \min\{w_{pj}, w_{qj}\} \geq L_W(\varepsilon),$$

where the last inequality follows from the definition of $L_W(\varepsilon)$ in (21). Finally, our assumption that $\varepsilon \leq L_W(\varepsilon)$ implies that $\sum_{j\in[k]} \min\{f_j(U), f_j(V)\} \geq \varepsilon$ as desired. $\square$

Therefore, Lemma 8, along with Corollary 3, establish that $f_1, \dots, f_k$ and g are submodular when $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$.

5.2 Polymatroid inequalities and aggregated mixing inequalities

Consider $\mathcal{P}(W, \mathbf{0}, \varepsilon)$ with $W \in \mathbb{R}_+^{n\times k}$ and $\varepsilon \in \mathbb{R}_+$. Then, from Lemma 5 we deduce that

$$\operatorname{conv}(\mathcal{P}(W, \mathbf{0}, \varepsilon)) \subseteq \left\{(y, z) \in \mathbb{R}^k \times [0, 1]^n : \ (y_j, z) \in \operatorname{conv}(Q_{f_j}), \ \forall j \in [k], \ \ (y_1 + \cdots + y_k, z) \in \operatorname{conv}(Q_g)\right\},$$

where f_j, g are as defined in (18). In this section we will prove that in fact equality holds in the above relation when g is submodular, i.e., by Lemma 8, when $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$. Then, consequently, if $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$, then the separation problem over $\operatorname{conv}(\mathcal{P}(W, \mathbf{0}, \varepsilon))$ (equivalently, $\operatorname{conv}(\mathcal{M}(W, \mathbf{0}, \varepsilon))$) can be solved in $O(kn \log n)$ time by a simple greedy algorithm. To this end, we first characterize the $\mathcal{V}$-polyhedral, or inner, description of $\operatorname{conv}(\mathcal{P}(W, \mathbf{0}, \varepsilon))$. For notational purposes, we define a specific set of binary solutions as follows:

$$S(\varepsilon) := \left\{ z \in \{0, 1\}^n : \sum_{j\in[k]} \max_{i\in[n]} \{w_{ij} z_i\} > \varepsilon \right\}. \tag{24}$$

Lemma 9. *The extreme rays of* $\mathrm{conv}(\mathcal{P}(W, \mathbf{0}, \varepsilon))$ *are* $(e^j, \mathbf{0})$ *for* $j \in [k]$*, and the extreme points are precisely the following:*

- $A(z) = (y^z, z)$ *for* $z \in S(\varepsilon)$ *where* $y_j^z = \max_{i \in [n]} \{w_{ij} z_i\}$ *for* $j \in [k]$*,*
- $B(z, d) = (y^{z,d}, z)$ *for* $z \in \{0,1\}^n \setminus S(\varepsilon)$ *and* $d \in [k]$ *where*

$$y_j^{z,d} = \begin{cases} \max_{i \in [n]} \{w_{ij} z_i\}, & \text{if } j \neq d, \\ \max_{i \in [n]} \{w_{id} z_i\} + \left(\varepsilon - \sum_{j \in [k]} \max_{i \in [n]} \{w_{ij} z_i\}\right), & \text{if } j = d. \end{cases}$$

Proof. It is clear that $(e^j, \mathbf{0})$ for $j \in [k]$ are the extreme rays of $\mathrm{conv}(\mathcal{P}(W, \mathbf{0}, \varepsilon))$. Let $(\bar{y}, \bar{z})$ be an extreme point of $\mathrm{conv}(\mathcal{P}(W, \mathbf{0}, \varepsilon))$. Then $\bar{z} \in \{0,1\}^n$, and constraints (8) become $\bar{y}_j \geq \max_{i \in [n]} \{w_{ij} \bar{z}_i\}$ for $j \in [k]$. If $\bar{z} \in S(\varepsilon)$, then $\sum_{j \in [k]} \max_{i \in [n]} \{w_{ij} \bar{z}_i\} > \varepsilon$, so $(\bar{y}, \bar{z})$ automatically satisfies (9)–(10). As $(\bar{y}, \bar{z})$ is an extreme point, it follows that $\bar{y}_j = \max_{i \in [n]} \{w_{ij} \bar{z}_i\}$ for $j \in [k]$, and therefore, $(\bar{y}, \bar{z}) = A(\bar{z})$. If $\bar{z} \notin S(\varepsilon)$, then $\sum_{j \in [k]} \max_{i \in [n]} \{w_{ij} \bar{z}_i\} \leq \varepsilon$. Since $(\bar{y}, \bar{z})$ satisfies $\bar{y}_1 + \cdots + \bar{y}_k \geq \varepsilon$ and $(\bar{y}, \bar{z})$ cannot be expressed as a convex combination of two distinct points, it follows that $\bar{y}_1 + \cdots + \bar{y}_k \geq \varepsilon$ and constraints $\bar{y}_j \geq \max_{i \in [n]} \{w_{ij} \bar{z}_i\}, j \in [k] \setminus \{d\}$ are tight at $(\bar{y}, \bar{z})$ for some $d \in [k]$, so $(\bar{y}, \bar{z}) = B(z, d)$. $\square$

Based on the definition of $S(\varepsilon)$ and (18), we have

$$g(z) = \max\left\{\varepsilon, \sum_{j \in [k]} f_j(z)\right\} = \begin{cases} \sum_{j \in [k]} f_j(z), & \text{if } z \in S(\varepsilon) \\ \varepsilon, & \text{if } z \notin S(\varepsilon). \end{cases}$$

Remember the definition of $\bar{I}(\varepsilon)$ in (20) and the conditions for $\bar{I}(\varepsilon)$ to be ε-negligible. Recall the definition of $L_W(\varepsilon)$ in (21) as well. Based on these definitions and Proposition 2, we are now ready to give the explicit inequality characterization of the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$.

Proposition 5. *Let* $W = \{w_{ij}\} \in \mathbb{R}_+^{n \times k}$ *and* $\varepsilon \in \mathbb{R}_+$*. If* $\bar{I}(\varepsilon)$ *is* ε*-negligible and* $\varepsilon \leq L_W(\varepsilon)$*, then the convex hull of* $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ *is given by*

$$\left\{(y, z) \in \mathbb{R}^k \times [0,1]^n : \ (y_j, \mathbf{1} - z) \in \mathrm{conv}(Q_{f_j}), \ \forall j \in [k], \ \ (y_1 + \cdots + y_k, \mathbf{1} - z) \in \mathrm{conv}(Q_g)\right\}.$$

Proof. We will show that $y_1, \ldots, y_k$ and $\sum_{j \in [k]} y_j$ are weakly independent with respect to submodular functions $f_1, \ldots, f_k$ and g (recall Definition 1). Consider $\alpha \in \mathbb{R}_+^k \setminus \{\mathbf{0}\}$, and let $\alpha_{\min}$ denote the smallest coordinate value of α. Then α and $\alpha^\top y$ can be written as $\alpha = \alpha_{\min} \mathbf{1} + \sum_{j \in [k]} (\alpha_j - \alpha_{\min}) e^j$ and $\alpha^\top y = \alpha_{\min} \sum_{j \in [k]} y_j + \sum_{j \in [k]} (\alpha_j - \alpha_{\min}) y_j$. Let f_α be defined as $f_\alpha(z) := \min\left\{\alpha^\top y : \ (y, z) \in \mathcal{P}(W, \mathbf{0}, \varepsilon)\right\}$ for $z \in \{0,1\}^n$. Then, it is sufficient to show that $f_\alpha = \alpha_{\min} g + \sum_{j \in [k]} (\alpha_j - \alpha_{\min}) f_j$.

Let $\bar{z} \in \{0,1\}^n$. For any y with $(y, \bar{z}) \in \mathcal{P}(W, \mathbf{0}, \varepsilon)$, we have $y_j \geq f_j(\bar{z})$ for $j \in [k]$ and $\sum_{j \in [k]} y_j \geq g(\bar{z})$ by Lemma 5, implying in turn that

$$f_\alpha(\bar{z}) = \min\left\{\alpha^\top y : (y, \bar{z}) \in \mathcal{P}(W, \mathbf{0}, \varepsilon)\right\} \geq \alpha_{\min} g(\bar{z}) + \sum_{j \in [k]} (\alpha_j - \alpha_{\min}) f_j(\bar{z}). \tag{25}$$

Recall the definition of $S(\varepsilon)$ in (24). If $\bar{z} \in S(\varepsilon)$, then $g(\bar{z}) = \sum_{j\in[k]} f_j(\bar{z})$, and therefore, $A(\bar{z}) = (y^{\bar{z}}, \bar{z})$ defined in Lemma 9 satisfies (25) at equality. If $\bar{z} \notin S(\varepsilon)$, then $g(\bar{z}) = \varepsilon$. Let $d \in [k]$ be the index satisfying $\alpha_d = \alpha_{\min}$. Then $B(\bar{z}, d) = (y^{\bar{z},d}, \bar{z})$ defined in Lemma 9 satisfies (25) at equality. Therefore, we deduce that $f_\alpha = \alpha_{\min} g + \sum_{j\in[k]}(\alpha_j - \alpha_{\min}) f_j$.

From Proposition 2 applied to (19), we obtain that $\mathrm{conv}(\mathcal{P}(W, \mathbf{0}, \varepsilon))$ is equal to

$$\left\{(y,z) \in \mathbb{R}^k \times [0,1]^n : \ (y_j, z) \in \mathrm{conv}(Q_{f_j}),\ \forall j \in [k],\ \ (y_1 + \cdots + y_k, z) \in \mathrm{conv}(Q_g)\right\}.$$

After complementing the z variables, we obtain the desired description of $\mathrm{conv}(\mathcal{M}(W, \mathbf{0}, \varepsilon))$. This finishes the proof. $\square$

Proposition 5 indicates that if $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \le L_W(\varepsilon)$, then the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ is described by the polymatroid inequalities of f_j with left-hand side y_j for $j \in [k]$ and the polymatroid inequalities of g with left-hand side $\sum_{j\in[k]} y_j$. We have seen in Section 3 that the polymatroid inequalities of f_j with left-hand side y_j for $j \in [k]$ are nothing but the mixing inequalities. In fact, it turns out that an extremal polymatroid inequality of g with left-hand side $\sum_{j\in[k]} y_j$ is either the linking constraint $y_1 + \cdots + y_k \ge \varepsilon$ or an aggregated mixing inequality, depending on whether or not $\bar{I}(\varepsilon) = [n]$. We consider the $\bar{I}(\varepsilon) = [n]$ case first.

Proposition 6. *Assume that $\bar{I}(\varepsilon) = [n]$ and $\bar{I}(\varepsilon)$ is ε-negligible. Then for every extreme point π of $EP_{\tilde{g}}$, the corresponding polymatroid inequality $\sum_{j\in[k]} y_j + \sum_{i\in[n]} \pi_i z_i \ge \varepsilon + \sum_{i\in[n]} \pi_i$ is equivalent to the linking constraint.*

Proof. By Theorem 2.2, there exists a permutation σ of $[n]$ such that $\pi_{\sigma(t)} = g(V_t) - g(V_{t-1})$ where $V_t = \{\sigma(1), \ldots, \sigma(t)\}$ for $t \in [n]$ and $V_0 = \emptyset$. Since $\bar{I}(\varepsilon) = [n]$ and $\bar{I}(\varepsilon)$ is ε-negligible, it follows that $g(U) = \varepsilon$ for every $U \subseteq [n]$, so $\pi_{\sigma(t)} = 0$ for all t. Therefore, $\sum_{j\in[k]} y_j + \sum_{i\in[n]} \pi_i z_i \ge \varepsilon + \sum_{i\in[n]} \pi_i$ equals $\sum_{j\in[k]} y_j \ge \varepsilon$, as required. $\square$

The $\bar{I}(\varepsilon) \ne [n]$ case is more interesting; the following proposition is similar to Proposition 3:

Proposition 7. *Assume that $\bar{I}(\varepsilon) \ne [n]$ is ε-negligible and $\varepsilon \le L_W(\varepsilon)$. Then for every extreme point π of $EP_{\tilde{g}}$, there exists a sequence $\Theta = \{i_1 \to \cdots \to i_\theta\}$ contained in $[n] \setminus \bar{I}(\varepsilon)$ that satisfies the following:*

(1) *the j-mixing-subsequence $\{j_1 \to \cdots \to j_{\tau_j}\}$ of Θ satisfies $w_{j_1 j} = \max\{w_{ij} : i \in [n]\}$ for each $j \in [k]$,*

(2) *the corresponding polymatroid inequality $\sum_{j\in[k]} y_j + \sum_{i\in[n]} \pi_i z_i \ge \varepsilon + \sum_{i\in[n]} \pi_i$ is equivalent to the aggregated mixing inequality (A-Mix) derived from Θ.*

In particular, the polymatroid inequality is of the form

$$\sum_{j\in[k]} \left(y_j + \sum_{s\in[\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) z_{j_s} \right) - \varepsilon z_{i_\theta} \ge \sum_{j\in[k]} \max\{w_{ij} : i \in [n]\}. \tag{A-Mix*}$$

Proof. By Theorem 2.2, there exists a permutation σ of $[n]$ such that $\pi_{\sigma(t)} = g(V_t) - g(V_{t-1})$ where $V_t = \{\sigma(1), \ldots, \sigma(t)\}$ for $t \in [n]$ and $V_0 = \emptyset$. By Lemma 7, $g(V_t) - g(V_{t-1}) = g(V_t \setminus \bar{I}(\varepsilon)) - g(V_{t-1} \setminus \bar{I}(\varepsilon))$, so $\pi_{\sigma(t)}$ is nonzero only if $\sigma(t) \notin \bar{I}(\varepsilon)$. This in turn implies that at most $|n \setminus \bar{I}(\varepsilon)|$ coordinates of π are nonzero. Let $\{t_1, \ldots, t_\theta\}$ be the collection of t's such that $\pi_{\sigma(t)} \neq 0$. Then $1 \leq \theta \leq |n \setminus \bar{I}(\varepsilon)|$. Without loss of generality, we may assume that $t_1 > \cdots > t_\theta$. Let $i_1 = \sigma(t_1), i_2 = \sigma(t_2), \ldots, i_\theta = \sigma(t_\theta)$, and Θ denote the sequence $\{i_1 \to \cdots \to i_\theta\}$. We will show that Θ satisfies conditions **(1)** and **(2)** of the proposition.

(1): For $j \in [k]$, let $\Theta_j = \left\{j_1 \to \cdots \to j_{\tau_j}\right\}$ denote the j-mixing-subsequence of Θ. By definition of the j-mixing-subsequence of Θ, we have $w_{j_1 j} = \max\{w_{ij} : i \in \Theta\}$. By our choice of $\{t_1, \ldots, t_\theta\}$ and assumption that $t_1 > \cdots > t_\theta$, it follows that $g(V_{t_1}) = g([n])$, which means that $f_j(V_{t_1}) = f_j([n])$ for each $j \in [k]$. Therefore, we deduce that $\max\{w_{ij} : i \in \Theta\} = \max\{w_{ij} : i \in [n]\}$, as required.

(2): By convention, we have $w_{i_{\theta+1} j} = w_{j_{\tau_j+1} j} = 0$ for $j \in [k]$. In addition, due to our choice of $\{t_1, \ldots, t_\theta\}$, we have $g(V_{t_s}) > g(V_{t_s - 1}) = \cdots = g(V_{t_{s+1}})$ holds for $s < \theta$. Then, we obtain

$$\begin{aligned} \pi_{i_s} = \pi_{\sigma(t_s)} = g(V_{t_s}) - g(V_{t_{s+1}}) &= \sum_{j \in [k]} f_j(V_{t_s}) - \sum_{j \in [k]} f_j(V_{t_{s+1}}) \\ &= \sum_{j \in [k]} f_j(\{i_\theta, i_{\theta-1}, \ldots, i_s\}) - \sum_{j \in [k]} f_j(\{i_\theta, i_{\theta-1}, \ldots, i_{s+1}\}). \end{aligned}$$

We observed before that $g(V_{t_s}) > g(V_{t_s - 1}) = \cdots = g(V_{t_{s+1}})$, so it follows that $f_j(V_{t_s}) \geq f_j(V_{t_s - 1}) = \cdots = f_j(V_{t_{s+1}})$, implying in turn that

$$f_j(\{i_\theta, i_{\theta-1}, \ldots, i_s\}) - f_j(\{i_\theta, i_{\theta-1}, \ldots, i_{s+1}\}) = \left(w_{i_s j} - \max\left\{w_{ij} :\ i_s \text{ precedes } i \text{ in } \Theta\right\}\right)_+ .$$

This means that for $s < \theta$,

$$\pi_{i_s} = \sum_{j \in [k]} \left(w_{i_s j} - \max\left\{w_{ij} :\ i_s \text{ precedes } i \text{ in } \Theta\right\}\right)_+ . \tag{26}$$

Note that

$$\pi_{i_\theta} = \pi_{\sigma(t_\theta)} = g(V_{t_\theta}) - g(V_0) = \sum_{j \in [k]} f_j(V_{t_\theta}) - \varepsilon = \sum_{j \in [k]} f_j(\{i_\theta\}) - \varepsilon.$$

Since $f_j(\{i_\theta\}) = w_{i_\theta j}$ and $\max\left\{w_{ij} :\ i_\theta \text{ precedes } i \text{ in } \Theta\right\}$ was set to $w_{j_{\tau_j+1} j} = 0$, it follows that

$$\pi_{i_\theta} = \sum_{j \in [k]} \left(w_{i_\theta j} - \max\left\{w_{ij} :\ i_\theta \text{ precedes } i \text{ in } \Theta\right\}\right)_+ - \varepsilon. \tag{27}$$

Therefore, by (26) and (27), it follows that the polymatroid inequality $\sum_{j \in [k]} y_j + \sum_{i \in [n]} \pi_i z_i \geq \varepsilon + \sum_{i \in [n]} \pi$ is precisely (A-Mix*). Since $\varepsilon \leq L_W(\varepsilon)$ by our assumption and $L_W(\varepsilon) \leq L_{W,\Theta}$ by Lemma 6, $\min\{\varepsilon,\ L_{W,\Theta}\} = \varepsilon$, and thus the inequality (A-Mix*) is identical to the aggregated mixing inequality (A-Mix) derived from Θ, as required. $\square$

By Propositions 5, 6, and 7, if $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$, then the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ can be described by the mixing and the aggregated mixing inequalities together with the linking constraint

$y_1 + \cdots + y_k \geq \varepsilon$ and the bounds $\mathbf{0} \leq z \leq \mathbf{1}$. Another implication is that if $\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$, then the aggregated mixing inequalities other than the ones of the form (A-Mix*) are not necessary.

5.3 Necessary conditions for obtaining the convex hull by the mixing and the aggregated mixing inequalities

In Example 4, $\bar{I}(\varepsilon)$ is ε-negligible but $\varepsilon > L_W(\varepsilon)$ (see Examples 5 and 6). In Example 7, $\bar{I}(\varepsilon)$ is not ε-negligible in each of the two joint mixing sets with lower bounds. These examples already demonstrate that the mixing and the aggregated mixing inequalities are not sufficient whenever ε-negligibility condition or the condition $\varepsilon \leq L_W(\varepsilon)$ does not hold. In this section, we will formally show in general that $\bar{I}(\varepsilon)$ being ε-negligible and $\varepsilon \leq L_W(\varepsilon)$ are necessary for the mixing and the aggregated mixing inequalities to be sufficient to describe the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$. The following is the main result of this section.

Theorem 5.1. *Let $W = \{w_{ij}\} \in \mathbb{R}_+^{n\times k}$ and $\varepsilon \geq 0$. Let $\bar{I}(\varepsilon)$ and $L_W(\varepsilon)$ be defined as in (20) and (21), respectively. Then the following statements are equivalent:*

(i) *$\bar{I}(\varepsilon)$ is ε-negligible and $\varepsilon \leq L_W(\varepsilon)$,*

(ii) *the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ can be described by the mixing inequalities (Mix) and the aggregated mixing inequalities (A-Mix) together with the linking constraint $y_1 + \cdots + y_k \geq \varepsilon$ and the bounds $\mathbf{0} \leq z \leq \mathbf{1}$, and*

(iii) *the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$ can be described by the mixing inequalities of the form (Mix*) and the aggregated mixing inequalities of the form (A-Mix*) together with the linking constraint $y_1 + \cdots + y_k \geq \varepsilon$ and the bounds $\mathbf{0} \leq z \leq \mathbf{1}$.*

Proof. Propositions 5, 6 and 7 already prove that **(i)⇒(iii)**, and the direction **(iii)⇒(ii)** is trivial. Thus, what remains is to show **(ii)⇒(i)**. We will prove the contrapositive of this direction. It is sufficient to exhibit a point $(\bar{y}, \bar{z})$ with $\sum_{j\in[k]} \bar{y}_j \geq \varepsilon$ and $\mathbf{0} \leq \bar{z} \leq \mathbf{1}$ that satisfies the mixing and the aggregated mixing inequalities but is not contained in the convex hull of $\mathcal{M}(W, \mathbf{0}, \varepsilon)$.

Assume first that $\bar{I}(\varepsilon)$ is not ε-negligible. Then $\bar{I}(\varepsilon)$ is nonempty and either (C1) or (C2) is violated. First, consider the case when (C2) is violated. Take a minimal subset U of $\bar{I}(\varepsilon)$ satisfying $\sum_{j\in[k]} \max_{i\in U}\{w_{ij}\} > \varepsilon$. Note that by definition of $\bar{I}(\varepsilon)$, we have for every $i \in \bar{I}(\epsilon)$ that $\sum_{j\in[k]} w_{ij} \leq \varepsilon$. Then by the assumption that $\sum_{j\in[k]} \max_{i\in U}\{w_{ij}\} > \varepsilon$, we deduce that $|U| \geq 2$. Moreover, by our minimal choice of U, we have $\sum_{j\in[k]} \max_{i\in V}\{w_{ij}\} \leq \varepsilon$ for any $V \subset U$ such that $|V| \leq |U| - 1$. Moreover, for each $j \in [k]$, the largest element of $\{w_{ij} : i \in U\}$ is contained in $|U| - 1$ subsets of $\{w_{ij} : i \in U\}$ of size $|U| - 1$, while the second largest element of U is the largest in another subset of size $|U| - 1$. From these observations, we deduce that

$$(|U| - 1) \sum_{j\in[k]} \max_{i\in U}\{w_{ij}\} + \sum_{j\in[k]} \text{second-max}\{w_{ij} : i \in U\} = \sum_{\substack{V\subset U\\ |V|=|U|-1}} \left(\sum_{j\in[k]} \max_{i\in V}\{w_{ij}\} \right) \leq |U|\,\varepsilon \quad (28)$$

where second-max$\{w_{ij} : i \in U\}$ denotes the second largest value in $\{w_{ij} : i \in U\}$ for $j \in [k]$. Let us consider $(\bar{y}, \bar{z})$ where

$$\bar{z}_i = \begin{cases} \frac{1}{|U|} & \text{if } i \in U \\ 1 & \text{if } i \notin U \end{cases} \quad \text{and} \quad \bar{y}_j = \begin{cases} \frac{|U|-1}{|U|} \max\limits_{i \in U} \{w_{ij}\}, & \text{if } j \in [k-1] \\ \frac{|U|-1}{|U|} \max\limits_{i \in U} \{w_{ik}\} + \left(\varepsilon - \frac{|U|-1}{|U|} \sum\limits_{j \in [k]} \max\limits_{i \in U} \{w_{ij}\}\right), & \text{if } j = k \end{cases}$$

Then, we always have $\sum_{j \in [k]} \bar{y}_j = \varepsilon$. This, together with (28), implies that

$$\sum_{j \in [k]} \bar{y}_j = \varepsilon \geq \frac{|U|-1}{|U|} \sum_{j \in [k]} \max_{i \in U} \{w_{ij}\} + \frac{1}{|U|} \sum_{j \in [k]} \text{second-max}\{w_{ij} : i \in U\}. \tag{29}$$

Then, from $W \in \mathbb{R}_+^{n \times k}$ we deduce $\sum_{j \in [k]}$ second-max$\{w_{ij} : i \in U\} \geq 0$, and hence $\bar{y}_k \geq \frac{|U|-1}{|U|} \max\limits_{i \in U} \{w_{ik}\}$. Let us argue that $(\bar{y}, \bar{z})$ satisfies the mixing and the aggregated mixing inequalities. Take a j-mixing-sequence $\{j_1 \to \cdots \to j_{\tau_j}\}$. Since $\sum_{s \in [\tau_j]} (w_{j_s j} - w_{j_{s+1} j}) = w_{j_1 j}$, $(\bar{y}, \bar{z})$ satisfies (Mix) if and only if

$$\bar{y}_j \geq \frac{|U|-1}{|U|} \sum_{j_s \in U} \left(w_{j_s j} - w_{j_{s+1} j}\right).$$

As $\sum_{j_s \in U} \left(w_{j_s j} - w_{j_{s+1} j}\right) \leq \max\limits_{i \in U} \{w_{ij}\}$, it follows that $(\bar{y}, \bar{z})$ satisfies (Mix). Now we argue that $(\bar{y}, \bar{z})$ satisfies every aggregated mixing inequality. By Lemma 4, it is sufficient to argue this for only the sequences $\Theta = \{i_1 \to \cdots \to i_\theta\}$ that are contained in U. By (15), $(\bar{y}, \bar{z})$ satisfies (A-Mix) for Θ if and only if

$$\sum_{j \in [k]} \left(\bar{y}_j + \sum_{t \in [\Theta]} \left(w_{i_t j} - \max\{w_{ij} : i_t \text{ precedes } i \text{ in } \Theta\}\right)_+ \bar{z}_{i_t} \right) - \min\{\varepsilon, L_{W,\Theta}\} \bar{z}_{i_\theta} \geq \sum_{j \in [k]} \max\{w_{ij} : i \in \Theta\}. \tag{30}$$

Since $\Theta \subseteq U$, we have $\bar{z}_{i_1} = \cdots = \bar{z}_{i_\theta} = \frac{1}{|U|}$. Then, (30) is exactly

$$\sum_{j \in [k]} \bar{y}_j \geq \frac{|U|-1}{|U|} \sum_{j \in [k]} \max\{w_{ij} : i \in \Theta\} + \frac{1}{|U|} \min\{\varepsilon, L_{W,\Theta}\}. \tag{31}$$

Recall that $\sum_{j \in [k]} \bar{y}_j = \varepsilon$. If $|\Theta| = 1$, then because $|U| \geq 2$ we deduce $\Theta \neq U$. Moreover, because $|\Theta| = 1$ and Θ is a proper subset of $\bar{I}(\epsilon)$, we deduce from the definition of $\bar{I}(\epsilon)$ that $\sum_{j \in [k]} \max\{w_{ij} : i \in \Theta\} \leq \varepsilon$. Hence, when $|\Theta| = 1$, we also have $\min\{\varepsilon, L_{W,\Theta}\} \leq \varepsilon$, and thus (31) clearly holds. So, we may assume that $|\Theta| \geq 2$. By definition of $L_{W,\Theta}$ in (16), we have $L_{W,\Theta} \leq \sum_{j \in [k]}$ second-max$\{w_{ij} : i \in \Theta\}$ where second-max$\{w_{ij} : i \in \Theta\}$ denotes the second largest element in $\{w_{ij} : i \in \Theta\}$. Since $\max\limits_{i \in \Theta} \{w_{ij}\} \leq \max\limits_{i \in U} \{w_{ij}\}$ and second-max$\{w_{ij} : i \in \Theta\} \leq$ second-max$\{w_{ij} : i \in U\}$ hold because $\Theta \subseteq U$, we deduce from (29) that (31) holds. Consequently, Lemma 4 implies that $(\bar{y}, \bar{z})$ satisfies the aggregated mixing inequalities (A-Mix) for all sequences as well. Let us now show that $(\bar{y}, \bar{z})$ is not contained in $\mathrm{conv}(\mathcal{M}(W, \mathbf{0}, \varepsilon))$. Observe that $(\bar{y}, \bar{z})$ satisfies the constraints $z_i \leq 1$ for $i \notin U$ at equality. For $j \in [k-1]$, let $\{j_1 \to \cdots \to j_{|U|}\}$ be an ordering of the indices in U such that $w_{j_1 j} \geq \cdots \geq w_{j_{|U|} j}$. Then $\{j_1\}, \{j_1 \to j_2\}, \ldots, \{j_1 \to \cdots \to j_{|U|}\}$ are all j-mixing-sequences, and notice that $(\bar{y}, \bar{z})$ satisfies the mixing inequalities corresponding to all these j-mixing-sequences at equality. In particular, it follows that $(\bar{y}, \bar{z})$ satisfies $z_{j_1} = z_{j_2} = \cdots = z_{j_{|U|}}$ at equality. There are only two points in $\{0,1\}^n$ that satisfy both of the constraints $z_i \le 1$ for $i \notin U$ and $z_{j_1} = z_{j_2} = \cdots = z_{j_{|U|}}$ at equality; these points are $\mathbf{1}$ and $\mathbf{1}_{[n]\setminus U}$. Let $y^1, y^2 \in \mathbb{R}^k$ be such that $(y^1, \mathbf{1}), (y^2, \mathbf{1}_{[n]\setminus U}) \in \mathcal{M}(W, \mathbf{0}, \varepsilon)$. Then we have

$$\sum_{j\in[k]} y_j^1 \ge \varepsilon \quad \text{and} \quad \sum_{j\in[k]} y_j^2 \ge \sum_{j\in[k]} \max_{i\in U}\{w_{ij}\}.$$

As $\sum_{j\in[k]} \max\limits_{i\in U}\{w_{ij}\} > \varepsilon$ by our assumption and $\sum_{j\in[k]} \bar{y}_j = \varepsilon$, $(\bar{y}, \bar{z})$ cannot be a convex combination of $(y^1, \mathbf{1})$ and $(y^2, \mathbf{1}_{[n]\setminus U})$, implying in turn that $(\bar{y}, \bar{z})$ does not belong to $\operatorname{conv}(\mathcal{M}(W, \mathbf{0}, \varepsilon))$.

Now consider the case when (C1) is violated. Then there exist $p \in [n] \setminus \bar{I}(\varepsilon)$ and $q \in \bar{I}(\varepsilon)$ such that $w_{qj} > w_{pj}$ for some $j \in [k]$. In particular, $\sum_{j\in[k]} w_{pj} < \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\}$. Let us consider the point $(\bar{y}, \bar{z})$ where

$$\bar{z}_i = \begin{cases} \frac{1}{2} & \text{if } i \in \{p,q\} \\ 1 & \text{if } i \notin \{p,q\} \end{cases}, \quad \text{and}$$

$$\bar{y}_j = \begin{cases} \frac{1}{2}\max\{w_{pj},\ w_{qj}\}, & \text{if } j \in [k-1] \\ \frac{1}{2}\max\{w_{pk},\ w_{qk}\} + \frac{1}{2}\left(\varepsilon + \sum\limits_{j\in[k]} w_{pj} - \sum\limits_{j\in[k]} \max\{w_{pj},\ w_{qj}\}\right), & \text{if } j = k \end{cases}$$

By definition of $\bar{y}$, we always have $\sum_{j\in[k]} \bar{y}_j = \frac{1}{2}\left(\varepsilon + \sum\limits_{j\in[k]} w_{pj}\right) > \varepsilon$, where the inequality follows from $p \notin \bar{I}(\epsilon)$. Moreover, as $p \in [n] \setminus \bar{I}(\varepsilon)$ and $q \in \bar{I}(\varepsilon)$, we have $\sum_{j\in[k]} w_{pj} > \varepsilon \ge \sum_{j\in[k]} w_{qj}$, and hence

$$\varepsilon + \sum_{j\in[k]} w_{pj} - \sum_{j\in[k]} \max\{w_{pj},\ w_{qj}\} \ge \sum_{j\in[k]} w_{qj} + \sum_{j\in[k]} w_{pj} - \sum_{j\in[k]} \max\{w_{pj},\ w_{qj}\} = \sum_{j\in[k]} \min\{w_{pj},\ w_{qj}\} \ge 0,$$

where the last inequality follows from the fact that $w_{ij} \ge 0$ for all $i \in [n]$ and $j \in [k]$. So, it follows that

$$\bar{y}_k \ge \frac{1}{2}\max\{w_{pk},\ w_{qk}\}.$$

As before, we can argue that $(\bar{y}, \bar{z})$ satisfies the mixing inequalities. Now we argue that $(\bar{y}, \bar{z})$ satisfies every aggregated mixing inequality. By Lemma 4, it is sufficient to consider only the sequences $\Theta = \{i_1 \to \cdots \to i_\theta\}$ that are contained in $\{p, q\}$. Since $\Theta \subseteq \{p, q\}$, we know that $\bar{z}_{i_1} = \cdots = \bar{z}_{i_\theta} = \frac{1}{2}$. Then, the following inequality (32) implies (30).

$$\sum_{j\in[k]} \bar{y}_j = \frac{1}{2}\left(\varepsilon + \sum_{j\in[k]} w_{pj}\right) \ge \frac{1}{2}\min\{\varepsilon,\ L_{W,\Theta}\} + \frac{1}{2}\sum_{j\in[k]} \max\{w_{ij} :\ i \in \Theta\}. \tag{32}$$

When Θ contains both p and q, we have $L_{W,\Theta} = \sum_{j\in[k]} \min\{w_{pj}, w_{qj}\} \le \sum_{j\in[k]} w_{qj} \le \varepsilon$ (since $q \in \bar{I}(\epsilon)$) and $\sum_{j\in[k]} \max\{w_{ij} :\ i \in \Theta\} = \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\}$. Then the right-hand side of (32) is

$$\frac{1}{2}\left(\sum_{j\in[k]} \min\{w_{pj}, w_{qj}\} + \sum_{j\in[k]} \min\{w_{pj}, w_{qj}\}\right) = \frac{1}{2}\sum_{j\in[k]} w_{pj} + \frac{1}{2}\sum_{j\in[k]} w_{qj},$$

so inequality (32) holds in this case since $q \in \bar{I}(\epsilon)$. If $\Theta = \{p\}$ or $\Theta = \{q\}$, inequality (32) clearly holds. Consequently, Lemma 4 implies that $(\bar{y}, \bar{z})$ satisfies the aggregated mixing inequalities (A-Mix) for all sequences as well. Suppose for a contradiction that $(\bar{y}, \bar{z})$ is a convex combination of two points (y^1, z^1) and (y^2, z^2) in $\mathcal{M}(W, \mathbf{0}, \varepsilon)$. As the previous case, we can argue that z^1 and z^2 satisfy $z_p = z_q$ and $z_i \leq 1$ for $i \notin \{p, q\}$ at equality, and therefore, $z^1 = \mathbf{1}$ and $z^2 = \mathbf{1}_{[n]\setminus\{p,q\}}$. Then we have

$$\sum_{j\in[k]} y_j^1 \geq \varepsilon, \quad \sum_{j\in[k]} y_j^2 \geq \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\} \quad \text{and} \quad (\bar{y}, \bar{z}) = \frac{1}{2}(y^1, z^1) + \frac{1}{2}(y^2, z^2),$$

which implies that

$$\frac{1}{2}\left(\varepsilon + \sum_{j\in[k]} w_{pj}\right) = \sum_{j\in[k]} \bar{y}_j = \frac{1}{2}\sum_{j\in[k]}(y_j^1 + y_j^2) \geq \frac{1}{2}\varepsilon + \frac{1}{2}\sum_{j\in[k]} \max\{w_{pj}, w_{qj}\}.$$

This is a contradiction, because we assumed $\sum_{j\in[k]} w_{pj} < \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\}$. Therefore, $(\bar{y}, \bar{z})$ is not contained in $\mathrm{conv}(\mathcal{M}(W, \mathbf{0}, \varepsilon))$, as required.

In order to finish the proof we consider the case of $\varepsilon > L_W(\varepsilon)$. Based on the previous parts of the proof, we may assume that $\bar{I}(\varepsilon)$ is ε-negligible. Then, $L_W(\varepsilon)$ is finite, and thus there exist distinct $p, q \in [n] \setminus \bar{I}(\varepsilon)$ such that $\varepsilon > \sum_{j\in[k]} \min\{w_{pj},\ w_{qj}\} = L_W(\varepsilon)$. Let us consider the point $(\bar{y}, \bar{z})$ where

$$\bar{z}_i = \begin{cases} \frac{1}{2} & \text{if } i \in \{p, q\} \\ 1 & \text{if } i \notin \{p, q\} \end{cases} \quad \text{and} \quad \bar{y}_j = \begin{cases} \frac{1}{2}\max\{w_{pj},\ w_{qj}\} & \text{if } j \in [k-1] \\ \frac{1}{2}\max\{w_{pk},\ w_{qk}\} + \frac{1}{2}L_W(\varepsilon) & \text{if } j = k \end{cases}.$$

Then

$$\sum_{j\in[k]} \bar{y}_j = \sum_{j\in[k]} \frac{1}{2}\max\{w_{pj},\ w_{qj}\} + \frac{1}{2}\sum_{j\in[k]} \min\{w_{pj},\ w_{qj}\} = \frac{1}{2}\sum_{j\in[k]} w_{pj} + \frac{1}{2}\sum_{j\in[k]} w_{qj} > \varepsilon,$$

where the first equation follows from the properties of $L_W(\varepsilon)$ in this case, and the inequality follows from our assumption that $p, q \in [n] \setminus \bar{I}(\varepsilon)$. Similar to the previous cases, we can argue that $(\bar{y}, \bar{z})$ satisfies the mixing inequalities. Now we argue that $(\bar{y}, \bar{z})$ satisfies every aggregated mixing inequality. By Lemma 4, it is sufficient to consider only the sequences $\Theta = \{i_1 \to \cdots \to i_\theta\}$ contained in $\{p, q\}$. Since $\Theta \subseteq \{p, q\}$, we know that $\bar{z}_{i_1} = \cdots = \bar{z}_{i_\theta} = \frac{1}{2}$. Then the following inequality (33) implies (30).

$$\sum_{j\in[k]} \bar{y}_j = \frac{1}{2}\sum_{j\in[k]} w_{pj} + \frac{1}{2}\sum_{j\in[k]} w_{qj} \geq \frac{1}{2}\min\{\varepsilon,\ L_{W,\Theta}\} + \frac{1}{2}\sum_{j\in[k]} \max\{w_{ij}\ :\ i \in \Theta\}. \tag{33}$$

When Θ contains both p and q, we have

$$L_{W,\Theta} = \sum_{j\in[k]} \min\{w_{pj}, w_{qj}\} \quad \text{and} \quad \sum_{j\in[k]} \max\{w_{ij}\ :\ i \in \Theta\} = \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\}.$$

Therefore, (33) holds in this case. (33) clearly holds if $\Theta = \{p\}$ or $\Theta = \{q\}$, because ε is smaller than $\sum_{j\in[k]} w_{pj}$ and $\sum_{j\in[k]} w_{qj}$ (this follows from $p, q \notin \bar{I}(\varepsilon)$). Consequently, Lemma 4 implies that $(\bar{y}, \bar{z})$ satisfies the aggregated mixing inequalities (A-Mix) for all sequences as well. Suppose for a contradiction that $(\bar{y}, \bar{z})$ is a convex combination of two points (y^1, z^1) and (y^2, z^2) in $\mathcal{M}(W, \mathbf{0}, \varepsilon)$. As in the previous cases, we can argue that z^1 and z^2 satisfy the constraints $z_i \le 1$ for $i \notin \{p, q\}$ and $z_p = z_q$ at equality. Therefore, $z^1 = \mathbf{1}$ and $z^2 = \mathbf{1}_{[n]\setminus\{p,q\}}$. Then we have

$$\sum_{j\in[k]} y_j^1 \ge \varepsilon, \quad \sum_{j\in[k]} y_j^2 \ge \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\} \quad \text{and} \quad (\bar{y}, \bar{z}) = \frac{1}{2}(y^1, z^1) + \frac{1}{2}(y^2, z^2),$$

which implies that

$$\sum_{j\in[k]} \frac{1}{2} \max\{w_{pj},\ w_{qj}\} + \frac{1}{2} \sum_{j\in[k]} \min\{w_{pj},\ w_{qj}\} = \sum_{j\in[k]} \bar{y}_j = \frac{1}{2} \sum_{j\in[k]} (y_j^1 + y_j^2) \ge \frac{1}{2}\varepsilon + \frac{1}{2} \sum_{j\in[k]} \max\{w_{pj}, w_{qj}\}.$$

This is a contradiction to our assumption that $\varepsilon > \sum_{j\in[k]} \min\{w_{pj},\ w_{qj}\}$. Therefore, $(\bar{y}, \bar{z})$ is not contained in $\operatorname{conv}(\mathcal{M}(W, \mathbf{0}, \varepsilon))$, as required. □

6 Two-sided chance-constrained programs

Liu et al. [19] considered the mixed-integer set defined by

$$\begin{aligned}
& y_p + y_d + w_i z_i \ge w_i, && \forall i \in [n], && \text{(34a)}\\
& y_p - y_d + (v_i + u_d) z_i \ge v_i, && \forall i \in [n], && \text{(34b)}\\
& u_d \ge y_d \ge 0, && && \text{(34c)}\\
& y_p \ge 0, && && \text{(34d)}\\
& z \in \{0,1\}^n, && && \text{(34e)}
\end{aligned}$$

where u_d is a positive constant satisfying $u_d \ge \max\{w_i : i \in [n]\}$, $w_i \ge v_i \ge 0$ for $i \in [n]$. After setting $y_p + y_d = y_1$ and $y_p - y_d + u_d = y_2$, the set (34) is equivalent to the following system:

$$\begin{aligned}
& y_1 + w_i z_i \ge w_i, && \forall i \in [n], && \text{(35a)}\\
& y_2 + (v_i + u_d) z_i \ge (v_i + u_d), && \forall i \in [n], && \text{(35b)}\\
& u_d \ge y_1 - y_2 \ge -u_d, && && \text{(35c)}\\
& y_1 + y_2 \ge u_d, && && \text{(35d)}\\
& z \in \{0,1\}^n. && && \text{(35e)}
\end{aligned}$$

Note that the set defined by (35a), (35b), (35d), and (35e) is a joint mixing set with lower bounds of the form $\mathcal{M}(W, \mathbf{0}, u_d)$ with $k = 2$. Moreover,

$$w_i + (v_i + u_d) \ge u_d \ \text{ for all } \ i \in [n] \quad \text{and} \quad \min\{w_i : i \in [n]\} + \min\{v_i + u_d : i \in [n]\} \ge u_d,$$

implying in turn that the convex hull of the joint mixing set with lower bounds can be obtained after applying the mixing and the aggregated mixing inequalities by Theorem 5.1.

In particular, given a sequence $\{i_1 \to \cdots \to i_\theta\}$ of indices in $[n]$, the corresponding aggregated mixing inequality (A-Mix) is of the following form:

$$y_1 + y_2 + \sum_{s\in[\tau_R]} (w_{r_s} - w_{r_{s+1}})z_{r_s} + \sum_{s\in[\tau_G]} (v_{g_s} - v_{g_{s+1}})z_{g_s} - u_d z_{i_\theta} \geq w_{r_1} + (v_{g_1} + u_d), \tag{36}$$

where $\{r_1 \to \cdots \to r_{\tau_R}\}$ and $\{g_1 \to \cdots \to g_{\tau_G}\}$ are the 1-mixing-subsequence and the 2-mixing-subsequence of Θ, respectively, and $w_{r_{\tau_R+1}} := 0$, $v_{g_{\tau_G+1}} := -u_d$. By Lemma 3, we know that $z_{g_{\tau_G}} = z_{i_\theta}$, so $(v_{g_{\tau_G}} - v_{g_{\tau_G+1}})z_{g_{\tau_G}} - u_d z_{i_\theta} = v_{g_{\tau_G}} z_{g_{\tau_G}}$. Since $y_1 + y_2 = 2y_p + u_d$, (36) is equivalent to the following inequality:

$$2y_p + \sum_{s\in[\tau_R]} (w_{r_s} - w_{r_{s+1}})z_{r_s} + \sum_{s\in[\tau_G]} (v_{g_s} - v_{g_{s+1}})z_{g_s} \geq w_{r_1} + v_{g_1}, \tag{37}$$

where $w_{r_{\tau_R+1}} := 0$ as before but $v_{g_{\tau_G+1}}$ is now set to 0.

In [19], the inequality (37) is called the *generalized mixing inequality from* Θ, so the aggregated mixing inequalities generalize the generalized mixing inequalities to arbitrary k. Furthermore, Theorem 5.1 can be extended slightly to recover the following main result of [19]:

Theorem 6.1 ([19], Theorem 3.1)**.** *Let $\mathcal{P}$ be the mixed-integer set defined by* (35a)–(35e)*. Then the convex hull of $\mathcal{P}$ can be described by the mixing inequalities for y_1, y_2, the aggregated mixing inequalities of the form* (36) *together with* (35c) *and the bounds $\mathbf{0} \leq z \leq \mathbf{1}$ under the assumption that $w_i \geq v_i \geq 0$ for $i \in [n]$ and $u_d \geq \max\{w_i : i \in [n]\}$.*

Proof. Let $\mathcal{R}$ be the mixed-integer set defined by (35a), (35b), (35d), and (35e). Then $\mathcal{P} \subseteq \mathcal{R}$ and, by Theorem 5.1, $\operatorname{conv}(\mathcal{R})$ is described by the mixing inequalities for y_1, y_2 and the generalized mixing inequalities of the form (36) together with $\mathbf{0} \leq z \leq \mathbf{1}$. We will argue that adding constraint (35c), that is $u_d \geq y_1 - y_2 \geq -u_d$, to the description of $\operatorname{conv}(\mathcal{R})$ does not affect integrality of the resulting system.

By Lemma 9, the extreme rays of $\operatorname{conv}(\mathcal{R})$ are $(e^j, \mathbf{0})$ for $j \in [k]$, and the extreme points are

- $A(z) = (y_1, y_2, z)$ for $z \in \{0,1\}^n \setminus \{\mathbf{1}\}$ where

$$y_1 = \max_{i\in[n]} \{w_i(1 - z_i)\} \quad \text{and} \quad y_2 = \max_{i\in[n]} \{(v_i + u_d)(1 - z_i)\},$$

- $B(1) = (u_d, 0, \mathbf{1})$ and $B(2) = (0, u_d, \mathbf{1})$.

It follows from the assumption that $w_i \geq v_i \geq 0$ for $i \in [n]$ and $u_d \geq \max\{w_i : i \in [n]\}$ that all extreme points of $\operatorname{conv}(\mathcal{R})$ satisfy $u_d \geq y_1 - y_2 \geq -u_d$. Observe that two hyperplanes $\{(y,z) : u_d = y_1 - y_2\}$ and $\{(y,z) : y_1 - y_2 = -u_d\}$ are parallel. So, each of the new extreme points created after adding $u_d \geq y_1 - y_2 \geq -u_d$ is obtained as the intersection of one of the two hyperplanes and a ray emanating from an extreme point of $\operatorname{conv}(\mathcal{R})$. Since every extreme ray of $\operatorname{conv}(\mathcal{R})$ has $\mathbf{0}$ in its z component and every extreme point of $\operatorname{conv}(\mathcal{R})$ has integral z component, the z component of every new extreme point is also integral, as required.

Therefore, the convex hull of $\mathcal{P}$ is equal to $\{(y,z) \in \text{conv}(\mathcal{R}) : (y,z) \text{ satisfies (35c)}\}$, implying in turn that $\text{conv}(\mathcal{P})$ can be described by the mixing inequalities for y_1, y_2, the aggregated mixing inequalities of the form (36) together with (35c) and the bounds $\mathbf{0} \le z \le \mathbf{1}$, as required. $\square$

Acknowledgments

This research is supported, in part, by ONR grant N00014-19-1-2321 and the Institute for Basic Science, No. IBS-R029-C1.

References

[1] Ahmad Abdi and Ricardo Fukasawa. On the mixing set with a knapsack constraint. *Mathematical Programming*, 157:191–217, 2016.

[2] Shabbir Ahmed, James Luedtke, Song Yongjia, and Weijun Xie. Nonanticipative duality, relaxations, and formulations for chance-constrained stochastic programs. *Mathematical Programming*, 162:51–81, 2017.

[3] Alper Atamtürk and Andrés Gómez. Submodularity in conic quadratic mixed 0-1 optimization. *Operations Research, to appear*.

[4] Alper Atamtürk and Vishnu Narayanan. Polymatroids and mean-risk minimization in discrete optimization. *Operations Research Letters*, 36(5):618–622, 2008.

[5] Alper Atamtürk, George L. Nemhauser, and Martin W.P. Savelsbergh. The mixed vertex packing problem. *Mathematical Programming*, 89(1):35–53, 2000.

[6] Frank Baumann, Sebastian Berckey, and Christoph Buchheim. Exact algorithms for combinatorial optimization problems with submodular objective functions. In Michael Jünger and Gerhard Reinelt, editors, *Facets of Combinatorial Optimization: Festschrift for Martin Grötschel*, pages 271–294. Springer, 2013.

[7] G. C. Calafiore and M. C. Campi. Uncertain convex programs: Randomized solutions and confidence levels. *Mathematical Programming*, 102(1):25–46, 2005.

[8] G. C. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2006.

[9] Michele Conforti, Marco Di Summa, and Laurence A. Wolsey. The mixing set with flows. *SIAM Journal on Discrete Mathematics*, 21(2):396–407, 2007.

[10] Miguel Constantino, Andrew J. Miller, and Mathieu Van Vyve. Mixing MIR inequalities with two divisible coefficients. *Mathematical Programming*, 123:451–483, 2010.

[11] Jack Edmonds. Submodular functions, matroids, and certain polyhedra. In *Combinatorial Structures and Their Applications*, pages 69–87. Gordon and Breach, 1970.

[12] Oktay Günlük and Yves Pochet. Mixing MIR inequalities for mixed integer programs. *Mathematical Programming*, 90:429–457, 2001.

[13] R. Henrion and A. Möller. Optimization of a continuous distillation process under random inflow rate. *Computers & Mathematics with Applications*, 45(1):247 – 262, 2003.

[14] René Henrion, Pu Li, Andris Möller, Marc C. Steinbach, Moritz Wendt, and Günter Wozny. *Stochastic Optimization for Operating Chemical Processes under Uncertainty*, pages 457–478. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[15] Simge Küçükyavuz. On mixing sets arising in chance-constrained programming. *Mathematical Programming*, 132(1-2):31–56, 2012.

[16] M. A. Lejeune and A. Ruszczyński. An efficient trajectory method for probabilistic inventory-production-distribution problems. *Operations Research*, 55(2):378–394, 2007.

[17] X. Liu, S. Küçükyavuz, and J. Luedtke. Decomposition algorithms for two-stage chance-constrained programs. *Mathematical Programming*, 157(1):219–243, 2016.

[18] Xiao Liu and Simge Küçükyavuz. A polyhedral study of the static probabilistic lot-sizing problem. *Annals of Operations Research*, 261(1-2):233–254, 2018.

[19] Xiao Liu, Fatma Kılınç-Karzan, and Simge Küçükyavuz. On intersection of two mixing sets with applications to joint chance-constrained programs. *Mathematical Programming*, 175:29–68, 2019.

[20] László Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art: Bonn 1982*, pages 235–257. Springer, Berlin, Heidelberg, 1983.

[21] James Luedtke. A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support. *Mathematical Programming*, 146:219–244, 2014.

[22] James Luedtke and Shabbir Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19:674–699, 2008.

[23] James Luedtke, Shabbir Ahmed, and George L. Nemhauser. An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122(2):247–272, 2010.

[24] Andrew J. Miller and Laurence A. Wolsey. Tight formulations for some simple mixed integer programs and convex objective integer programs. *Mathematical Programming*, 98(1-3):73–88, 2003.

[25] Michael R. Murr and András Prékopa. *Solution of a Product Substitution Problem Using Stochastic Programming*, pages 252–271. Springer US, Boston, MA, 2000.

[26] Y. Pochet and L. A. Wolsey. Polyhedra for lot-sizing with Wagner-Whitin costs. *Mathematical Programming*, 67:297–323, 1994.

[27] András Prékopa. Probabilistic programming. In A. Ruszczyński and A. Shapiro, editors, *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 267–351. Elsevier, 2003.

[28] Feng Qiu, Shabbir Ahmed, Santanu S. Dey, and Laurence A. Wolsey. Covering linear programming with violations. *INFORMS Journal on Computing*, 26(3):531–546, 2014.

[29] A. Ruszczyński. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming*, 93:195–215, 2002.

[30] Yongjia Song, James Luedtke, and Simge Küçükyavuz. Chance-constrained binary packing problems. *INFORMS Journal on Computing*, 26(4):735–747, 2014.

[31] Andrews K. Takyi and Barbara J. Lence. Surface water quality management using a multiple-realization chance constraint method. *Water Resources Research*, 35(5):1657–1670, 1999.

[32] Q. Wang, Y. Guan, and J. Wang. A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output. *IEEE Transactions on Power Systems*, 27(1):206–215, 2012.

[33] Laurence A. Wolsey. Mixing sets. In J.J. Cochran, L.A. Cox, Kharoufeh P.J. Keskinocak, P., and J.C. Smith, editors, *Wiley Encyclopedia of Operations Research and Management Science*. Wiley, Hoboken, 2010.

[34] Weijun Xie. On distributionally robust chance constrained programs with Wasserstein distance. *Working paper*, 2018.

[35] Weijun Xie and Shabbir Ahmed. On quantile cuts and their closure for chance constrained optimization problems. *Mathematical Programming*, 172:621–646, 2018.

[36] Jiajin Yu and Shabbir Ahmed. Polyhedral results for a class of cardinality constrained submodular minimization problems. *Discrete Optimization*, 24:87–102, 2017.

[37] M. Zhang, S. Küçükyavuz, and S. Goel. A branch-and-cut method for dynamic decision making under joint chance constraints. *Management Science*, 60(5):1317–1333, 2014.

[38] Yiling Zhang, Ruiwei Jiang, and Siqian Shen. Ambiguous chance-constrained binary programs under mean-covariance information. *SIAM Journal on Optimization*, 28(4):2922–2944, 2018.

[39] Ming Zhao and I.R. de Farias. The mixing MIR set with divisible capacities. *Mathematical Programming*, 115(1):73–103, 2008.

[40] Ming Zhao, Kai Huang, and Bo Zeng. A polyhedral study on chance constrained program with random right-hand side. *Mathematical Programming*, 166:19–64, 2017.

A Proof of Lemma 4

Let $(\bar{y}, \bar{z}) \in \mathbb{R}^k_+ \times [0,1]^n$ be a point satisfying (13a)–(13c), and assume that $(\bar{y}, \bar{z})$ satisfies (A-Mix) for all sequences contained in $\{i \in [n] : \bar{z}_i < 1\}$. Then we need to prove that $(\bar{y}, \bar{z})$ satisfies (A-Mix) for all the other sequences as well.

For a sequence Θ, we denote by $N(\Theta)$ the set $\{i \in \Theta : \bar{z}_i = 1\}$. We argue by induction on $|N(\Theta)|$ that $(\bar{y}, \bar{z})$ satisfies (A-Mix) for Θ. If $|N(\Theta)| = 0$, then $(\bar{y}, \bar{z})$ satisfies (A-Mix) by the assumption. For the induction step, we assume that $(\bar{y}, \bar{z})$ satisfies (A-Mix) for every sequence Θ with $|N(\Theta)| < N$ for some $N \geq 1$. Now we take a sequence $\Theta = \{i_1 \to \cdots \to i_\theta\}$ with $|N(\Theta)| = N$. Notice that $(\bar{y}, \bar{z})$ satisfies (A-Mix) if and only if $(\bar{y}, \bar{z})$ satisfies

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{t\in[\theta]} \left(w_{i_t j} - \max\left\{w_{ij} : \ i_t \text{ precedes } i \text{ in } \Theta\right\}\right)_+ \bar{z}_{i_t} \right) - \min\left\{\varepsilon, \ L_{W,\Theta}\right\} \bar{z}_{i_\theta} \geq \sum_{j\in[k]} \max\left\{w_{ij} : \ i \in \Theta\right\}. \tag{38}$$

Hence, it is sufficient to show that $(\bar{y}, \bar{z})$ satisfies (38). We consider two cases $\bar{z}_{i_\theta} = 1$ and $\bar{z}_{i_\theta} \neq 1$ separately.

First, consider the case when $\bar{z}_{i_\theta} \neq 1$. Since $|N(\Theta)| \geq 1$, we have $\bar{z}_{i_p} = 1$ for some $p \in [\theta - 1]$. Let Θ' denote the subsequence of Θ obtained by removing i_p. Then $|N(\Theta')| = |N(\Theta)| - 1$, so it follows from the induction hypothesis that (A-Mix) for Θ' is valid for $(\bar{y}, \bar{z})$:

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{t\in[\theta]\setminus\{p\}} \left(w_{i_t j} - \max\left\{w_{ij} : \ i_t \text{ precedes } i \text{ in } \Theta'\right\}\right)_+ \bar{z}_{i_t} \right) - \min\left\{\varepsilon, \ L_{W,\Theta'}\right\} \bar{z}_{i_\theta} \geq \sum_{j\in[k]} \max\left\{w_{ij} : \ i \in \Theta'\right\}. \tag{39}$$

Since Θ' is a subsequence of Θ, it follows that for any $t \neq p$.

$$\left(w_{i_t j} - \max\left\{w_{ij} : \ i_t \text{ precedes } i \text{ in } \Theta'\right\}\right)_+ \geq \left(w_{i_t j} - \max\left\{w_{ij} : \ i_t \text{ precedes } i \text{ in } \Theta\right\}\right)_+. \tag{40}$$

Since $-\bar{z}_{i_t} \geq -1$ is valid for each t, we deduce the following inequality from (39):

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{t\in[\theta]\setminus\{p\}} \left(w_{i_t j} - \max\left\{w_{ij} : \ i_t \text{ precedes } i \text{ in } \Theta\right\}\right)_+ \bar{z}_{i_t} \right) - \min\left\{\varepsilon, \ L_{W,\Theta'}\right\} \bar{z}_{i_\theta} \geq \sum_{j\in[k]} \max\left\{w_{ij} : \ i \in \Theta\right\} - \sum_{j\in[k]} \left(w_{i_p j} - \max\left\{w_{ij} : \ i_p \text{ precedes } i \text{ in } \Theta\right\}\right)_+, \tag{41}$$

because

$$\sum_{t\in[\theta]\setminus\{p\}} \left(w_{i_t j} - \max\left\{w_{ij} : \ i_t \text{ precedes } i \text{ in } \Theta'\right\}\right)_+ = \sum_{j\in[k]} \max\left\{w_{ij} : \ i \in \Theta'\right\}$$

and

$$\sum_{t\in[\theta]} \left(w_{i_t j} - \max\left\{w_{ij} : \ i_t \text{ precedes } i \text{ in } \Theta\right\}\right)_+ = \sum_{j\in[k]} \max\left\{w_{ij} : \ i \in \Theta\right\}. \tag{42}$$

Moreover, notice that $L_{W,\Theta'} \geq L_{W,\Theta}$ due to (40). So, it follows that (41) implies (38) since $\bar{z}_{i_p} = 1$. This in turn implies that $(\bar{y}, \bar{z})$ satisfies (A-Mix) for Θ, as required.

Next we consider the $\bar{z}_{i_\theta} = 1$ case. In this case, (38) is equivalent to

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{t\in[\theta-1]} \left(w_{i_t j} - \max\left\{ w_{ij} \,:\, i_t \text{ precedes } i \text{ in } \Theta \right\} \right)_+ \bar{z}_{i_t} \right)$$
$$\geq \min\left\{\varepsilon,\ L_{W,\Theta}\right\} - \sum_{j\in[k]} w_{i_\theta j} + \sum_{j\in[k]} \max\left\{ w_{ij} \,:\, i \in \Theta \right\}. \quad (43)$$

Take the subsequence Θ' of Θ obtained by removing i_θ. As before, we have $|N(\Theta')| = |N(\Theta)| - 1$, and the induction hypothesis implies that (A-Mix) for Θ' is valid for $(\bar{y}, \bar{z})$:

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{t\in[\theta-2]} \left(w_{i_t j} - \max\left\{ w_{ij} \,:\, i_t \text{ precedes } i \text{ in } \Theta' \right\} \right)_+ \bar{z}_{i_t} \right)$$
$$+ \left(\sum_{j\in[k]} w_{i_{\theta-1} j} - \min\left\{\varepsilon,\ L_{W,\Theta'}\right\} \right) \bar{z}_{i_{\theta-1}} \geq \sum_{j\in[k]} \max\left\{ w_{ij} \,:\, i \in \Theta' \right\}. \quad (44)$$

We will deduce from (44) that (43) is valid for $(\bar{y}, \bar{z})$. As Θ' is a subsequence of Θ, (40) holds for $t \in [\theta - 2]$. So, as $(\bar{y}, \bar{z})$ satisfies $-\bar{z}_{i_t} \geq -1$ for $t \in [\theta - 2]$, we obtain the following from (44):

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{t\in[\theta-2]} \left(w_{i_t j} - \max\left\{ w_{ij} : i_t \text{ precedes } i \text{ in } \Theta \right\} \right)_+ \bar{z}_{i_t} \right) + \left(\sum_{j\in[k]} w_{i_{\theta-1} j} - \min\left\{\varepsilon, L_{W,\Theta'}\right\} \right) \bar{z}_{i_{\theta-1}}$$
$$\geq \sum_{j\in[k]} \min\left\{ w_{i_{\theta-1} j}, w_{i_\theta j} \right\} - \sum_{j\in[k]} w_{i_\theta j} + \sum_{j\in[k]} \max\left\{ w_{ij} \,:\, i \in \Theta \right\}, \quad (45)$$

because (42) holds,

$$\sum_{t\in[\theta-1]} \left(w_{i_t j} - \max\left\{ w_{ij} \,:\, i_t \text{ precedes } i \text{ in } \Theta' \right\} \right)_+ = \sum_{j\in[k]} \max\left\{ w_{ij} \,:\, i \in \Theta' \right\},$$

and

$$\sum_{j\in[k]} w_{i_{\theta-1} j} - \sum_{j\in[k]} \left(w_{i_{\theta-1} j} - \max\left\{ w_{ij} \,:\, i_{\theta-1} \text{ precedes } i \text{ in } \Theta \right\} \right)_+ = \sum_{j\in[k]} \min\left\{ w_{i_{\theta-1} j}, w_{i_\theta j} \right\}. \quad (46)$$

Now let us compare the coefficient of $\bar{z}_{i_{\theta-1}}$ in (45) and that of $\bar{z}_{i_{\theta-1}}$ in (43). If the coefficient in (45) is less than the coefficient in (43), then (45) implies that (43) is valid, because we can add an appropriate scalar multiple of $\bar{z}_{i_{\theta-1}} \geq 0$ to (45) in order to achieve the coefficient in (43) and the term $\sum_{j\in[k]} \min\left\{ w_{i_{\theta-1} j}, w_{i_\theta j} \right\}$ in the right-hand side of (45) is at least $L_{W,\Theta}$. If not, then by adding an appropriate scalar multiple of $-\bar{z}_{i_{\theta-1}} \geq -1$ to (45), we deduce the following inequality:

$$\sum_{j\in[k]} \left(\bar{y}_j + \sum_{t\in[\theta-1]} \left(w_{i_t j} - \max\left\{ w_{ij} \,:\, i_t \text{ precedes } i \text{ in } \Theta \right\} \right)_+ \bar{z}_{i_t} \right)$$
$$\geq \min\left\{\varepsilon,\ L_{W,\Theta'}\right\} - \sum_{j\in[k]} w_{i_\theta j} + \sum_{j\in[k]} \max\left\{ w_{ij} \,:\, i \in \Theta \right\}, \quad (47)$$

because (46) holds. Since Θ' is a subsequence of Θ, we have $L_{W,\Theta'} \geq L_{W,\Theta}$, so it follows that the term $\min\{\varepsilon,\ L_{W,\Theta'}\}$ in the right-hand side of (47) is at least $\min\{\varepsilon,\ L_{W,\Theta}\}$. Therefore, (47) implies that (43) is valid for $(\bar{y}, \bar{z})$. In summary, when $\bar{z}_{i_\theta} = 1$, $(\bar{y}, \bar{z})$ satisfies (43), thereby proving that $(\bar{y}, \bar{z})$ satisfies (A-Mix). This finishes the proof of this lemma.