Submitted to *Operations Research*
manuscript (Please, provide the manuscript number!)

Preservation of Supermodularity in Parametric Optimization: Necessary and Sufficient Conditions on Constraint Structures

Xin Chen
Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, xinchen@illinois.edu

Daniel Zhuoyu Long
Department of Systems Engineering & Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong, zylong@se.cuhk.edu.hk

Jin Qi
Department of Industrial Engineering and Logistics Management, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, jinqi@ust.hk

This paper presents a systematic study of the preservation of supermodularity under parametric optimization, that allows us to derive complementarity among parameters and monotone structural properties of optimal policies in many operations models. We introduce new concepts of mostly-lattice and additive mostly-lattice, which significantly generalize the commonly imposed lattice condition, and use them to establish the necessary and sufficient conditions on the feasible set so that supermodularity can be preserved under various assumptions on the objective functions. We further identify some classes of polyhedral sets which satisfy these concepts. Finally, we illustrate how our results can be used on assemble-to-order systems.

Key words: supermodularity, parametric optimization, necessary and sufficient conditions, assemble-to-order, dynamic programming

1. Introduction

The concept of supermodularity has received considerable attention in the economics and operations research literature. It is closely related to the concept of complementarity in economics and has very strong economic implications (see Topkis 1998). It has also been proved to be an important tool to derive monotone comparative statics in parametric optimization problems and game theory models, i.e., how the optimal decisions or equilibria vary monotonically with respect to the parameters (Topkis 1998).

In many operations problems, it is often of great interest to study whether a set of products are complements. For example, in an assemble-to-order system, a firm assembles products from components, and has to decide the ordering plan of all components with the aim of maximizing the profit. In this case, a key structural analysis is whether the components are complements and hence the optimal order-up-to level of one component would be monotonically increasing with the initial stock level of other components (Song and Zipkin 2003, Quah 2007). In Markovian decision processes, we face a sequence of parametric optimization problems. To derive monotone structural properties of optimal policies, it is critical that supermodularity can be carried over under dynamic programming recursions (e.g., Topkis 1998 and Chen et al. 2013). A classical example is the multi-product inventory control problem under the multi-period setting (e.g., Chen et al. 2013).

However, to investigate whether supermodularity can be carried over is not always a trivial task and our purpose is to study constraint structures in parametric optimization problems systematically so that supermodularity can be preserved. In particular, we consider the following parametric optimization problem.

$$g(\boldsymbol{t}) = \max\left\{f(\boldsymbol{x},\boldsymbol{t}) \;:\; \boldsymbol{x}\in\mathcal{S}_t\right\}, \tag{1}$$

where $f:\mathcal{X}\times\mathcal{T}\to\Re$, $\mathcal{X},\mathcal{T}$ are subsets of corresponding Euclidean spaces, and $\mathcal{S}_t=\{\boldsymbol{x}\in\mathcal{X}:(\boldsymbol{x},\boldsymbol{t})\in\mathcal{S}\}$ and $\mathcal{S}$, the graph of the constraint set, is a subset of $\mathcal{X}\times\mathcal{T}$. In the assemble-to-order problem, $\boldsymbol{t}$ represents the amount of components while $\boldsymbol{x}$ is the amount of products to assemble, and $\mathcal{S}_t$ includes all the assemble plans for given input $\boldsymbol{t}$. In dynamic programming problem (e.g., multi-product inventory control), the equation (1) is the dynamic recursion where $g(\boldsymbol{t})$ is the profit-to-go function with state variables being $\boldsymbol{t}$.

A fundamental approach in verifying the supermodularity of g is the preservation of supermodularity in parametric optimization (1), i.e., under certain condition on the graph $\mathcal{S}$, g is supermodular when f is so. In a single stage problem (e.g., assemble-to-order), the profit function f typically takes an explicit form and its supermodularity can be easily shown; hence, we can prove the supermodularity of g if the supermodularity can be preserved in the problem (1). In a multi-period problem (e.g., multi-product inventory control), more often than not, the profit-to-go function in the last period also takes explicit form and is supermodular. Therefore, if the dynamic recursion with the form (1) can preserve supermodularity, the profit-to-go function in all periods can be proved to be supermodular. The following result by Topkis (1998) establishes conditions under which supermodularity can be preserved under the optimization operation (1).

Theorem 1 (Theorem 2.7.6, Topkis 1998) *The function g is supermodular on the projection of $\mathcal{S}$ on $\mathcal{T}$ whenever f is supermodular on $\mathcal{X} \times \mathcal{T}$ if $\mathcal{S}$ is a sublattice of $\mathcal{X} \times \mathcal{T}$.*

Although powerful and widely used in the literature, Topkis' result imposes the lattice requirement on $\mathcal{S}$, which may often be too restrictive for a variety of applications in operations (e.g., Zhu and Thonemann 2009, Chen et al. 2013). To see this, consider a commonly encountered case in which $\mathcal{S}$ is defined by linear inequalities:

$$\mathcal{S}_{\boldsymbol{c}}^P = \{(\boldsymbol{x}, \boldsymbol{t}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{t} \le \boldsymbol{c}\} \tag{2}$$

for some nonzero matrices $\boldsymbol{A} \in \Re^{m \times n_1}$, $\boldsymbol{B} \in \Re^{m \times n_2}$, and a vector $\boldsymbol{c} \in \Re^m$. In general, to allow that the supermodularity can be carried over under dynamic programming recursions, we have to investigate the case for each $\boldsymbol{c} \in \Re^m$.

Theorem 2 (Page 26, Topkis 1998) *The polyhedron $\mathcal{S}_{\boldsymbol{c}}^P$ defined by the equation (2) is a lattice for each $\boldsymbol{c} \in \Re^m$ if and only if each coefficient vector $(a_{i1}, \ldots, a_{in_1}, b_{i1}, \ldots, b_{in_2})$, $i \in \{1, \ldots, m\}$ with more than one nonzero component has exactly two nonzero components with opposite signs, where $(a_{i1}, \ldots, a_{in_1})$ and $(b_{i1}, \ldots, b_{in_2})$ are the ith rows of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively.*

For ease of discussion, we refer to a matrix with the property that every row vector with more than one nonzero component has exactly two nonzero components with opposite signs as a *lattice-matrix*. Hence, the condition in the above theorem is equivalent to say that the matrix $(\boldsymbol{A}|\boldsymbol{B})$ concatenating $\boldsymbol{A}$ and $\boldsymbol{B}$ is a lattice-matrix.

The goal of this paper is to relax the lattice requirement and hence provide more powerful tools in proving the supermodularity in many operations models. As we show in the next section, this may not always be fruitful if we merely impose supermodularity on f. However, in most operations applications, f possesses other salient properties in addition to being supermodular. For example, f may be linear, concave, or independent of $\boldsymbol{t}$ in its formulation. The type of questions we want to address is: under what conditions on $\mathcal{S}$, for any f with property $\mathcal{P}_f$, g has property $\mathcal{P}_g$?

We provide a host of answers to these questions. Specifically, we identify necessary and sufficient conditions on $\mathcal{S}$, under which

(a) g is supermodular for any supermodular function f;

(b) g is concave and supermodular for any concave and supermodular function f;

(c) g is supermodular for any supermodular function f independent of $\boldsymbol{t}$;

(d) g is concave and supermodular for any concave and supermodular function f independent of $\boldsymbol{t}$;

(e) g is supermodular for any linear function f; and

(f) g is concave and supermodular for any linear function f.

Being both necessary and sufficient, the conditions we provide can serve as the theoretically least restrictive condition to preserve supermodularity. In addition, to preserve concavity and supermodularity jointly is not necessarily equivalent to imposing both the condition of preserving concavity and that of preserving supermodularity simultaneously. Hence, it is of interest to study the condition of joint preservation as in (b), (d), and (f).

Our results illustrate that if one imposes (a) supermodularity or (b) concavity and supermodularity on f, we require $\mathcal{S}$ to be a mostly-lattice, i.e., for any unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$ and any $\boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$, we have $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$ and $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$. For (c), (d), (e) and (f), our results require $\mathcal{S}$ to have certain additive mostly-lattice conditions, i.e., for any unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$, $\boldsymbol{x}' \in \mathcal{S}_{t'}$, $\boldsymbol{x}'' \in \mathcal{S}_{t''}$, there exist $\boldsymbol{y} \in Conv(\mathcal{W} \cap \mathcal{S}_{t' \wedge t''})$ and $\boldsymbol{z} \in Conv(\mathcal{W} \cap \mathcal{S}_{t' \vee t''})$ such that $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$ for some set $\mathcal{W}$ depending on $\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''$.

We then focus on $\mathcal{S}$ with the polyhedral structure (2), and derive conditions on $\boldsymbol{A}$ and $\boldsymbol{B}$ such that for any $\boldsymbol{c}$, $\mathcal{S}^P_{\boldsymbol{c}}$ satisfies these identified conditions. Without loss of generality (WLOG), we assume $\boldsymbol{A}, \boldsymbol{B}$ have no zero column, $\boldsymbol{A}$ has no zero row, and $n_2 \geq 2$. Interestingly, we provide a complete characterization of the structures of $\boldsymbol{A}$ and $\boldsymbol{B}$ in several settings. Specifically, for (a) and (b), $\boldsymbol{A}$ and $\boldsymbol{B}$ are matrices such that the corresponding $\mathcal{S}_{\boldsymbol{c}}^P$ should be a lattice. We give a detailed description for (c) in Section 3. For (d), (e) and (f), the key requirement is that given any $\boldsymbol{\beta}$, $\boldsymbol{B}\boldsymbol{\beta}^+$ should be in the column space of $\boldsymbol{A}$ if $\boldsymbol{B}\boldsymbol{\beta}$ is so.

Table 1 provides an overview of our preservation results under various settings. The first column lists the properties of f, $\mathcal{P}_f$, and the second column lists the properties preserved by g, $\mathcal{P}_g$. The last column lists the conditions of the preservation results (under different assumptions on $\mathcal{S}$).

$\mathcal{P}_f$	$\mathcal{P}_g$	Set $\mathcal{S}$		
		General $\mathcal{S}$	Convex $\mathcal{S}$	Polyhedron $\mathcal{S}$
Supermodular	Supermodular	mostly-lattice		$(\boldsymbol{A}\mid\boldsymbol{B})$ lattice-matrix
Concave & supermodular	Concave & supermodular	mostly-lattice and convex	mostly-lattice	$(\boldsymbol{A}\mid\boldsymbol{B})$ lattice-matrix
Supermodular & independent of $\boldsymbol{t}$	Supermodular	additive mostly-lattice $\mathcal{W} = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$		Theorem 7
Concave & supermodular & independent of $\boldsymbol{t}$	Concave & supermodular	additive mostly-lattice with $\mathcal{W} = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$, $\boldsymbol{x}_\lambda \in Conv(\mathcal{W} \bigcap \mathcal{S}_{t_\lambda})$[1]	additive mostly-lattice with $\mathcal{W} = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$	Theorem 9
Linear	Supermodular	additive mostly-lattice with $\mathcal{W} = \mathcal{X}$		Theorem 11
	Concave & Supermodular	additive mostly-lattice with $\mathcal{W} = \mathcal{X}$, $\boldsymbol{x}_\lambda \in Conv(\mathcal{S}_{t_\lambda})$	additive mostly-lattice with $\mathcal{W} = \mathcal{X}$	Remark 1

$Conv(\cdot)$ denotes convex hull, $\boldsymbol{x}_\lambda = \lambda \boldsymbol{x}' + (1-\lambda)\boldsymbol{x}'', \boldsymbol{t}_\lambda = \lambda \boldsymbol{t}' + (1-\lambda)\boldsymbol{t}'', \lambda \in [0,1]$.

Table 1 A summary for the preservation result.

As an application, we use our results to analyze an assemble-to-order system where the number of outputs is a linear mapping from the inputs. We derive an efficient way to check any given assemble-to-order system whether a subset of components are complementary in the sense that the profit margin of one component's inventory increases with other components' inventories. In particular, we show that for the whole assemble-to-order system to have such a property, whenever two products share common components, their usage of all common components must follow the same proportion. This type of systems often appears in settings where products differ from each other mainly in size, and the needs for common components are proportional to the size of product.

There are very few papers deriving supermodularity preservation properties when the graph of the constraint set $\mathcal{S}$ is not a lattice. Gale and Politof (1981) and Granot and Veinott Jr (1985) derive supermodularity preservation properties for network flow optimization problems. Zipkin (2003) focuses on an assemble-to-order problem and models it as a linear program. Chen et al. (2013) investigate a sufficient condition for preservation of supermodularity in an optimization problem parameterized by a two-dimensional vector. Their main results cover several preservation properties scattered in the literature including those in Zipkin (2003) and Chao et al. (2009) and provide a tool to analyze a variety of operations models (e.g., Ceryan et al. 2013, Chen et al. 2016, Song and Xue 2007, Yang 2004). Compared with these papers, ours provides a systematic study on supermodularity preservation without lattice conditions. Our necessary and sufficient condition for preserving supermodularity covers many theoretical tools in Gale and Politof (1981), Granot and Veinott Jr (1985), Zipkin (2003) and Chen et al. (2013) as special cases. A closely related paper is Quah (2007), which shares certain flavor of this paper. Specifically, Quah (2007) considers similar parametric optimization problems. However, he focuses on identifying conditions on the feasible sets under which the optimal solutions are monotone for any concave and supermodular objective functions.

In Section 2, we discuss the condition for preserving supermodularity and then extend the result to study the condition for preserving both concavity and supermodularity simultaneously. In Section 3, we study a special class of the parametric optimization problem when function f is independent of parameters $\boldsymbol{t}$. Then, we pay particular attention to the case when function f is linear. We apply our results to study the assemble-to-order system in Section 4 and conclude the paper in Section 5.

Notation and convention: Vectors and matrices are represented by lower- and upper-case boldface characters, respectively. Given any vector $\boldsymbol{x} \in \Re^n$, we assume it is a column vector unless otherwise specified and denote by x_i its i^{th} element. We use parentheses to construct row vectors with comma separated each element as $\boldsymbol{x}^T = (x_1, \ldots, x_n)$, and $\boldsymbol{x} \leq \boldsymbol{y}$ if and only if $x_i \leq y_i$ for all i. The ith unit vector, whose components are zero except the ith component being one, is denoted as $\boldsymbol{e}_i$. For a matrix $\boldsymbol{A} = (a_{ij})_{i=1,\ldots,m;j=1,\ldots,n} \in \Re^{m\times n}$, define $\boldsymbol{a}_i^T$ and $\boldsymbol{A}_j$ as its ith row vector and jth column vector, respectively. We denote by $\boldsymbol{A}_{\mathcal{I}}$ the submatrix consisting of all rows $\boldsymbol{a}_i^T$, $i \in \mathcal{I} \subseteq \{1, \ldots, m\}$, and denote by $C(\boldsymbol{A})$ the column space of $\boldsymbol{A}$. Given any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ with the same number of rows, their concatenated matrix is denoted by $(\boldsymbol{A}|\boldsymbol{B})$. We denote the meet and joint operations by "$\wedge$" and "$\vee$", respectively, i.e., $\boldsymbol{x}' \wedge \boldsymbol{x}'' = (\min\{x_i', x_i''\})_{i=1,\ldots,n}$ and $\boldsymbol{x}' \vee \boldsymbol{x}'' = (\max\{x_i', x_i''\})_{i=1,\ldots,n}$ for any vectors $\boldsymbol{x}', \boldsymbol{x}'' \in \Re^n$. The positive part of x is denoted by $x^+ = \max\{x, 0\}$. We use $Conv(\cdot)$ to represent the convex hull of a given set and $\mathcal{S}|_{\mathcal{T}}$ to denote the projection of $\mathcal{S}$ on its component $\mathcal{T}$. For the sake of brevity, we present only selected proofs in the paper and all remaining proofs in the appendix.

2. Main Results for General Problems

In this section, we derive conditions on $\mathcal{S}$ so that certain properties can be preserved under the optimization operation (1). For simplicity, we focus on cases in which $\mathcal{X}, \mathcal{T}$ are sublattices of Euclidean spaces with the commonly used partial order "$\leq$". We also assume that $\mathcal{S}$ is a nonempty closed set, and the maximization problem in the equation (1) is well defined for all $\boldsymbol{t} \in \mathcal{T}$ with nonempty $\mathcal{S}_t$. Furthermore, if $\mathcal{S}|_{\mathcal{T}}$ is a chain (i.e., a fully ordered set), $g(\boldsymbol{t})$ is always supermodular. Hence, unless otherwise specified we assume that $\mathcal{S}|_{\mathcal{T}}$ is not a chain, or equivalently contains an unordered pair, to avoid the trivial case.

We first study the necessary and sufficient condition for preserving the property of supermodularity. For this purpose, we introduce a new concept of mostly-lattice.

Definition 1 *A set $\mathcal{S} \subseteq \mathcal{X} \times \mathcal{T}$ is a mostly-lattice if for any unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$ and any $\boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$, we have $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$ and $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$.*

Different from a lattice, the definition of mostly-lattice only imposes requirements for unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$. It automatically holds if $\mathcal{S}|_{\mathcal{T}}$ does not contain any unordered pair. On the other hand, if $\mathcal{S}|_{\mathcal{T}} = \{\boldsymbol{t}', \boldsymbol{t}'', \boldsymbol{t}' \vee \boldsymbol{t}'', \boldsymbol{t}' \wedge \boldsymbol{t}''\}$ for an unordered pair $\boldsymbol{t}'$ and $\boldsymbol{t}''$, the only condition for $\mathcal{S}$ to be a mostly-lattice is that $\{\boldsymbol{x}' \vee \boldsymbol{x}'' : \boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}\} \subseteq \mathcal{S}_{t' \vee t''}$ and $\{\boldsymbol{x}' \wedge \boldsymbol{x}'' : \boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}\} \subseteq \mathcal{S}_{t' \wedge t''}$, which does not even require $\mathcal{S}_{t'}, \mathcal{S}_{t''}, \mathcal{S}_{t' \vee t''}, \mathcal{S}_{t' \wedge t''}$ to be lattices in $\mathcal{X}$, the very basic requirement for $\mathcal{S}$ to be a lattice. We next provide two examples to illustrate this definition.

Example 1. We let $\mathcal{X} = \mathcal{T} = \Re^2$, and $\mathcal{S} = \{(1,0,0,0), (0,1,1,1)\}$. Obviously, neither $\mathcal{S}$ nor $Conv(\mathcal{S})$ is a lattice. However, both of them are mostly-lattices since there does not exist any unordered pair $\boldsymbol{t}', \boldsymbol{t}''$ in their projections on $\mathcal{T}$.

Example 2. We let $\mathcal{X} = \mathcal{T} = \Re^2$, $\mathcal{S} = \{(0,1,0,1), (1,0,1,0), (0,0,0,0), (1,1,1,1), (0.5,1.5,1,1)\}$. $\mathcal{S}$ is not a lattice since $\mathcal{S}_{(1,1)}$ is already not a lattice. However, $\mathcal{S}$ is a mostly-lattice.

Interestingly, $\mathcal{S}$ being a mostly-lattice is both necessary and sufficient for the preservation of supermodularity in problem (1).

Theorem 3 *The function g is supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is supermodular on $\mathcal{X} \times \mathcal{T}$ if and only if $\mathcal{S}$ is a mostly-lattice.*

Proof. The "if" part can be proved by a similar argument for Theorem 2.7.6 in Topkis (1998). For completeness, we present its proof here. For any unordered $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$, and any $\boldsymbol{x}' \in \mathcal{S}_{t'}$ and $\boldsymbol{x}'' \in \mathcal{S}_{t''}$,

$$f(\boldsymbol{x}', \boldsymbol{t}') + f(\boldsymbol{x}'', \boldsymbol{t}'') \leq f(\boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{t}' \vee \boldsymbol{t}'') + f(\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{t}' \wedge \boldsymbol{t}'') \leq g(\boldsymbol{t}' \vee \boldsymbol{t}'') + g(\boldsymbol{t}' \wedge \boldsymbol{t}''),$$

where the first inequality follows from the supermodularity of f on $\mathcal{X} \times \mathcal{T}$ and the second inequality holds by the definition of g and mostly-lattice. Taking supremum on the left hand side of the inequalities over $\boldsymbol{x}' \in \mathcal{S}_{t'}$ and $\boldsymbol{x}'' \in \mathcal{S}_{t''}$, we have that

$$g(\boldsymbol{t}') + g(\boldsymbol{t}'') \leq g(\boldsymbol{t}' \vee \boldsymbol{t}'') + g(\boldsymbol{t}' \wedge \boldsymbol{t}'').$$

We now prove the "only if" part. Assume to the contrary that $\mathcal{S}$ is not a mostly-lattice, i.e., there exists an unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$, and $\boldsymbol{x}' \in \mathcal{S}_{t'}$ and $\boldsymbol{x}'' \in \mathcal{S}_{t''}$ such that $(\boldsymbol{x}', \boldsymbol{t}') \vee (\boldsymbol{x}'', \boldsymbol{t}'') \notin \mathcal{S}$ or $(\boldsymbol{x}', \boldsymbol{t}') \wedge (\boldsymbol{x}'', \boldsymbol{t}'') \notin \mathcal{S}$. Let $\mathcal{W} = \{(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}''), (\boldsymbol{x}', \boldsymbol{t}') \vee (\boldsymbol{x}'', \boldsymbol{t}''), (\boldsymbol{x}', \boldsymbol{t}') \wedge (\boldsymbol{x}'', \boldsymbol{t}'')\}$ and define

$$f(\boldsymbol{x}, \boldsymbol{t}) = \max_{\boldsymbol{w} \in \mathcal{W}} \{-\|(\boldsymbol{x}, \boldsymbol{t}) - \boldsymbol{w}\|_1\}. \tag{3}$$

From Theorem 2.7.6 in Topkis (1998), $f(\boldsymbol{x}, \boldsymbol{t})$ is supermodular on $\mathcal{X} \times \mathcal{T}$. Since $f(\boldsymbol{w}) = 0$ for any $\boldsymbol{w} \in \mathcal{W}$ and $f(\boldsymbol{w}) < 0$ for any $\boldsymbol{w} \notin \mathcal{W}$, we have $g(\boldsymbol{t}) \leq 0$ for any $\boldsymbol{t} \in \mathcal{T}$, and $g(\boldsymbol{t}') = g(\boldsymbol{t}'') = 0$. In addition, $g(\boldsymbol{t}' \vee \boldsymbol{t}'') < 0$ if $(\boldsymbol{x}', \boldsymbol{t}') \vee (\boldsymbol{x}'', \boldsymbol{t}'') \notin \mathcal{S}$ or $g(\boldsymbol{t}' \wedge \boldsymbol{t}'') < 0$ if $(\boldsymbol{x}', \boldsymbol{t}') \wedge (\boldsymbol{x}'', \boldsymbol{t}'') \notin \mathcal{S}$ (the strict inequality follows from the closedness assumption of $\mathcal{S}_t$). In either case,

$$g(\boldsymbol{t}') + g(\boldsymbol{t}'') > g(\boldsymbol{t}' \vee \boldsymbol{t}'') + g(\boldsymbol{t}' \wedge \boldsymbol{t}''),$$

which implies that g is not supermodular. The "only if" part is now completed. Q.E.D.

The mostly-lattice requirement in Theorem 3 implies that to preserve supermodularity, there is no requirement for the ordered pair $\boldsymbol{t}', \boldsymbol{t}''$, since for ordered pair $\boldsymbol{t}', \boldsymbol{t}''$, $\{\boldsymbol{t}' \wedge \boldsymbol{t}'', \boldsymbol{t}' \vee \boldsymbol{t}''\} = \{\boldsymbol{t}', \boldsymbol{t}''\}$, and hence $g(\boldsymbol{t}') + g(\boldsymbol{t}'') \leq g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'')$ automatically holds.

Corollary 1 *When the set $\mathcal{S}$ is convex, the function g is supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is supermodular on $\mathcal{X} \times \mathcal{T}$ if and only if $\mathcal{S}$ is a mostly-lattice.*

To see the above result, note that the same condition in Theorem 3 must be a sufficient condition for the case when $\mathcal{S}$ is convex. In addition, the counter example constructed to prove the "necessary" direction in Theorem 3, is still a valid counter example when $\mathcal{S}$ is convex. Hence, $\mathcal{S}$ being a mostly-lattice is also a necessary condition for the case when $\mathcal{S}$ is convex.

We now investigate the conditions on $\boldsymbol{A} \in \Re^{m \times n_1}$ and $\boldsymbol{B} \in \Re^{m \times n_2}$ under which the corresponding polyhedron $\mathcal{S}_{\boldsymbol{c}}^P$ defined in the equation (2) is a mostly-lattice for any $\boldsymbol{c} \in \Re^m$. Recall that Theorem 2 implies the polyhedron $\mathcal{S}_{\boldsymbol{c}}^P$ is a lattice for any $\boldsymbol{c}$ if and only if $(\boldsymbol{A}|\boldsymbol{B})$ is a lattice-matrix. Interestingly, the lattice-matrix requirement also applies to the mostly-lattice.

Theorem 4 *The polyhedron $\mathcal{S}_{\boldsymbol{c}}^P$ defined by the equation (2) is a mostly-lattice for each $\boldsymbol{c} \in \Re^m$ if and only if $(\boldsymbol{A}|\boldsymbol{B})$ is a lattice-matrix.*

Theorem 4 implies that for $\mathcal{S}_{\boldsymbol{c}}^P$ to be a mostly-lattice for any $\boldsymbol{c}$, we need to guarantee that $\mathcal{S}_{\boldsymbol{c}}^P$ is a lattice for any $\boldsymbol{c}$. However, as we illustrated earlier, in general, a lattice imposes much stronger conditions than a mostly-lattice.

We now present the preservation results when $f(\boldsymbol{x}, \boldsymbol{t})$ is both concave and supermodular.

Theorem 5 *The function g is concave and supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is concave and supermodular on $\mathcal{X} \times \mathcal{T}$ if and only if the set $\mathcal{S}$ is a convex mostly-lattice.*

In Theorem 5, in addition to the mostly-lattice condition to preserve supermodularity (stated in Theorem 3), we need convexity of the feasible set $\mathcal{S}$ to preserve concavity of the function. The result seems to be intuitive. As mostly-lattice and convexity are needed to preserve the properties of supermodularity and concavity, respectively, we have to satisfy both requirements for preserving the two properties at the same time. Nevertheless, in the next section, we will show that, when f has some special properties such as independence from $\boldsymbol{t}$, the condition for preserving the two properties simultaneously is no longer a simple addition of the two individual conditions.

From Theorem 5, we can also easily derive the condition for preserving concavity and supermodularity when $\mathcal{S}$ is already convex.

Corollary 2 *Consider the case that $\mathcal{S}$ is convex. The function g is concave and supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is concave and supermodular on $\mathcal{X} \times \mathcal{T}$ if and only if $\mathcal{S}$ is a mostly-lattice.*

Since a polyhedral set is convex, we immediately get that a sufficient and necessary condition to preserve both concavity and supermodularity under the optimization operation (1) when $\mathcal{S}$ is specified by (2) is $\mathcal{S}_{\boldsymbol{c}}^P$ being a mostly-lattice. The preservation holds for any $\boldsymbol{c} \in \Re^m$ if and only if $(\boldsymbol{A}|\boldsymbol{B})$ is a lattice-matrix.

3. Main Results for A Special Class of Problems

In this section, we consider the case where the function f is independent of the parameters $\boldsymbol{t}$, i.e., $f(\boldsymbol{x},\boldsymbol{t}') = f(\boldsymbol{x},\boldsymbol{t}'')$, $\forall \boldsymbol{x} \in \mathcal{X}, \boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$. To simplify the notation, in the rest of the paper, we henceforth abuse the notation and let f be a function mapping from $\mathcal{X}$ to $\Re$. Therefore, the problem (1) becomes

$$g(\boldsymbol{t}) = \max\{f(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{S}_t\}. \tag{4}$$

This special class of the parametric optimization problems is widely used in many operations models (see Chen et al. 2013). Its special structure allows us to derive less restrictive conditions to preserve supermodularity. We first introduce a definition of additive mostly-lattice.

Definition 2 *Given a mapping $\mathcal{W} : \mathcal{X} \times \mathcal{X} \to 2^{\mathcal{X}}$ satisfying the condition $\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'')$, a set $\mathcal{S} \subseteq \mathcal{X} \times \mathcal{T}$ is an additive mostly-lattice with $\mathcal{W}$ if and only if for any unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$ and any $\boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$, there exist $\boldsymbol{y} \in Conv(\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') \cap \mathcal{S}_{t' \wedge t''})$ and $\boldsymbol{z} \in Conv(\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') \cap \mathcal{S}_{t' \vee t''})$ such that $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$. If $\mathcal{W} = \mathcal{X}$, we simply say $\mathcal{S}$ is an additive mostly-lattice.*

Mostly-lattices are special cases of additive mostly-lattices with any $\mathcal{W}$ satisfying $\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x} \vee \boldsymbol{x}'' \in \mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'')$. To show that, consider any mostly-lattice $\mathcal{S}$ and any unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$, $\boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$. Based on the definition, we have $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}, \boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$. For any mapping $\mathcal{W}$ satisfying the condition $\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'')$, let $\boldsymbol{y} = \boldsymbol{x}' \wedge \boldsymbol{x}''$ and $\boldsymbol{z} = \boldsymbol{x}' \vee \boldsymbol{x}''$. Clearly, $\boldsymbol{y} \in Conv(\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') \cap \mathcal{S}_{t' \wedge t''})$, $\boldsymbol{z} \in Conv(\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') \cap \mathcal{S}_{t' \vee t''})$, and $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' \wedge \boldsymbol{x}'' + \boldsymbol{x}' \vee \boldsymbol{x}'' = \boldsymbol{x}' + \boldsymbol{x}''$. Hence, $\mathcal{S}$ is an additive mostly-lattice with any $\mathcal{W}$ satisfying $\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'')$. We now provide an example which is an additive mostly-lattice but not a mostly-lattice.

Example 3. Let $\mathcal{X} = \Re, \mathcal{T} = \Re^2$, $\mathcal{S} = \{(0,0,1),(0,1,1),(1,0,0),(1,1,0)\}$. Note that the only unordered pair of $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{S}|_{\mathcal{T}}$ is $\boldsymbol{t}' = (0,1)$, $\boldsymbol{t}'' = (1,0)$, and the corresponding $x' = 0 \in \mathcal{S}_{t'}$, $x'' = 1 \in \mathcal{S}_{t''}$. However, since $x' \wedge x'' = 0 \notin \mathcal{S}_{t' \wedge t''} = \mathcal{S}_{(0,0)} = \{1\}$, set $\mathcal{S}$ is not a mostly-lattice. Let $\mathcal{W}(x', x'') = \{x', x'', x' \wedge x'', x' \vee x''\}$, and $y = 1 \in \mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''}$, $z = 0 \in \mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''}$. We have $y + z =$

$x' + x''$. Hence, $\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}$. We can also verify that $Conv(\mathcal{S})$ is not a mostly-lattice but is an additive mostly-lattice with $\mathcal{W}$.

For the optimization problem (4) with objective functions independent of $\boldsymbol{t}$, we are now ready to present a necessary and sufficient condition to preserve supermodularity.

Theorem 6 *The function g is supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is supermodular on $\mathcal{X}$ if and only if set $\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$.*

Proof. We first prove the "if" part. Suppose set $\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$. Consider any unordered pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$ and any $\boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$, and denote a set $\mathcal{W}^o = \mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'')$.

If $\boldsymbol{x}', \boldsymbol{x}''$ are ordered, then $\mathcal{W}^o = \{\boldsymbol{x}', \boldsymbol{x}''\}$. Since there exists $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})$, we can find $\alpha \in [0, 1]$ such that $\boldsymbol{y} = \alpha \boldsymbol{x}' + (1-\alpha)\boldsymbol{x}''$. Moreover, there exists $\boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}'' - \boldsymbol{y} = (1-\alpha)\boldsymbol{x}' + \alpha \boldsymbol{x}''$ and $\boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \vee t''})$. For the case of $\alpha = 0$, it is easy to observe that $\boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$ and $\boldsymbol{x}' \in \mathcal{S}_{t' \vee t''}$. Consider the case where $\alpha \neq 0$. If $\boldsymbol{x}' \notin \mathcal{S}_{t' \wedge t''}$, we have $\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''} \subseteq \{\boldsymbol{x}''\}$ which contradicts with the fact $\alpha \boldsymbol{x}' + (1-\alpha)\boldsymbol{x}'' = \boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})$ and $\alpha \neq 0$. Hence, we have $\boldsymbol{x}' \in \mathcal{S}_{t' \wedge t''}$. Similarly, $\boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$. Therefore, no matter whether $\alpha = 0$ or $\alpha \neq 0$, we always have

$$g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') = \max_{x \in \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) + \max_{x \in \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) \geq f(\boldsymbol{x}') + f(\boldsymbol{x}''). \tag{5}$$

If $\boldsymbol{x}', \boldsymbol{x}''$ are unordered, then $\mathcal{W}^o = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$. Since $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})$, there exist $\alpha, \beta, \gamma, \lambda \geq 0, \alpha + \beta + \gamma + \lambda = 1$ and $\lambda \times \gamma = 0$ such that $\boldsymbol{y} = \alpha \boldsymbol{x}' + \beta \boldsymbol{x}'' + \gamma(\boldsymbol{x}' \wedge \boldsymbol{x}'') + \lambda(\boldsymbol{x}' \vee \boldsymbol{x}'')$. WLOG, we assume $\lambda = 0$ (the case with $\lambda \neq 0$ implies $\gamma = 0$ and can be proved similarly). Hence, there exists $\boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}'' - \boldsymbol{y} = (1 - \gamma - \alpha)\boldsymbol{x}' + (1 - \gamma - \beta)\boldsymbol{x}'' + \gamma(\boldsymbol{x}' \vee \boldsymbol{x}'') = \beta \boldsymbol{x}' + \alpha \boldsymbol{x}'' + \gamma(\boldsymbol{x}' \vee \boldsymbol{x}'')$ and $\boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \vee t''})$.

In the case of $\gamma = 0$, $\boldsymbol{y} = \alpha \boldsymbol{x}' + (1-\alpha)\boldsymbol{x}''$ and $\boldsymbol{z} = (1-\alpha)\boldsymbol{x}' + \alpha \boldsymbol{x}''$. Similar to the case with $\boldsymbol{x}', \boldsymbol{x}''$ being ordered, we always have either 1) $\boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$ and $\boldsymbol{x}' \in \mathcal{S}_{t' \vee t''}$ if $\alpha = 0$, or 2) $\boldsymbol{x}' \in \mathcal{S}_{t' \wedge t''}$ and $\boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$ if $\alpha \neq 0$, and thus the inequality (5) holds.

In the case of $\gamma \neq 0$, we have $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$, and correspondingly, $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$. Hence,

$$g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') = \max_{x \in \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) + \max_{x \in \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) \geq f(\boldsymbol{x}' \wedge \boldsymbol{x}'') + f(\boldsymbol{x}' \vee \boldsymbol{x}'') \geq f(\boldsymbol{x}') + f(\boldsymbol{x}''), \quad (6)$$

where the last inequality follows from the supermodularity of f.

For both inequalities (5) and (6), taking the maximum of the right hand side over all $\boldsymbol{x}' \in \mathcal{S}_{t'}$ and $\boldsymbol{x}'' \in \mathcal{S}_{t''}$, we have $g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') \geq g(\boldsymbol{t}') + g(\boldsymbol{t}'')$, i.e., $g(\boldsymbol{t})$ is supermodular.

We next prove the "only if" part. Suppose that there exists an unordered pair $\boldsymbol{t}', \boldsymbol{t}''$, and $\boldsymbol{x}' \in \mathcal{S}_{t'}$, $\boldsymbol{x}'' \in \mathcal{S}_{t''}$, such that there does not exist $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})$, and $\boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \vee t''})$ with $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$. Here we denote the set $\mathcal{W}^o = \mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'')$. Let

$$\mathcal{H} = \{\boldsymbol{y} + \boldsymbol{z} \ : \ \boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}), \ \boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \vee t''})\}.$$

Observing that $\mathcal{H}$ is a closed convex set and $(\boldsymbol{x}' + \boldsymbol{x}'') \notin \mathcal{H}$. By the separating hyperplane theorem (e.g., Bertsimas and Tsitsiklis 1997, Theorem 4.11), there exists a vector $\boldsymbol{a} \in \Re^{dim(\boldsymbol{x}')}$ and a scalar $b \in \Re$ such that $\boldsymbol{a}^T(\boldsymbol{x}' + \boldsymbol{x}'') > b > \boldsymbol{a}^T \boldsymbol{h}$, $\forall \boldsymbol{h} \in \mathcal{H}$. We now construct a function $f : \mathcal{X} \to \Re$ as

$$f(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{x} + K \times \max_{\boldsymbol{w} \in \mathcal{W}^o} \{-||\boldsymbol{x} - \boldsymbol{w}||_1\},$$

where K is positive constant chosen based on Lemma 2 in the appendix such that

$$\max_{\boldsymbol{x} \in \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) = \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}), \ \max_{\boldsymbol{x} \in \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) = \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}).$$

Clearly, f is supermodular. However,

$$\begin{aligned}
g(\boldsymbol{t}') + g(\boldsymbol{t}'') &\geq f(\boldsymbol{x}') + f(\boldsymbol{x}'') = \boldsymbol{a}^T(\boldsymbol{x}' + \boldsymbol{x}'') \ > \ b, \\
g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') &= \max_{\boldsymbol{x} \in \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) + \max_{\boldsymbol{x} \in \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) \\
&= \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) + \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) \\
&= \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}} \boldsymbol{a}^T \boldsymbol{x} + \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \vee t''}} \boldsymbol{a}^T \boldsymbol{x} \\
&= \max_{\boldsymbol{x} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})} \boldsymbol{a}^T \boldsymbol{x} + \max_{\boldsymbol{x} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \vee t''})} \boldsymbol{a}^T \boldsymbol{x} \\
&= \max_{\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}), \boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \vee t''})} \boldsymbol{a}^T(\boldsymbol{y} + \boldsymbol{z}) \\
&= \max_{\boldsymbol{h} \in \mathcal{H}} \boldsymbol{a}^T \boldsymbol{h} \\
&< b,
\end{aligned}$$

which implies $g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') < g(\boldsymbol{t}') + g(\boldsymbol{t}'')$, i.e., $g(\boldsymbol{t})$ is not supermodular. Q.E.D.

As discussed before, all mostly-lattices are additive mostly-lattices with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$ while an additive mostly-lattice may fail to be a mostly-lattice. Hence, Theorem 6 implies that when function f is independent of the decision variables $\boldsymbol{t}$, the condition for preserving supermodularity is less restrictive than that for the general case stated in Theorem 3.

Similar to the general case, the convexity of $\mathcal{S}$ does not play any role in preserving supermodularity even when f is independent from $\boldsymbol{t}$. For completeness, we formalize the result as follows.

Corollary 3 *When the set $\mathcal{S}$ is convex, the function g is supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is supermodular on $\mathcal{X}$ if and only if set $\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$.*

We now investigate when the polyhedron defined by (2) is an additive mostly-lattice.

Theorem 7 *The polyhedron $\mathcal{S}^P_{\boldsymbol{c}}$ defined by the equation (2) is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$ for each $\boldsymbol{c} \in \Re^m$ if and only if at least one of the following conditions holds.*

1. *Either $(\boldsymbol{A}|\boldsymbol{B})$ or $(-\boldsymbol{A}|\boldsymbol{B})$ is a lattice-matrix.*
2. *Denote $\mathcal{I} = \{i : \boldsymbol{b}_i \neq \boldsymbol{0}\} \subseteq \{1, \ldots, m\}$. There exist $\boldsymbol{k} \in \Re^{|\mathcal{I}|}$, $\boldsymbol{d} \in \Re^{n_1}$ such that $\boldsymbol{A}_{\mathcal{I}} = \boldsymbol{k}\boldsymbol{d}^T$, and $(\boldsymbol{k}|\boldsymbol{B}_{\mathcal{I}})$ is a lattice-matrix.*

Recall that the condition on $\boldsymbol{A}, \boldsymbol{B}$ imposed by mostly-lattice is that $(\boldsymbol{A}|\boldsymbol{B})$ has to be a lattice-matrix. Here, for $\mathcal{S}^P_{\boldsymbol{c}}$ to be an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$, the first condition in Theorem 7 implies that other than $(\boldsymbol{A}|\boldsymbol{B})$ being a lattice-matrix, $(-\boldsymbol{A}|\boldsymbol{B})$ being so also works. This is due to the additional flexibility of the additive mostly-lattice over the mostly-lattice. Consider any unordered pair $\boldsymbol{t}', \boldsymbol{t}''$, and $\boldsymbol{x}' \in \mathcal{S}_{t'}$, $\boldsymbol{x}'' \in \mathcal{S}_{t''}$. Note that mostly-lattice requires $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$, $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$. The additive mostly-lattice, however, is fine with either the same condition (i.e., $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$ and $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$) or its reverse: $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$ and $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$

(e.g., Example 3). Essentially, while the former leads to $(\boldsymbol{A}|\boldsymbol{B})$ being a lattice-matrix, the latter results in $(-\boldsymbol{A}|\boldsymbol{B})$ being a lattice-matrix.

We illustrate the condition by two examples: $\mathcal{S}^P_{\boldsymbol{c},1} = \{(\boldsymbol{x},\boldsymbol{t}) : x_1 + t_1 \le c_1, -x_2 - t_2 \le c_2\}$ and $\mathcal{S}^P_{\boldsymbol{c},2} = \{(\boldsymbol{x},\boldsymbol{t}) : -\sum_{i=1}^n x_i + t_1 \le c_1, -\sum_{i=1}^n x_i + t_2 \le c_2\}$. Specifically, $\mathcal{S}^P_{\boldsymbol{c},1}$ satisfies the first condition while $\mathcal{S}^P_{\boldsymbol{c},2}$ satisfies the second condition in Theorem 7. Hence, the optimization problem (4) preserves supermodularity in both examples. However, they do not satisfy the condition in Theorem 4. Hence, there exist $\boldsymbol{c}$ and a (concave and) supermodular function $f(\boldsymbol{x},\boldsymbol{t})$ such that g is not (concave and) supermodular under the optimization operation (1) when $\mathcal{S}$ is specified by $\mathcal{S}^P_{\boldsymbol{c},1}$ or $\mathcal{S}^P_{\boldsymbol{c},2}$.

We now derive a necessary and sufficient condition to preserve both concavity and supermodularity under the optimization operation (4).

Theorem 8 *The function g is concave and supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is concave and supermodular on $\mathcal{X}$ if and only if the following two conditions hold simultaneously:*

1. *$\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'')$.*
2. *For any $\boldsymbol{t}',\boldsymbol{t}'' \in \mathcal{T}$, $\boldsymbol{x}' \in \mathcal{S}_{t'}$, $\boldsymbol{x}'' \in \mathcal{S}_{t''}$ and $\lambda \in [0,1]$, we have $\boldsymbol{x}_\lambda \in Conv(\mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'') \cap \mathcal{S}_{t_\lambda})$.*

Here $\boldsymbol{t}_\lambda = \lambda\boldsymbol{t}' + (1-\lambda)\boldsymbol{t}''$, $\boldsymbol{x}_\lambda = \lambda\boldsymbol{x}' + (1-\lambda)\boldsymbol{x}''$ and $\mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'') = Conv(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}'')$.

Proof. We first prove the "sufficient" direction. Suppose f is supermodular and concave, and both conditions 1 and 2 hold. Consider any $(\boldsymbol{x}',\boldsymbol{t}'),(\boldsymbol{x}'',\boldsymbol{t}'') \in \mathcal{S}$, and denote the set $\mathcal{W}^o = \mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'') = Conv(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}'')$.

To prove the concavity of $g(\boldsymbol{t})$, we need to show $g(\boldsymbol{t}_\lambda) \ge \lambda g(\boldsymbol{t}') + (1-\lambda)g(\boldsymbol{t}'')$. Since $\boldsymbol{x}_\lambda \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t_\lambda})$, following Lemma 4 in the appendix, we have

$$g(\boldsymbol{t}_\lambda) = \max_{\boldsymbol{x}\in\mathcal{S}_{t_\lambda}} f(\boldsymbol{x}) \ge \max_{\boldsymbol{x}\in\mathcal{W}^o\cap\mathcal{S}_{t_\lambda}} f(\boldsymbol{x}) \ge \lambda f(\boldsymbol{x}') + (1-\lambda)f(\boldsymbol{x}'').$$

Taking the maximum on the right hand side over all $\boldsymbol{x}' \in \mathcal{S}_{t'}$ and $\boldsymbol{x}'' \in \mathcal{S}_{t''}$, we get $g(\boldsymbol{t}_\lambda) \ge \lambda g(\boldsymbol{t}') + (1-\lambda)g(\boldsymbol{t}'')$. Equivalently, the function $g(\boldsymbol{t})$ is concave.

To prove the supermodularity of the function $g(\boldsymbol{t})$, it suffices to prove $g(\boldsymbol{t}'\wedge\boldsymbol{t}'') + g(\boldsymbol{t}'\vee\boldsymbol{t}'') \ge g(\boldsymbol{t}') + g(\boldsymbol{t}'')$ for any unordered pair $\boldsymbol{t}',\boldsymbol{t}''$. Since set $\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}^o$, we can find $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t'\wedge t''})$, $\boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t'\vee t''})$ such that $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$.

If $\boldsymbol{x}', \boldsymbol{x}''$ are ordered, $\mathcal{W}^o = Conv(\boldsymbol{x}', \boldsymbol{x}'')$. In this case, $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})$ implies that $\exists \beta \in [0,1]$ such that $\boldsymbol{y} = \beta \boldsymbol{x}' + (1-\beta)\boldsymbol{x}''$. Hence, $\boldsymbol{z} = (\boldsymbol{x}' + \boldsymbol{x}'') - \boldsymbol{y} = (1-\beta)\boldsymbol{x}' + \beta \boldsymbol{x}''$. Since $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})$, there exist $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}$ such that $\boldsymbol{y} \in Conv(\boldsymbol{y}_1, \boldsymbol{y}_2)$. Similarly, there exist $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{W}^o \cap \mathcal{S}_{t' \vee t''}$ such that $\boldsymbol{z} \in Conv(\boldsymbol{z}_1, \boldsymbol{z}_2)$. Therefore,

$$\begin{aligned}
g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') &= \max_{\boldsymbol{x} \in \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) + \max_{\boldsymbol{x} \in \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) \\
&\geq \max\{f(\boldsymbol{y}_1), f(\boldsymbol{y}_2)\} + \max\{f(\boldsymbol{z}_1), f(\boldsymbol{z}_2)\} \\
&\geq (\beta f(\boldsymbol{x}') + (1-\beta) f(\boldsymbol{x}'')) + ((1-\beta) f(\boldsymbol{x}') + \beta f(\boldsymbol{x}'')) \\
&= f(\boldsymbol{x}') + f(\boldsymbol{x}''),
\end{aligned}$$

where the second inequality follows from Lemma 5 in the appendix.

If $\boldsymbol{x}', \boldsymbol{x}''$ are unordered, there exist $\lambda_i, \mu_i, \nu_i, \beta_i \in [0,1]$, $\lambda_i + \mu_i + \nu_i + \beta_i = 1$ for $i = 1,2$ such that $\boldsymbol{y} = \lambda_1 \boldsymbol{x}' + \mu_1 \boldsymbol{x}'' + \nu_1 (\boldsymbol{x}' \wedge \boldsymbol{x}'') + \beta_1 (\boldsymbol{x}' \vee \boldsymbol{x}'')$, $\boldsymbol{z} = \lambda_2 \boldsymbol{x}' + \mu_2 \boldsymbol{x}'' + \nu_2 (\boldsymbol{x}' \wedge \boldsymbol{x}'') + \beta_2 (\boldsymbol{x}' \vee \boldsymbol{x}'')$. Moreover, WLOG, we can have $\nu_i \times \beta_i = 0, i = 1,2$ as $\boldsymbol{x}' \wedge \boldsymbol{x}'' + \boldsymbol{x}' \vee \boldsymbol{x}'' = \boldsymbol{x}' + \boldsymbol{x}''$. Since $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''})$, $\boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t' \vee t''})$, following Lemma 4 in the appendix, we have

$$g(\boldsymbol{t}' \wedge \boldsymbol{t}'') = \max_{\boldsymbol{x} \in \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) \geq \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}} f(\boldsymbol{x}) \geq \lambda_1 f(\boldsymbol{x}') + \mu_1 f(\boldsymbol{x}'') + \nu_1 f(\boldsymbol{x}' \wedge \boldsymbol{x}'') + \beta_1 f(\boldsymbol{x}' \vee \boldsymbol{x}''), \quad (7)$$

$$g(\boldsymbol{t}' \vee \boldsymbol{t}'') = \max_{\boldsymbol{x} \in \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) \geq \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t' \vee t''}} f(\boldsymbol{x}) \geq \lambda_2 f(\boldsymbol{x}') + \mu_2 f(\boldsymbol{x}'') + \nu_2 f(\boldsymbol{x}' \wedge \boldsymbol{x}'') + \beta_2 f(\boldsymbol{x}' \vee \boldsymbol{x}''). \quad (8)$$

Since $\boldsymbol{x}', \boldsymbol{x}''$ are unordered, based on Lemma 6 in the appendix, we can get $\mu_1 + \nu_1 + \mu_2 + \nu_2 = 1, \lambda_1 + \nu_1 + \lambda_2 + \nu_2 = 1, \beta_1 + \beta_2 = \nu_1 + \nu_2$. Therefore,

$$\begin{aligned}
& g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') \\
\geq\ & (\lambda_1 + \lambda_2) f(\boldsymbol{x}') + (\mu_1 + \mu_2) f(\boldsymbol{x}'') + (\nu_1 + \nu_2) f(\boldsymbol{x}' \wedge \boldsymbol{x}'') + (\beta_1 + \beta_2) f(\boldsymbol{x}' \vee \boldsymbol{x}'') \\
=\ & (\lambda_1 + \lambda_2) f(\boldsymbol{x}') + (\mu_1 + \mu_2) f(\boldsymbol{x}'') + (\nu_1 + \nu_2)(f(\boldsymbol{x}' \wedge \boldsymbol{x}'') + f(\boldsymbol{x}' \vee \boldsymbol{x}'')) \\
\geq\ & (\lambda_1 + \lambda_2 + \nu_1 + \nu_2) f(\boldsymbol{x}') + (\mu_1 + \mu_2 + \nu_1 + \nu_2) f(\boldsymbol{x}'') \\
=\ & f(\boldsymbol{x}') + f(\boldsymbol{x}''),
\end{aligned}$$

where the first inequality is a result of the inequalities (7) and (8) and the second inequality follows from the supermodularity of f. Therefore, $g(\boldsymbol{t}' \vee \boldsymbol{t}'') + g(\boldsymbol{t}' \wedge \boldsymbol{t}'') \geq f(\boldsymbol{x}') + f(\boldsymbol{x}'')$ for all $\boldsymbol{x}' \in \mathcal{S}_{t'}$

and $\boldsymbol{x}'' \in \mathcal{S}_{t''}$. Taking the maximum of the right hand side over all $\boldsymbol{x}' \in \mathcal{S}_{t''}$ and $\boldsymbol{x}'' \in \mathcal{S}_{t''}$, we have $g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}'') \geq g(\boldsymbol{t}') + g(\boldsymbol{t}'')$, i.e., $g(\boldsymbol{t})$ is supermodular.

We now prove the "only if" part. Suppose condition 2 does not hold, i.e., there exists $\lambda \in [0,1]$ such that $\boldsymbol{x}_\lambda \notin Conv(\mathcal{W}^o \cap \mathcal{S}_{t_\lambda})$ for some $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$, $\boldsymbol{x}' \in \mathcal{S}_{t'}$, $\boldsymbol{x}'' \in \mathcal{S}_{t''}$. Therefore, according to the separating hyperplane theorem, there exist a vector $\boldsymbol{a} \in \Re^{dim(\boldsymbol{x}')}$ and a scalar $b \in \Re$ such that $\boldsymbol{a}^T \boldsymbol{x}_\lambda > b > \boldsymbol{a}^T \boldsymbol{w}$, $\forall \boldsymbol{w} \in Conv(\mathcal{W}^o \cap \mathcal{S}_{t_\lambda})$. We define

$$f(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{x} + K \times \max_{\boldsymbol{w} \in \mathcal{W}^o} \{-||\boldsymbol{x} - \boldsymbol{w}||_1\}.$$

with a large $K > 0$ such that $\max_{\boldsymbol{x} \in \mathcal{S}_{t_\lambda}} f(\boldsymbol{x}) = \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t_\lambda}} f(\boldsymbol{x})$ (see Lemma 2 in the appendix). Since $\mathcal{W}^o$ is a lattice and convex set (see Lemma 1 in the appendix), the function $f(\boldsymbol{x})$ is concave and supermodular. Observe that $f(\boldsymbol{x}') = \boldsymbol{a}^T \boldsymbol{x}'$, $f(\boldsymbol{x}'') = \boldsymbol{a}^T \boldsymbol{x}''$, $\lambda f(\boldsymbol{x}') + (1-\lambda) f(\boldsymbol{x}'') = \boldsymbol{a}^T \boldsymbol{x}_\lambda > b$,

$$g(\boldsymbol{t}_\lambda) = \max_{\boldsymbol{x} \in \mathcal{S}_{t_\lambda}} f(\boldsymbol{x}) = \max_{\boldsymbol{x} \in \mathcal{W}^o \cap \mathcal{S}_{t_\lambda}} f(\boldsymbol{x}) < b < \lambda f(\boldsymbol{x}') + (1-\lambda) f(\boldsymbol{x}'') \leq \lambda g(\boldsymbol{t}') + (1-\lambda) g(\boldsymbol{t}'').$$

Therefore, the function $g(\boldsymbol{t})$ is not concave.

Suppose condition 1 does not hold. The proof is similar to the proof in Theorem 6 for the "only if" part. The only difference is that in constructing the function f, now the set $\mathcal{W}^o$ is chosen as $\mathcal{W}^o = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'')$ instead of $\mathcal{W}^o = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}''\}$ in Theorem 6. Q.E.D.

When function f is independent of $\boldsymbol{t}$, comparing the condition to preserve supermodularity (Theorem 6) and that to preserve both concavity and supermodularity (Theorem 8), we observe two differences. Firstly, for any $\boldsymbol{x}', \boldsymbol{x}''$, the set $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'')$ in Theorem 6 is a subset of that in Theorem 8, i.e., $\{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}''\} \subseteq Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'')$. According to Definition 2, any additive mostly-lattice in Theorem 6 must be an additive mostly-lattice in Theorem 8. It implies that the additional property of concavity leads to a less restrictive condition. Secondly, Theorem 8 requires condition 2 to ensure the preservation of concavity. This condition is weaker than requiring the set $\mathcal{S}$ to be convex, which is commonly imposed for preservation of concavity under optimization operations. Therefore, the conditions to preserve concavity and supermodularity is not a simple addition of that to preserve supermodularity and that to preserve concavity. In the following, we provide an example which satisfies the conditions in Theorem 8 but is not convex.

EXAMPLE 4. Let $\mathcal{X} = \Re, \mathcal{T} = \Re^2, \mathcal{S} = \{(x, t_1, t_2) : x \in \{0,1\}, t_1, t_2 \in [0,1]\}$. For any pair $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{S}|_{\mathcal{T}}$ and $\lambda \in [0,1]$, we have $\mathcal{S}_{t'} = \mathcal{S}_{t''} = \mathcal{S}_{t' \wedge t''} = \mathcal{S}_{t' \vee t''} = \mathcal{S}_{t_\lambda} = \{0,1\}$. Hence, if $x' = x''$, we have $\mathcal{W}(x', x'') = Conv(x', x'', x' \wedge x'', x' \vee x'') = \{x'\} \subseteq \{0,1\}$ and $Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''}) = Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''}) = Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t_\lambda}) = \{x'\}$. Define $y = z = x'$. We have $y \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''}), z \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''})$ and $y + z = x' + x''$. In addition, $x_\lambda = x' \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t_\lambda})$. If $x' \neq x''$, WLOG, let $x' = 0, x'' = 1$. Then, we have $\mathcal{W}(x', x'') = [0,1]$ and $Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''}) = Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''}) = Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t_\lambda}) = [0,1]$. Let $y = 0 \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''}), z = 1 \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''})$. We have $y + z = x' + x''$. In addition, $x_\lambda = \lambda x' + (1 - \lambda) x'' = 1 - \lambda \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t_\lambda})$. Hence, set $\mathcal{S}$ satisfies both conditions in Theorem 8 but is not a convex set.

When set $\mathcal{S}$ is already convex, the conditions in Theorem 8 can be simplified.

Corollary 4 *Consider the case that $\mathcal{S}$ is convex. The function g is concave and supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is concave and supermodular on $\mathcal{X}$ if and only if set $\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$.*

The condition in Corollary 4 is less restrictive than that in Corollary 2 due to the independence of f from $\boldsymbol{t}$ as illustrated by the following example. Interestingly, this example is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$ (and thus satisfies the condition in Corollary 4) but not an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$ (and thus violates the condition in Corollary 3).

EXAMPLE 5. Let $\mathcal{X} = \Re, \mathcal{T} = \Re^2, \mathcal{S} = \{(x, t_1, t_2) : x - t_1 + t_2 = 1, t_1, t_2 \in [0,1]\}$ and consider any $(x', t_1', t_2'), (x'', t_1'', t_2'') \in \mathcal{S}$ and $(t_1', t_2'), (t_1'', t_2'')$ are unordered. WLOG, we assume that $t_1' < t_1''$ and $t_2' > t_2''$. Hence, we have $x' = 1 + t_1' - t_2' < x'' = 1 + t_1'' - t_2''$ and $\boldsymbol{t}' \wedge \boldsymbol{t}'' = (t_1', t_2''), \boldsymbol{t}' \vee \boldsymbol{t}'' = (t_1'', t_2')$. It is easy to get that $\mathcal{S}_{t' \wedge t''} = \{1 + t_1' - t_2''\}$ and $\mathcal{S}_{t' \vee t''} = \{1 + t_1'' - t_2'\}$. Clearly, $\mathcal{S}$ is not a mostly-lattice.

To show that $\mathcal{S}$ is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$, let $\mathcal{W}(x', x'') = Conv(x', x'', x' \wedge x'', x' \vee x'') = Conv(x', x'') = [1 + t_1' - t_2', 1 + t_1'' - t_2'']$. We can define $y = 1 + t_1' - t_2'' \in [1 + t_1' - t_2', 1 + t_1'' - t_2'']$ and $z = 1 + t_1'' - t_2' \in [1 + t_1' - t_2', 1 + t_1'' - t_2'']$. It is easy to verify that $y \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''})$, $z \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''})$ and $y + z = x' + x''$.

To show $\mathcal{S}$ is not an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$, let $\mathcal{W}(x', x'') = \{x', x'', x' \wedge x'', x' \vee x''\} = \{x', x''\} = \{1 + t_1' - t_2', 1 + t_1'' - t_2''\}$. Since $1 + t_1' - t_2' < 1 + t_1' - t_2'' < 1 + t_1'' - t_2''$ and $1 + t_1' - t_2' < 1 + t_1'' - t_2' < 1 + t_1'' - t_2''$, we can easily verify that $\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''} = \mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''} = \emptyset$. Hence, there do not exist $y \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''})$ and $z \in Conv(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''})$ such that $y + z = x' + x''$.

We now focus on the polyhedron defined by the equation (2).

Theorem 9 *The polyhedron $\mathcal{S}_{\boldsymbol{c}}^P$ defined by the equation (2) is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$ for any $\boldsymbol{c} \in \Re^m$ only if for any $\boldsymbol{\alpha} \in \Re^{n_1}$, $\boldsymbol{\beta} \in \Re^{n_2}$, $\mathcal{I} \subseteq \{1, \ldots, m\}$, and $\boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta} = \boldsymbol{A}_{\mathcal{I}} \boldsymbol{\alpha}$, there exist $\lambda_1, \lambda_2 \in [0, 1]$ such that*

$$\boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}^+ = \boldsymbol{A}_{\mathcal{I}} \left(\lambda_1 \boldsymbol{\alpha}^+ - \lambda_2 (-\boldsymbol{\alpha})^+ \right),$$

which in turn implies that $\boldsymbol{B}$ is a lattice-matrix.

Note that the condition stated in Theorem 9 is only a necessary condition. We conjecture that the condition is a sufficient condition as well. In the following, we show the conjecture is valid for two special cases: 1) $rank(\boldsymbol{B}) = 1$; 2) $m = 2$ (i.e., $\boldsymbol{A}$ and $\boldsymbol{B}$ have only two rows).

Proposition 1 *Suppose $rank(\boldsymbol{B}) = 1$. The polyhedron $\mathcal{S}_{\boldsymbol{c}}^P$ defined by the equation (2) is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$ for any $\boldsymbol{c} \in \Re^m$ if and only if $\boldsymbol{B}$ is a lattice-matrix.*

Proposition 2 *Suppose both $\boldsymbol{A}$ and $\boldsymbol{B}$ have two rows (i.e., $m = 2$). The following three statements are equivalent.*

- **S1**. *The polyhedron* $\mathcal{S}^P_{\boldsymbol{c}}$ *defined by the equation (2) is an additive mostly-lattice with* $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$ *for any* $\boldsymbol{c} \in \Re^m$.
- **S2**. *For any* $\boldsymbol{\alpha} \in \Re^{n_1}$, $\boldsymbol{\beta} \in \Re^{n_2}$, $\mathcal{I} \subseteq \{1,2\}$, *and* $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta} = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{\alpha}$, *there exist* $\lambda_1, \lambda_2 \in [0,1]$ *such that*

$$\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+ = \boldsymbol{A}_{\mathcal{I}}\left(\lambda_1 \boldsymbol{\alpha}^+ - \lambda_2(-\boldsymbol{\alpha})^+\right).$$

- **S3**. $\boldsymbol{B}$ *is a lattice-matrix. When* $rank(\boldsymbol{B}) = 2$, $\boldsymbol{B}$ *has either two or three columns (i.e.,* $n_2 = 2$ *or* $n_2 = 3$*) and satisfies the following conditions:*

When $n_2 = 2$, *the matrix* $\boldsymbol{D} = \boldsymbol{B}^{-1}\boldsymbol{A}$ *should satisfy 1)* $d_{11}d_{21} \geq 0$ *in case of* $rank(\boldsymbol{D}) = 1$, *and 2)* $\boldsymbol{D} \geq \boldsymbol{0}$ *or* $\boldsymbol{D} \leq \boldsymbol{0}$ *in case of* $rank(\boldsymbol{D}) = 2$; *When* $n_2 = 3$, *there exists* $i \in \{1,2,3\}$ *such that* $\boldsymbol{B}_i$ *has no zero element, and* $\boldsymbol{A} = \boldsymbol{B}_i\boldsymbol{k}^T$ *for some* $\boldsymbol{k} \in \Re^{n_1}$.

In Proposition 2, the statement **S2** indicates that the necessary condition provided by Theorem 9 is also sufficient in this particular case. Meanwhile, the statement **S3** explicitly characterizes the structure of such matrices $\boldsymbol{A}, \boldsymbol{B}$. When $n_2 = 2$ and $rank(\boldsymbol{D}) = 2$, this condition implies that either all column vectors in $\boldsymbol{A}$ are conic combination of $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$, or all column vectors in $\boldsymbol{A}$ are conic combination of $(-\boldsymbol{B}_1)$ and $(-\boldsymbol{B}_2)$. Proposition 2 can be slightly extended.

Proposition 3 *For* $m = 2$, *given any nonempty convex lattice* $\mathcal{D}$, *both set* $\{(\boldsymbol{x}, \boldsymbol{t}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{t} = \boldsymbol{c}, \boldsymbol{x} \in \mathcal{D}\}$ *and set* $\{(\boldsymbol{x}, \boldsymbol{t}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{t} \leq \boldsymbol{c}, \boldsymbol{x} \in \mathcal{D}\}$ *are additive mostly-lattice with* $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$ *for all* $\boldsymbol{c} \in \Re^2$ *if* $\boldsymbol{A}, \boldsymbol{B}$ *satisfy the condition in Statement* **S3** *of Proposition 2.*

Proposition 3 covers an existing result by Chen et al. (2013) as a special case. Specifically, Chen et al. (2013) consider $\mathcal{S} = \{(\boldsymbol{x}, \boldsymbol{t}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{t} = \boldsymbol{0}, \boldsymbol{x} \in \mathcal{D}\}$, with $\boldsymbol{B} \in \Re^{2\times 2}$, and $\mathcal{D}$ being a nonempty closed convex lattice. They prove that when $\boldsymbol{B}^T\boldsymbol{B}$ is an $\mathcal{L}_0$-matrix (i.e., with non-negative diagonal entries and non-positive off-diagonal entries) and $\boldsymbol{B}^T\boldsymbol{A} \leq 0$, $g(\boldsymbol{t}) = \max\{f(\boldsymbol{x}) : (\boldsymbol{x}, \boldsymbol{t}) \in \mathcal{S}\}$ is concave and supermodular if f is so. Here we show their result is an immediate corollary of Proposition 3. In particular, when $rank(\boldsymbol{B}) = 1$, $\boldsymbol{B}^T\boldsymbol{B}$ being $\mathcal{L}_0$-matrix implies that $\boldsymbol{B}$ is a lattice-matrix.

Hence, the condition in Proposition 3 is satisfied. When $rank(\boldsymbol{B}) = 2$, we can equivalently represent $\mathcal{S} = \{(\boldsymbol{x}, \boldsymbol{t}) : (\boldsymbol{B}^T \boldsymbol{A})\boldsymbol{x} + (\boldsymbol{B}^T \boldsymbol{B})\boldsymbol{t} = 0, \boldsymbol{x} \in \mathcal{D}\}$ and it can be checked that the condition in Proposition 3 is also satisfied. Thus, any set $\mathcal{S}$ satisfying the condition in Chen et al. (2013) is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$. Nevertheless, the reverse is not true. We now provide examples which satisfy the condition in Proposition 3 but cannot be analyzed using the result in Chen et al. (2013).

Example 6. We consider the following sets,

$$\mathcal{S}_1 = \left\{ (\boldsymbol{x}, \boldsymbol{t}) : \begin{bmatrix} 1 \\ 1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix} \boldsymbol{t} = \boldsymbol{0} \right\}.$$

$$\mathcal{S}_2 = \left\{ (\boldsymbol{x}, \boldsymbol{t}) : \begin{bmatrix} 1 & 2 & 3 \\ -2 & -4 & -6 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{t} = \boldsymbol{0} \right\},$$

$$\mathcal{S}_3 = \left\{ (\boldsymbol{x}, \boldsymbol{t}) : \begin{bmatrix} 1 & 3 & -2 \\ 2 & 6 & -4 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & -1 \end{bmatrix} \boldsymbol{t} = \boldsymbol{0} \right\}.$$

For the first two cases, we do not have $\boldsymbol{B}^T \boldsymbol{A} \leq 0$. The third case does not fall into the framework of Chen et al. (2013) since $n_2 = 3$.

From linearity to supermodularity

In this subsection, we consider the case that the objective function $f(\boldsymbol{x})$ is a linear function, i.e., $f(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{x}$ for some vector $\boldsymbol{a}$. We provide a necessary and sufficient condition on $\mathcal{S}$ under which the function g defined by the optimization problem (4) is supermodular.

Theorem 10 *The function g is supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is linear on $\mathcal{X}$ if and only if the set $\mathcal{S}$ is an additive mostly-lattice.*

Proof. We first prove the "if" part. Let $f(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{x}$ for some vector $\boldsymbol{a}$. Given any unordered $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$ and $\boldsymbol{x}' \in \mathcal{S}_{t'}$, $\boldsymbol{x}'' \in \mathcal{S}_{t''}$, we can find $\boldsymbol{y} \in Conv(\mathcal{S}_{t' \wedge t''})$, $\boldsymbol{z} \in Conv(\mathcal{S}_{t' \vee t''})$ such that $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$. Hence,

$$f(\boldsymbol{x}') + f(\boldsymbol{x}'') = \boldsymbol{a}^T (\boldsymbol{x}' + \boldsymbol{x}'') = \boldsymbol{a}^T (\boldsymbol{y} + \boldsymbol{z}) = f(\boldsymbol{y}) + f(\boldsymbol{z}) \leq g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}''), \tag{9}$$

where the second equality holds since $\boldsymbol{y}+\boldsymbol{z}=\boldsymbol{x}'+\boldsymbol{x}''$ and the inequality holds from Lemma 8 in the appendix. Taking supremum on the left hand side over $\boldsymbol{x}'\in\mathcal{S}_{t'}$ and $\boldsymbol{x}''\in\mathcal{S}_{t''}$, we get

$$g(\boldsymbol{t}')+g(\boldsymbol{t}'')\le g(\boldsymbol{t}'\wedge\boldsymbol{t}'')+g(\boldsymbol{t}'\vee\boldsymbol{t}'').$$

Now we prove the "only if" part. Suppose there exists an unordered pair $\boldsymbol{t}',\boldsymbol{t}''\in\mathcal{T}$ and $\boldsymbol{x}'\in\mathcal{S}_{t'},\boldsymbol{x}''\in\mathcal{S}_{t''}$, such that $(\boldsymbol{x}'+\boldsymbol{x}'')\notin\mathcal{H}$, where $\mathcal{H}=\{\boldsymbol{y}+\boldsymbol{z}:\boldsymbol{y}\in Conv(\mathcal{S}_{t'\wedge t''}),\boldsymbol{z}\in Conv(\mathcal{S}_{t'\vee t''})\}$ is a nonempty closed convex set. By the separating hyperplane theorem, there exists a vector $\boldsymbol{\eta}$ and a scalar $\lambda\in\Re$ such that $\boldsymbol{\eta}^T(\boldsymbol{x}'+\boldsymbol{x}'')>\lambda>\boldsymbol{\eta}^T\boldsymbol{w}$ for all $\boldsymbol{w}\in\mathcal{H}$. Define a linear function $f(\boldsymbol{x})=\boldsymbol{\eta}^T\boldsymbol{x}$. We have

$$g(\boldsymbol{t}')+g(\boldsymbol{t}'')\ge f(\boldsymbol{x}')+f(\boldsymbol{x}'')=\boldsymbol{\eta}^T(\boldsymbol{x}'+\boldsymbol{x}'')>\lambda.$$

Moreover, since

$$\begin{aligned}
g(\boldsymbol{t}'\wedge\boldsymbol{t}'')+g(\boldsymbol{t}'\vee\boldsymbol{t}'') &= \max\{\boldsymbol{\eta}^T\boldsymbol{y}:\boldsymbol{y}\in\mathcal{S}_{t'\wedge t''}\}+\max\{\boldsymbol{\eta}^T\boldsymbol{z}:\boldsymbol{z}\in\mathcal{S}_{t'\vee t''}\}\\
&= \max\{\boldsymbol{\eta}^T(\boldsymbol{y}+\boldsymbol{z}):\boldsymbol{y}\in\mathcal{S}_{t'\wedge t''},\boldsymbol{z}\in\mathcal{S}_{t'\vee t''}\}\\
&\le \max\{\boldsymbol{\eta}^T\boldsymbol{w}:\boldsymbol{w}\in\mathcal{H}\}\\
&\le \lambda,
\end{aligned}$$

we have

$$g(\boldsymbol{t}')+g(\boldsymbol{t}'')>g(\boldsymbol{t}'\wedge\boldsymbol{t}'')+g(\boldsymbol{t}'\vee\boldsymbol{t}''),$$

which implies that g is not supermodular. Q.E.D.

Not surprisingly, as we have more specific properties for f, the condition in Theorem 10 is less restrictive than those in Theorems 3 and 6. We now give an example which satisfies the condition in Theorem 10 but not those in Theorems 3 and 6.

EXAMPLE 7. Let $\mathcal{X}=\Re$, $\mathcal{T}=\Re^2$, $\mathcal{S}=\{(0,0,1),(1,1,0),(-1,1,1),(2,0,0)\}$. For $x'=0,\boldsymbol{t}'=(0,1)$, $x''=1,\boldsymbol{t}''=(1,0)$, we have $\mathcal{S}_{t'\wedge t''}=\mathcal{S}_{(0,0)}=\{2\}$, $\mathcal{S}_{t'\vee t''}=\mathcal{S}_{(1,1)}=\{-1\}$. Since $Conv(x',x'',x'\wedge x'',x'\vee x'')=[0,1]$, which has no intersection with $\mathcal{S}_{t'\wedge t''}$, $\mathcal{S}$ is not an additive mostly-lattice with $\mathcal{W}(x',x'')=Conv(x',x'',x'\wedge x'',x'\vee x'')$. However, if $\mathcal{W}(x',x'')=\mathcal{X}=\Re$, we can set $y=2\in(\mathcal{W}(x',x'')\cap\mathcal{S}_{t'\wedge t''})$, $z=-1\in(\mathcal{W}(x',x'')\cap\mathcal{S}_{t'\vee t''})$, which gives $x'+x''=y+z$. Therefore, $\mathcal{S}$ is an additive mostly-lattice. Furthermore, we can easily verify that $Conv(\mathcal{S})$ is an additive mostly-lattice but is not an additive mostly-lattice with $\mathcal{W}(x',x'')=Conv(x',x'',x'\wedge x'',x'\vee x'')$.

When the set $\mathcal{S}$ is convex, it is easy to derive the preservation condition from Theorem 10.

Corollary 5 *Consider the case that $\mathcal{S}$ is convex. The function g is supermodular on $\mathcal{S}|_\mathcal{T}$ whenever f is linear on $\mathcal{X}$ if and only if the set $\mathcal{S}$ is an additive mostly-lattice.*

We can also derive the following condition on $\mathcal{S}^P_{\boldsymbol{c}}$ under which the linearity of function f leads to the supermodularity of function g.

Theorem 11 *The polyhedron $\mathcal{S}^P_{\boldsymbol{c}}$ defined by the equation (2) is an additive mostly-lattice for any $\boldsymbol{c}\in\Re^m$ if and only if either 1) $n=m$, or 2) for any $\boldsymbol{\beta}\in\Re^{n_2}$ and $\mathcal{I}\subseteq\{1,\ldots,m\}$ with $|\mathcal{I}|=n+1$, $Rank(\boldsymbol{A}_\mathcal{I})=n$, and $\boldsymbol{B}_\mathcal{I}\boldsymbol{\beta}\in C(\boldsymbol{A}_\mathcal{I})$ we have $\boldsymbol{B}_\mathcal{I}\boldsymbol{\beta}^+\in C(\boldsymbol{A}_\mathcal{I})$. Here $n=Rank(\boldsymbol{A})$.*

Given any specific matrices $\boldsymbol{A}$, $\boldsymbol{B}$ with the same number of rows, we can use the following algorithm to check whether it satisfies the condition proposed in Theorem 11.

Algorithm 1: Check Theorem 11

Input: $\boldsymbol{A}\in\Re^{m\times n_1}$, $\boldsymbol{B}\in\Re^{m\times n_2}$, $n=\text{Rank}(\boldsymbol{A})$, and indicator s.

1. If $n=m$, set the indicator $s=1$, and the algorithm terminates. Otherwise, arbitrarily remove columns in $\boldsymbol{A}$, if any, until $\boldsymbol{A}$ contains only n linearly independent columns.

2. Enumerate all index sets $\hat{\mathcal{I}}\subseteq\{1,\ldots,m\}$ satisfying the condition that $|\hat{\mathcal{I}}|=n$ and $\boldsymbol{A}_{\hat{\mathcal{I}}}$ is invertible.

3. For each of index set $\hat{\mathcal{I}}$, calculate the vector $\boldsymbol{d}_k^T=\boldsymbol{b}_k^T-\boldsymbol{a}_k^T\boldsymbol{A}_{\hat{\mathcal{I}}}^{-1}\boldsymbol{B}_{\hat{\mathcal{I}}},\forall k\in\{1,\ldots,m\}\setminus\hat{\mathcal{I}}$. If $\exists k\in\{1,\ldots,m\}\setminus\hat{\mathcal{I}}$ such that $\boldsymbol{d}_k$ contains two nonzero elements with the same sign, set the indicator $s=0$, and the algorithm terminates. Otherwise, set the indicator $s=1$, and the algorithm terminates after all enumeration.

Output: s.

Theorem 12 *The polyhedron $\mathcal{S}^P_{\boldsymbol{c}}$ defined by the equation (2) is an additive mostly-lattice for any $\boldsymbol{c} \in \Re^m$ if and only if the Algorithm 1 returns $s = 1$.*

For illustration, we provide two examples. They satisfy the condition under which supermodularity is preserved when the function is linear (Theorem 12), but may easily violate the conditions for preserving supermodularity when the function is a general function (Theorems 2 and 7).

EXAMPLE 8. $\boldsymbol{A}$ is any matrix with full row rank, i.e., $n = m$, and $\boldsymbol{B}$ is any matrix. More specific examples of $\boldsymbol{A}$ can be $\boldsymbol{A} = \boldsymbol{I}$ (it cannot satisfy the condition in Theorem 2 if $\boldsymbol{B}$ has a strictly positive element), or $\boldsymbol{A} = [\boldsymbol{I} \;\; -\boldsymbol{I}]$.

EXAMPLE 9. $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_n \\ \boldsymbol{e}_i^T \end{bmatrix}$ with any $i \in \{1, \ldots, n\}$, $\boldsymbol{B}$ is any matrix such that $(\boldsymbol{b}_{n+1} - \boldsymbol{b}_i)$ does not have two nonzero elements of the same sign.

The above procedure requires the enumeration of all non-singular sub-matrices, whose computational complexity is exponential in the rank of $\boldsymbol{A}$ and thus in m and n_1. It remains unclear whether it can be done in a polynomial time.

We now investigate the conditions for preserving both concavity and supermodularity.

Theorem 13 *The function g is concave and supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is linear on $\mathcal{X}$ if and only if the following conditions hold:*

1. *$\mathcal{S}$ is an additive mostly-lattice.*
2. *For any $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}$, $\boldsymbol{x}' \in \mathcal{S}_{t'}$, $\boldsymbol{x}'' \in \mathcal{S}_{t''}$ and $\lambda \in (0,1)$, we have $\boldsymbol{x}_\lambda \in Conv(\mathcal{S}_{t_\lambda})$.*

Here $\boldsymbol{t}_\lambda = \lambda \boldsymbol{t}' + (1-\lambda)\boldsymbol{t}''$ and $\boldsymbol{x}_\lambda = \lambda \boldsymbol{x}' + (1-\lambda)\boldsymbol{x}''$.

Remark 1 *When set $\mathcal{S}$ is convex, the function g must be concave when f is linear. Hence, the function g is concave and supermodular on $\mathcal{S}|_{\mathcal{T}}$ whenever f is linear on $\mathcal{X}$ if and only if $\mathcal{S}$ is an* *additive mostly-lattice. This condition is exactly the same as that in Theorem 10. Therefore, for* $\mathcal{S}_{\boldsymbol{c}}^P$ *to be an additive mostly-lattice for every* $\boldsymbol{c}$*, the requirement for* $\boldsymbol{A},\boldsymbol{B}$ *is the same as that in Theorem 11, and can be checked by* **Algorithm 1**.

We next discuss how our results relate to some existing ones in the literature. Zipkin (2003) analyzes the preservation of supermodularity for a series of problems with linear structure. Specifically, he considers the problem $g(\boldsymbol{t}) = \max\{\boldsymbol{p}^T\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{t}, \boldsymbol{C}\boldsymbol{x} \le \boldsymbol{0}, \boldsymbol{x} \ge \boldsymbol{0}\}$ and proves that $g(\boldsymbol{t})$ is supermodular in $\boldsymbol{t} \in \Re^2_+$. Corollary 5 implies that it suffices to show that the graph of the constraint mapping is an additive mostly-lattice.

Proposition 4 *The set* $\mathcal{S} = \{(\boldsymbol{x},\boldsymbol{t}) : \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{t}, \boldsymbol{C}\boldsymbol{x} \le \boldsymbol{0}, \boldsymbol{x} \ge \boldsymbol{0}, \boldsymbol{t} \in \Re^2_+\}$ *is an additive mostly-lattice for any* $\boldsymbol{A}$ *and* $\boldsymbol{C}$.

Proof. We consider any $(\boldsymbol{x}',\boldsymbol{t}'),(\boldsymbol{x}'',\boldsymbol{t}'') \in \Re^n \times \Re^2_+$ such that $A\boldsymbol{x}' \le \boldsymbol{t}'$, $C\boldsymbol{x}' \le \boldsymbol{0}$, $A\boldsymbol{x}'' \le \boldsymbol{t}''$, $C\boldsymbol{x}'' \le \boldsymbol{0}$, $\boldsymbol{x}',\boldsymbol{x}'' \ge \boldsymbol{0}$. Since $\boldsymbol{t}',\boldsymbol{t}'' \in \Re^2_+$, there exist $\alpha,\beta \in [0,1]$ such that

$$\boldsymbol{t}' \wedge \boldsymbol{t}'' = \alpha\boldsymbol{t}' + \beta\boldsymbol{t}''.$$

Let $\boldsymbol{y} = \alpha\boldsymbol{x}' + \beta\boldsymbol{x}'' \ge \boldsymbol{0}$, $\boldsymbol{z} = (1-\alpha)\boldsymbol{x}' + (1-\beta)\boldsymbol{x}'' \ge \boldsymbol{0}$. Obviously, we have

$$C\boldsymbol{y} = \alpha C\boldsymbol{x}' + \beta C\boldsymbol{x}'' \le \boldsymbol{0},\ C\boldsymbol{z} = (1-\alpha)C\boldsymbol{x}' + (1-\beta)C\boldsymbol{x}'' \le \boldsymbol{0},$$

and $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$. Moreover,

$$A\boldsymbol{y} = \alpha A\boldsymbol{x}' + \beta A\boldsymbol{x}'' \le \alpha\boldsymbol{t}' + \beta\boldsymbol{t}'' = \boldsymbol{t}' \wedge \boldsymbol{t}'',$$

$$A\boldsymbol{z} = (1-\alpha)A\boldsymbol{x}' + (1-\beta)A\boldsymbol{x}'' \le (1-\alpha)\boldsymbol{t}' + (1-\beta)\boldsymbol{t}'' = \boldsymbol{t}' \vee \boldsymbol{t}''.$$

Hence, $\boldsymbol{y} \in \mathcal{S}_{t' \wedge t''}$ and $\boldsymbol{z} \in \mathcal{S}_{t' \vee t''}$, set $\mathcal{S}$ is an additive mostly-lattice. Q.E.D.

Some important supermodularity properties in network flow problems developed by Gale and Politof (1981) and Granot and Veinott Jr (1985) can also be studied by Corollary 5. Consider any directed graph $\mathcal{G} = (\mathcal{V},\mathcal{A})$ where $\mathcal{V}$ is the vertex set and $\mathcal{A}$ is the arc set. Let $\boldsymbol{N} \in \{0,+1,-1\}^{|\mathcal{V}|\times|\mathcal{A}|}$

be the vertex-arc incidence matrix. Let q be a mapping from $\mathcal{A}$ to $\Re^2$ defined as $q(\gamma) = (\underline{c}(\gamma), \overline{c}(\gamma))$, where $\underline{c}(\gamma)$ and $\overline{c}(\gamma)$ represent the lower and upper bound of the flow on the arc γ, respectively. Gale and Politof (1981) and Granot and Veinott Jr (1985) are concerned with the case that q is constant on all arcs except arcs α and β. In this case, let $\mu(q(\alpha), q(\beta))$ be the maximal weight over feasible circulations for given values $q(\alpha), q(\beta)$, i.e.,

$$\mu\left(q(\alpha), q(\beta)\right) = \max\left\{\sum_{\gamma\in\mathcal{A}} w(\gamma)x(\gamma) \ : \ \boldsymbol{N}\boldsymbol{x} = \boldsymbol{0},\ \underline{c}(\gamma) \le x(\gamma) \le \overline{c}(\gamma),\ \forall \gamma \in \mathcal{A}\right\}, \tag{10}$$

where $w(\gamma)$ denotes the weight of the flow on the arc γ.

Gale and Politof (1981) and Granot and Veinott Jr (1985) establish that $\mu\left(q(\alpha), q(\beta)\right)$ is supermodular if arcs α and β are in series, i.e., there is no simple cycle contains one as forward, the other as backward. Since their result holds for any linear weights $w(\gamma)$, from Corollary 5 we immediately have the following result.

Proposition 5 *The set* $\mathcal{S} = \{(\boldsymbol{x}, \boldsymbol{t}) : \boldsymbol{N}\boldsymbol{x} = \boldsymbol{0},\ (q(\alpha), q(\beta)) = \boldsymbol{t},\ \underline{c}(\gamma) \le x(\gamma) \le \overline{c}(\gamma), \forall \gamma \in \mathcal{A}\}$ *is an additive mostly-lattice.*

It is noteworthy to point out that although sets $\mathcal{S}$ in both Propositions 4 and 5 are polyhedra, Theorem 12 cannot be directly used to check the preservation of supermodularity. The reason is that Theorem 12 explores the condition such that for any $\boldsymbol{c}$, $\mathcal{S}_c^P$ is an additive mostly-lattice, while in Propositions 4 and 5, we only require the set $\mathcal{S}$ to be an additive mostly-lattice for some specific $\boldsymbol{c}$.

4. An Application on Assemble-to-Order Systems

Song and Zipkin (2003) describe the ATO system, which assembles n types of products from m types of components, as a stochastic programming problem. Specifically, let $\hat{g}(\boldsymbol{x})$ be the maximum expected profit with the initial inventory levels of components being $\boldsymbol{x} \in \Re_+^m$. Assuming unsatisfied demand is lost, we have

$$\hat{g}(\boldsymbol{x}) = \max_{\boldsymbol{y} \ge \boldsymbol{x}} \left\{-\boldsymbol{c}^T\left(\boldsymbol{y} - \boldsymbol{x}\right) + \mathrm{E}\left[g\left(\boldsymbol{y}, \tilde{\boldsymbol{d}}\right)\right]\right\}, \tag{11}$$

where

$$g(\boldsymbol{y},\boldsymbol{d}) = \max_{(\boldsymbol{v},\boldsymbol{u},\boldsymbol{w})\in\mathcal{P}_A(\boldsymbol{y},\boldsymbol{d})} \boldsymbol{r}^T\boldsymbol{v} - \boldsymbol{p}^T\boldsymbol{w} - \boldsymbol{h}^T\boldsymbol{u} \tag{12}$$

and

$$\mathcal{P}_A(\boldsymbol{y},\boldsymbol{d}) = \left\{ (\boldsymbol{v},\boldsymbol{u},\boldsymbol{w}) \in \Re_+^n \times \Re_+^m \times \Re_+^n \ : \ \begin{array}{c} \boldsymbol{A}\boldsymbol{v} + \boldsymbol{u} = \boldsymbol{y}, \\ \boldsymbol{v} + \boldsymbol{w} = \boldsymbol{d} \end{array} \right\}.$$

Here $\boldsymbol{y}$ denotes the order-up-to level of components, $\tilde{\boldsymbol{d}}$ and $\boldsymbol{d}$ represent the random demands for products and their realizations, $\boldsymbol{v},\boldsymbol{w},\boldsymbol{u}$ denote the sales quantities of products, unmet demands of products, and amounts of leftover components, respectively. Parameters $\boldsymbol{c}$ and $\boldsymbol{h}$ represent the components' per-unit ordering cost and holding cost, and $\boldsymbol{r}$, $\boldsymbol{p}$ represent the products' per-unit revenue from selling and penalty from shortage. In addition, $\boldsymbol{A} = (a_{ij})_{i=1,\dots,m;j=1,\dots,n}$, where a_{ij} is the number of units of component i required to assemble one unit of product j. To be consistent with the notation of Song and Zipkin (2003), we use $\boldsymbol{A}$ and n again here but with different meanings from those in previous sections.

Consider any nonempty index set $\mathcal{I} \subseteq \{1,\dots,m\}$. In the ATO system, one interesting question is whether the function g is supermodular in $(y_i)_{i\in\mathcal{I}}$ for any $\boldsymbol{d} \in \Re_+^n$, or equivalently whether the inventories of all components $i \in \mathcal{I}$ are complementary (Zipkin 2003). We first present the result for the case with $|\mathcal{I}| = 2$.

Theorem 14 *Consider any distinct indexes $i,j \in \{1,\dots,m\}$. The function $g(\boldsymbol{y},\boldsymbol{d})$ is supermodular in (y_i,y_j) for any non-negative $y_l, l \in \{1,\dots,m\}\setminus\{i,j\}$, $\boldsymbol{d},\boldsymbol{r},\boldsymbol{p},\boldsymbol{h}$ if and only if for any $k = 2,\dots,\min\{m-1,n\}$, any $(k+1)\times k$ submatrix $\boldsymbol{Q}$ is either 1) rank$(\boldsymbol{Q}) < k$, or 2) for any $\boldsymbol{\lambda} \in \Re^{k+1}$ with $\boldsymbol{Q}^T\boldsymbol{\lambda} = \boldsymbol{0}$, we have $\lambda_1\lambda_2 \le 0$. Here $\boldsymbol{Q}$ is obtained from $\boldsymbol{A}$ by deleting any $(n-k)$ columns and any $(m-(k+1))$ rows except rows i,j which are repositioned to the first two rows in $\boldsymbol{Q}$.*

The above result allows us to derive a condition under which g is supermodular in the whole vector $\boldsymbol{y}$.

Theorem 15 *The function $g(\boldsymbol{y},\boldsymbol{d})$ is supermodular in $\boldsymbol{y}$ for any non-negative $\boldsymbol{d},\boldsymbol{r},\boldsymbol{p},\boldsymbol{h}$ if and only if every 3×2 submatrix of the associated matrix $\boldsymbol{A}$ contains at least two row vectors which are linearly dependent.*

Proof. We first prove the "only if" direction. Suppose $\boldsymbol{A}$ has a 3×2 submatrix $\hat{\boldsymbol{A}}$ which contains no pair of rows that are linearly dependent. In this case, $\exists \boldsymbol{\lambda}\in\Re^3$ with $\hat{\boldsymbol{A}}^T\boldsymbol{\lambda}=\boldsymbol{0}$, $\lambda_i\neq 0\ \forall i\in\{1,2,3\}$. Hence, there are two elements in $\boldsymbol{\lambda}$ with the same sign. WLOG, let $\lambda_1\lambda_2>0$. We also let $\hat{\boldsymbol{A}}$ be the submatrix in the upper left of $\boldsymbol{A}$, i.e., it is obtained from $\boldsymbol{A}$ by deleting all rows except the first three, and deleting all columns except the first two. However, $g(\boldsymbol{y},\boldsymbol{d})$ is not supermodular in (y_1,y_2) since the condition in Theorem 14 is violated when we choose $k=2$ and $\boldsymbol{Q}=\hat{\boldsymbol{A}}$.

We now prove the "if" direction by showing that for all pairs of distinct indexes $i,j\in\{1,\dots,n\}$, the function g is supermodular in (y_i,y_j). To this end, following Theorem 14, we consider any $(k+1)\times k$ submatrix of $\boldsymbol{Q}$. With $rank(\boldsymbol{Q})=k$, it suffices to show that for any $\boldsymbol{\lambda}\in\Re^{k+1}$ satisfying $\boldsymbol{Q}^T\boldsymbol{\lambda}=\boldsymbol{0}$, we have $\lambda_i\lambda_j\leq 0$ for all distinct indexes $i,j\in\{1,\dots,k+1\}$. By Lemma 11 in the appendix, for such $\boldsymbol{Q}$ (i.e., it is with rank k and its every 3×2 submatrix contains at least two row vectors which are linearly dependent), we always have $\gamma\boldsymbol{q}_s=\boldsymbol{q}_t$ for some $\gamma\in\Re^+$ and distinct indexes $s,t\in\{1,\dots,k+1\}$, where $\boldsymbol{q}_s^T$ ($\boldsymbol{q}_t^T$) is the sth (tth) row vector in $\boldsymbol{Q}$. WLOG, let $s=1$, $t=k+1$, i.e., $\gamma\boldsymbol{q}_1=\boldsymbol{q}_{k+1}$. We then have

$$\boldsymbol{0}=\boldsymbol{Q}^T\boldsymbol{\lambda}=\sum_{i=1}^{k+1}\lambda_i\boldsymbol{q}_i^T=\sum_{i=1}^{k}\hat{\lambda}_i\boldsymbol{q}_i^T, \tag{13}$$

where $\hat{\lambda}_1=\lambda_1+\gamma\lambda_{k+1}$, $\hat{\lambda}_j=\lambda_j$, $j=2,\dots,k$. Since $rank(\boldsymbol{Q})=k$ and $\gamma\boldsymbol{q}_1=\boldsymbol{q}_{k+1}$, vectors $\boldsymbol{q}_1,\dots,\boldsymbol{q}_k$ are linearly independent. Hence, the equality (13) implies $\hat{\boldsymbol{\lambda}}=\boldsymbol{0}$, i.e., $\lambda_2=\cdots=\lambda_k=0$, $\lambda_1+\gamma\lambda_{k+1}=0$. By $\gamma\geq 0$, we have $\lambda_1\lambda_{k+1}\leq 0$. In addition, we have $\lambda_i\lambda_j=0$ for all pairs of indexes $i,j\in\{1,\dots,k+1\}$ such that $\{i,j\}\neq\{1,k+1\}$. Therefore, we have $\lambda_i\lambda_j\leq 0$ for all pairs of distinct indexes $i,j\in\{1,\dots,k+1\}$. The proof is complete. Q.E.D.

Theorems 14 and 15 can help us analyze several classical assemble-to-order systems, which include many settings discussed in the literature as special cases and some configurations new to the literature.

- **W system:** $\boldsymbol{A}=\begin{bmatrix}\boldsymbol{D}\\ \boldsymbol{c}^T\end{bmatrix}\in\Re_+^{(n+1)\times n}$ where $\boldsymbol{D}$ is a diagonal matrix and $D_{ii}>0,\ c_i>0\ \forall i=1,\dots,n$. In this system, there are n products and $n+1$ components. The last component is used in all products; for all other components, each one is specific to a single product.

We now show that function $g(\boldsymbol{y},\boldsymbol{d})$ is not supermodular for any two distinct components $i,j \in \{1,\ldots,n\}$ using Theorem 14. Let the submatrix $\boldsymbol{Q}$ be a 3×2 matrix as

$$\boldsymbol{Q} = \begin{bmatrix} D_{ii} & 0 \\ 0 & D_{jj} \\ c_i & c_j \end{bmatrix}.$$

It is easy to verify that $\mathrm{rank}(\boldsymbol{Q}) = 2$ and there exists $\boldsymbol{\lambda} \in \Re^3$ with $\sum_{i=1}^{3} \lambda_i \boldsymbol{q}_i = \boldsymbol{0}$ such that $\lambda_1\lambda_2 > 0$. Hence, the condition in Theorem 14 is not satisfied.

We then show that function $g(\boldsymbol{y},\boldsymbol{d})$ is supermodular for any component $i \in \{1,\ldots,n\}$, and component $n+1$ using Theorem 14. Notice that the submatrix containing rows i, $n+1$, and any other row j can be written as

$$\begin{bmatrix} 0 & \cdots & D_{ii} & \cdots & 0 & \cdots & 0 \\ c_1 & \cdots & c_i & \cdots & c_j & \cdots & c_n \\ 0 & \cdots & 0 & \cdots & D_{jj} & \cdots & 0 \end{bmatrix}.$$

Hence, for any 3×2 submatrix $\boldsymbol{Q}$ obtained from the above matrix, $\sum_{i=1}^{3} \lambda_i \boldsymbol{q}_i = \boldsymbol{0}$ only if $\lambda_1\lambda_2 \leq 0$.

- **M system:** $\boldsymbol{A} = [\boldsymbol{D}\ \boldsymbol{c}] \in \Re_+^{n\times(n+1)}$ where $\boldsymbol{D}$ is a diagonal matrix with $D_{ii} > 0, \forall i = 1,\ldots,n$, i.e., there are $n+1$ products and n components. The last product uses all n components. It is straightforward to verify that every 3×2 submatrix of $\boldsymbol{A}$ contains at least one row dependent on another. Hence, $\boldsymbol{A}$ satisfies Theorem 15 and function $g(\boldsymbol{y},\boldsymbol{d})$ is supermodular in $\boldsymbol{y}$. We remark that for the special W system and M system, where all elements in $\boldsymbol{A}$ are either 0 or 1, the same result is derived by Zipkin (2003).

- **Chained System:** for any two products i,i' with $S_i \cap S_{i'} \neq \emptyset$, either $S_i \subseteq S_{i'}$ or $S_{i'} \subseteq S_i$. Here for any product k, we denote $S_k = \{j : a_{jk} \neq 0\}$ which represents the set of components used by product k. In this chained system, if two products use a common component, then the set of components used by one product must be a subset of the other. This system is proposed by Doğru et al. (2017). In general, it does not satisfy the condition in Theorem 15. For example,

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 0 \end{bmatrix}$$

characterizes a chained system but fails to meet the condition in Theorem 15. We now investigate two special cases of the chained system.

—**Binary Chained System:** all elements in $\boldsymbol{A}$ is either 0 or 1. It is a generalization of the nested system defined in ElHafsi et al. (2008). Doğru et al. (2017) show the supermodularity of $g(\boldsymbol{y},\boldsymbol{d})$ in $\boldsymbol{y}$, which can also be easily derived by Theorem 15.

—**Proportional Chained System:** for any two products i, i' with the set of common components $C = S_i \cap S_{i'} \neq \emptyset$, $a_{ji}/a_{ji'}$ is a constant independent of $j \in C$. It implies that for any two products sharing at least one common component, the usage of all common components has the same proportion. This is often the case when products differ from each other mainly in size, so their consumptions for common components are proportional to the size of product. Indeed, the proportional chained system includes the binary chained system as a special case. We can verify that the condition in Theorem 15 is satisfied in this system. For example,

$$\boldsymbol{A} = \begin{bmatrix} 2 & 6 & 4 \\ 1 & 3 & 2 \\ 2 & 0 & 4 \\ 0 & 0 & 3 \end{bmatrix}.$$

We remark that any proportional chained system can be converted to a binary chained system by multiplying positive scalars to rows/columns. Hence, this extends the result of binary chained system by Doğru et al. (2017).

- $\boldsymbol{A} \in \Re_+^{2\times n}$, i.e., there are only two components in the system. $\boldsymbol{A}$ satisfies Theorem 15 and hence function $g(\boldsymbol{y},\boldsymbol{d})$ is supermodular in $\boldsymbol{y}$. The same result can also be derived from Remark 2 in Chen et al. (2013).

5. Conclusion

In this paper, we present a systematic study on the constraint structures which allow for the preservation of supermodularity in parametric optimization problems. Depending on the assumptions imposed on the objective functions, we introduce the concepts of mostly-lattice and additive mostly-lattice, which generalize the concept of lattice, and illustrate that they provide the needed sufficient and necessary conditions for such a preservation. We then characterize some classes of polyhedral sets which satisfy these generalized lattice-like concepts. Finally, these characterizations are used to analyze assemble-to-order inventory models.

Our analysis and results illustrate the importance of the newly introduced concepts of mostly-lattice and additive mostly-lattice, which we hope can find applications in other settings.

References

Bertsekas DP (2015) *Convex Optimization Algorithms* (Athena Scientific Belmont, MA).

Bertsimas D, Tsitsiklis JN (1997) *Introduction to Linear Optimization* (Athena Scientific Belmont, MA).

Ceryan O, Sahin O, Duenyas I (2013) Dynamic pricing of substitutable products in the presence of capacity flexibility. *Manufacturing & Service Operations Management* 15(1):86–101.

Chao X, Chen H, Zheng S (2009) Dynamic capacity expansion for a service firm with capacity deterioration and supply uncertainty. *Operations Research* 57(1):82–93.

Chen X, Hu P, He S (2013) Technical note-preservation of supermodularity in parametric optimization problems with nonlattice structures. *Operations Research* 61(5):1166–1173.

Chen X, Hu P, Shum SW, Zhang Y (2016) Dynamic stochastic inventory management with reference price effects. *Operations Research* 64(6):1529–1536.

Doğru MK, Reiman MI, Wang Q (2017) Assemble-to-order inventory management via stochastic programming: Chained boms and the m-system. *Production and Operations Management* 26(3):446–468.

ElHafsi M, Camus H, Craye E (2008) Optimal control of a nested-multiple-product assemble-to-order system. *International Journal of Production Research* 46(19):5367–5392.

Gale D, Politof T (1981) Substitutes and complements in network flow problems. *Discrete Applied Mathematics* 3(3):175–186.

Granot F, Veinott Jr AF (1985) Substitutes, complements and ripples in network flows. *Mathematics of Operations Research* 10(3):471–497.

Quah JKH (2007) The comparative statics of constrained optimization problems. *Econometrica* 75(2):401–431.

Simchi-Levi D, Chen X, Bramel J (2014) The logic of logistics. *Theory, Algorithms, and Applications for Logistics and Supply Chain Management* .

Song JS, Xue Z (2007) Demand management and inventory control for substitutable products. *Working paper* .

Song JS, Zipkin P (2003) Supply chain operations: Assemble-to-order systems. *Handbooks in Operations Research and Management Science* 11:561–596.

Topkis DM (1998) *Supermodularity and Complementarity* (Princeton University Press).

Yang J (2004) Production control in the face of storable raw material, random supply, and an outside market. *Operations Research* 52(2):293–311.

Zhu K, Thonemann UW (2009) Coordination of pricing and inventory control across products. *Naval Research Logistics (NRL)* 56(2):175–190.

Zipkin P (2003) Some modularity properties of linear programs. *Working paper, Duke University, Durham, NC* .

Appendix

Proof for Theorem 4

The "if" part is straightforward since all lattices are mostly-lattices. We now prove the "only if" part by contradiction. Suppose $\exists i \in \{1,\dots,m\}$ such that $(a_{i1},\dots,a_{in_1},b_{i1},\dots,b_{in_2}),\ i \in \{1,\dots,m\}$ has two nonzero components with the same sign. For all $j \neq i$, choose c_j large enough such that the jth constraint in defining $\mathcal{S}^P_{\boldsymbol{c}}$ can always be satisfied for the vectors to be constructed. Hence, we specifically focus on the ith constraint. With a slight abuse of notation, denote $\boldsymbol{a} = \boldsymbol{a}_i$, $\boldsymbol{b} = \boldsymbol{b}_i$, $c = c_i$. We discuss the following scenarios and choose $(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'')$ accordingly.

1. $\exists j,k\ (j \neq k)$ such that $b_j b_k > 0$. WLOG, we assume $b_j, b_k > 0$. Let $\boldsymbol{x}' = \boldsymbol{x}'' = \boldsymbol{0}$, $\boldsymbol{t}' = \frac{1}{b_j}\boldsymbol{e}_j$, $\boldsymbol{t}'' = \frac{1}{b_k}\boldsymbol{e}_k$, $c = 1$.
2. $\exists j,k\ (j \neq k)$ such that $a_j a_k > 0$. Choose two different indexes $l_1, l_2 \in \{1,\dots,n_2\}$. Let $\boldsymbol{x}' = \frac{1}{a_j}\boldsymbol{e}_j$, $\boldsymbol{x}'' = \frac{1}{a_k}\boldsymbol{e}_k$, $\boldsymbol{t}' = \epsilon \boldsymbol{e}_{l_1}$, $\boldsymbol{t}'' = \epsilon \boldsymbol{e}_{l_2}$, $c = 1.5$, where ϵ is a strictly positive number such that $\epsilon \max\{b_{l_1}, b_{l_2}, -(b_{l_1} + b_{l_2})\} < 0.5$.
3. Both scenarios 1 and 2 are false. We then have j,k such that $a_j b_k > 0$. WLOG, we assume $a_j, b_k > 0$. Consider any $l \in \{1,\dots,n_2\} \setminus \{k\}$. Let $\boldsymbol{x}' = \frac{1}{a_j}\boldsymbol{e}_j$, $\boldsymbol{x}'' = \boldsymbol{0}$, $\boldsymbol{t}' = \epsilon \boldsymbol{e}_l$, $\boldsymbol{t}'' = \frac{1}{b_k}\boldsymbol{e}_k$, $c = 1.5$, where ϵ is a strictly positive number such that $\epsilon |b_l| < 0.5$.

In all of the above scenarios, we can verify that $(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'') \in \mathcal{S}^P_{\boldsymbol{c}}$ and $\boldsymbol{t}', \boldsymbol{t}''$ are unordered, but $(\boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{t}' \vee \boldsymbol{t}'') \notin \mathcal{S}^P_{\boldsymbol{c}}$ (specifically, the ith constraint is violated). Hence, the polyhedron is not a mostly-lattice. Q.E.D.

Proof for Theorem 5

The "if" part is straightforward. We consider any function f which is concave and supermodular. The supermodularity of g is immediately implied by Theorem 3 and the assumption that $\mathcal{S}$ is a mostly-lattice. The concavity of g can be obtained easily from the convexity of $\mathcal{S}$.

We now prove the "only if" part. We first show the necessity of the mostly-lattice requirement. Assume to the contrary that $\mathcal{S}$ is not a mostly-lattice. It implies that $\exists \boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$ with $\boldsymbol{t}', \boldsymbol{t}''$

being unordered, $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}_{t' \wedge t''}$ and $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}_{t' \vee t''}$ cannot be true simultaneously. We define a function $f : \mathcal{X} \times \mathcal{T} \to \Re$ as

$$f(\boldsymbol{x}, \boldsymbol{t}) = \max_{\boldsymbol{w} \in \mathcal{W}} \{-\|(\boldsymbol{x}, \boldsymbol{t}) - \boldsymbol{w}\|_1\}, \tag{EC.1}$$

where

$$\mathcal{W} = Conv\left((\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}''), (\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{t}' \wedge \boldsymbol{t}''), (\boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{t}' \vee \boldsymbol{t}'')\right). \tag{EC.2}$$

Since set $\mathcal{W}$ is convex and a lattice (Lemma 1), f is concave and supermodular. In addition, we have $g(\boldsymbol{t}') = g(\boldsymbol{t}'') = 0$ and $g(\cdot)$ is always non-positive. If $\boldsymbol{x}' \wedge \boldsymbol{x}'' \notin \mathcal{S}_{t' \wedge t''}$, we have $g(\boldsymbol{t}' \wedge \boldsymbol{t}'') < 0$, since $\boldsymbol{t}'$ and $\boldsymbol{t}''$ are unordered and $(\boldsymbol{x}, \boldsymbol{t}' \wedge \boldsymbol{t}'') \in \mathcal{W}$ if and only if $\boldsymbol{x} = \boldsymbol{x}' \wedge \boldsymbol{x}''$. Similarly, $g(\boldsymbol{t}' \vee \boldsymbol{t}'') < 0$ if $\boldsymbol{x}' \vee \boldsymbol{x}'' \notin \mathcal{S}_{t' \vee t''}$. Therefore,

$$g(\boldsymbol{t}') + g(\boldsymbol{t}'') > g(\boldsymbol{t}' \wedge \boldsymbol{t}'') + g(\boldsymbol{t}' \vee \boldsymbol{t}''),$$

which implies that g is not supermodular.

We now show the necessity of the convexity of $\mathcal{S}$. Given any $(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'') \in \mathcal{S}$, $\lambda \in (0,1)$, we need to show $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \in \mathcal{S}$. Here we denote $\boldsymbol{x}_\alpha = \alpha \boldsymbol{x}' + (1-\alpha)\boldsymbol{x}''$, $\boldsymbol{t}_\alpha = \alpha \boldsymbol{t}' + (1-\alpha)\boldsymbol{t}''$, $\forall \alpha \in [0,1]$. We will finish the proof by discussing three scenarios.

Scenario 1: $\boldsymbol{t}'$ and $\boldsymbol{t}''$ are unordered.

Since the order of dimension does not affect the operations of convex combination, join, and meet, WLOG, we can reorder the dimension on $\mathcal{X}$ and $\mathcal{T}$ such that

$$\boldsymbol{x}' = (\boldsymbol{x}_1', \boldsymbol{x}_2'),\ \boldsymbol{x}'' = (\boldsymbol{x}_1'', \boldsymbol{x}_2''),\ \boldsymbol{x}_1' \geq \boldsymbol{x}_1'',\ \boldsymbol{x}_2' \leq \boldsymbol{x}_2'',$$
$$\boldsymbol{t}' = (\boldsymbol{t}_1', \boldsymbol{t}_2'),\ \boldsymbol{t}'' = (\boldsymbol{t}_1'', \boldsymbol{t}_2''),\ \boldsymbol{t}_1' \geq \boldsymbol{t}_1'',\ \boldsymbol{t}_2' \leq \boldsymbol{t}_2''.$$

As $\boldsymbol{t}'$ and $\boldsymbol{t}''$ are unordered, we know that $\boldsymbol{t}_1' \neq \boldsymbol{t}_1''$ and $\boldsymbol{t}_2' \neq \boldsymbol{t}_2''$. We define function f and set $\mathcal{W}$ as equations (EC.1) and (EC.2). For any $(\boldsymbol{x}, \boldsymbol{t}_\lambda) \in \mathcal{W}$, there exist $\mu_i \geq 0$, $i = 1,2,3,4$, with $\sum_{i=1}^{4} \mu_i = 1$,

$$\begin{aligned}
(\boldsymbol{x}, \boldsymbol{t}_\lambda) &= \mu_1(\boldsymbol{x}', \boldsymbol{t}') + \mu_2(\boldsymbol{x}'', \boldsymbol{t}'') + \mu_3(\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{t}' \wedge \boldsymbol{t}'') + \mu_4(\boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{t}' \vee \boldsymbol{t}'') \\
&= \mu_1(\boldsymbol{x}_1', \boldsymbol{x}_2', \boldsymbol{t}_1', \boldsymbol{t}_2') + \mu_2(\boldsymbol{x}_1'', \boldsymbol{x}_2'', \boldsymbol{t}_1'', \boldsymbol{t}_2'') + \mu_3(\boldsymbol{x}_1'', \boldsymbol{x}_2', \boldsymbol{t}_1'', \boldsymbol{t}_2') + \mu_4(\boldsymbol{x}_1', \boldsymbol{x}_2'', \boldsymbol{t}_1', \boldsymbol{t}_2'') \\
&= \Big((\mu_1 + \mu_4)\boldsymbol{x}_1' + (\mu_2 + \mu_3)\boldsymbol{x}_1'',\ (\mu_1 + \mu_3)\boldsymbol{x}_2' + (\mu_2 + \mu_4)\boldsymbol{x}_2'', \\
&\qquad (\mu_1 + \mu_4)\boldsymbol{t}_1' + (\mu_2 + \mu_3)\boldsymbol{t}_1'',\ (\mu_1 + \mu_3)\boldsymbol{t}_2' + (\mu_2 + \mu_4)\boldsymbol{t}_2''\Big).
\end{aligned}$$

By definition, $\boldsymbol{t}_\lambda = (\lambda \boldsymbol{t}'_1 + (1-\lambda)\boldsymbol{t}''_1, \lambda \boldsymbol{t}'_2 + (1-\lambda)\boldsymbol{t}''_2)$, $\boldsymbol{t}'_1 \neq \boldsymbol{t}''_1$, $\boldsymbol{t}'_2 \neq \boldsymbol{t}''_2$, and thus we have $\mu_1 + \mu_4 = \lambda$ and $\mu_1 + \mu_3 = \lambda$. Hence,

$$\boldsymbol{x} = (\lambda \boldsymbol{x}'_1 + (1-\lambda)\boldsymbol{x}''_1, \lambda \boldsymbol{x}'_2 + (1-\lambda)\boldsymbol{x}''_2) = \lambda \boldsymbol{x}' + (1-\lambda)\boldsymbol{x}'' = \boldsymbol{x}_\lambda.$$

That is, $(\boldsymbol{x}, \boldsymbol{t}_\lambda) \in \mathcal{W}$ if and only if $\boldsymbol{x} = \boldsymbol{x}_\lambda$. If $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \notin \mathcal{S}$, there does not exist $(\boldsymbol{x}, \boldsymbol{t}_\lambda) \in \mathcal{W} \cap \mathcal{S}$, and thus

$$g(\boldsymbol{t}_\lambda) < 0 = \lambda g(\boldsymbol{t}') + (1-\lambda) g(\boldsymbol{t}''),$$

which implies that g is not concave and we have a contradiction. Hence, $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \in \mathcal{S}$.

Scenario 2: $\boldsymbol{t}'$ and $\boldsymbol{t}''$ are ordered but unequal. WLOG, let $\boldsymbol{t}' \le \boldsymbol{t}''$ and $\boldsymbol{t}' \neq \boldsymbol{t}''$.

If $\boldsymbol{x}' \le \boldsymbol{x}''$, define the function f and set $\mathcal{W}$ by the equations (EC.1) and (EC.2). We know that f is concave and supermodular on $\mathcal{X} \times \mathcal{T}$, $g(\boldsymbol{t}') = g(\boldsymbol{t}'') = 0$. Since $(\boldsymbol{x}', \boldsymbol{t}') \le (\boldsymbol{x}'', \boldsymbol{t}'')$, $\mathcal{W} = Conv((\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}''))$. Since $\boldsymbol{t}' \neq \boldsymbol{t}''$, $(\boldsymbol{x}, \boldsymbol{t}_\lambda) \in \mathcal{W}$ if and only if $\boldsymbol{x} = \boldsymbol{x}_\lambda$. If $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \notin \mathcal{S}$, there does not exist $(\boldsymbol{x}, \boldsymbol{t}_\lambda) \in \mathcal{W} \cap \mathcal{S}$, and thus

$$g(\boldsymbol{t}_\lambda) < 0 = \lambda g(\boldsymbol{t}') + (1-\lambda) g(\boldsymbol{t}''),$$

which implies that g is not concave and we have a contradiction. Hence, $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \in \mathcal{S}$.

Now let us move to the case of $\boldsymbol{x}' \not\le \boldsymbol{x}''$. Assume to the contrary that $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \notin \mathcal{S}$. We get the contradiction with the following steps.

- First, we show that $\forall \alpha \in (0,1)$, $\mathcal{S}_{t_\alpha}$ must have intersection with both of the two sets

$$\mathcal{K}_1^\alpha = Conv(\alpha(\boldsymbol{x}' \wedge \boldsymbol{x}'') + (1-\alpha)\boldsymbol{x}'', \boldsymbol{x}_\alpha), \qquad \mathcal{K}_2^\alpha = Conv(\boldsymbol{x}_\alpha, \alpha \boldsymbol{x}' + (1-\alpha)(\boldsymbol{x}' \vee \boldsymbol{x}''))$$

by contradiction.

We start by assuming $\mathcal{K}_2^\alpha \cap \mathcal{S}_{t_\alpha} = \emptyset$. Let $\boldsymbol{c} = \boldsymbol{x}' - (\boldsymbol{x}' \wedge \boldsymbol{x}'')$, we have $\boldsymbol{c}^T(\boldsymbol{x}' - (\boldsymbol{x}' \wedge \boldsymbol{x}'')) = ||\boldsymbol{x}' - (\boldsymbol{x}' \wedge \boldsymbol{x}'')||_2^2 > 0$, and hence $\boldsymbol{c}^T \boldsymbol{x}' > \boldsymbol{c}^T(\boldsymbol{x}' \wedge \boldsymbol{x}'')$. Denote $\mathcal{A} = \{(\boldsymbol{x}, \boldsymbol{t}) \in \mathcal{S} : \boldsymbol{t} = \boldsymbol{t}_\alpha\} \cup \{(\alpha(\boldsymbol{x}' \wedge \boldsymbol{x}'') + (1-\alpha)\boldsymbol{x}'', \boldsymbol{t}_\alpha)\}$, $\mathcal{W} = Conv((\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}''), (\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{t}'), (\boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{t}''))$. We consider any $(\boldsymbol{x}, \boldsymbol{t}) \in \mathcal{A} \cap \mathcal{W}$ and notice that: 1) $(\boldsymbol{x}, \boldsymbol{t}) = \gamma_1(\boldsymbol{x}', \boldsymbol{t}') + \gamma_2(\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{t}') + \gamma_3(\boldsymbol{x}'', \boldsymbol{t}'') + \gamma_4(\boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{t}'')$ for some $\gamma_i \in [0,1]$, $i = 1, \ldots, 4$, with $\sum_{i=1}^4 \gamma_i = 1$; 2) $\boldsymbol{t} = \boldsymbol{t}_\alpha$; 3) $\boldsymbol{t}' \neq \boldsymbol{t}''$. Hence, $\gamma_1 + \gamma_2 = \alpha$, $\gamma_3 + \gamma_4 = 1 - \alpha$. WLOG, we can have $\gamma_2 \cdot \gamma_4 = 0$ since $(\boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{t}') + (\boldsymbol{x}' \vee \boldsymbol{x}'', \boldsymbol{t}'') = (\boldsymbol{x}', \boldsymbol{t}') + (\boldsymbol{x}'', \boldsymbol{t}'')$. We now prove $\gamma_2 > 0$ by contradiction. Assume the contrary, i.e., $\gamma_2 = 0$. It implies

$$\boldsymbol{x} = \gamma_1 \boldsymbol{x}' + \gamma_3 \boldsymbol{x}'' + \gamma_4 (\boldsymbol{x}' \vee \boldsymbol{x}'') = \frac{\gamma_3}{\gamma_3 + \gamma_4} \boldsymbol{x}_\alpha + \frac{\gamma_4}{\gamma_3 + \gamma_4} (\alpha \boldsymbol{x}' + (1-\alpha)(\boldsymbol{x}' \vee \boldsymbol{x}'')) \in \mathcal{K}_2^\alpha,$$

which contradicts with $\mathcal{K}_2^\alpha \cap \mathcal{S}_{t_\alpha} = \emptyset$. Hence, $\gamma_2 > 0$, $\gamma_4 = 0$,

$$\boldsymbol{c}^T \boldsymbol{x} = \gamma_1 \boldsymbol{c}^T \boldsymbol{x}' + \gamma_2 \boldsymbol{c}^T (\boldsymbol{x}' \wedge \boldsymbol{x}'') + \gamma_3 \boldsymbol{c}^T \boldsymbol{x}'' < (\gamma_1 + \gamma_2) \boldsymbol{c}^T \boldsymbol{x}' + \gamma_3 \boldsymbol{c}^T \boldsymbol{x}'' = \alpha \boldsymbol{c}^T \boldsymbol{x}' + (1-\alpha) \boldsymbol{c}^T \boldsymbol{x}'',$$

where the inequality follows from $\gamma_2 > 0$ and $\boldsymbol{c}^T \boldsymbol{x}' > \boldsymbol{c}^T (\boldsymbol{x}' \wedge \boldsymbol{x}'')$. Observe that $(\alpha(\boldsymbol{x}' \wedge \boldsymbol{x}'') + (1-\alpha)\boldsymbol{x}'', \boldsymbol{t}_\alpha) \in \mathcal{A} \cap \mathcal{W} \neq \emptyset$. Following Lemma 2, we can find $K \geq 0$ such that $\max_{(\boldsymbol{x},\boldsymbol{t}) \in \mathcal{A}} f(\boldsymbol{x}, \boldsymbol{t}) = \max_{(\boldsymbol{x},\boldsymbol{t}) \in \mathcal{A} \cap \mathcal{W}} f(\boldsymbol{x}, \boldsymbol{t})$, where $f(\boldsymbol{x}, \boldsymbol{t}) = \boldsymbol{c}^T \boldsymbol{x} + K \max_{\boldsymbol{w} \in \mathcal{W}} \{-||(\boldsymbol{x}, \boldsymbol{t}) - \boldsymbol{w}||_1\}$. In addition, by Lemma 1, $\mathcal{W}$ is a lattice, and hence the function f is concave and supermodular. We get

$$g(\boldsymbol{t}_\alpha) = \max_{(\boldsymbol{x},\boldsymbol{t}) \in \mathcal{S}, \boldsymbol{t} = \boldsymbol{t}_\alpha} f(\boldsymbol{x}, \boldsymbol{t}) \leq \max_{(\boldsymbol{x},\boldsymbol{t}) \in \mathcal{A}} f(\boldsymbol{x}, \boldsymbol{t}) = \max_{(\boldsymbol{x},\boldsymbol{t}) \in \mathcal{A} \cap \mathcal{W}} \boldsymbol{c}^T \boldsymbol{x} < \alpha \boldsymbol{c}^T \boldsymbol{x}' + (1-\alpha) \boldsymbol{c}^T \boldsymbol{x}'' \leq \alpha g(\boldsymbol{t}') + (1-\alpha) g(\boldsymbol{t}''),$$

which implies that g is not concave. Therefore, we have a contradiction. Hence, $\mathcal{K}_2^\alpha \cap \mathcal{S}_{t_\alpha} \neq \emptyset$.

With a similar logic, we can show $\mathcal{K}_1^\alpha \cap \mathcal{S}_{t_\alpha} = \emptyset$. The proof is similar, except that here we choose $\boldsymbol{c} = \boldsymbol{x}'' - (\boldsymbol{x}' \vee \boldsymbol{x}'')$ and define the set $\mathcal{A} = \{(\boldsymbol{x}, \boldsymbol{t}) \in \mathcal{S} : \boldsymbol{t} = \boldsymbol{t}_\alpha\} \cup \{(\alpha \boldsymbol{x}' + (1-\alpha)(\boldsymbol{x}' \vee \boldsymbol{x}''), \boldsymbol{t}_\alpha)\}$. We can again get the contradiction that g is not concave.

Therefore, $\mathcal{S}_{t_\alpha}$ has nonempty intersection with both $\mathcal{K}_1^\alpha$ and $\mathcal{K}_2^\alpha$.

- Secondly, we prove that $\exists \underline{\lambda}, \bar{\lambda}$ with $0 \leq \underline{\lambda} < \bar{\lambda} \leq 1$, such that $\forall \alpha \in [\underline{\lambda}, \bar{\lambda}]$, $(\boldsymbol{x}_\alpha, \boldsymbol{t}_\alpha) \notin \mathcal{S}$. Assume to the contrary that $\forall \underline{\lambda}, \bar{\lambda}$ with $0 \leq \underline{\lambda} < \bar{\lambda} \leq 1$, we can find $\alpha \in [\underline{\lambda}, \bar{\lambda}]$ such that $(\boldsymbol{x}_\alpha, \boldsymbol{t}_\alpha) \in \mathcal{S}$. In this case, $\forall n = 1, 2, \ldots$, we choose $\underline{\lambda}_n = \lambda$, $\bar{\lambda}_n = \lambda + \Delta/n$ where Δ are chosen such that $\lambda + \Delta \leq 1$. We then have $0 \leq \underline{\lambda}_n < \bar{\lambda}_n \leq 1$, and can find $\alpha_n \in [\underline{\lambda}_n, \bar{\lambda}_n]$ such that $(\boldsymbol{x}_{\alpha_n}, \boldsymbol{t}_{\alpha_n}) \in \mathcal{S}$. Due to the closedness of $\mathcal{S}$, $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) = \lim_{n \to} (\boldsymbol{x}_{\alpha_n}, \boldsymbol{t}_{\alpha_n}) \in \mathcal{S}$ since $(\boldsymbol{x}_{\alpha_n}, \boldsymbol{t}_{\alpha_n}) \in \mathcal{S}$ $\forall n$. Therefore, it contradicts with the assumption $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \notin \mathcal{S}$.

- Finally, we get the contradiction as follows. For any $\lambda_1 \in (\underline{\lambda}, \bar{\lambda})$, by the conclusion in the first step, we can find $\boldsymbol{y} \in \mathcal{S}_{t_{\lambda_1}} \cap \mathcal{K}_1^{\lambda_1}$. Since $\boldsymbol{y} \in \mathcal{K}_1^{\lambda_1}$, $\exists \beta \in [0, 1]$ such that

$$\boldsymbol{y} = \beta \left(\lambda_1 (\boldsymbol{x}' \wedge \boldsymbol{x}'') + (1 - \lambda_1) \boldsymbol{x}'' \right) + (1 - \beta) \boldsymbol{x}_{\lambda_1}.$$

Since $\boldsymbol{y} \in \mathcal{S}_{t_{\lambda_1}}$ and $\boldsymbol{x}_{\lambda_1} \notin \mathcal{S}_{t_{\lambda_1}}$ (conclusion from the second step), we get $\beta > 0$.

Denote the index set $\mathcal{I} = \{i : x'_i > x''_i\}$. Since $\boldsymbol{x}' \not\leq \boldsymbol{x}''$, $\mathcal{I} \neq \emptyset$. We note that

$$\begin{aligned} \forall i \in \mathcal{I} \ : \ y_i &= \beta\left(\lambda_1(x'_i \wedge x''_i) + (1-\lambda_1)x''_i\right) + (1-\beta)x_{\lambda_1 i} = \beta x''_i + (1-\beta)x_{\lambda_1 i} < x_{\lambda_1 i}, \\ \forall i \notin \mathcal{I} \ : \ y_i &= \beta\left(\lambda_1(x'_i \wedge x''_i) + (1-\lambda_1)x''_i\right) + (1-\beta)x_{\lambda_1 i} = \beta x_{\lambda_1 i} + (1-\beta)x_{\lambda_1 i} = x_{\lambda_1 i}. \end{aligned} \tag{EC.3}$$

Within $(\underline{\lambda}, \lambda_1)$, we now choose λ_2. By the conclusion in the first step, we can find $\boldsymbol{z} \in \mathcal{S}_{t_{\lambda_2}} \cap \mathcal{K}_2^{\lambda_2}$. Similarly, $\exists \gamma > 0$ such that

$$\boldsymbol{z} = \gamma\left(\lambda_2 \boldsymbol{x}' + (1-\lambda_2)(\boldsymbol{x}' \vee \boldsymbol{x}'')\right) + (1-\gamma)\boldsymbol{x}_{\lambda_2}.$$

We note that

$$\forall i \notin \mathcal{I} \ : \ z_i = \gamma\left(\lambda_2 x'_i + (1-\lambda_2)(x'_i \vee x''_i)\right) + (1-\gamma)x_{\lambda_2 i} = \gamma x_{\lambda_2 i} + (1-\gamma)x_{\lambda_2 i} = x_{\lambda_2 i} \geq x_{\lambda_1 i} = y_i, \tag{EC.4}$$

where the inequality follows from $\lambda_2 < \lambda_1$ and hence

$$x_{\lambda_2 i} - x_{\lambda_1 i} = (\lambda_2 x'_i + (1-\lambda_2)x''_i) - (\lambda_1 x'_i + (1-\lambda_1)x''_i) = (\lambda_2 - \lambda_1)(x'_i - x''_i) \geq 0,$$

and the last equality is due to the equality in (EC.3).

Moreover, by the inequality (EC.3), $y_i < x_{\lambda_1 i}$ $\forall i \in \mathcal{I}$, we can choose λ_2 close to λ_1 enough such that $\boldsymbol{x}_{\lambda_1}$ and $\boldsymbol{x}_{\lambda_2}$ are also very close, and hence $\forall i \in \mathcal{I}$, $y_i < x_{\lambda_2 i}$. Therefore,

$$\forall i \in \mathcal{I} \ : \ z_i = \gamma\left(\lambda_2 x'_i + (1-\lambda_2)(x'_i \vee x''_i)\right) + (1-\gamma)x_{\lambda_2 i} = \gamma x'_i + (1-\gamma)x_{\lambda_2 i} > x_{\lambda_2 i} > y_i, \tag{EC.5}$$

where the first inequality is due to $\gamma > 0, \lambda_2 \in (0,1)$ and $x'_i > x''_i$ $\forall i \in \mathcal{I}$. Combining (EC.4) and (EC.5), we get $\boldsymbol{z} \geq \boldsymbol{y}$. Moreover, recall that $\boldsymbol{y} \in \mathcal{S}_{t_{\lambda_1}}$, $\boldsymbol{z} \in \mathcal{S}_{t_{\lambda_2}}$, $\boldsymbol{t}_{\lambda_1} \leq \boldsymbol{t}_{\lambda_2}$ since

$$\boldsymbol{t}_{\lambda_1} - \boldsymbol{t}_{\lambda_2} = (\lambda_1 \boldsymbol{t}' + (1-\lambda_1)\boldsymbol{t}'') - (\lambda_2 \boldsymbol{t}' + (1-\lambda_2)\boldsymbol{t}'') = (\lambda_1 - \lambda_2)(\boldsymbol{t}' - \boldsymbol{t}'') \leq 0.$$

Therefore, based on the analysis at the beginning of this **Scenario 2**, we conclude that $\forall \alpha \in [0,1]$, $(\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{z}, \alpha \boldsymbol{t}_{\lambda_1} + (1-\alpha)\boldsymbol{t}_{\lambda_2}) \in \mathcal{S}$. Choose $\alpha = \frac{\gamma(1-\lambda_2)}{\beta\lambda_1 + \gamma(1-\lambda_2)}$, we can verify that

$$\begin{aligned}
\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{z} &= \alpha \left(\beta \left(\lambda_1(\boldsymbol{x}' \wedge \boldsymbol{x}'') + (1-\lambda_1)\boldsymbol{x}''\right) + (1-\beta)\boldsymbol{x}_{\lambda_1}\right) \\
&\quad + (1-\alpha)\left(\gamma\left(\lambda_2 \boldsymbol{x}' + (1-\lambda_2)(\boldsymbol{x}' \vee \boldsymbol{x}'')\right) + (1-\gamma)\boldsymbol{x}_{\lambda_2}\right) \\
&= \delta \boldsymbol{x}' + (1-\delta)\boldsymbol{x}'' \\
&= \boldsymbol{x}_\delta, \\
\alpha \boldsymbol{t}_{\lambda_1} + (1-\alpha)\boldsymbol{t}_{\lambda_2} &= \alpha\left(\lambda_1 \boldsymbol{t}' + (1-\lambda_1)\boldsymbol{t}''\right) + (1-\alpha)\left(\lambda_2 \boldsymbol{t}' + (1-\lambda_2)\boldsymbol{t}''\right) \\
&= \delta \boldsymbol{t}' + (1-\delta)\boldsymbol{t}'' \\
&= \boldsymbol{t}_\delta,
\end{aligned}$$

where

$$\delta = \frac{\beta\lambda_1\lambda_2 + \gamma\lambda_1(1-\lambda_2)}{\beta\lambda_1 + \gamma(1-\lambda_2)}.$$

We also note that $\delta \in (\lambda_2, \lambda_1)$ since

$$\lambda_2\left(\beta\lambda_1 + \gamma(1-\lambda_2)\right) < \beta\lambda_1\lambda_2 + \gamma\lambda_1(1-\lambda_2) < \lambda_1\left(\beta\lambda_1 + \gamma(1-\lambda_2)\right).$$

Therefore, we have $\delta \in (\lambda_2, \lambda_1) \subset [\underline{\lambda}, \bar{\lambda}]$ such that $(\boldsymbol{x}_\delta, \boldsymbol{t}_\delta) \in \mathcal{S}$, which contradicts with the conclusion in the second step.

Therefore, we conclude that when $\boldsymbol{x}' \not\leq \boldsymbol{x}''$, we also have $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) \in \mathcal{S}$.

Scenario 3: $\boldsymbol{t}' = \boldsymbol{t}''$.

Denote $\boldsymbol{t} = \boldsymbol{t}' = \boldsymbol{t}'' = \boldsymbol{t}_\lambda$. Since $\mathcal{T}$ is not a singleton, we can find $(\boldsymbol{x}^o, \boldsymbol{t}^o) \in \mathcal{S}$ with $\boldsymbol{t}^o \neq \boldsymbol{t}$. Based on the conclusion from **Scenario 1** and **Scenario 2**, $\alpha(\boldsymbol{x}^o, \boldsymbol{t}^o) + (1-\alpha)(\boldsymbol{x}'', \boldsymbol{t}) \in \mathcal{S}$ $\forall \alpha \in [0,1]$. Therefore, $\forall n = 1, 2, \ldots$, $(\boldsymbol{x}^{o,n}, \boldsymbol{t}^{o,n}) \in \mathcal{S}$, where we denote

$$(\boldsymbol{x}^{o,n}, \boldsymbol{t}^{o,n}) = (\boldsymbol{x}'', \boldsymbol{t}) + \frac{1}{n}\left((\boldsymbol{x}^o, \boldsymbol{t}^o) - (\boldsymbol{x}'', \boldsymbol{t})\right).$$

Again, based on the conclusion from **Scenario 1** and **Scenario 2**, $(\boldsymbol{x}^n, \boldsymbol{t}^n) \in \mathcal{S}$, where we let

$$(\boldsymbol{x}^n, \boldsymbol{t}^n) \in \mathcal{S} = \lambda(\boldsymbol{x}', \boldsymbol{t}) + (1-\lambda)(\boldsymbol{x}^{o,n}, \boldsymbol{t}^{o,n}).$$

Hence, $(\boldsymbol{x}_\lambda, \boldsymbol{t}_\lambda) = (\boldsymbol{x}_\lambda, \boldsymbol{t}) = \lim_{n\to\infty}(\boldsymbol{x}^n, \boldsymbol{t}^n) \in \mathcal{S}$. Q.E.D.

Proof for Theorem 7

We first prove the "if" part. Consider any $(\boldsymbol{x}',\boldsymbol{t}'),(\boldsymbol{x}'',\boldsymbol{t}'')\in\mathcal{S}_{\boldsymbol{c}}^P$ with $\boldsymbol{t}',\boldsymbol{t}''$ unordered.

Assume the first condition holds. If $(\boldsymbol{A}|\boldsymbol{B})$ is a lattice-matrix, as shown by Theorem 4, $(\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{t}'\wedge\boldsymbol{t}''),(\boldsymbol{x}'\vee\boldsymbol{x}'',\boldsymbol{t}'\vee\boldsymbol{t}'')\in\mathcal{S}_{\boldsymbol{c}}^P$. If $(-\boldsymbol{A}|\boldsymbol{B})$ is a lattice-matrix, with a similar logic, we can show that $(\boldsymbol{x}'\vee\boldsymbol{x}'',\boldsymbol{t}'\wedge\boldsymbol{t}''),(\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{t}'\vee\boldsymbol{t}'')\in\mathcal{S}_{\boldsymbol{c}}^P$. Therefore, we can either choose $\boldsymbol{y}=\boldsymbol{x}'\wedge\boldsymbol{x}''$, $\boldsymbol{z}=\boldsymbol{x}'\vee\boldsymbol{x}''$ in the first case, or $\boldsymbol{y}=\boldsymbol{x}'\vee\boldsymbol{x}''$, $\boldsymbol{z}=\boldsymbol{x}'\wedge\boldsymbol{x}''$ in the second case. We always have that $\boldsymbol{y},\boldsymbol{z}$ satisfy the requirement for additive mostly-lattice with the corresponding $\mathcal{W}$.

Assume the second condition holds. Let $\mathcal{L}=\{1,\dots,m\}\setminus\mathcal{I}$, and denote scalars $r'=\boldsymbol{d}^T\boldsymbol{x}'$, $r''=\boldsymbol{d}^T\boldsymbol{x}''$. WLOG, let $r'\le r''$. By $(\boldsymbol{x}',\boldsymbol{t}'),(\boldsymbol{x}'',\boldsymbol{t}'')\in\mathcal{S}_{\boldsymbol{c}}^P$, we have

$$\begin{aligned}\boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'+\boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}'=\boldsymbol{k}r'+\boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}' &\le \boldsymbol{c}_{\mathcal{I}}\\ \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}''+\boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}''=\boldsymbol{k}r''+\boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}'' &\le \boldsymbol{c}_{\mathcal{I}}\end{aligned} \tag{EC.6}$$

and

$$\begin{aligned}\boldsymbol{A}_{\mathcal{L}}\boldsymbol{x}' &\le \boldsymbol{c}_{\mathcal{L}}\\ \boldsymbol{A}_{\mathcal{L}}\boldsymbol{x}'' &\le \boldsymbol{c}_{\mathcal{L}}\end{aligned}, \tag{EC.7}$$

where the inequalities (EC.6) follow from the fact that $\boldsymbol{B}_{\mathcal{L}}=\boldsymbol{0}$. If $(\boldsymbol{k}|\boldsymbol{B}_{\mathcal{I}})$ is a lattice-matrix, by the inequalities (EC.6) and Theorem 4, we have $\boldsymbol{k}(r'\wedge r'')+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\wedge\boldsymbol{t}'')\le\boldsymbol{c}_{\mathcal{I}}$ and $\boldsymbol{k}(r'\vee r'')+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\vee\boldsymbol{t}'')\le\boldsymbol{c}_{\mathcal{I}}$. Hence,

$$\begin{aligned}\boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\wedge\boldsymbol{t}'')=\boldsymbol{k}r'+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\wedge\boldsymbol{t}'')=\boldsymbol{k}(r'\wedge r'')+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\wedge\boldsymbol{t}'') &\le \boldsymbol{c}_{\mathcal{I}}\\ \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}''+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\vee\boldsymbol{t}'')=\boldsymbol{k}r''+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\vee\boldsymbol{t}'')=\boldsymbol{k}(r'\vee r'')+\boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'\vee\boldsymbol{t}'') &\le \boldsymbol{c}_{\mathcal{I}}\end{aligned}.$$

The above inequalities, together with (EC.7), imply that $(\boldsymbol{x}',\boldsymbol{t}'\wedge\boldsymbol{t}''),(\boldsymbol{x}'',\boldsymbol{t}'\vee\boldsymbol{t}'')\in\mathcal{S}_{\boldsymbol{c}}^P$, and we can choose $\boldsymbol{y}=\boldsymbol{x}',\boldsymbol{z}=\boldsymbol{x}''$ for the requirement of the additive mostly-lattice with the corresponding $\mathcal{W}$.

We now prove the "only if" part. Assuming that $\mathcal{S}_{\boldsymbol{c}}^P$ is an additive mostly-lattice with $\mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'')=\{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\}$ $\forall\boldsymbol{c}\in\Re^m$ but the second condition in Theorem 7 is violated. It suffices to show that the first condition in Theorem 7 is satisfied. To this end, denote set $\mathcal{W}^o=\mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'')=\{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\}$ for given $\boldsymbol{x}',\boldsymbol{x}''$, and $\mathcal{S}_{c,t}^P=\{\boldsymbol{x}:(\boldsymbol{x},\boldsymbol{t})\in\mathcal{S}_{\boldsymbol{c}}^P\}$ for given $\boldsymbol{t}$. In addition, WLOG, let $\mathcal{I}=\{i:\boldsymbol{b}_i\neq\boldsymbol{0}\}=\{1,\dots,s\}$ where $s\ge 2$ since the second condition in Theorem 7 is violated. we first prove four statements as follows.

Statement 1—Each row of $\boldsymbol{B}$ has at most one nonzero. Assume to the contrary that, WLOG, $b_{11}b_{12} \neq 0$. Since $\boldsymbol{a}_1 \neq \boldsymbol{0}$ ($\boldsymbol{A}$ has no zero rows), we can assume, WLOG, $a_{11} \neq 0$. For simplicity of presentation, we normalize the coefficients such that $|a_{11}| = |b_{11}| = |b_{12}| = 1$. Fix $c_i, i \geq 2$ with very large values such that the constraints $\boldsymbol{a}_i^T \boldsymbol{x} + \boldsymbol{b}_i^T \boldsymbol{t} \leq c_i$ would not be violated. Let $\boldsymbol{t}' = \boldsymbol{e}_1$, $\boldsymbol{t}'' = \boldsymbol{e}_2$. Hence, $\boldsymbol{t}' \wedge \boldsymbol{t}'' = \boldsymbol{0}$, $\boldsymbol{t}' \vee \boldsymbol{t}'' = \boldsymbol{e}_1 + \boldsymbol{e}_2$.

- If $b_{11} = b_{12}$, let $c_1 = b_{11} = b_{12}$. Choosing $\boldsymbol{x}' = \boldsymbol{x}'' = \boldsymbol{0}$, we then have $(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'') \in \mathcal{S}_{\boldsymbol{c}}^P$ and $\mathcal{W}^o = \{\boldsymbol{0}\}$. However, $\boldsymbol{0} \notin \mathcal{S}_{c,t' \vee t''}^P$ if $b_{11} = 1$, and $\boldsymbol{0} \notin \mathcal{S}_{c,t' \wedge t''}^P$ if $b_{11} = -1$.
- If $b_{11} = -b_{12}$, let $c_1 = 1$. WLOG, let $b_{11} = 1$ and $b_{12} = -1$. Choosing $\boldsymbol{x}' = \boldsymbol{0}$, $\boldsymbol{x}'' = 2a_{11}\boldsymbol{e}_1$, we then have $(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'') \in \mathcal{S}_{\boldsymbol{c}}^P$ and $\mathcal{W}^o = \{\boldsymbol{0}, 2a_{11}\boldsymbol{e}_1\}$. However, $\mathcal{W}^o \cap \mathcal{S}_{c,t' \wedge t''}^P = \mathcal{W}^o \cap \mathcal{S}_{c,t' \vee t''}^P = \{\boldsymbol{0}\}$.

In either case, we cannot find $\boldsymbol{y}, \boldsymbol{z}$ satisfying the requirement of additive mostly-lattice with $\mathcal{W}$.

Statement 2—There does not have $i \neq j, l_i \neq l_j$, and k such that $b_{il_i} b_{jl_j} \neq 0$, $a_{ik} b_{il_i} > 0$, $a_{jk} b_{jl_j} < 0$. Assume the contrary, WLOG, $a_{11} b_{11} > 0$, $a_{21} b_{22} < 0$. In addition, since each $\boldsymbol{b}_i$ has at most one nonzero, $b_{1q} = 0 \ \forall q \neq 1$, $b_{2q} = 0 \ \forall q \neq 2$. We normalize the coefficients such that $a_{11} = b_{11}$, $a_{21} = -b_{22}$. Set $c_1 = b_{11}$, $c_2 = 0$ and $c_i, i \geq 3$ to take large values such that the constraints $\boldsymbol{a}_i^T \boldsymbol{x} + \boldsymbol{b}_i^T \boldsymbol{t} \leq c_i$ would not be violated. Let $\boldsymbol{x}' = \boldsymbol{0}$, $\boldsymbol{x}'' = \boldsymbol{e}_1$, $\boldsymbol{t}' = \boldsymbol{e}_1$, $\boldsymbol{t}'' = \boldsymbol{e}_2$. We then have $\boldsymbol{x}' \in \mathcal{S}_{c,t'}^P$, $\boldsymbol{x}'' \in \mathcal{S}_{c,t''}^P$, and $\mathcal{W}^o = \{\boldsymbol{0}, \boldsymbol{e}_1\}$. If $\exists \boldsymbol{y}, \boldsymbol{z}$ satisfying the requirement of additive mostly-lattice with $\mathcal{W}$, by Lemma 3, $\begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} a_{11} \\ 0 \end{bmatrix}$. However, the structure of $\mathcal{W}^o$ implies that $\boldsymbol{y}_i = 0 \ \forall i \geq 2$ and hence $\begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{bmatrix} \boldsymbol{y} = y_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$, which is a contradiction as $a_{11} a_{21} \neq 0$.

Statement 3—$\exists i \neq j, l_i \neq l_j$ such that $b_{il_i} b_{jl_j} \neq 0$, $\boldsymbol{a}_i, \boldsymbol{a}_j$ are linearly independent. Suppose that all rows of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_s$ are linearly dependent. Recall that we assume the second condition in Theorem 7 is violated. So there must exist $i \neq j$, $l_i \neq l_j$, k such that $b_{il_i} b_{jl_j} \neq 0$, $a_{ik} b_{il_i}$ and $a_{jk} b_{jl_j}$ are of different sign, which contradicts with Statement 2. Hence, there must exist two rows $i, j \leq s$ such that $\boldsymbol{a}_i, \boldsymbol{a}_j$ are linearly independent. Denote the indexes of nonzero element of $\boldsymbol{b}_i, \boldsymbol{b}_j$ by l_i, l_j. If $l_i \neq l_j$, then it is done. If $l_i = l_j$, since $\boldsymbol{B}$ has at least two nonzero columns, there must exist another row k such that the index of the nonzero element of $\boldsymbol{b}_k$ does not equal to l_i. In addition, as $\boldsymbol{a}_i, \boldsymbol{a}_j$

are linearly independent, $\boldsymbol{a}_k$ must be linearly independent with at least one of $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$. WLOG, let $\boldsymbol{a}_k, \boldsymbol{a}_i$ be linearly independent. Since we already have $b_{kl_k} b_{il_i} \neq 0$ and $l_k \neq l_i$, the statement is proved.

Statement 4—For any two rows i, j with $l_i \neq l_j$ such that $b_{il_i} b_{jl_j} \neq 0$, $\boldsymbol{a}_i, \boldsymbol{a}_j$ are linearly independent, we have that each row of $\boldsymbol{a}_i, \boldsymbol{a}_j$ has only one nonzero; in addition, if the two nonzeros in the two rows are a_{ik_i}, a_{jk_j}, we have $a_{ik_i} b_{il_i}$, $a_{jk_j} b_{jl_j}$ are of the same sign.

To prove Statement 4, WLOG, suppose we have $b_{11}, b_{22} \neq 0$ and $\boldsymbol{a}_1, \boldsymbol{a}_2$ are linearly independent. Since each $\boldsymbol{b}_i$ has at most one nonzero, $b_{1q} = 0\ \forall q \neq 1$, $b_{2q} = 0\ \forall q \neq 2$. Sine $\boldsymbol{a}_1, \boldsymbol{a}_2$ are linearly independent, WLOG, we can assume the following submatrix

$$\bar{\boldsymbol{A}} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

has full rank. Since $\bar{\boldsymbol{A}}$ is of full rank, WLOG, we let $a_{11} a_{22} \neq 0$. In addition, since $\bar{\boldsymbol{A}}$ is of full rank, $\exists$ unique $\bar{\boldsymbol{x}}', \bar{\boldsymbol{x}}'' \in \Re^2$ with

$$\bar{\boldsymbol{A}} \bar{\boldsymbol{x}}' = \begin{bmatrix} -b_{11} \\ 0 \end{bmatrix}, \qquad \bar{\boldsymbol{A}} \bar{\boldsymbol{x}}'' = \begin{bmatrix} 0 \\ -b_{22} \end{bmatrix}. \tag{EC.8}$$

Let $\boldsymbol{x}' = (\bar{x}'_1, \bar{x}'_2, 0, \ldots, 0)$, $\boldsymbol{x}'' = (\bar{x}''_1, \bar{x}''_2, 0, \ldots, 0)$, $\boldsymbol{t}' = \boldsymbol{e}_1$, $\boldsymbol{t}'' = \boldsymbol{e}_2$, $c_1 = c_2 = 0$, c_i is of large value for all $i \geq 3$ and hence it suffices to focus on the first two constraints. We then have

$$\bar{\boldsymbol{A}} \bar{\boldsymbol{x}}' + \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} \boldsymbol{t}' = \bar{\boldsymbol{A}} \bar{\boldsymbol{x}}'' + \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} \boldsymbol{t}'' = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

and hence $\boldsymbol{x}' \in \mathcal{S}^P_{c,t'}$, $\boldsymbol{x}'' \in \mathcal{S}^P_{c,t''}$. As $\mathcal{S}^P_{\boldsymbol{c}}$ is an additive mostly-lattice with $\mathcal{W}$, we can find $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}^P_{c,t' \wedge t''})$, $\boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}^P_{c,t' \vee t''})$, $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$. Since $\boldsymbol{y}, \boldsymbol{z} \in Conv(\mathcal{W}^o)$, we have $y_i = z_i = 0$, $\forall i \geq 3$. Let $\bar{\boldsymbol{y}}$ be the column vector of y_1, y_2; and denote $\bar{\boldsymbol{z}}$ similarly. Since $\boldsymbol{y} \in Conv(\mathcal{S}^P_{c,t' \wedge t''}) = \mathcal{S}^P_{c,t' \wedge t''}$, $\boldsymbol{z} \in Conv(\mathcal{S}^P_{c,t' \vee t''}) = \mathcal{S}^P_{c,t' \vee t''}$, by Lemma 3 we have

$$\bar{\boldsymbol{A}} \bar{\boldsymbol{y}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} (\boldsymbol{t}' \wedge \boldsymbol{t}'') = \boldsymbol{0} - \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} \boldsymbol{0} = \boldsymbol{0}, \tag{EC.9}$$

which implies $\bar{\boldsymbol{y}} = \boldsymbol{0}$ since $\bar{\boldsymbol{A}}$ is of full rank. Note that it also implies $\boldsymbol{y} = \boldsymbol{0}$ since $y_i = 0, \forall i \geq 3$.

We then demonstrate that $\{\boldsymbol{y},\boldsymbol{z}\}=\{\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\}$. Note that since $\boldsymbol{y}\in Conv(\mathcal{W}^o)$, we have

$$\boldsymbol{y}=\lambda_1\boldsymbol{x}'+\lambda_2\boldsymbol{x}''+\lambda_3(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\lambda_4(\boldsymbol{x}'\vee\boldsymbol{x}''),$$

for some $\boldsymbol{\lambda}\geq\mathbf{0}$ with $\sum_{i=1}^4\lambda_i=1$. Since $\boldsymbol{x}'+\boldsymbol{x}''=\boldsymbol{x}'\wedge\boldsymbol{x}''+\boldsymbol{x}'\vee\boldsymbol{x}'$, we can find $\boldsymbol{\lambda}$ with $\lambda_1\lambda_2=0$.

- If $\lambda_1\neq0$, $\lambda_2=0$,

$$\boldsymbol{y}=\lambda_1\boldsymbol{x}'+\lambda_3(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\lambda_4(\boldsymbol{x}'\vee\boldsymbol{x}''), \tag{EC.10}$$

$$\boldsymbol{z}=\boldsymbol{x}'+\boldsymbol{x}''-\boldsymbol{y}=\lambda_1\boldsymbol{x}''+(1-\lambda_1-\lambda_3)(\boldsymbol{x}'\wedge\boldsymbol{x}'')+(1-\lambda_1-\lambda_4)(\boldsymbol{x}'\vee\boldsymbol{x}''). \tag{EC.11}$$

Since $\boldsymbol{y}\in Conv(\mathcal{W}^o\cap\mathcal{S}^P_{c,t'\wedge t''})$ where $\mathcal{W}^o=\{\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\}$, the expression (EC.10) implies $\boldsymbol{x}'\in\mathcal{S}^P_{c,t'\wedge t''}$. Otherwise, $\boldsymbol{y}\in Conv(\{\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\})$ and cannot be written by expression (EC.10) with $\lambda_1\neq0$. Similarly, we can get $\boldsymbol{x}''\in\mathcal{S}^P_{c,t'\vee t''}$ since $\boldsymbol{z}\in Conv(\mathcal{W}^o\cap\mathcal{S}^P_{c,t'\vee t''})$.

Hence, by the first constraint of $\boldsymbol{A}\boldsymbol{x}'+\boldsymbol{B}(\boldsymbol{t}'\wedge\boldsymbol{t}'')\leq\boldsymbol{c}$, and that of $\boldsymbol{A}\boldsymbol{x}''+\boldsymbol{B}(\boldsymbol{t}'\vee\boldsymbol{t}'')\leq\boldsymbol{c}$, we have

$$-b_{11}\leq0,\qquad b_{11}\leq0,$$

which contradicts with $b_{11}\neq0$.

- If $\lambda_1=0$, $\lambda_2\neq0$, similar to the previous case ($\lambda_1\neq0$, $\lambda_2=0$), we can get a contradiction.

Therefore, we must have $\lambda_1=\lambda_2=0$. To show that $\{\boldsymbol{y},\boldsymbol{z}\}=\{\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\}$, we still need to prove $\lambda_3\lambda_4=0$. Assume to the contrary that $\lambda_3\lambda_4\neq0$. We have

$$\boldsymbol{y}=\lambda_3(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\lambda_4(\boldsymbol{x}'\vee\boldsymbol{x}''),\qquad \boldsymbol{z}=\lambda_4(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\lambda_3(\boldsymbol{x}'\vee\boldsymbol{x}'').$$

That implies both $\boldsymbol{x}'\wedge\boldsymbol{x}''$ and $\boldsymbol{x}'\vee\boldsymbol{x}''$ are elements of $\mathcal{S}^P_{c,t'\wedge t''}$ and $\mathcal{S}^P_{c,t'\vee t''}$. Hence, $\boldsymbol{y}=\boldsymbol{x}'\wedge\boldsymbol{x}''$ and $\boldsymbol{z}=\boldsymbol{x}'\vee\boldsymbol{x}''$ satisfy the requirement in the additive mostly-lattice, that implies $\boldsymbol{x}'\wedge\boldsymbol{x}''=\boldsymbol{y}=\mathbf{0}$ based on our analysis after the equation (EC.9). Similarly, $\boldsymbol{y}=\boldsymbol{x}'\vee\boldsymbol{x}''$ and $\boldsymbol{z}=\boldsymbol{x}'\wedge\boldsymbol{x}''$ also meet the requirement in the additive mostly-lattice, which implies $\boldsymbol{x}'\vee\boldsymbol{x}''=\boldsymbol{y}=\mathbf{0}$. Therefore, we have $\boldsymbol{x}'=\boldsymbol{x}''=\mathbf{0}$, which contradicts with the equation (EC.8). Hence, we conclude that $\lambda_3\lambda_4=0$ and $\{\boldsymbol{y},\boldsymbol{z}\}=\{\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\}$.

Therefore, by $\boldsymbol{y} = \boldsymbol{0}$ we conclude that either $\bar{\boldsymbol{x}}' = (0, x_2'), \bar{\boldsymbol{x}}'' = (x_1'', 0)$, or $\bar{\boldsymbol{x}}' = (x_1', 0), \bar{\boldsymbol{x}}'' = (0, x_2'')$ is true. If the first scenario is true, by equation (EC.8) we have

$$\bar{\boldsymbol{A}} = \begin{bmatrix} 0 & -\frac{b_{11}}{x_2'} \\ -\frac{b_{22}}{x_1''} & 0 \end{bmatrix},$$

which contradicts with $a_{11}a_{22} \neq 0$. Therefore, we must have $\bar{\boldsymbol{x}}' = (x_1', 0), \bar{\boldsymbol{x}}'' = (0, x_2'')$, and

$$\bar{\boldsymbol{A}} = \begin{bmatrix} -\frac{b_{11}}{x_1'} & 0 \\ 0 & -\frac{b_{22}}{x_2''} \end{bmatrix}.$$

In addition, $a_{11}b_{11} = -b_{11}^2/x_1'$, $a_{22}b_{22} = -b_{22}^2/x_2''$, which are of the same sign since x_1', x_2'' are of the same sign otherwise $\bar{\boldsymbol{x}}', \bar{\boldsymbol{x}}''$ cannot be unordered.

Now it remains to show that there is no other nonzero element in the first two rows of $\boldsymbol{A}$. Assume to the contrary and let (a_{13}, a_{23}) be a nonzero vector. Since $\bar{\boldsymbol{A}}$ is of full rank, WLOG, we can assume (a_{13}, a_{23}) is linearly independent with (a_{11}, a_{21}). With the same logic as above, we can show that

$$\begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{23} \end{bmatrix}$$

with $a_{11}b_{11}$ and $a_{23}b_{22}$ having the same sign. Set $\boldsymbol{t}' = \boldsymbol{e}_1$, $\boldsymbol{t}'' = \boldsymbol{e}_2$, $c_1 = c_2 = 0$, and let $c_i, i \geq 3$ take values large enough such that the constraints $\boldsymbol{a}_i^T \boldsymbol{x} + \boldsymbol{b}_i^T \boldsymbol{t} \leq c_i$ would not be violated. Normalize the coefficients such that $|a_{11}| = |a_{22}| = |a_{23}| = |b_{11}| = |b_{22}| = 1$ and denote $m = a_{11}b_{11} = a_{22}b_{22} = a_{23}b_{22}$. We then choose $\boldsymbol{x}' = (-m, 2, -2, 0, \dots, 0)^T$, $\boldsymbol{x}'' = -m\boldsymbol{e}_2$, and hence $\boldsymbol{x}' \in \mathcal{S}_{c,t'}^P$, $\boldsymbol{x}'' \in \mathcal{S}_{c,t''}^P$ and we can find $\boldsymbol{y}, \boldsymbol{z}$ satisfying the requirement in the additive mostly-lattice. Therefore, we can find $\boldsymbol{\lambda} \geq \boldsymbol{0}$ with $\sum_{i=1}^4 \lambda_i = 1$ such that $\boldsymbol{y} = \lambda_1 \boldsymbol{x}' + \lambda_2 \boldsymbol{x}'' + \lambda_3(\boldsymbol{x}' \wedge \boldsymbol{x}'') + \lambda_4(\boldsymbol{x}' \vee \boldsymbol{x}'')$. We can show $\lambda_1 = \lambda_2 = 0$ with similar logic as in the above proof: if $\lambda_1 \neq \lambda_2 = 0$, we have $\boldsymbol{x}' \in \mathcal{S}_{c,t' \wedge t''}^P$ and $\boldsymbol{x}'' \in \mathcal{S}_{c,t' \vee t''}^P$, which implies $-b_{11} \leq 0$ and $b_{11} \leq 0$, contradicts with $b_{11} \neq 0$; if $\lambda_2 \neq \lambda_1 = 0$, we have $\boldsymbol{x}'' \in \mathcal{S}_{c,t' \wedge t''}^P$ and $\boldsymbol{x}' \in \mathcal{S}_{c,t' \vee t''}^P$, which implies $-b_{22} \leq 0$ and $b_{22} \leq 0$, contradicts with $b_{22} \neq 0$. Therefore, $\boldsymbol{y} = \lambda_3(\boldsymbol{x}' \wedge \boldsymbol{x}'') + \lambda_4(\boldsymbol{x}' \vee \boldsymbol{x}')$. In addition, by Lemma 3, we have

$$\begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{bmatrix} \boldsymbol{y} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} (\boldsymbol{t}' \wedge \boldsymbol{t}'') = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} \boldsymbol{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{EC.12}$$

Nevertheless, if $m=1$, $\boldsymbol{x}'\wedge\boldsymbol{x}''=(-1,-1,-2,0,\ldots,0)^T$, $\boldsymbol{x}'\vee\boldsymbol{x}''=(0,2,0,0,\ldots,0)^T$; if $m=-1$, $\boldsymbol{x}'\wedge\boldsymbol{x}''=(0,-1,-2,0,\ldots,0)^T$, $\boldsymbol{x}'\vee\boldsymbol{x}''=(1,2,0,0,\ldots,0)^T$. In either case, we cannot have $\boldsymbol{y}$ as a convex combination of $\boldsymbol{x}'\wedge\boldsymbol{x}''$ and $\boldsymbol{x}'\vee\boldsymbol{x}''$ such that the equation (EC.12) holds. Therefore, the assumption is false. There is no other nonzero element in the first two rows of $\boldsymbol{A}$. Hence, the proof for Statement 4 is completed.

With all Statements proved, we are ready to show that the first condition in Theorem 7 is satisfied. We start from the first s rows. By Statement 3, WLOG, let $\boldsymbol{a}_1,\boldsymbol{a}_2$ be linearly independent and $b_{11}b_{22}\neq 0$. By Statement 4, each vector of $\boldsymbol{a}_1,\boldsymbol{a}_2$ has only one nonzero; in addition, if the nonzeros are a_{1k_1},a_{2k_2}, then $a_{1k_1}b_{11}$ and $a_{2k_2}b_{22}$ are of the same sign. Consider any $i=3,\ldots,s$, and denote the index of the nonzero in $\boldsymbol{b}_i$ as l_i, i.e., $b_{il_i}\neq 0$. If $\boldsymbol{a}_i$ has more than one nonzero component, then $\boldsymbol{a}_i$ is linearly independent with both $\boldsymbol{a}_1,\boldsymbol{a}_2$, which contradicts with Statement 4. Hence, $\boldsymbol{a}_i$ has only one nonzero component, say a_{ik_i}. By Statement 2 and 4, we can conclude $a_{ik_i}b_{il_i}$ has the same sign as $a_{1k_1}b_{11}$.

We now prove the result for the other rows. Assume the contrary, i.e., $\exists k>s$ such that the row vector $(a_{k1},\ldots,a_{kn_1})$ has two nonzero elements of the same sign. With the result on the first s rows, together with Statement 3, we can find $i,j\leq s$ such that the nonzeros in $\boldsymbol{a}_i,\boldsymbol{a}_j$ are of different indexes, and the nonzeros in $\boldsymbol{b}_i,\boldsymbol{b}_j$ are also of different indexes. WLOG, let $\{i,j\}=\{1,2\}$, first two constraints of $\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{t}\leq\boldsymbol{c}$ are

$$a_{11}x_1+b_{11}t_1\leq c_1,\qquad a_{22}x_2+b_{22}t_2\leq c_2,$$

where $a_{11}b_{11},a_{22}b_{22}$ are of the same sign. WLOG, we normalize $|a_{11}|=|a_{22}|=|b_{11}|=|b_{22}|=1$. Here we only prove in the case that $a_{11}b_{11}=1$, the case of $a_{11}b_{11}=-1$ can be proved similarly. In constructing a counter example to the condition of the additive mostly-lattice, it suffices to consider the case that the only remaining constraint is the third constraint, where $\boldsymbol{b}_3=\boldsymbol{0}$, and the row vector $(a_{31},\ldots,a_{3n_1})$ has two nonzeros with the same sign. Choose $\boldsymbol{t}'=-\boldsymbol{e}_1$, $\boldsymbol{t}''=-\boldsymbol{e}_2$, $c_1=c_2=0$. We now discuss in three scenarios.

1. $a_{31}a_{32}>0$. We let $\boldsymbol{x}'=\boldsymbol{e}_1$, $\boldsymbol{x}''=\boldsymbol{e}_2$, $c_3=\max\{a_{31},a_{32}\}$.

2. $a_{31}a_{32} \le 0$, $\exists i \in \{1,2\}, j \ge 3$ such that $a_{3i}a_{3j} > 0$. WLOG, let $a_{31}a_{33} > 0$. We let $\boldsymbol{x}' = \boldsymbol{e}_1$, $\boldsymbol{x}'' = \boldsymbol{e}_2 + l\boldsymbol{e}_3$, $c_3 = \max\{a_{31}, a_{33}\}$. Since $a_{32}a_{33} \le 0$, we can choose $l \ge 1$ such that $a_{32} + la_{33} = a_{33}$.

3. $a_{31}a_{32} \le 0$, $a_{3i}a_{3j} \le 0$ for all $i \in \{1,2\}, j \ge 3$. In this case, WLOG, we have $a_{33}a_{34} > 0$. Let $\boldsymbol{x}' = \boldsymbol{e}_1 + l_1\boldsymbol{e}_3$, $\boldsymbol{x}'' = \boldsymbol{e}_2 + l_2\boldsymbol{e}_4$, $c_3 = \max\{a_{33}, a_{34}\}$. Since $a_{31}a_{33}, a_{32}a_{34} \le 0$, we can choose $l_1, l_2 \ge 1$ such that $a_{31} + l_1a_{33} = a_{33}$, $a_{32} + l_2a_{34} = a_{34}$.

In all scenarios, we have

$$\begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{bmatrix} \boldsymbol{x}' + \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} \boldsymbol{t}' = \begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{bmatrix} \boldsymbol{x}'' + \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} \boldsymbol{t}'' = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

and $\boldsymbol{x}' \in \mathcal{S}^P_{c,t'}$, $\boldsymbol{x}'' \in \mathcal{S}^P_{c,t''}$. If the condition for the additive mostly-lattice is satisfied, we can find $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}^P_{c,t'\wedge t''}), \boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}^P_{c,t'\vee t''})$ such that $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$. By Lemma 3, we have

$$\begin{bmatrix} \boldsymbol{a}_1^T \\ \boldsymbol{a}_2^T \end{bmatrix} \boldsymbol{z} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{b}_1^T \\ \boldsymbol{b}_2^T \end{bmatrix} (\boldsymbol{t}' \vee \boldsymbol{t}'') = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{EC.13}$$

Since $\boldsymbol{a}_1 = a_{11}\boldsymbol{e}_1, \boldsymbol{a}_2 = a_{22}\boldsymbol{e}_2$, the equality (EC.13) implies $z_1 = z_2 = 0$. Recall that $\boldsymbol{z} \in Conv(\mathcal{W}^o)$. In all the four elements of $\mathcal{W}^o = \{\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}''\}$, except $\boldsymbol{x}' \wedge \boldsymbol{x}''$, all other elements have the first two components not equal to but no less than z_1, z_2. Therefore, $\boldsymbol{z} = \boldsymbol{x}' \wedge \boldsymbol{x}''$ and hence $\boldsymbol{y} = \boldsymbol{x}' + \boldsymbol{x}'' - \boldsymbol{z} = \boldsymbol{x}' \vee \boldsymbol{x}''$. Therefore, by $\boldsymbol{y} \in Conv(\mathcal{W}^o \cap \mathcal{S}^P_{c,t'\wedge t''})$ and $\boldsymbol{z} \in Conv(\mathcal{W}^o \cap \mathcal{S}^P_{c,t'\vee t''})$, we have $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}^P_{c,t'\wedge t''}$, $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}^P_{c,t'\vee t''}$. Nevertheless, if $c_3 > 0$, in all scenarios, we cannot get $\boldsymbol{x}' \vee \boldsymbol{x}'' \in \mathcal{S}^P_{c,t'\wedge t''}$. If $c_3 < 0$, in all scenarios, we cannot get $\boldsymbol{x}' \wedge \boldsymbol{x}'' \in \mathcal{S}^P_{c,t'\vee t''}$. Therefore, we always have a contradiction. Hence, the first condition in Theorem 7 holds. Q.E.D.

Proof for Theorem 9

Consider any $\boldsymbol{\beta} \in \Re^{n_2}$ and $\boldsymbol{\alpha} \in \Re^{n_1}$ with $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta} = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{\alpha}$ for some $\mathcal{I} \subseteq \{1, \dots, m\}$. Assume that $\boldsymbol{\beta}$ and $\boldsymbol{0}$ are unordered; otherwise, the statement of the theorem is trivially true. Let $\boldsymbol{x}' = \boldsymbol{\alpha}$, $\boldsymbol{t}' = \boldsymbol{0}$, $\boldsymbol{x}'' = \boldsymbol{0}$, $\boldsymbol{t}'' = \boldsymbol{\beta}$. Define the vector $\boldsymbol{c}$ as follows:

$$\boldsymbol{c}_{\mathcal{I}} = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}' + \boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}' = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{\alpha} = \boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta} = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'' + \boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}'';$$

$c_i, i \notin \mathcal{I}$ are set to very large values so that $(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'') \in \mathcal{S}^P_{\boldsymbol{c}}$.

Since $\mathcal{S}_{\boldsymbol{c}}^P$ is an additive mostly-lattice with $\mathcal{W}$, we denote $\mathcal{W}^o = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$ and can find $\boldsymbol{y} \in \mathcal{W}^o \cap \mathcal{S}_{t' \wedge t''}$, $\boldsymbol{z} \in \mathcal{W}^o \cap \mathcal{S}_{t' \vee t''}$ with $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}'' = \boldsymbol{\alpha}$. Hence, by Lemma 3, we get

$$\boldsymbol{A}_{\mathcal{I}} \boldsymbol{y} = \boldsymbol{c}_{\mathcal{I}} - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{0} \wedge \boldsymbol{\beta}) = \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta} - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{0} \wedge \boldsymbol{\beta}) = \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}^+.$$

We recall that $\boldsymbol{y} \in \mathcal{W}^o$, i.e., $\exists \mu_1, \mu_2, \mu_3, \mu_4 \in [0,1]$ with $\sum_{i=1}^4 \mu_i = 1$, and

$$\boldsymbol{y} = \mu_1 \boldsymbol{0} + \mu_2 \boldsymbol{\alpha} + \mu_3(\boldsymbol{0} \wedge \boldsymbol{\alpha}) + \mu_4(\boldsymbol{0} \vee \boldsymbol{\alpha}) = \mu_2 \boldsymbol{\alpha} - \mu_3(-\boldsymbol{\alpha})^+ + \mu_4 \boldsymbol{\alpha}^+ = (\mu_2 + \mu_4)\boldsymbol{\alpha}^+ - (\mu_2 + \mu_3)(-\boldsymbol{\alpha})^+.$$

Denoting $\lambda_1 = \mu_2 + \mu_4 \in [0,1]$, $\lambda_2 = \mu_2 + \mu_3 \in [0,1]$, we have

$$\boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}^+ = \boldsymbol{A}_{\mathcal{I}} \boldsymbol{y} = \boldsymbol{A}_{\mathcal{I}} \left(\lambda_1 \boldsymbol{\alpha}^+ - \lambda_2(-\boldsymbol{\alpha})^+ \right).$$

We next show that $\boldsymbol{B}$ is a lattice-matrix. Assume to the contrary that $\boldsymbol{B}$ is not a lattice-matrix and thus $b_{ij_1} b_{ij_2} > 0$ for some i, j_1, j_2. Choosing $\mathcal{I} = \{i\}$, $\boldsymbol{\alpha} = \boldsymbol{0}$, $\boldsymbol{\beta} = \frac{1}{b_{ij_1}} \boldsymbol{e}_{j_1} - \frac{1}{b_{ij_2}} \boldsymbol{e}_{j_2}$, we have $\boldsymbol{A}_{\mathcal{I}} \boldsymbol{\alpha} = \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta} = 0$. However, $\forall \lambda_1, \lambda_2 \in [0,1]$,

$$\boldsymbol{A}_{\mathcal{I}} \left(\lambda_1 \boldsymbol{\alpha}^+ - \lambda_2(-\boldsymbol{\alpha})^+ \right) = \boldsymbol{A}_{\mathcal{I}} \boldsymbol{0} = 0 \neq \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}^+.$$

Q.E.D.

Proof for Proposition 1

The "necessary" direction follows from Theorem 9. Hence, we focus on the "sufficient" direction. By Theorem 8, it suffices to show that $g(\boldsymbol{t}) = \max\{f(\boldsymbol{x}) : (\boldsymbol{x}, \boldsymbol{t}) \in \mathcal{S}_{\boldsymbol{c}}^P\}$ is concave and supermodular when $f(\boldsymbol{x})$ is so. From the lattice-matrix requirement and $rank(\boldsymbol{B}) = 1$, $\boldsymbol{B}$ has at most two columns (recall that we assume $\boldsymbol{B}$ does not contain any column with all zeros). Since we assume $n_2 \geq 2$, $\boldsymbol{B}$ has exactly two columns, which are linearly dependent. Let $\boldsymbol{B}_2 = -k\boldsymbol{B}_1$ with $k > 0$. Hence, $g(\boldsymbol{t}) = \max\{f(\boldsymbol{x}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{t} \leq \boldsymbol{c}\} = \max\{f(\boldsymbol{x}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}_1(t_1 - kt_2) \leq \boldsymbol{c}\} = \hat{g}(t_1 - kt_2)$, where the function $\hat{g} : \Re \to \Re$ is defined as $\hat{g}(z) = \max\{f(\boldsymbol{x}) : \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}_1 z \leq \boldsymbol{c}\}$. Noticing that $\hat{g}$ is concave and supermodular, we can get $g(\boldsymbol{t}) = \hat{g}(t_1 - kt_2)$ is also concave and supermodular (e.g., part b) of Theorem 2.2.6 in Simchi-Levi et al. 2014). Q.E.D.

Proof for Proposition 2

The case with $rank(\boldsymbol{B}) = 1$ has been analyzed by Proposition 1. Thus, we focus on cases with $rank(\boldsymbol{B}) = 2$.

"**S1**→**S2**". It is the result from Theorem 9.

"**S2**→**S3**". $\boldsymbol{B}$ being a lattice-matrix follows directly from Theorem 9. Assume **S3** is false. Since $\boldsymbol{B}$ is a lattice-matrix with two rows, it has at most four columns.

We first consider the case that $\boldsymbol{B}$ has two columns. As in the proposition, we denote $\boldsymbol{D} = \boldsymbol{B}^{-1}\boldsymbol{A}$, which does not have zero column since $\boldsymbol{A}$ has no zero column. If $rank(\boldsymbol{D}) = 1$, **S3** being false implies that $d_{11}d_{21} < 0$. By the definition of $\boldsymbol{D}$, since $\boldsymbol{A}_1 = \boldsymbol{B}\boldsymbol{D}_1 = d_{11}\boldsymbol{B}_1 + d_{21}\boldsymbol{B}_2$, Lemma 7 implies that **S2** is false. If $rank(\boldsymbol{D}) = 2$, **S3** being false implies that $\boldsymbol{D}$ has two nonzero elements with opposite sign. If there exists $i \in \{1, \ldots, n_1\}$ such that $d_{1i}d_{2i} < 0$, considering $\boldsymbol{A}_i = \boldsymbol{B}\boldsymbol{D}_i = d_{1i}\boldsymbol{B}_1 + d_{2i}\boldsymbol{B}_2$, again from Lemma 7, **S2** is false. If for any $i \in \{1, \ldots, n_1\}$, $d_{1i}d_{2i} \geq 0$, together with $rank(\boldsymbol{D}) = 2$ and $\boldsymbol{D}$ having elements with opposite signs, we can find $i, j \in \{1, \ldots, n_1\}$ such that $\boldsymbol{D}_i \geq 0$, $\boldsymbol{D}_j \leq 0$, and $\boldsymbol{D}_i, \boldsymbol{D}_j$ are linearly independent. We can choose $\eta, \lambda > 0$ such that $\eta\boldsymbol{D}_i + \lambda\boldsymbol{D}_j \in \Re^2$ and $\boldsymbol{0}$ are unordered. Let $\boldsymbol{\alpha} = \eta\boldsymbol{e}_i + \lambda\boldsymbol{e}_j \geq \boldsymbol{0}$. We have $\boldsymbol{A}\boldsymbol{\alpha} = \eta\boldsymbol{A}_i + \lambda\boldsymbol{A}_j = \boldsymbol{B}(\eta\boldsymbol{D}_i + \lambda\boldsymbol{D}_j)$, which implies that **S2** is false based on Lemma 7.

We now consider the case that $\boldsymbol{B}$ has three columns. WLOG, let $\boldsymbol{b}_1^T = (b_{11}, b_{12}, 0)$, $\boldsymbol{b}_2^T = (0, b_{22}, b_{23})$ where $b_{11}b_{12} < 0, b_{23} \neq 0$. If **S3** is false, we first prove that there exists $\boldsymbol{\alpha} \geq \boldsymbol{0}$ such that $\boldsymbol{A}\boldsymbol{\alpha}$ is independent from any column of $\boldsymbol{B}$. We show this by contradiction. Suppose for any $\boldsymbol{\alpha} \geq \boldsymbol{0}$, $\boldsymbol{A}\boldsymbol{\alpha}$ is linearly dependent on one column in $\boldsymbol{B}$, which implies that $\boldsymbol{A} = \boldsymbol{B}_i\boldsymbol{k}^T$ for some $i \in \{1, 2, 3\}$ and $\boldsymbol{k} \in \Re^{n_1}$. Since **S3** is false, we have either $b_{22} = 0$ or $\boldsymbol{A} = \boldsymbol{B}_1\boldsymbol{k}^T$ or $\boldsymbol{A} = \boldsymbol{B}_3\boldsymbol{k}^T$. In any case, $\boldsymbol{A}$ has a zero row, which contradicts with the original assumption on $\boldsymbol{A}$. Hence, there exists a vector $\boldsymbol{\alpha} \geq \boldsymbol{0}$ such that $\boldsymbol{A}\boldsymbol{\alpha}$ is independent from any column of $\boldsymbol{B}$. Observing that $b_{11}b_{12} < 0, b_{23} \neq 0, b_{22}b_{23} \leq 0$, by plotting vectors $\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{B}_3$ and $\boldsymbol{A}\boldsymbol{\alpha}$ on a Cartesian coordinate system, we can get $\boldsymbol{A}\boldsymbol{\alpha} = \delta\boldsymbol{B}_i + \gamma\boldsymbol{B}_j$ for some $i, j \in \{1, 2, 3\}$ and $\delta\gamma < 0$. According to Lemma 7, **S2** is false.

We now show that $\boldsymbol{B}$ cannot have four columns. Assume to the contraty that $\boldsymbol{B}$ has four columns. Since we assume it is a lattice-matrix and does not contain zero column, WLOG, we can let

$\boldsymbol{b}_1^T = (b_{11}, b_{12}, 0, 0)$, $\boldsymbol{b}_2^T = (0, 0, b_{23}, b_{24})$ where $b_{11}b_{12} < 0$, $b_{23}b_{24} < 0$. The proof is similar to the case with three columns. We can show that there exists $\boldsymbol{\alpha} \geq \boldsymbol{0}$ such that $\boldsymbol{A\alpha}$ is independent from any column of $\boldsymbol{B}$. Hence, there exist some $i \in \{1,2\}$, $j \in \{3,4\}$ and $\delta\gamma < 0$ such that $\boldsymbol{A\alpha} = \delta \boldsymbol{B}_i + \gamma \boldsymbol{B}_j$. According to Lemma 7, **S2** is false.

"**S3**→**S1**". The case of $rank(\boldsymbol{B}) = 1$ has been shown in Proposition 1 and hence we only focus on the case of $rank(\boldsymbol{B}) = 2$. To show $\mathcal{S}_{\boldsymbol{c}}^P$ is an additive most-lattice with $\mathcal{W}(\boldsymbol{x}', \boldsymbol{x}'') = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$, by Theorem 8, it suffices to show that $g(\boldsymbol{t}) = \max\{f(\boldsymbol{x}) : (\boldsymbol{x}, \boldsymbol{t}) \in \mathcal{S}_{\boldsymbol{c}}^P\}$ is concave supermodular when $f(\boldsymbol{x})$ is so. We start from the case when $\boldsymbol{B}$ has two columns and denote $\boldsymbol{D} = \boldsymbol{B}^{-1}\boldsymbol{A}$ as in the proposition. Since $\boldsymbol{B} \in \Re^{2\times 2}$ is a lattice-matrix, for each column in $\boldsymbol{B}^{-1}$, the elements in it are with the same sign. Hence, we can denote $\boldsymbol{B}^{-1} = \boldsymbol{M} \times \begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix}$, where $\boldsymbol{M} \geq \boldsymbol{0}$, u, v are either 1 or -1. In that case,

$$
\begin{aligned}
g(\boldsymbol{t}) &= \max\{f(\boldsymbol{x}) \ : \ \boldsymbol{Ax} + \boldsymbol{Bt} \leq \boldsymbol{c}\} \\
&= \max\left\{ f(\boldsymbol{x}) \ : \ \boldsymbol{Ax} + \begin{bmatrix} -u & 0 \\ 0 & -v \end{bmatrix} \boldsymbol{p} + \boldsymbol{Bt} = \boldsymbol{c}, \begin{bmatrix} -u & 0 \\ 0 & -v \end{bmatrix} \boldsymbol{p} \in \Re_+^2 \right\} \\
&= \max\left\{ f(\boldsymbol{x}) \ : \ -\boldsymbol{Dx} - \boldsymbol{M} \begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix} \begin{bmatrix} -u & 0 \\ 0 & -v \end{bmatrix} \boldsymbol{p} - \boldsymbol{t} = -\boldsymbol{B}^{-1}\boldsymbol{c}, \begin{bmatrix} -u & 0 \\ 0 & -v \end{bmatrix} \boldsymbol{p} \in \Re_+^2 \right\} \\
&= \max\left\{ f(\boldsymbol{x}) \ : \ -\boldsymbol{Dx} + \boldsymbol{Mp} + \boldsymbol{q} = \boldsymbol{t}, (\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}) \in \Re^{n_1} \times \hat{\Re}^2 \times \{\boldsymbol{B}^{-1}\boldsymbol{c}\} \right\},
\end{aligned}
$$

where $\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}$ are decision variables and the set $\hat{\Re}^2 = \{\boldsymbol{p} \in \Re_2 : -up_1 \geq 0, -vp_2 \geq 0\}$ is a sublattice.

- Consider the case of $rank(\boldsymbol{D}) = 2$. If $\boldsymbol{D} \leq \boldsymbol{0}$, we have $g(\boldsymbol{t}) = \max\{\hat{f}(\boldsymbol{z}) : \hat{\boldsymbol{A}}\boldsymbol{z} = \boldsymbol{t}, \boldsymbol{z} \in \mathcal{D}\}$, where $\hat{\boldsymbol{A}} = [-\boldsymbol{D} \ \boldsymbol{M} \ \boldsymbol{I}] \geq \boldsymbol{0}$, $\mathcal{D} = \Re^{n_1} \times \hat{\Re}^2 \times \{\boldsymbol{B}^{-1}\boldsymbol{c}\}$ is a sublattice, $\hat{f}(\boldsymbol{z}) = \hat{f}(\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{q}) = f(\boldsymbol{x})$ is concave and supermodular. Therefore, following Theorem 1 in Chen et al. (2013), $g(\boldsymbol{t})$ is concave and supermodular. For the case of $\boldsymbol{D} \geq \boldsymbol{0}$, we let $\boldsymbol{x} = -\boldsymbol{y}$ and $\boldsymbol{Ax} = (-\boldsymbol{A})\boldsymbol{y}$ with $\boldsymbol{B}^{-1}(-\boldsymbol{A}) = -\boldsymbol{D} \leq \boldsymbol{0}$. Hence, the problem becomes $g(\boldsymbol{t}) = \max\{f(-\boldsymbol{y}) : \boldsymbol{Dy} + \boldsymbol{Mp} + \boldsymbol{q} = \boldsymbol{t}, (\boldsymbol{y}, \boldsymbol{p}, \boldsymbol{q}) \in \Re^{n_1} \times \hat{\Re}^2 \times \{\boldsymbol{B}^{-1}\boldsymbol{c}\}\}$. Since the objective function $f(-\boldsymbol{y})$ is also concave and supermodular in $\boldsymbol{y}$, we have $g(\boldsymbol{t})$ is concave and supermodular.

- Consider the case of $rank(\boldsymbol{D})=1$. We then have $\boldsymbol{k}\in\Re^{n_1}$ such that $\boldsymbol{D}=\boldsymbol{D}_1\boldsymbol{k}^T$. Denote the function $\bar{f}(w)=\max\{f(\boldsymbol{x}):-\boldsymbol{k}^T\boldsymbol{x}=w\}$. It is concave since f is concave and the set $\{\boldsymbol{x}:-\boldsymbol{k}^T\boldsymbol{x}=w\}$ is convex. Moreover, it is supermodular since it is a function on a scalar. Hence,

$$\begin{aligned} g(\boldsymbol{t}) &= \max\left\{f(\boldsymbol{x}) \ : \ -\boldsymbol{D}_1\boldsymbol{k}^T\boldsymbol{x}+\boldsymbol{M}\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{t},(\boldsymbol{x},\boldsymbol{p},\boldsymbol{q})\in\Re^{n_1}\times\hat{\Re}^2\times\{\boldsymbol{B}^{-1}\boldsymbol{c}\}\right\} \\ &= \max\left\{\bar{f}(w) \ : \ \boldsymbol{D}_1 w+\boldsymbol{M}\boldsymbol{p}+\boldsymbol{q}=\boldsymbol{t},(w,\boldsymbol{p},\boldsymbol{q})\in\Re\times\hat{\Re}^2\times\{\boldsymbol{B}^{-1}\boldsymbol{c}\}\right\}. \end{aligned}$$

Similar to the case of $rank(\boldsymbol{D})=2$, here $g(\boldsymbol{t})$ is also concave and supermodular.

We now move to the case when $\boldsymbol{B}$ has three columns. WLOG, assume $\boldsymbol{B}_3$ does not contain any zero element. We then have $\boldsymbol{k}\in\Re^{n_1}$ such that $\boldsymbol{A}=\boldsymbol{B}_3\boldsymbol{k}^T$. Similar to the discussion in the previous case, let $\bar{f}(w)=\max\{f(\boldsymbol{x}):\boldsymbol{k}^T\boldsymbol{x}=w\}$, which is concave and supermodular. In addition,

$$\begin{aligned} g(\boldsymbol{t}) &= \max\{f(\boldsymbol{x}) \ : \ \boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{t}\le\boldsymbol{c}\} \\ &= \max\left\{f(\boldsymbol{x}) \ : \ \boldsymbol{B}_3\boldsymbol{k}^T\boldsymbol{x}+\boldsymbol{B}_1 t_1+\boldsymbol{B}_2 t_2+\boldsymbol{B}_3 t_3\le\boldsymbol{c}\right\} \\ &= \max\left\{\bar{f}(z-t_3) \ : \ \boldsymbol{B}_1 t_1+\boldsymbol{B}_2 t_2+\boldsymbol{B}_3 z\le\boldsymbol{c}\right\}, \end{aligned}$$

where $\bar{f}(z-t_3)$ is concave and supermodular in (z,t_3) since $\bar{f}(w)$ is concave and supermodular (see, for instance, part b) of Theorem 2.2.6 in Simchi-Levi et al. 2014). Since the matrix $\boldsymbol{B}$ is a lattice-matrix, $g(\boldsymbol{t})$ is concave and supermodular. Q.E.D.

Proof for Proposition 3

Assume that **S3** of Proposition 2 holds. Here we prove that set $\mathcal{S}_{\boldsymbol{c}}=\{(\boldsymbol{x},\boldsymbol{t}):\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{t}=\boldsymbol{c},\boldsymbol{x}\in\mathcal{D}\}$ is an additive mostly-lattice with $\mathcal{W}$. Consider any $(\boldsymbol{x}',\boldsymbol{t}'),(\boldsymbol{x}'',\boldsymbol{t}'')\in\mathcal{S}_{\boldsymbol{c}}\subseteq\mathcal{S}^P_{\boldsymbol{c}}$, where $\mathcal{S}^P_{\boldsymbol{c}}$ is the polyhedron defined by the equation (2). By Statement **S1** of Proposition 2, there exist $\boldsymbol{y},\boldsymbol{z}\in\mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'')$ such that $\boldsymbol{y}+\boldsymbol{z}=\boldsymbol{x}'+\boldsymbol{x}''$, and $(\boldsymbol{y},\boldsymbol{t}'\wedge\boldsymbol{t}''),(\boldsymbol{z},\boldsymbol{t}'\vee\boldsymbol{t}'')\in\mathcal{S}^P_{\boldsymbol{c}}$. By Lemma 3, we have

$$\begin{aligned} \boldsymbol{A}\boldsymbol{y}+\boldsymbol{B}(\boldsymbol{t}'\wedge\boldsymbol{t}'') &= \boldsymbol{c}, \\ \boldsymbol{A}\boldsymbol{z}+\boldsymbol{B}(\boldsymbol{t}'\vee\boldsymbol{t}'') &= \boldsymbol{c}. \end{aligned}$$

Meanwhile, since $\mathcal{D}$ is a lattice, $\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}''\in\mathcal{D}$. Therefore, $\boldsymbol{y},\boldsymbol{z}\in\mathcal{W}(\boldsymbol{x}',\boldsymbol{x}'')\subseteq\mathcal{D}$ as $\mathcal{D}$ is convex. Hence, $(\boldsymbol{y},\boldsymbol{t}'\wedge\boldsymbol{t}''),(\boldsymbol{z},\boldsymbol{t}'\vee\boldsymbol{t}'')\in\mathcal{S}_{\boldsymbol{c}}$. The case for set $\{(\boldsymbol{x},\boldsymbol{t}):\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{t}\le\boldsymbol{c},\boldsymbol{x}\in\mathcal{D}\}$ can be proved similarly. Q.E.D.

Proof for Theorem 11

We first prove the "only if" part. Assume to the contrary, i.e., both conditions in the theorem are not satisfied. In other words, $n > m$, and we can find $\mathcal{I} \subseteq \{1, \ldots, m\}$ and $\boldsymbol{\beta} \in \Re^{n_2}$ with $|\mathcal{I}| = n+1$, $\text{Rank}(\boldsymbol{A}_{\mathcal{I}}) = n$, $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta} \in C(\boldsymbol{A}_{\mathcal{I}})$, but $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+ \notin C(\boldsymbol{A}_{\mathcal{I}})$. In this case, there exists $\boldsymbol{\alpha} \in \Re^{n_1}$ satisfying $\boldsymbol{A}_{\mathcal{I}}\boldsymbol{\alpha} = \boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}$.

Choose any $\boldsymbol{x}'' \in \Re^{n_1}, \boldsymbol{t}'' \in \Re^{n_2}$ and denote $\boldsymbol{x}' = \boldsymbol{x}'' + \boldsymbol{\alpha}, \boldsymbol{t}' = \boldsymbol{t}'' - \boldsymbol{\beta}$, $\boldsymbol{c}_{\mathcal{I}} = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'' + \boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}''$. We have

$$\boldsymbol{A}_{\mathcal{I}}(\boldsymbol{x}' - \boldsymbol{x}'') = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{\alpha} = \boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta} = \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'' - \boldsymbol{t}'),$$

and hence $\boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}' + \boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}' = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'' + \boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}'' = \boldsymbol{c}_{\mathcal{I}}$. Choosing $c_i, i \notin \mathcal{I}$ to be values large enough, we have $(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'') \in \mathcal{S}_{\boldsymbol{c}}^P$. Moreover, neither $\boldsymbol{\beta} \geq \boldsymbol{0}$ nor $\boldsymbol{\beta} \leq \boldsymbol{0}$ is true since $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta} \in C(\boldsymbol{A}_{\mathcal{I}})$ and $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+ \notin C(\boldsymbol{A}_{\mathcal{I}})$, which implies that $\boldsymbol{t}', \boldsymbol{t}''$ are unordered.

If we can find $\boldsymbol{y}, \boldsymbol{z}$ with $(\boldsymbol{y}, \boldsymbol{t}' \wedge \boldsymbol{t}''), (\boldsymbol{z}, \boldsymbol{t}' \vee \boldsymbol{t}'') \in \mathcal{S}_{\boldsymbol{c}}^P$ and $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$, by Lemma 3 in the appendix, we have

$$\boldsymbol{A}_{\mathcal{I}}\boldsymbol{y} = \boldsymbol{c}_{\mathcal{I}} - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'') = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'' + \boldsymbol{B}_{\mathcal{I}}\boldsymbol{t}'' - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'') = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'' + \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')) = \boldsymbol{A}_{\mathcal{I}}\boldsymbol{x}'' + \boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+.$$

It contradicts $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+ \notin C(\boldsymbol{A}_{\mathcal{I}})$. Hence, no such $\boldsymbol{y}, \boldsymbol{z}$ exist and $\mathcal{S}_{\boldsymbol{c}}^P$ is not an additive mostly-lattice.

Now we prove the "if" part by contradiction. Suppose there exists $\boldsymbol{c}$ such that the set $\mathcal{S}_{\boldsymbol{c}}^P$ is not an additive mostly-lattice. That is, there exist $\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{t}', \boldsymbol{t}''$ such that $\boldsymbol{A}\boldsymbol{x}' + \boldsymbol{B}\boldsymbol{t}' \leq \boldsymbol{c}$, $\boldsymbol{A}\boldsymbol{x}'' + \boldsymbol{B}\boldsymbol{t}'' \leq \boldsymbol{c}$, but there do not exist $\boldsymbol{y}, \boldsymbol{z}$ satisfying $\boldsymbol{A}\boldsymbol{y} + \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'') \leq \boldsymbol{c}$, $\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}(\boldsymbol{t}' \vee \boldsymbol{t}'') \leq \boldsymbol{c}$, and $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$. Let $\boldsymbol{u}' = \boldsymbol{A}\boldsymbol{x}' + \boldsymbol{B}\boldsymbol{t}'$, $\boldsymbol{u}'' = \boldsymbol{A}\boldsymbol{x}'' + \boldsymbol{B}\boldsymbol{t}''$. We then have the set

$$\mathcal{W} = \{\boldsymbol{y} \in \Re^{n_1} \ : \ (\boldsymbol{u}' \wedge \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'') \leq \boldsymbol{A}\boldsymbol{y} \leq (\boldsymbol{u}' \vee \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'')\} = \emptyset;$$

otherwise, we can have $\boldsymbol{y}$ with $(\boldsymbol{u}' \wedge \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'') \leq \boldsymbol{A}\boldsymbol{y} \leq (\boldsymbol{u}' \vee \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'')$, which implies

$$\boldsymbol{A}\boldsymbol{y} + \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'') \leq \boldsymbol{u}' \vee \boldsymbol{u}'' \leq \boldsymbol{c},$$

$$\boldsymbol{A}(\boldsymbol{x}' + \boldsymbol{x}'' - \boldsymbol{y}) + \boldsymbol{B}(\boldsymbol{t}' \vee \boldsymbol{t}'') = \boldsymbol{A}\boldsymbol{x}' + \boldsymbol{B}\boldsymbol{t}' + \boldsymbol{A}\boldsymbol{x}'' + \boldsymbol{B}\boldsymbol{t}'' - (\boldsymbol{A}\boldsymbol{y} + \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) \leq \boldsymbol{u}' + \boldsymbol{u}'' - (\boldsymbol{u}' \wedge \boldsymbol{u}'') = \boldsymbol{u}' \vee \boldsymbol{u}'' \leq \boldsymbol{c},$$

and contradicts with the previous assumption of the nonexistence of $\boldsymbol{y}, \boldsymbol{z}$. By $\mathcal{W} = \emptyset$, the following problem

$$\begin{aligned} \max \ & 0 \\ \text{s.t.} \ & \begin{bmatrix} \boldsymbol{A} \\ -\boldsymbol{A} \end{bmatrix} \boldsymbol{y} \leq \begin{bmatrix} (\boldsymbol{u}' \vee \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'') \\ -((\boldsymbol{u}' \wedge \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) \end{bmatrix} \end{aligned} \tag{EC.14}$$

is infeasible. If $n = m$, we can solve $\boldsymbol{y}$ with $\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{u}' \vee \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'') \geq (\boldsymbol{u}' \wedge \boldsymbol{u}'') - \boldsymbol{B}(\boldsymbol{t}' \wedge \boldsymbol{t}'')$, contradicting with that the problem (EC.14) is infeasible. Hence, $n < m$. By Lemma 9 in the appendix, there exists $\mathcal{I} \subseteq \{1, \ldots, m\}$ such that $|\mathcal{I}| = n+1$, $\text{Rank}(\boldsymbol{A}_{\mathcal{I}}) = n$, and

$$\begin{aligned} \max \ & 0 \\ \text{s.t.} \ & \begin{bmatrix} \boldsymbol{A}_{\mathcal{I}} \\ -\boldsymbol{A}_{\mathcal{I}} \end{bmatrix} \boldsymbol{y} \leq \begin{bmatrix} (\boldsymbol{u}'_{\mathcal{I}} \vee \boldsymbol{u}''_{\mathcal{I}}) - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'') \\ -((\boldsymbol{u}'_{\mathcal{I}} \wedge \boldsymbol{u}''_{\mathcal{I}}) - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) \end{bmatrix} \end{aligned} \tag{EC.15}$$

is infeasible.

Therefore, the dual of problem (EC.15)

$$\begin{aligned} \min \ & \boldsymbol{p}_1^T ((\boldsymbol{u}'_{\mathcal{I}} \vee \boldsymbol{u}''_{\mathcal{I}}) - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) - \boldsymbol{p}_2^T ((\boldsymbol{u}'_{\mathcal{I}} \wedge \boldsymbol{u}''_{\mathcal{I}}) - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) \\ \text{s.t.} \ & \boldsymbol{A}_{\mathcal{I}}^T (\boldsymbol{p}_1 - \boldsymbol{p}_2) = \boldsymbol{0} \\ & \boldsymbol{p}_1, \boldsymbol{p}_2 \geq \boldsymbol{0} \end{aligned} \tag{EC.16}$$

is unbounded. Hence, there exist $\boldsymbol{p}_1, \boldsymbol{p}_2 \geq \boldsymbol{0}$ such that $\boldsymbol{A}_{\mathcal{I}}^T (\boldsymbol{p}_1 - \boldsymbol{p}_2) = \boldsymbol{0}$, and

$$h = \boldsymbol{p}_1^T ((\boldsymbol{u}'_{\mathcal{I}} \vee \boldsymbol{u}''_{\mathcal{I}}) - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) - \boldsymbol{p}_2^T ((\boldsymbol{u}'_{\mathcal{I}} \wedge \boldsymbol{u}''_{\mathcal{I}}) - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) < 0.$$

Moreover,

$$\begin{aligned} 0 &> h \\ &= h - (\boldsymbol{p}_1 - \boldsymbol{p}_2)^T \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}' \\ &\geq \boldsymbol{p}_1^T (\boldsymbol{u}'_{\mathcal{I}} - \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}' - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) - \boldsymbol{p}_2^T (\boldsymbol{u}'_{\mathcal{I}} - \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}' - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) \\ &= (\boldsymbol{p}_1 - \boldsymbol{p}_2)^T (\boldsymbol{u}'_{\mathcal{I}} - \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}' - \boldsymbol{B}_{\mathcal{I}} \boldsymbol{t}' + \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' - (\boldsymbol{t}' \wedge \boldsymbol{t}''))) \\ &= (\boldsymbol{p}_1 - \boldsymbol{p}_2)^T \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')), \end{aligned} \tag{EC.17}$$

$$
\begin{aligned}
0 &> h \\
&= h - (\boldsymbol{p}_1 - \boldsymbol{p}_2)^T \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}'' \\
&\geq \boldsymbol{p}_1^T (\boldsymbol{u}''_{\mathcal{I}} - \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}'' - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) - \boldsymbol{p}_2^T (\boldsymbol{u}''_{\mathcal{I}} - \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}'' - \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' \wedge \boldsymbol{t}'')) \qquad \text{(EC.18)} \\
&= (\boldsymbol{p}_1 - \boldsymbol{p}_2)^T (\boldsymbol{u}''_{\mathcal{I}} - \boldsymbol{A}_{\mathcal{I}} \boldsymbol{x}'' - \boldsymbol{B}_{\mathcal{I}} \boldsymbol{t}'' + \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')) \\
&= (\boldsymbol{p}_1 - \boldsymbol{p}_2)^T \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')),
\end{aligned}
$$

where in both equations the first equality follows from $\boldsymbol{A}_{\mathcal{I}}^T(\boldsymbol{p}_1 - \boldsymbol{p}_2) = \boldsymbol{0}$, and the second inequality follows from $\boldsymbol{p}_1, \boldsymbol{p}_2 \geq \boldsymbol{0}$. Denote $\boldsymbol{q} = \boldsymbol{p}_1 - \boldsymbol{p}_2$, we have $\boldsymbol{A}_{\mathcal{I}}^T \boldsymbol{q} = \boldsymbol{0}$. Let $\Delta_1 = \boldsymbol{q}^T \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')) < 0$, $\Delta_2 = \boldsymbol{q}^T \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}'' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')) < 0$. We have

$$
0 = \frac{\Delta_1}{\Delta_1} - \frac{\Delta_2}{\Delta_2} = \boldsymbol{q}^T \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}, \qquad \text{(EC.19)}
$$

where $\boldsymbol{\beta} = \frac{\boldsymbol{t}' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')}{\Delta_1} - \frac{\boldsymbol{t}'' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')}{\Delta_2}$. Note that $\boldsymbol{q} \neq \boldsymbol{0}$; otherwise $\boldsymbol{q}^T \boldsymbol{B}_{\mathcal{I}}(\boldsymbol{t}' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')) = 0$, which contradicts with the inequality (EC.17). Moreover, since $\text{Rank}(\boldsymbol{A}_{\mathcal{I}}^T) = \text{Rank}(\boldsymbol{A}_{\mathcal{I}}) = n$, the solution space of $\boldsymbol{A}_{\mathcal{I}}^T \boldsymbol{v} = \boldsymbol{0}$ is of dimension 1. Hence, for any $\boldsymbol{v}$ with $\boldsymbol{A}_{\mathcal{I}}^T \boldsymbol{v} = \boldsymbol{0}$, we have $\boldsymbol{v} = k\boldsymbol{q}$ for some $k \in \Re$ and hence $\boldsymbol{v}^T \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta} = k\boldsymbol{q}^T \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta} = 0$, which implies that 0 is the optimal value of the problem

$$
\begin{aligned}
\min \quad & \boldsymbol{v}^T \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta} \\
\text{s.t.} \quad & \boldsymbol{A}_{\mathcal{I}}^T \boldsymbol{v} = \boldsymbol{0},
\end{aligned}
$$

and its dual problem

$$
\begin{aligned}
\max \quad & 0 \\
\text{s.t.} \quad & \boldsymbol{A}_{\mathcal{I}} \boldsymbol{\alpha} = \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}
\end{aligned}
$$

is feasible. Hence, we can find $\boldsymbol{\alpha}$ satisfying $\boldsymbol{A}_{\mathcal{I}} \boldsymbol{\alpha} = \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}$, i.e., $\boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta} \in C(\boldsymbol{A}_{\mathcal{I}})$. Nevertheless, since the inequality (EC.18) still holds when multiplying $\boldsymbol{q}$ by any positive value and $\boldsymbol{\beta}^+ = -\frac{\boldsymbol{t}'' - (\boldsymbol{t}' \wedge \boldsymbol{t}'')}{\Delta_2}$. the problem

$$
\begin{aligned}
\min \quad & \boldsymbol{v}^T \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}^+ \\
\text{s.t.} \quad & \boldsymbol{A}_{\mathcal{I}}^T \boldsymbol{v} = \boldsymbol{0}.
\end{aligned}
$$

is unbounded. Therefore, its dual problem

$$
\begin{aligned}
\max \quad & 0 \\
\text{s.t.} \quad & \boldsymbol{A}_{\mathcal{I}} \boldsymbol{\alpha} = \boldsymbol{B}_{\mathcal{I}} \boldsymbol{\beta}^+,
\end{aligned}
$$

is infeasible, i.e., $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+ \notin C(\boldsymbol{A}_{\mathcal{I}})$. Hence, we show that if there do not exist $\boldsymbol{y},\boldsymbol{z}$ such that $\boldsymbol{A}\boldsymbol{y}+\boldsymbol{B}(\boldsymbol{t}'\wedge\boldsymbol{t}'')\le\boldsymbol{c}, \boldsymbol{A}\boldsymbol{z}+\boldsymbol{B}(\boldsymbol{t}'\vee\boldsymbol{t}'')\le\boldsymbol{c}$, and $\boldsymbol{y}+\boldsymbol{z}=\boldsymbol{x}'+\boldsymbol{x}''$, then $n<m$, and there exists $\mathcal{I}\subseteq\{1,\dots,m\}$ with $|\mathcal{I}|=n+1$, $\boldsymbol{\beta}\in\Re^{n_2}$ such that $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}\in C(A_{\mathcal{I}})$ and $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+\notin C(A_{\mathcal{I}})$. The proof of the "if" part is completed. Q.E.D.

Proof for Theorem 12

The case with $n=m$ is straightforward. Here we just consider the case with $n<m$. Note that the condition in Theorem 11 only depends on the relationship between $\boldsymbol{B}$ and $C(\boldsymbol{A})$. Therefore, removing the dependent columns of $\boldsymbol{A}$ would not change the satisfaction/violation of the condition.

If Algorithm 1 returns $s=0$, then we stop at step 2 with $\hat{\mathcal{I}}$, and i,j,k such that $k\in\{1,\dots,m\}\setminus\hat{\mathcal{I}}$ and $d_{ki}d_{kj}>0$. Let

$$\boldsymbol{\beta}=\frac{1}{d_{ki}}\boldsymbol{e}_i-\frac{1}{d_{kj}}\boldsymbol{e}_j,\qquad \boldsymbol{\alpha}=\boldsymbol{A}_{\hat{\mathcal{I}}}^{-1}\left(\frac{1}{d_{ki}}\boldsymbol{B}_{\hat{\mathcal{I}},i}-\frac{1}{d_{kj}}\boldsymbol{B}_{\hat{\mathcal{I}},j}\right),$$

where $\boldsymbol{B}_{\hat{\mathcal{I}},i},\boldsymbol{B}_{\hat{\mathcal{I}},j}$ are the i,jth columns of the submatrix $\boldsymbol{B}_{\hat{\mathcal{I}}}$, respectively. Denote $\mathcal{I}=\hat{\mathcal{I}}\cup\{k\}$. We have

$$\begin{bmatrix}\boldsymbol{B}_{\hat{\mathcal{I}}}\\ \boldsymbol{b}_k^T\end{bmatrix}\boldsymbol{\beta}=\begin{bmatrix}\frac{\boldsymbol{B}_{\hat{\mathcal{I}},i}}{d_{ki}}-\frac{\boldsymbol{B}_{\hat{\mathcal{I}},j}}{d_{kj}}\\ \frac{b_{ki}}{d_{ki}}-\frac{b_{kj}}{d_{kj}}\end{bmatrix}=\begin{bmatrix}\frac{\boldsymbol{B}_{\hat{\mathcal{I}},i}}{d_{ki}}-\frac{\boldsymbol{B}_{\hat{\mathcal{I}},j}}{d_{kj}}\\ \frac{d_{ki}+\boldsymbol{a}_k^T\boldsymbol{A}_{\hat{\mathcal{I}}}^{-1}\boldsymbol{B}_{\hat{\mathcal{I}},i}}{d_{ki}}-\frac{d_{kj}+\boldsymbol{a}_k^T\boldsymbol{A}_{\hat{\mathcal{I}}}^{-1}\boldsymbol{B}_{\hat{\mathcal{I}},j}}{d_{kj}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{\hat{\mathcal{I}}}\\ \boldsymbol{a}_k^T\end{bmatrix}\boldsymbol{\alpha},$$

where the last equality follows from the definition of $\boldsymbol{\alpha}$. Hence, $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}=\boldsymbol{A}_{\mathcal{I}}\boldsymbol{\alpha}\in C(\boldsymbol{A}_{\mathcal{I}})$. However, if $d_{ki}>0$, we have $(-d_{kj})<0$, and

$$\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+=\begin{bmatrix}\boldsymbol{B}_{\hat{\mathcal{I}}}\\ \boldsymbol{b}_k^T\end{bmatrix}\frac{\boldsymbol{e}_i}{d_{k_i}}=\begin{bmatrix}\frac{\boldsymbol{B}_{\hat{\mathcal{I}},i}}{d_{ki}}\\ \frac{b_{ki}}{d_{ki}}\end{bmatrix}=\begin{bmatrix}\frac{\boldsymbol{B}_{\hat{\mathcal{I}},i}}{d_{ki}}\\ \frac{\boldsymbol{a}_k^T\boldsymbol{A}_{\hat{\mathcal{I}}}^{-1}\boldsymbol{B}_{\hat{\mathcal{I}},i}}{d_{ki}}+1\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{\hat{\mathcal{I}}}\\ \boldsymbol{a}_k^T\end{bmatrix}\frac{\boldsymbol{A}_{\hat{\mathcal{I}}}^{-1}\boldsymbol{B}_{\hat{\mathcal{I}},i}}{d_{ki}}+\begin{bmatrix}\boldsymbol{0}\\ 1\end{bmatrix}.$$

On the right hand side, the first term is in $C(\boldsymbol{A}_{\mathcal{I}})$ but the second term is not (as $\boldsymbol{A}_{\hat{\mathcal{I}}}$ is invertible, there does not exist $\boldsymbol{x}$ satisfying $\begin{bmatrix}\boldsymbol{A}_{\hat{\mathcal{I}}}\\ \boldsymbol{a}_k^T\end{bmatrix}\boldsymbol{x}=\begin{bmatrix}\boldsymbol{0}\\ 1\end{bmatrix}$). Similarly, we can also prove $\boldsymbol{B}_{\mathcal{I}}\boldsymbol{\beta}^+\notin C(\boldsymbol{A}_{\mathcal{I}})$ in the case of $d_{ki}<0$. Therefore, the condition in Theorem 11 cannot be satisfied.

We now prove the other direction. Assume that the condition in Theorem 11 is violated, i.e., there exist $\boldsymbol{\beta} \in \Re^{n_2}$ and $\mathcal{I} \subseteq \{1,\ldots,m\}$ with $|\mathcal{I}| = n+1$, $\text{Rank}(\boldsymbol{A}_\mathcal{I}) = n$, $\boldsymbol{B}_\mathcal{I}\boldsymbol{\beta} \in C(\boldsymbol{A}_\mathcal{I})$ and $\boldsymbol{B}_\mathcal{I}\boldsymbol{\beta}^+ \notin C(\boldsymbol{A}_\mathcal{I})$. Let set $\hat{\mathcal{I}} \subset \mathcal{I}$ be the set such that $\boldsymbol{A}_{\hat{\mathcal{I}}}$ is invertible, then $|\hat{\mathcal{I}}| = n$ and $\{k\} = \mathcal{I} \setminus \hat{\mathcal{I}}$. We have

$$\boldsymbol{B}_\mathcal{I} = \begin{bmatrix} \boldsymbol{B}_{\hat{\mathcal{I}}} \\ \boldsymbol{b}_k^T \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{b}_k^T - \boldsymbol{a}_k^T \boldsymbol{A}_{\hat{\mathcal{I}}}^{-1} \boldsymbol{B}_{\hat{\mathcal{I}}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{A}_{\hat{\mathcal{I}}} \\ \boldsymbol{a}_k^T \end{bmatrix} \boldsymbol{A}_{\hat{\mathcal{I}}}^{-1} \boldsymbol{B}_{\hat{\mathcal{I}}} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{d}_k^T \end{bmatrix} + \boldsymbol{A}_\mathcal{I} \boldsymbol{A}_{\hat{\mathcal{I}}}^{-1} \boldsymbol{B}_{\hat{\mathcal{I}}},$$

where we denote $\boldsymbol{d}_k^T = \boldsymbol{b}_k^T - \boldsymbol{a}_k^T \boldsymbol{A}_{\hat{\mathcal{I}}}^{-1} \boldsymbol{B}_{\hat{\mathcal{I}}}$ as in the algorithm. Therefore, for any $\boldsymbol{\gamma}$, we have

$$\boldsymbol{B}_\mathcal{I}\boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{d}_k^T \end{bmatrix} \boldsymbol{\gamma} + \boldsymbol{A}_\mathcal{I} \boldsymbol{A}_{\hat{\mathcal{I}}}^{-1} \boldsymbol{B}_{\hat{\mathcal{I}}} \boldsymbol{\gamma}.$$

Since the second term is always in $C(\boldsymbol{A}_\mathcal{I})$, $\boldsymbol{B}_\mathcal{I}\boldsymbol{\gamma} \in C(\boldsymbol{A}_\mathcal{I})$ if and only if $\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{d}_k^T \end{bmatrix} \boldsymbol{\gamma} \in C(\boldsymbol{A}_\mathcal{I}) = C\left(\begin{bmatrix} \boldsymbol{A}_{\hat{\mathcal{I}}} \\ \boldsymbol{a}_k^T \end{bmatrix}\right)$. It is equivalent to $\boldsymbol{d}_k^T \boldsymbol{\gamma} = 0$, since $\boldsymbol{A}_{\hat{\mathcal{I}}}$ is invertible. Therefore, we have $\boldsymbol{d}_k^T \boldsymbol{\beta} = 0$, and $\boldsymbol{d}_k^T \boldsymbol{\beta}^+ \neq 0$. It can only happen when $\boldsymbol{d}_k$ contains at least two nonzero elements with the same sign. So the algorithm returns $s = 0$. Q.E.D.

Proof for Theorem 13

We first prove the "if" direction by assuming that both conditions hold. Similar to the proof in Theorem 10, we can show the preservation of supermodularity. Furthermore, we consider any $\boldsymbol{t}', \boldsymbol{t}'' \in \mathcal{T}, \boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$ and any $\lambda \in (0,1)$. Since $\boldsymbol{x}_\lambda \in Conv(S_{t_\lambda})$, there exists an integer k such that $\boldsymbol{x}_\lambda = \sum_{i=1}^k \theta_i \boldsymbol{x}_i$, where $\sum_{i=1}^k \theta_i = 1, \theta_i \geq 0$ and $\boldsymbol{x}_i \in \mathcal{S}_{t_\lambda}, i = 1, \ldots, k$. Hence, we have

$$g(\mathcal{S}_{t_\lambda}) = \max_{\boldsymbol{x} \in \mathcal{S}_{t_\lambda}} \boldsymbol{a}^T \boldsymbol{x} \geq \sum_{i=1}^k \theta_i \boldsymbol{a}^T \boldsymbol{x}_i = \boldsymbol{a}^T \boldsymbol{x}_\lambda = \boldsymbol{a}^T (\lambda \boldsymbol{x}' + (1-\lambda)\boldsymbol{x}'') = \lambda \boldsymbol{a}^T \boldsymbol{x}' + (1-\lambda) \boldsymbol{a}^T \boldsymbol{x}''.$$

Since this inequality holds for every $\boldsymbol{x}' \in \mathcal{S}_{t'}, \boldsymbol{x}'' \in \mathcal{S}_{t''}$, we have that

$$g(\mathcal{S}_{t_\lambda}) \geq \lambda \max_{\boldsymbol{x} \in \mathcal{S}_{t'}} \boldsymbol{a}^T \boldsymbol{x}' + (1-\lambda) \max_{\boldsymbol{x} \in \mathcal{S}_{t''}} \boldsymbol{a}^T \boldsymbol{x}'' = \lambda g(\boldsymbol{t}') + (1-\lambda) g(\boldsymbol{t}'').$$

Equivalently, the function $g(\boldsymbol{t})$ is concave on $\boldsymbol{t} \in \mathcal{T}$.

We next prove the "only if" direction. Suppose condition 1 does not hold. Similar to the proof of Theorem 10, we can construct a counter example leading to a contradiction. Suppose condition 2 does not hold, i.e., $\exists(\boldsymbol{x}', \boldsymbol{t}'), (\boldsymbol{x}'', \boldsymbol{t}'') \in \mathcal{S}$ and $\lambda \in (0,1)$ such that $\boldsymbol{x}_\lambda \notin Conv(\mathcal{S}_{t_\lambda})$. By the separating hyperplane theorem, there exist a vector $\boldsymbol{\eta}$ and a scalar $\gamma \in \Re$ such that $\boldsymbol{\eta}^T \boldsymbol{x}_\lambda > \gamma > \boldsymbol{\eta}^T \boldsymbol{w}\ \forall \boldsymbol{w} \in Conv(\mathcal{S}_{t_\lambda})$. Let $f(\boldsymbol{x}) = \boldsymbol{\eta}^T \boldsymbol{x}$. We then have

$$g(\boldsymbol{t}_\lambda) = \max_{\boldsymbol{x} \in \mathcal{S}_{t_\lambda}} \boldsymbol{a}^T \boldsymbol{x} \le \max_{\boldsymbol{x} \in Conv(\mathcal{S}_{t_\lambda})} \boldsymbol{a}^T \boldsymbol{x} < \gamma < \boldsymbol{\eta}^T \boldsymbol{x}_\lambda = \lambda \boldsymbol{\eta}^T \boldsymbol{x}' + (1-\lambda) \boldsymbol{\eta}^T \boldsymbol{x}'' \le \lambda g(\boldsymbol{t}') + (1-\lambda) g(\boldsymbol{t}'').$$

Hence, g is not concave. Q.E.D.

Proof for Proposition 5

Let function $q'(\gamma) \in \Re^3$ and $q'(\alpha) \ge q(\alpha), q'(\beta) \ge q(\beta)$, and $\mathcal{G}$ be the graph with parameters $q(\alpha), q(\beta)$ on the arcs α, β, respectively. We denote the graphs $\mathcal{G}^\alpha$ with $q(\alpha)$ replaced by $q'(\alpha)$, $\mathcal{G}^\beta$ with $q(\beta)$ replaced by $q'(\beta)$ and $\mathcal{G}^{\alpha\beta}$ with both replacements. The quantities $\underline{c}^\alpha$, $\overline{c}^\alpha$, $\underline{c}^\beta$, $\overline{c}^\beta$, $\underline{c}^{\alpha\beta}$, $\overline{c}^{\alpha\beta}$ are defined correspondingly. To prove supermodularity of μ, by Theorem 10, it suffices to prove that for any $\boldsymbol{x}^\alpha, \boldsymbol{x}^\beta$ feasible on $\mathcal{G}^\alpha, \mathcal{G}^\beta$, respectively, we can find $\boldsymbol{y}, \boldsymbol{z}$ feasible on $\mathcal{G}, \mathcal{G}^{\alpha\beta}$, respectively, such that $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}^\alpha + \boldsymbol{x}^\beta$.

Case 1. $x^\alpha(\beta) \ge \underline{c}^\beta(\beta)$ and $x^\beta(\beta) \le \overline{c}(\beta)$. In this case, we have

$$\underline{c}^{\alpha\beta}(\alpha) = \underline{c}^\alpha(\alpha) \le x^\alpha(\alpha) \le \overline{c}^\alpha(\alpha) = \overline{c}^{\alpha\beta}(\alpha), \qquad \underline{c}^{\alpha\beta}(\beta) = \underline{c}^\beta(\beta) \le x^\alpha(\beta) \le \overline{c}(\beta) \le \overline{c}^{\alpha\beta}(\beta),$$

$$\underline{c}(\alpha) = \underline{c}^\beta(\alpha) \le x^\beta(\alpha) \le \overline{c}^\beta(\alpha) = \overline{c}(\alpha), \qquad \underline{c}(\beta) \le \underline{c}^\beta(\beta) \le x^\beta(\beta) \le \overline{c}(\beta).$$

Hence, $\boldsymbol{x}^\alpha$ is feasible on $\mathcal{G}^{\alpha\beta}$ and $\boldsymbol{x}^\beta$ is feasible on $\mathcal{G}$. We just choose $\boldsymbol{y} = \boldsymbol{x}^\beta$, $\boldsymbol{z} = \boldsymbol{x}^\alpha$.

Case 2. $x^\beta(\alpha) \ge \underline{c}^\alpha(\alpha)$ and $x^\alpha(\alpha) \le \overline{c}(\alpha)$. It can be proved symmetrically with *Case 1.*

Case 3. Neither *Case 1* nor *Case 2* is true. We define a circulation $\boldsymbol{v} = \boldsymbol{x}^\alpha - \boldsymbol{x}^\beta$. Since *Case 1* is false, we get either $x^\alpha(\beta) < \underline{c}^\beta(\beta)$ or $x^\beta(\beta) > \overline{c}(\beta)$, and in either case

$$v(\beta) = x^\alpha(\beta) - x^\beta(\beta) < 0.$$

Similarly, as *Case 2* is false, we have either $x^\beta(\alpha) < \underline{c}^\alpha(\alpha)$ or $x^\alpha(\alpha) > \overline{c}(\alpha)$, and in either case

$$v(\alpha) = x^\alpha(\alpha) - x^\beta(\alpha) > 0.$$

By Lemma 10, there exist circulations $\boldsymbol{v}^\alpha, \boldsymbol{v}^\beta$ such that

$$\boldsymbol{v} = \boldsymbol{v}^\alpha + \boldsymbol{v}^\beta, \tag{EC.20}$$

$$v^\alpha(\gamma) \cdot v^\beta(\gamma) \geq 0, \qquad \forall \gamma \in \mathcal{A}, \tag{EC.21}$$

$$v^\alpha(\beta) = v^\beta(\alpha) = 0,$$

$$v^\alpha(\alpha) = v(\alpha) - v^\beta(\alpha) = v(\alpha) = x^\alpha(\alpha) - x^\beta(\alpha),$$

$$v^\beta(\beta) = v(\beta) - v^\alpha(\beta) = v(\beta) = x^\alpha(\beta) - x^\beta(\beta).$$

We first define two circulations $\boldsymbol{y} = \boldsymbol{x}^\alpha - \boldsymbol{v}^\alpha$, $\boldsymbol{z} = \boldsymbol{x}^\beta + \boldsymbol{v}^\alpha$. It is easy to observe that $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}^\alpha + \boldsymbol{x}^\beta$, hence, $\boldsymbol{y}, \boldsymbol{z}$ are feasible on $\mathcal{G}, \mathcal{G}^{\alpha\beta}$, respectively.

For any arc $\gamma \in \mathcal{A} \backslash \{\alpha, \beta\}$, following equations (EC.20) and (EC.21), $v^\alpha(\gamma)$ is with the same sign as $v(\gamma)$ and has no larger absolute value than $v(\gamma)$. If $v(\gamma) \geq 0$, we have $0 \leq v^\alpha(\gamma) \leq v(\gamma)$, and hence

$$y(\gamma) = x^\alpha(\gamma) - v^\alpha(\gamma) \in [x^\alpha(\gamma) - v(\gamma), x^\alpha(\gamma)] = \left[x^\beta(\gamma), x^\alpha(\gamma)\right],$$

$$z(\gamma) = x^\beta(\gamma) + v^\alpha(\gamma) \in \left[x^\beta(\gamma), x^\beta(\gamma) + v(\gamma)\right] = \left[x^\beta(\gamma), x^\alpha(\gamma)\right].$$

If $v(\gamma) \leq 0$, we have $v(\gamma) \leq v^\alpha(\gamma) \leq 0$, and hence

$$y(\gamma) = x^\alpha(\gamma) - v^\alpha(\gamma) \in [x^\alpha(\gamma), x^\alpha(\gamma) - v(\gamma)] = \left[x^\alpha(\gamma), x^\beta(\gamma)\right],$$

$$z(\gamma) = x^\beta(\gamma) + v^\alpha(\gamma) \in \left[x^\beta(\gamma) + v(\gamma), x^\beta(\gamma)\right] = \left[x^\alpha(\gamma), x^\beta(\gamma)\right].$$

In either case, we get $y(\gamma), z(\gamma) \in [\underline{c}(\gamma), \overline{c}(\gamma)] = [\underline{c}^{\alpha\beta}(\gamma), \overline{c}^{\alpha\beta}(\gamma)]$.

Moreover, we observe that

$$y(\alpha) = x^\alpha(\alpha) - v^\alpha(\alpha) = x^\beta(\alpha) \in [\underline{c}(\alpha), \overline{c}(\alpha)],$$

$$y(\beta) = x^\alpha(\beta) - v^\alpha(\beta) = x^\alpha(\beta) \in [\underline{c}(\beta), \overline{c}(\beta)],$$

$$z(\alpha) = x^\beta(\alpha) + v^\alpha(\alpha) = x^\alpha(\alpha) \in \left[\underline{c}^{\alpha\beta}(\alpha), \overline{c}^{\alpha\beta}(\alpha)\right],$$

$$z(\beta) = x^\beta(\beta) + v^\alpha(\beta) = x^\beta(\beta) \in \left[\underline{c}^{\alpha\beta}(\beta), \overline{c}^{\alpha\beta}(\beta)\right].$$

Therefore, $\boldsymbol{y}$ is feasible on $\mathcal{G}$, and $\boldsymbol{z}$ is feasible on $\mathcal{G}^{\alpha\beta}$. Q.E.D.

Proof for Theorem 14

WLOG, let $i=1$, $j=2$, i.e., we prove the condition on the supermodularity with respect to the order-up-to level of components 1 and 2. Correspondingly, $\boldsymbol{Q}$ is the $(k+1)\times k$ submatrix obtained from $\boldsymbol{A}$ by deleting any $(n-k)$ columns, and any $(m-(k+1))$ rows except the first two rows.

We first prove the "if" part by using Theorem 11. To this end, the first step is to convert the function g in (y_1,y_2) to the format in Theorem 11 as follows.

$$\begin{aligned} g_{12}(y_1,y_2) \;=\; -\textstyle\sum_{i=3}^m h_i y_i + \max \;\; & \left(\boldsymbol{r}+\boldsymbol{p}+\boldsymbol{A}^T\boldsymbol{h}\right)^T \boldsymbol{v} - \textstyle\sum_{i=1}^2 h_i y_i - \boldsymbol{p}^T\boldsymbol{d} \\ & \boldsymbol{A}^o\boldsymbol{v}+\boldsymbol{B}^o\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \le \boldsymbol{c}^o, \end{aligned}$$

where

$$\boldsymbol{A}^o=\begin{bmatrix}\boldsymbol{A}\\ \boldsymbol{I}_n\\ -\boldsymbol{I}_n\end{bmatrix},\quad \boldsymbol{B}^o=\begin{bmatrix}-\boldsymbol{I}_2\\ \boldsymbol{0}_{(2n+m-2)\times 2}\end{bmatrix},\quad \boldsymbol{c}^o=\begin{bmatrix}\boldsymbol{0}_{2\times 1}\\ y_3\\ \vdots\\ y_m\\ \boldsymbol{d}\\ \boldsymbol{0}_{n\times 1}\end{bmatrix},$$

and we use $\boldsymbol{0}_{i\times j}$ to represent the zero matrix with size $i\times j$.

We now prove that $\boldsymbol{A}^o$, $\boldsymbol{B}^o$ satisfy the condition in Theorem 11 and hence the "if" part can be completed. Observe that $\text{rank}(\boldsymbol{A}^o)=n<m+2n$. Hence, it suffices to prove that for any $\mathcal{I}\subset\{1,\dots,2n+m\}$ with $|\mathcal{I}|=n+1$ and $\text{rank}(\boldsymbol{A}^o_{\mathcal{I}})=n$, and for any $\boldsymbol{\beta}\in\Re^2$ satisfying $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}\in C(\boldsymbol{A}^o_{\mathcal{I}})$, we must have $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}^+\in C(\boldsymbol{A}^o_{\mathcal{I}})$.

The case that $\{1,2\}\not\subset\mathcal{I}$ is trivial, since in that case, $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}^+$ is either $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}$ or $\boldsymbol{0}$, both of which are in $C(\boldsymbol{A}^o_{\mathcal{I}})$. Hence, it remains to consider the case with $\{1,2\}\subset\mathcal{I}$.

We represent $\mathcal{I}$ by a union of two disjoint subsets, $\mathcal{I}=\mathcal{I}_1\cup\mathcal{I}_2$, where $\mathcal{I}_1\subseteq\{1,\dots,m\}$, $\mathcal{I}_2\subseteq\{m+1,\dots,2n+m\}$. That is, $\mathcal{I}_1$ contains the indexes of rows from $\boldsymbol{A}$, and $\mathcal{I}_2$ contains the indexes of other rows. In addition, let $k+1=|\mathcal{I}_1|\le\min\{m,n+1\}$, and $|\mathcal{I}_2|=|\mathcal{I}|-|\mathcal{I}_1|=n-k$.

- Consider the case that $\text{rank}(\boldsymbol{A}^o_{\mathcal{I}_2}) \le n-k-1$, i.e., there exits a row in $\boldsymbol{A}^o_{\mathcal{I}_2}$ such that it is linearly dependent on other rows. By the structure of $\boldsymbol{A}^o_{\mathcal{I}_2}$, we can find $i,j \in \mathcal{I}_2$ such that $\boldsymbol{a}^o_i = -\boldsymbol{a}^o_j$. Together with $\text{rank}(\boldsymbol{A}^o_{\mathcal{I}}) = n$, we have $\text{rank}(\boldsymbol{A}^o_{\mathcal{I}\setminus\{j\}}) = n$. Hence, we can always find $\boldsymbol{\alpha} \in \Re^n$ such that $\boldsymbol{A}^o_{\mathcal{I}\setminus\{j\}}\boldsymbol{\alpha} = \boldsymbol{B}^o_{\mathcal{I}\setminus\{j\}}\boldsymbol{\beta}^+$. Furthermore, since $\boldsymbol{b}^o_i = \boldsymbol{b}^o_j = \boldsymbol{0}$, we get $(\boldsymbol{b}^o_j)^T\boldsymbol{\beta}^+ = 0 = (\boldsymbol{b}^o_i)^T\boldsymbol{\beta}^+ = (\boldsymbol{a}^o_i)^T\boldsymbol{\alpha} = (\boldsymbol{a}^o_j)^T\boldsymbol{\alpha}$, where the last equality follows from $\boldsymbol{a}^o_i = -\boldsymbol{a}^o_j$ and $(\boldsymbol{a}^o_i)^T\boldsymbol{\alpha} = 0$. Therefore, we have $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}^+ = \boldsymbol{A}^o_{\mathcal{I}}\boldsymbol{\alpha} \in C(\boldsymbol{A}^o_{\mathcal{I}})$.
- Consider the case that $\text{rank}(\boldsymbol{A}^o_{\mathcal{I}_2}) = n-k$. By $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta} \in C(\boldsymbol{A}^o_{\mathcal{I}})$ we can find $\boldsymbol{\alpha} \in \Re^n$ such that $\boldsymbol{A}^o_{\mathcal{I}}\boldsymbol{\alpha} = \boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}$. For all $j \in \mathcal{I}_2$, we define the index $s(j) \in \{1,\dots,n\}$ such that $\boldsymbol{a}^o_j$ is either $\boldsymbol{e}_{s(j)}$ or $-\boldsymbol{e}_{s(j)}$. Since $\text{rank}(\boldsymbol{A}^o_{\mathcal{I}_2}) = n-k$, $s(j), j \in \mathcal{I}_2$ take distinct values. WLOG, let $\{s(j) : j \in \mathcal{I}_2\} = \{k+1,\dots,n\}$. In addition, for all $j \in \mathcal{I}_2$, we have $\alpha_{s(j)} = 0$ since $\boldsymbol{b}^o_j = \boldsymbol{0}$ and hence $0 = |(\boldsymbol{b}^o_j)^T\boldsymbol{\beta}| = |(\boldsymbol{a}^o_j)^T\boldsymbol{\alpha}| = |\alpha_{s(j)}|$. Therefore, $\alpha_j = 0$ for all $j \ge k+1$, and $\boldsymbol{A}^o_{\mathcal{I}_1}\boldsymbol{\alpha} = \bar{\boldsymbol{A}}^o_{\mathcal{I}_1}\bar{\boldsymbol{\alpha}}$ where $\bar{\boldsymbol{A}}^o$ refers to the matrix consisting of the first k columns of $\boldsymbol{A}^o$, and $\bar{\boldsymbol{\alpha}}$ refers to the vector consisting of the first k elements of $\boldsymbol{\alpha}$. Furthermore, since $\text{rank}(\boldsymbol{A}^o_{\mathcal{I}}) = n$, all columns in $\bar{\boldsymbol{A}}^o_{\mathcal{I}}$ are linearly independent. It implies that all columns in $\bar{\boldsymbol{A}}^o_{\mathcal{I}_1}$ are linearly independent since $\forall j \in \mathcal{I}_2$, $\boldsymbol{a}^o_j$ is either $\boldsymbol{e}_s$ or $(-\boldsymbol{e}_s)$ for some $s \ge k+1$ and hence $\bar{\boldsymbol{A}}^o_{\mathcal{I}_2} = \boldsymbol{0}$. Therefore, $\text{rank}(\bar{\boldsymbol{A}}^o_{\mathcal{I}_1}) = k$. Recall that $|\mathcal{I}_1| = k+1$. Hence, we can find $\lambda_i, i \in \mathcal{I}_1$ such that $\sum_{i\in\mathcal{I}_1}\lambda_i\bar{\boldsymbol{a}}^o_i = \boldsymbol{0}$ and $\lambda_i, i \in \mathcal{I}_1$ are not all zero.

—If at least one of λ_1 and λ_2 is nonzero, WLOG, let $\lambda_1 \ne 0$ and normalize it to $\lambda_1 = -1$. We then have $\bar{\boldsymbol{a}}^o_1 = \sum_{i\in\mathcal{I}_1\setminus\{1\}}\lambda_i\bar{\boldsymbol{a}}^o_i$. Notice that $\forall i \in \mathcal{I}_1 \setminus \{1,2\}$, $(\bar{\boldsymbol{a}}^o_i)^T\bar{\boldsymbol{\alpha}} = (\boldsymbol{a}^o_i)^T\boldsymbol{\alpha} = (\boldsymbol{b}^o_i)^T\boldsymbol{\beta} = 0$ since $\boldsymbol{b}^o_i = \boldsymbol{0}$, we have

$$-\beta_1 = (\boldsymbol{b}^o_1)^T\boldsymbol{\beta} = (\boldsymbol{a}^o_1)^T\boldsymbol{\alpha} = (\bar{\boldsymbol{a}}^o_1)^T\bar{\boldsymbol{\alpha}} = \sum_{i\in\mathcal{I}_1\setminus\{1\}}\lambda_i(\bar{\boldsymbol{a}}^o_i)^T\bar{\boldsymbol{\alpha}} = \lambda_2(\bar{\boldsymbol{a}}^o_2)^T\bar{\boldsymbol{\alpha}} = \lambda_2(\boldsymbol{a}^o_2)^T\boldsymbol{\alpha} = \lambda_2(\boldsymbol{b}^o_2)^T\boldsymbol{\beta} = -\lambda_2\beta_2.$$

Note that the matrix $\boldsymbol{A}$ satisfies the condition stated in Theorem 14. Consider its submatrix $\bar{\boldsymbol{A}}^o_{\mathcal{I}_1} \in \Re^{(k+1)\times k}$ whose rank is k. By the condition stated in Theorem 14, we have $\lambda_1\lambda_2 \le 0$. As $\lambda_1 = -1$, we have $\lambda_2 \ge 0$. Therefore, β_1 and β_2 have the same sign, which implies that$\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}^+$ is either $\boldsymbol{0}$ or $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}$, and always in $C(\boldsymbol{A}^o_{\mathcal{I}})$.

—If $\lambda_1=\lambda_2=0$, we can find $i\in\mathcal{I}_1\setminus\{1,2\}$ such that $\lambda_i\neq 0$ as $\lambda_j,\ j\in\mathcal{I}_1$ are not all zero. That implies $rank(\bar{\boldsymbol{A}}^o_{\mathcal{I}_1\setminus\{i\}})=k$ and hence we must have $\bar{\boldsymbol{\gamma}}\in\Re^k$ with $\bar{\boldsymbol{A}}^o_{\mathcal{I}_1\setminus\{i\}}\bar{\boldsymbol{\gamma}}=\boldsymbol{B}^o_{\mathcal{I}_1\setminus\{i\}}\boldsymbol{\beta}^+$. Normalizing $\lambda_i=-1$, we have $\bar{\boldsymbol{a}}^o_i=\sum_{j\in\mathcal{I}_1\setminus\{i\}}\lambda_j\bar{\boldsymbol{a}}^o_j=\sum_{j\in\mathcal{I}_1\setminus\{1,2,i\}}\lambda_j\bar{\boldsymbol{a}}^o_j$ since $\lambda_1=\lambda_2=0$, and hence

$$(\bar{\boldsymbol{a}}^o_i)^T\bar{\boldsymbol{\gamma}}=\sum_{j\in\mathcal{I}_1\setminus\{1,2,i\}}\lambda_j(\bar{\boldsymbol{a}}^o_j)^T\bar{\boldsymbol{\gamma}}=\sum_{j\in\mathcal{I}_1\setminus\{1,2,i\}}\lambda_j(\boldsymbol{b}^o_j)^T\boldsymbol{\beta}^+=0=(\boldsymbol{b}^o_i)^T\boldsymbol{\beta}^+,$$

where the last two equalities follow from $\boldsymbol{b}^o_j=\boldsymbol{0}\ \forall j\in\mathcal{I}_1\setminus\{1,2\}$. Hence, we have $\bar{\boldsymbol{A}}^o_{\mathcal{I}_1}\bar{\boldsymbol{\gamma}}=\boldsymbol{B}^o_{\mathcal{I}_1}\boldsymbol{\beta}^+$. Choosing $\boldsymbol{\gamma}\in\Re^n$ such that $\boldsymbol{\gamma}^T=(\bar{\gamma}_1,\ldots,\bar{\gamma}_k,0,\ldots,0)$, we have $\boldsymbol{A}^o_{\mathcal{I}_1}\boldsymbol{\gamma}=\bar{\boldsymbol{A}}^o_{\mathcal{I}_1}\bar{\boldsymbol{\gamma}}=\boldsymbol{B}^o_{\mathcal{I}_1}\boldsymbol{\beta}^+$. In addition, since $\bar{\boldsymbol{A}}^o_{\mathcal{I}_2}=\boldsymbol{0}$ and $\boldsymbol{B}^o_{\mathcal{I}_2}=\boldsymbol{0}$, $\boldsymbol{A}^o_{\mathcal{I}_2}\boldsymbol{\gamma}=\bar{\boldsymbol{A}}^o_{\mathcal{I}_2}\bar{\boldsymbol{\gamma}}=\boldsymbol{0}=\boldsymbol{B}^o_{\mathcal{I}_2}\boldsymbol{\beta}^+$. Therefore, we have $\boldsymbol{B}^o_{\mathcal{I}}\boldsymbol{\beta}^+=\boldsymbol{A}^o_{\mathcal{I}}\boldsymbol{\gamma}\in C(\boldsymbol{A}^o_{\mathcal{I}})$.

We now prove the "only if" part by contradiction. Suppose $\boldsymbol{A}$ does not satisfy the condition in the theorem, i.e., there exists a submatrix $\boldsymbol{Q}$ with rank$(\boldsymbol{Q})=k$ and $\boldsymbol{\lambda}\in\Re^{k+1}$ with $\lambda_1\lambda_2>0$ such that $\sum_{i=1}^{k+1}\lambda_i\boldsymbol{q}_i=0$. We will show that the function $g(\boldsymbol{y},\boldsymbol{d})$ is not supermodular. WLOG, normalize $\lambda_1=-1$. Hence $\boldsymbol{q}_1=\sum_{i=2}^{k+1}\lambda_i\boldsymbol{q}_i$, where $\lambda_2<0$. WLOG, let $\boldsymbol{Q}$ be the submatrix in the upper left of $\boldsymbol{A}$, i.e., it is obtained from $\boldsymbol{A}$ by deleting the last $(m-(k+1))$ rows and last $(n-k)$ columns.

Since rank$(\boldsymbol{Q})=k$, $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ cannot be both linearly dependent on $\boldsymbol{q}_3,\ldots,\boldsymbol{q}_{k+1}$. WLOG, let $\boldsymbol{q}_2$ be linearly independent from $\boldsymbol{q}_3,\ldots,\boldsymbol{q}_{k+1}$. Therefore, we can find $\boldsymbol{v}^P\in\Re^k$ such that it is perpendicular to $\boldsymbol{q}_3,\ldots,\boldsymbol{q}_{k+1}$ but not $\boldsymbol{q}_2$. WLOG, let $\boldsymbol{q}_2^T\boldsymbol{v}^P>0$. We can find $\boldsymbol{v}^o\in\Re^k_+$ such that both $\boldsymbol{v}^o+\boldsymbol{v}^P\geq\boldsymbol{0}$ and $\boldsymbol{v}^o+2\boldsymbol{v}^P\geq\boldsymbol{0}$. Define $\boldsymbol{v}',\boldsymbol{v}''\in\Re^n_+$ as vectors with $v'_i=v^o_i+v^P_i, v''_i=v^o_i+2v^P_i$ for any $i\leq k$, and $v'_i=v''_i=0$ for any $i>k$. We get $\forall i=3,\ldots,k+1$, $\boldsymbol{a}_i^T\boldsymbol{v}'=\boldsymbol{a}_i^T\boldsymbol{v}''=\boldsymbol{q}_i^T\boldsymbol{v}^o$. Denote $\boldsymbol{C}$ as the lower left submatrix of $\boldsymbol{A}$ obtained by deleting the first $(k+1)$ rows and the last $(n-k)$ columns. We can choose $\boldsymbol{s}'\in\Re^{m-(k+1)}_+$ appropriately such that

$$\boldsymbol{s}''=\boldsymbol{s}'-\boldsymbol{C}\boldsymbol{v}^P\in\Re^{m-(k+1)}_+.$$

Define $\boldsymbol{y}',\boldsymbol{y}''\in\Re^m_+$ as

$$\boldsymbol{y}'=\boldsymbol{A}\boldsymbol{v}'+\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{s}'\end{bmatrix}=\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{A}_{k+1} & \boldsymbol{A}_{k+2} & \cdots & \boldsymbol{A}_n\\ \boldsymbol{C} & & & & \end{bmatrix}\begin{bmatrix}\boldsymbol{v}^o+\boldsymbol{v}^P\\ \boldsymbol{0}\end{bmatrix}+\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{s}'\end{bmatrix}=\begin{bmatrix}\boldsymbol{Q}(\boldsymbol{v}^o+\boldsymbol{v}^P)\\ \boldsymbol{C}(\boldsymbol{v}^o+\boldsymbol{v}^P)+\boldsymbol{s}'\end{bmatrix},$$

$$\boldsymbol{y}''=\boldsymbol{A}\boldsymbol{v}''+\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{s}''\end{bmatrix}=\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{A}_{k+1} & \boldsymbol{A}_{k+2} & \cdots & \boldsymbol{A}_n\\ \boldsymbol{C} & & & & \end{bmatrix}\begin{bmatrix}\boldsymbol{v}^o+2\boldsymbol{v}^P\\ \boldsymbol{0}\end{bmatrix}+\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{s}''\end{bmatrix}=\begin{bmatrix}\boldsymbol{Q}(\boldsymbol{v}^o+2\boldsymbol{v}^P)\\ \boldsymbol{C}(\boldsymbol{v}^o+2\boldsymbol{v}^P)+\boldsymbol{s}''\end{bmatrix}.$$

Then $\boldsymbol{y}'$ and $\boldsymbol{y}''$ have the same elements except the first two. Here we have

$$y_1' = \boldsymbol{q}_1^T(\boldsymbol{v}^o + \boldsymbol{v}^P) = \sum_{i=2}^{k+1} \lambda_i \boldsymbol{q}_i^T(\boldsymbol{v}^o + \boldsymbol{v}^P) = \sum_{i=2}^{k+1} \lambda_i \boldsymbol{q}_i^T \boldsymbol{v}^o + \lambda_2 \boldsymbol{q}_2^T \boldsymbol{v}^P > \sum_{i=2}^{k+1} \lambda_i \boldsymbol{q}_i^T \boldsymbol{v}^o + 2\lambda_2 \boldsymbol{q}_2^T \boldsymbol{v}^P = y_1'',$$

$$y_2' = \boldsymbol{q}_2^T \boldsymbol{v}^o + \boldsymbol{q}_2^T \boldsymbol{v}^P < \boldsymbol{q}_2^T \boldsymbol{v}^o + 2\boldsymbol{q}_2^T \boldsymbol{v}^P = y_2'',$$

where the inequalities follow from $\boldsymbol{q}_2^T \boldsymbol{v}^P > 0, \lambda_2 < 0$.

Choose $\boldsymbol{r} = \boldsymbol{p} = \boldsymbol{0}$, $\boldsymbol{h} = \sum_{i=1}^{k+1} \boldsymbol{e}_i$, and $\boldsymbol{d} = \boldsymbol{v}' \vee \boldsymbol{v}''$. We then have $g(\boldsymbol{y}, \boldsymbol{d}) \leq \boldsymbol{0}$ for all $\boldsymbol{y}$. Moreover, $g(\boldsymbol{y}', \boldsymbol{d}) = 0$ since we can choose $\boldsymbol{v} = \boldsymbol{v}'$, $\boldsymbol{u} = \begin{bmatrix} \boldsymbol{0} \\ \\ \boldsymbol{s}' \end{bmatrix}$. Similarly, $g(\boldsymbol{y}'', \boldsymbol{d}) = 0$.

We now prove by contradiction that $g(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d}) < 0$. Assume to the contrary that $\exists (\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}) \in \mathcal{P}_A(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d})$ with the first $(k+1)$ elements in $\boldsymbol{u}$ being zero. Since $\boldsymbol{v} \leq \boldsymbol{d} = \boldsymbol{v}' \vee \boldsymbol{v}''$, we have $v_i = 0 \ \forall i > k$. Hence, we need to determine the first k elements of $\boldsymbol{v}$, which we denote by $\hat{\boldsymbol{v}}$. As the first $(k+1)$ elements of $\boldsymbol{u}$ are zero, the first $(k+1)$ equations of $\boldsymbol{y}' \wedge \boldsymbol{y}'' = \boldsymbol{A}\boldsymbol{v} + \boldsymbol{u}$ are

$$y_i' \wedge y_i'' = \boldsymbol{a}_i^T \boldsymbol{v} = \boldsymbol{q}_i^T \hat{\boldsymbol{v}}, \quad i = 1, \ldots, k+1. \tag{EC.22}$$

Since $\boldsymbol{q}_2, \ldots, \boldsymbol{q}_{k+1}$ are linearly independent, the last k equations of (EC.22) has a unique solution, which is exactly $\hat{\boldsymbol{v}} = \boldsymbol{v}^o + \boldsymbol{v}^P$. This solution, unfortunately, does not satisfy the first equation. Hence, (EC.22) is infeasible.Therefore, for every $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}) \in \mathcal{P}_A(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d})$, the first $(k+1)$ elements of $\boldsymbol{u}$ cannot be all 0. As the set $\mathcal{P}_A(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d})$ is closed, the minimal value of $\boldsymbol{h}^T \boldsymbol{u} = \sum_{i=1}^{k+1} u_i$ subject to $(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}) \in \mathcal{P}_A(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d})$ is strictly positive. Hence, $g(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d}) < 0$,

$$g(\boldsymbol{y}', \boldsymbol{d}) + g(\boldsymbol{y}'', \boldsymbol{d}) = 0 > g(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d}) \geq g(\boldsymbol{y}' \wedge \boldsymbol{y}'', \boldsymbol{d}) + g(\boldsymbol{y}' \vee \boldsymbol{y}'', \boldsymbol{d}),$$

which implies that g is not supermodular in (y_1, y_2). Q.E.D.

Lemma 1 *$Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$ is a lattice.*

Proof. The statement is obvious when $\boldsymbol{x}'$ and $\boldsymbol{x}''$ are ordered. Now we just focus on the case that $\boldsymbol{x}'$ and $\boldsymbol{x}''$ are unordered. Since the order of dimension does not affect operations of convex combination, join, and meet, we assume that $\boldsymbol{x}' = (\boldsymbol{x}_1', \boldsymbol{x}_2')$ and $\boldsymbol{x}'' = (\boldsymbol{x}_1'', \boldsymbol{x}_2'')$ such that $\boldsymbol{x}_1' > \boldsymbol{x}_1''$ and $\boldsymbol{x}_2' \leq \boldsymbol{x}_2''$, $\boldsymbol{x}_2' \neq \boldsymbol{x}_2''$. Hence, we have

$$\boldsymbol{x}' \wedge \boldsymbol{x}'' = (\boldsymbol{x}_1'', \boldsymbol{x}_2'), \quad \boldsymbol{x}' \vee \boldsymbol{x}'' = (\boldsymbol{x}_1', \boldsymbol{x}_2'').$$

Denote $\mathcal{W} = Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'') = Conv((\boldsymbol{x}_1', \boldsymbol{x}_2'), (\boldsymbol{x}_1'', \boldsymbol{x}_2''), (\boldsymbol{x}_1'', \boldsymbol{x}_2'), (\boldsymbol{x}_1', \boldsymbol{x}_2''))$.

If $\boldsymbol{y}', \boldsymbol{y}'' \in \mathcal{W}$, there exist $\lambda_i \geq 0$, $\mu_i \geq 0$, $i = 1, 2, 3, 4$, such that $\sum_{i=1}^{4} \lambda_i = \sum_{i=1}^{4} \mu_i = 1$, and

$$\boldsymbol{y}' = \lambda_1 \boldsymbol{x}' + \lambda_2 \boldsymbol{x}'' + \lambda_3 (\boldsymbol{x}' \wedge \boldsymbol{x}'') + \lambda_4 (\boldsymbol{x}' \vee \boldsymbol{x}'') = ((\lambda_1 + \lambda_4)\boldsymbol{x}_1' + (\lambda_2 + \lambda_3)\boldsymbol{x}_1'', (\lambda_1 + \lambda_3)\boldsymbol{x}_2' + (\lambda_2 + \lambda_4)\boldsymbol{x}_2''),$$

$$\boldsymbol{y}'' = \mu_1 \boldsymbol{x}' + \mu_2 \boldsymbol{x}'' + \mu_3 (\boldsymbol{x}' \wedge \boldsymbol{x}'') + \mu_4 (\boldsymbol{x}' \vee \boldsymbol{x}'') = ((\mu_1 + \mu_4)\boldsymbol{x}_1' + (\mu_2 + \mu_3)\boldsymbol{x}_1'', (\mu_1 + \mu_3)\boldsymbol{x}_2' + (\mu_2 + \mu_4)\boldsymbol{x}_2'').$$

Take the meet and get

$$\begin{aligned} \boldsymbol{y}' \wedge \boldsymbol{y}'' = (\quad & ((\lambda_1 + \lambda_4)\boldsymbol{x}_1' + (\lambda_2 + \lambda_3)\boldsymbol{x}_1'') \wedge ((\mu_1 + \mu_4)\boldsymbol{x}_1' + (\mu_2 + \mu_3)\boldsymbol{x}_1''), \\ & ((\lambda_1 + \lambda_3)\boldsymbol{x}_2' + (\lambda_2 + \lambda_4)\boldsymbol{x}_2'') \wedge ((\mu_1 + \mu_3)\boldsymbol{x}_2' + (\mu_2 + \mu_4)\boldsymbol{x}_2'') \qquad). \end{aligned}$$

WLOG, we assume that $\lambda_1 + \lambda_4 \geq \mu_1 + \mu_4$. Since $\sum_{i=1}^{4} \lambda_i = \sum_{i=1}^{4} \mu_i = 1$ and $\boldsymbol{x}_1' > \boldsymbol{x}_1''$, we have

$$((\lambda_1 + \lambda_4)\boldsymbol{x}_1' + (\lambda_2 + \lambda_3)\boldsymbol{x}_1'') \wedge ((\mu_1 + \mu_4)\boldsymbol{x}_1' + (\mu_2 + \mu_3)\boldsymbol{x}_1'') = (\mu_1 + \mu_4)\boldsymbol{x}_1' + (\mu_2 + \mu_3)\boldsymbol{x}_1''.$$

If $\lambda_1 + \lambda_3 \leq \mu_1 + \mu_3$, we can similarly get

$$((\lambda_1 + \lambda_3)\boldsymbol{x}_2' + (\lambda_2 + \lambda_4)\boldsymbol{x}_2'') \wedge ((\mu_1 + \mu_3)\boldsymbol{x}_2' + (\mu_2 + \mu_4)\boldsymbol{x}_2'') = (\mu_1 + \mu_3)\boldsymbol{x}_2' + (\mu_2 + \mu_4)\boldsymbol{x}_2'',$$

and $\boldsymbol{y}' \wedge \boldsymbol{y}'' = \boldsymbol{y}'' \in \mathcal{W}$. If $\lambda_1 + \lambda_3 > \mu_1 + \mu_3$, we have

$$((\lambda_1 + \lambda_3)\boldsymbol{x}_2' + (\lambda_2 + \lambda_4)\boldsymbol{x}_2'') \wedge ((\mu_1 + \mu_3)\boldsymbol{x}_2' + (\mu_2 + \mu_4)\boldsymbol{x}_2'') = (\lambda_1 + \lambda_3)\boldsymbol{x}_2' + (\lambda_2 + \lambda_4)\boldsymbol{x}_2'',$$

then

$$\boldsymbol{y}' \wedge \boldsymbol{y}'' = ((\mu_1 + \mu_4)\boldsymbol{x}_1' + (\mu_2 + \mu_3)\boldsymbol{x}_1'', (\lambda_1 + \lambda_3)\boldsymbol{x}_2' + (\lambda_2 + \lambda_4)\boldsymbol{x}_2'').$$

Let $m_1 = \min\{\lambda_1 + \lambda_3, \mu_1 + \mu_4\}$, and

$$\begin{aligned} m_2 &= m_1 - (\lambda_1 + \lambda_3) + \mu_2 + \mu_3, \\ m_3 &= -m_1 + \lambda_1 + \lambda_3, \\ m_4 &= -m_1 + \mu_1 + \mu_4. \end{aligned}$$

We can check that $m_i \in [0,1]$ for any $i = 1,2,3,4$, $\sum_{i=1}^{4} m_i = 1$, and

$$\boldsymbol{y}' \wedge \boldsymbol{y}'' = m_1(\boldsymbol{x}_1', \boldsymbol{x}_2') + m_2(\boldsymbol{x}_1'', \boldsymbol{x}_2'') + m_3(\boldsymbol{x}_1'', \boldsymbol{x}_2') + m_4(\boldsymbol{x}_1', \boldsymbol{x}_2'')$$

and we get $\boldsymbol{y}' \wedge \boldsymbol{y}'' \in \mathcal{W}$.

Similarly, we can show that $\boldsymbol{y}' \vee \boldsymbol{y}'' \in Conv(\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}' \wedge \boldsymbol{x}'', \boldsymbol{x}' \vee \boldsymbol{x}'')$. Q.E.D.

Lemma 2 *For any closed sets $\mathcal{A}, \mathcal{H} \subseteq \Re^n$ with $\mathcal{A} \cap \mathcal{H} \neq \emptyset$ and $\boldsymbol{c} \in \Re^n$, there exists a $K \geq 0$ such that*

$$\max_{\boldsymbol{x}\in\mathcal{A}} f(\boldsymbol{x}) = \max_{\boldsymbol{x}\in\mathcal{A}\cap\mathcal{H}} f(\boldsymbol{x}),$$

where the function $f : \mathcal{A} \to \Re$ is defined as

$$f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} - K \min_{\boldsymbol{h}\in\mathcal{H}} ||\boldsymbol{x} - \boldsymbol{h}||_1.$$

Proof. By optimality, we observe that for any $K > 0$,

$$\max_{\boldsymbol{x}\in\Re}\left\{\boldsymbol{c}^T\boldsymbol{x} - K\min_{\boldsymbol{h}\in\mathcal{H}}||\boldsymbol{x}-\boldsymbol{h}||_1 - K\min_{\boldsymbol{h}\in\mathcal{A}}||\boldsymbol{x}-\boldsymbol{h}||_1\right\} \geq \max_{\boldsymbol{x}\in\mathcal{A}}\left\{\boldsymbol{c}^T\boldsymbol{x} - K\min_{\boldsymbol{h}\in\mathcal{H}}||\boldsymbol{x}-\boldsymbol{h}||_1\right\} \geq \max_{\boldsymbol{x}\in\mathcal{A}\cap\mathcal{H}}\left\{\boldsymbol{c}^T\boldsymbol{x}\right\}. \tag{EC.23}$$

Note that both sets $\mathcal{A}, \mathcal{H} \in \Re^n$ are closed sets and $\mathcal{A} \cap \mathcal{H} \neq \emptyset$. According to Proposition 1.5.3 in Bertsekas (2015), there is a scalar $\bar{K} > 0$ such that for all $K \geq \bar{K}$,

$$\max_{\boldsymbol{x}\in\Re}\left\{\boldsymbol{c}^T\boldsymbol{x} - K\min_{\boldsymbol{h}\in\mathcal{H}}||\boldsymbol{x}-\boldsymbol{h}||_1 - K\min_{\boldsymbol{h}\in\mathcal{A}}||\boldsymbol{x}-\boldsymbol{h}||_1\right\} = \max_{\boldsymbol{x}\in\mathcal{A}\cap\mathcal{H}} \boldsymbol{c}^T\boldsymbol{x}. \tag{EC.24}$$

Based on relationships in (EC.23) and (EC.24), the lemma is proved. Q.E.D.

Lemma 3 *If $\boldsymbol{a}^T\boldsymbol{x}' + \boldsymbol{b}^T\boldsymbol{t}' = \boldsymbol{a}^T\boldsymbol{x}'' + \boldsymbol{b}^T\boldsymbol{t}'' = c$, $\max\{\boldsymbol{a}^T\boldsymbol{y} + \boldsymbol{b}^T(\boldsymbol{t}' \wedge \boldsymbol{t}''), \boldsymbol{a}^T\boldsymbol{z} + \boldsymbol{b}^T(\boldsymbol{t}' \vee \boldsymbol{t}'')\} \leq c$, $\boldsymbol{y} + \boldsymbol{z} = \boldsymbol{x}' + \boldsymbol{x}''$, we then have $\boldsymbol{a}^T\boldsymbol{y} = c - \boldsymbol{b}^T(\boldsymbol{t}' \wedge \boldsymbol{t}'')$, $\boldsymbol{a}^T\boldsymbol{z} = c - \boldsymbol{b}^T(\boldsymbol{t}' \vee \boldsymbol{t}'')$.*

Proof. We note that

$$\begin{aligned}
\boldsymbol{a}^T\boldsymbol{y} &= \boldsymbol{a}^T(\boldsymbol{x}' + \boldsymbol{x}'' - \boldsymbol{z}) + \boldsymbol{b}^T(\boldsymbol{t}' + \boldsymbol{t}'' - (\boldsymbol{t}' \vee \boldsymbol{t}'') - (\boldsymbol{t}' \wedge \boldsymbol{t}'')) \\
&= 2c - (\boldsymbol{a}^T\boldsymbol{z} + \boldsymbol{b}^T(\boldsymbol{t}' \vee \boldsymbol{t}'')) - \boldsymbol{b}^T(\boldsymbol{t}' \wedge \boldsymbol{t}'') \\
&\geq 2c - c - \boldsymbol{b}^T(\boldsymbol{t}' \wedge \boldsymbol{t}'') \\
&= c - \boldsymbol{b}^T(\boldsymbol{t}' \wedge \boldsymbol{t}''),
\end{aligned}$$

where the inequality follows from $\boldsymbol{a}^T\boldsymbol{z}+\boldsymbol{b}^T(\boldsymbol{t}'\vee\boldsymbol{t}'')\le c$. Since $\boldsymbol{a}^T\boldsymbol{y}+\boldsymbol{b}^T(\boldsymbol{t}'\wedge\boldsymbol{t}'')\le c$, we have $\boldsymbol{a}^T\boldsymbol{y}=c-\boldsymbol{b}^T(\boldsymbol{t}'\wedge\boldsymbol{t}'')$. Similarly, $\boldsymbol{a}^T\boldsymbol{z}=c-\boldsymbol{b}^T(\boldsymbol{t}'\vee\boldsymbol{t}'')$. Q.E.D.

Lemma 4 *Given $\boldsymbol{x}',\boldsymbol{x}''$, we assume that f is a concave and supermodular function on the set $\mathcal{W}=Conv(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}'')$. Consider any subset $\mathcal{N}\subseteq\mathcal{W}$, and $\boldsymbol{x}^o\in Conv(\mathcal{N})$ such that $\boldsymbol{x}^o=a\boldsymbol{x}'+b\boldsymbol{x}''+c(\boldsymbol{x}'\wedge\boldsymbol{x}'')+d(\boldsymbol{x}'\vee\boldsymbol{x}'')$ with $a,b,c,d\ge 0$, $a+b+c+d=1$, and $c\times d=0$, then we must have*

$$\max_{\boldsymbol{x}\in\mathcal{N}} f(\boldsymbol{x})\ge af(\boldsymbol{x}')+bf(\boldsymbol{x}'')+cf(\boldsymbol{x}'\wedge\boldsymbol{x}'')+df(\boldsymbol{x}'\vee\boldsymbol{x}'').$$

Proof. The case that $\boldsymbol{x}',\boldsymbol{x}''$ are ordered can be derived directly from Lemma 5. Now we just consider the case that $\boldsymbol{x}',\boldsymbol{x}''$ are unordered. Since $\boldsymbol{x}^o\in Conv(\mathcal{N})$, we have there exists $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\in\mathcal{N}$ and $\boldsymbol{\beta}\in\Re^n_+$ such that

$$\boldsymbol{x}^o=\sum_{i=1}^n\beta_i\boldsymbol{x}^i,\ \boldsymbol{\beta}^T\mathbf{1}=1,\ \boldsymbol{\beta}\ge\mathbf{0}.$$

Moreover, since $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\in\mathcal{N}\subseteq\mathcal{W}=Conv(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}'')$, we have

$$\begin{cases}\boldsymbol{x}_1=\alpha_1^1\boldsymbol{x}'+\alpha_2^1\boldsymbol{x}''+\alpha_3^1(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\alpha_4^1(\boldsymbol{x}'\vee\boldsymbol{x}''),\\ \vdots\\ \boldsymbol{x}_n=\alpha_1^n\boldsymbol{x}'+\alpha_2^n\boldsymbol{x}''+\alpha_3^n(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\alpha_4^n(\boldsymbol{x}'\vee\boldsymbol{x}''),\\ \boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4=\mathbf{1},\\ \boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4\ge\mathbf{0},\end{cases}$$

where $\boldsymbol{\alpha}_i$ is the column vector of $(\alpha_i^1,\ldots,\alpha_i^n)^T$. Therefore,

$$\boldsymbol{x}^o=\boldsymbol{\beta}^T\boldsymbol{\alpha}_1\boldsymbol{x}'+\boldsymbol{\beta}^T\boldsymbol{\alpha}_2\boldsymbol{x}''+\boldsymbol{\beta}^T\boldsymbol{\alpha}_3(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\boldsymbol{\beta}^T\boldsymbol{\alpha}_4(\boldsymbol{x}'\vee\boldsymbol{x}'')=a\boldsymbol{x}'+b\boldsymbol{x}''+c(\boldsymbol{x}'\wedge\boldsymbol{x}'')+d(\boldsymbol{x}'\vee\boldsymbol{x}'').$$

WLOG, let $\boldsymbol{x}'=(\boldsymbol{x}_1',\boldsymbol{x}_2')$, $\boldsymbol{x}''=(\boldsymbol{x}_1'',\boldsymbol{x}_2'')$ such that $\boldsymbol{x}_1'<\boldsymbol{x}_1''$, $\boldsymbol{x}_2'\ge\boldsymbol{x}_2''$ and $\boldsymbol{x}_2'\ne\boldsymbol{x}_2''$. Then we have $\boldsymbol{x}'\wedge\boldsymbol{x}''=(\boldsymbol{x}_1',\boldsymbol{x}_2'')$, $\boldsymbol{x}'\vee\boldsymbol{x}''=(\boldsymbol{x}_1'',\boldsymbol{x}_2')$, and

$$\begin{aligned}\mathbf{0}&=(\boldsymbol{\beta}^T\boldsymbol{\alpha}_1-a)\boldsymbol{x}'+(\boldsymbol{\beta}^T\boldsymbol{\alpha}_2-b)\boldsymbol{x}''+(\boldsymbol{\beta}^T\boldsymbol{\alpha}_3-c)(\boldsymbol{x}'\wedge\boldsymbol{x}'')+(\boldsymbol{\beta}^T\boldsymbol{\alpha}_4-d)(\boldsymbol{x}'\vee\boldsymbol{x}'')\\&=\Big(\big(\boldsymbol{\beta}^T\boldsymbol{\alpha}_1+\boldsymbol{\beta}^T\boldsymbol{\alpha}_3-a-c\big)\boldsymbol{x}_1'+\big(\boldsymbol{\beta}^T\boldsymbol{\alpha}_2+\boldsymbol{\beta}^T\boldsymbol{\alpha}_4-b-d\big)\boldsymbol{x}_1'',\\&\qquad\big(\boldsymbol{\beta}^T\boldsymbol{\alpha}_1+\boldsymbol{\beta}^T\boldsymbol{\alpha}_4-a-d\big)\boldsymbol{x}_2'+\big(\boldsymbol{\beta}^T\boldsymbol{\alpha}_2+\boldsymbol{\beta}^T\boldsymbol{\alpha}_3-b-c\big)\boldsymbol{x}_2''\Big)\\&=\Big(\big(\boldsymbol{\beta}^T\boldsymbol{\alpha}_1+\boldsymbol{\beta}^T\boldsymbol{\alpha}_3-a-c\big)(\boldsymbol{x}_1'-\boldsymbol{x}_1''),\big(\boldsymbol{\beta}^T\boldsymbol{\alpha}_1+\boldsymbol{\beta}^T\boldsymbol{\alpha}_4-a-d\big)(\boldsymbol{x}_2'-\boldsymbol{x}_2'')\Big),\end{aligned}$$

where the last equality holds since $\boldsymbol{\beta}^T(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4)=\boldsymbol{\beta}^T\mathbf{1}=1=a+b+c+d$. Hence, we have

$$\begin{cases}\boldsymbol{\beta}^T(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_3)=a+c,\\ \boldsymbol{\beta}^T(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_4)=a+d,\end{cases}\Longrightarrow\begin{cases}\boldsymbol{\beta}^T\boldsymbol{\alpha}_1=a+c-\boldsymbol{\beta}^T\boldsymbol{\alpha}_3,\\ \boldsymbol{\beta}^T\boldsymbol{\alpha}_2=b+c-\boldsymbol{\beta}^T\boldsymbol{\alpha}_3,\\ \boldsymbol{\beta}^T\boldsymbol{\alpha}_4=d-c+\boldsymbol{\beta}^T\boldsymbol{\alpha}_3.\end{cases}$$

Therefore, based on the concavity of function f, we can get

$$\begin{aligned}&\max\{f(\boldsymbol{x}_1),\ldots,f(\boldsymbol{x}_n)\}\\ \geq\ &\max\left\{\alpha_1^1f(\boldsymbol{x}')+\alpha_2^1f(\boldsymbol{x}'')+\alpha_3^1f(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\alpha_4^1f(\boldsymbol{x}'\vee\boldsymbol{x}''),\ldots,\right.\\ &\quad\left.\alpha_1^nf(\boldsymbol{x}')+\alpha_2^nf(\boldsymbol{x}'')+\alpha_3^nf(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\alpha_4^nf(\boldsymbol{x}'\vee\boldsymbol{x}'')\right\}\\ \geq\ &\beta_1(\alpha_1^1f(\boldsymbol{x}')+\alpha_2^1f(\boldsymbol{x}'')+\alpha_3^1f(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\alpha_4^1f(\boldsymbol{x}'\vee\boldsymbol{x}''))+\ldots\\ &+\beta_n(\alpha_1^nf(\boldsymbol{x}')+\alpha_2^nf(\boldsymbol{x}'')+\alpha_3^nf(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\alpha_4^nf(\boldsymbol{x}'\vee\boldsymbol{x}''))\\ =\ &\boldsymbol{\beta}^T\boldsymbol{\alpha}_1f(\boldsymbol{x}')+\boldsymbol{\beta}^T\boldsymbol{\alpha}_2f(\boldsymbol{x}'')+\boldsymbol{\beta}^T\boldsymbol{\alpha}_3f(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\boldsymbol{\beta}^T\boldsymbol{\alpha}_4f(\boldsymbol{x}'\vee\boldsymbol{x}'').\end{aligned}$$

Hence, we have

$$\begin{aligned}&\max\{f(\boldsymbol{x}_1),\ldots,f(\boldsymbol{x}_n)\}-(af(\boldsymbol{x}')+bf(\boldsymbol{x}'')+cf(\boldsymbol{x}'\wedge\boldsymbol{x}'')+df(\boldsymbol{x}'\vee\boldsymbol{x}''))\\ \geq\ &\left(\boldsymbol{\beta}^T\boldsymbol{\alpha}_1-a\right)f(\boldsymbol{x}')+\left(\boldsymbol{\beta}^T\boldsymbol{\alpha}_2-b\right)f(\boldsymbol{x}'')+\left(\boldsymbol{\beta}^T\boldsymbol{\alpha}_3-c\right)f(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\left(\boldsymbol{\beta}^T\boldsymbol{\alpha}_4-d\right)f(\boldsymbol{x}'\vee\boldsymbol{x}'')\\ =\ &\left(\boldsymbol{\beta}^T\boldsymbol{\alpha}_3-c\right)\left(f(\boldsymbol{x}'\wedge\boldsymbol{x}'')+f(\boldsymbol{x}'\vee\boldsymbol{x}'')-f(\boldsymbol{x}')-f(\boldsymbol{x}'')\right).\end{aligned}\tag{EC.25}$$

Recall that $c\times d=0$ and $\boldsymbol{\beta}^T\boldsymbol{\alpha}_4=d-c+\boldsymbol{\beta}^T\boldsymbol{\alpha}_3$. Therefore, we have $\boldsymbol{\beta}^T\boldsymbol{\alpha}_3-c\geq 0$. Together with the supermodularity of f, we know that the right hand side of the equation (EC.25) is non-negative. Hence,

$$\max_{\boldsymbol{x}\in\mathcal{N}}f(\boldsymbol{x})\geq\max\{f(\boldsymbol{x}_1),\ldots,f(\boldsymbol{x}_n)\}\geq af(\boldsymbol{x}')+bf(\boldsymbol{x}'')+cf(\boldsymbol{x}'\wedge\boldsymbol{x}'')+df(\boldsymbol{x}'\vee\boldsymbol{x}'').$$

Q.E.D.

Lemma 5 *Given* $\boldsymbol{x}_1,\boldsymbol{x}_2$*, we consider any concave function* $f(\boldsymbol{x})$ *defined on* $Conv(\boldsymbol{x}_1,\boldsymbol{x}_2)$*. Given* $\boldsymbol{x}_\lambda=\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2$*, where* $\lambda\in[0,1]$ *and* $\boldsymbol{y},\boldsymbol{z}\in Conv(\boldsymbol{x}_1,\boldsymbol{x}_2)$*, if* $\boldsymbol{x}_\lambda\in Conv(\boldsymbol{y},\boldsymbol{z})$*, then we must have* $\max\{f(\boldsymbol{y}),f(\boldsymbol{z})\}\geq\lambda f(\boldsymbol{x}_1)+(1-\lambda)f(\boldsymbol{x}_2)$*.*

Proof. Since $\boldsymbol{y},\boldsymbol{z}\in Conv(\boldsymbol{x}_1,\boldsymbol{x}_2)$, $\exists\beta,\mu\in[0,1]$ such that $\boldsymbol{y}=\beta\boldsymbol{x}_1+(1-\beta)\boldsymbol{x}_2$, $\boldsymbol{z}=\mu\boldsymbol{x}_1+(1-\mu)\boldsymbol{x}_2$. Moreover, as $\boldsymbol{x}_\lambda\in Conv(\boldsymbol{y},\boldsymbol{z})$, we get $\lambda\in Conv(\beta,\mu)$. Therefore,

$$\begin{aligned}\max\{f(\boldsymbol{y}),f(\boldsymbol{z})\} &= \max\{f(\beta\boldsymbol{x}_1+(1-\beta)\boldsymbol{x}_2),f(\mu\boldsymbol{x}_1+(1-\mu)\boldsymbol{x}_2)\}\\ &\geq \max\{\beta f(\boldsymbol{x}_1)+(1-\beta)f(\boldsymbol{x}_2),\mu f(\boldsymbol{x}_1)+(1-\mu)f(\boldsymbol{x}_2)\}\\ &\geq \lambda f(\boldsymbol{x}_1)+(1-\lambda)f(\boldsymbol{x}_2).\end{aligned}$$

Q.E.D.

Lemma 6 *For any unordered $\boldsymbol{x}',\boldsymbol{x}''$, if $\boldsymbol{y}+\boldsymbol{z}=\boldsymbol{x}'+\boldsymbol{x}''$ and $\boldsymbol{y},\boldsymbol{z}\in Conv(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}'\vee\boldsymbol{x}'')$ i.e., there exist $\lambda_i,\mu_i,\nu_i,\beta_i\in[0,1]$, $\lambda_i+\mu_i+\nu_i+\beta_i=1$ for $i\in\{1,2\}$ such that $\boldsymbol{y}=\lambda_1\boldsymbol{x}'+\mu_1\boldsymbol{x}''+\nu_1(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\beta_1(\boldsymbol{x}'\vee\boldsymbol{x}'')$, $\boldsymbol{z}=\lambda_2\boldsymbol{x}'+\mu_2\boldsymbol{x}''+\nu_2(\boldsymbol{x}'\wedge\boldsymbol{x}'')+\beta_2(\boldsymbol{x}'\vee\boldsymbol{x}'')$. We have*

$$\begin{aligned}&\mu_1+\nu_1+\mu_2+\nu_2=\lambda_1+\nu_1+\lambda_2+\nu_2=1,\\ &\beta_1+\beta_2=\nu_1+\nu_2.\end{aligned}$$

Proof. When $\boldsymbol{x}',\boldsymbol{x}''$ are unordered, we let $\boldsymbol{x}'=(\boldsymbol{x}'_1,\boldsymbol{x}'_2)$ and $\boldsymbol{x}''=(\boldsymbol{x}''_1,\boldsymbol{x}''_2)$ such that $\boldsymbol{x}'_1\geq\boldsymbol{x}''_1,\boldsymbol{x}'_1\neq\boldsymbol{x}''_1$ and $\boldsymbol{x}'_2<\boldsymbol{x}''_2$. Since $\boldsymbol{y},\boldsymbol{z}\in Conv(\boldsymbol{x}',\boldsymbol{x}'',\boldsymbol{x}'\wedge\boldsymbol{x}'',\boldsymbol{x}\vee\boldsymbol{x}'')$ and $\boldsymbol{x}'+\boldsymbol{x}''=\boldsymbol{x}'\wedge\boldsymbol{x}''+\boldsymbol{x}'\vee\boldsymbol{x}''$, we can get that

$$\begin{aligned}\boldsymbol{y}+\boldsymbol{z} &= (\lambda_1+\lambda_2)\boldsymbol{x}'+(\mu_1+\mu_2)\boldsymbol{x}''+(\nu_1+\nu_2)(\boldsymbol{x}'\wedge\boldsymbol{x}'')+(\beta_1+\beta_2)(\boldsymbol{x}'\vee\boldsymbol{x}'')\\ &= (\lambda_1+\lambda_2)\boldsymbol{x}'+(\mu_1+\mu_2)\boldsymbol{x}''+(\nu_1+\nu_2)(\boldsymbol{x}'+\boldsymbol{x}''-\boldsymbol{x}'\vee\boldsymbol{x}'')+(\beta_1+\beta_2)(\boldsymbol{x}'\vee\boldsymbol{x}'')\\ &= (\lambda_1+\lambda_2+\nu_1+\nu_2)\boldsymbol{x}'+(\mu_1+\mu_2+\nu_1+\nu_2)\boldsymbol{x}''+(\beta_1+\beta_2-\nu_1-\nu_2)(\boldsymbol{x}'\vee\boldsymbol{x}'')\\ &= (\lambda_1+\nu_1+\lambda_2+\nu_2)(\boldsymbol{x}'_1,\boldsymbol{x}'_2)+(\mu_1+\nu_1+\mu_2+\nu_2)(\boldsymbol{x}''_1,\boldsymbol{x}''_2)+(\beta_1-\nu_1+\beta_2-\nu_2)(\boldsymbol{x}'_1,\boldsymbol{x}''_2)\\ &= \Big(((\lambda_1+\nu_1+\lambda_2+\nu_2)+(\beta_1-\nu_1+\beta_2-\nu_2))\boldsymbol{x}'_1+(\mu_1+\nu_1+\mu_2+\nu_2)\boldsymbol{x}''_1,\\ &\qquad (\lambda_1+\nu_1+\lambda_2+\nu_2)\boldsymbol{x}'_2+((\mu_1+\nu_1+\mu_2+\nu_2)+(\beta_1-\nu_1+\beta_2-\nu_2))\boldsymbol{x}''_2\Big)\\ &= \Big((\lambda_1+\beta_1+\lambda_2+\beta_2)\boldsymbol{x}'_1+(\mu_1+\nu_1+\mu_2+\nu_2)\boldsymbol{x}''_1,\\ &\qquad (\lambda_1+\nu_1+\lambda_2+\nu_2)\boldsymbol{x}'_2+(\mu_1+\beta_1+\mu_1+\beta_2)\boldsymbol{x}''_2\Big).\end{aligned}$$

Since $\boldsymbol{y}+\boldsymbol{z}=\boldsymbol{x}'+\boldsymbol{x}''=(\boldsymbol{x}'_1+\boldsymbol{x}''_1,\boldsymbol{x}'_2+\boldsymbol{x}''_2)$, $\boldsymbol{x}'_1\geq\boldsymbol{x}''_1,\boldsymbol{x}'_1\neq\boldsymbol{x}''_1$, and $\boldsymbol{x}'_2<\boldsymbol{x}''_2$, we can easily get that

$$\mu_1+\nu_1+\mu_2+\nu_2=\lambda_1+\nu_1+\lambda_2+\nu_2=\lambda_1+\beta_1+\lambda_2+\beta_2=1.$$

In addition, $\beta_1+\beta_2=1-(\lambda_1+\lambda_2)=\nu_1+\nu_2$. Q.E.D.

Lemma 7 *The statement* **S2** *in Proposition 2 cannot hold if there exist* $j,k \in \{1,\dots,n_2\}, \boldsymbol{\alpha} \geq \mathbf{0}$ *such that* $\boldsymbol{B}_j, \boldsymbol{B}_k$ *are linearly independent, and* $\boldsymbol{A}\boldsymbol{\alpha} = \delta \boldsymbol{B}_j + \gamma \boldsymbol{B}_k$ *for some* $\delta\gamma < 0$.

Proof. We prove by contradiction. WLOG, assume $\boldsymbol{A}\boldsymbol{\alpha} = \delta \boldsymbol{B}_1 + \gamma \boldsymbol{B}_2$ with $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ linearly independent, and $\delta > 0, \gamma < 0$, $\boldsymbol{\alpha} \geq 0$. To show that **S2** is false, we choose $\boldsymbol{\beta} = \delta \boldsymbol{e}_1 + \gamma \boldsymbol{e}_2$. Then $\boldsymbol{A}\boldsymbol{\alpha} = \boldsymbol{B}\boldsymbol{\beta}$. If there exist $\lambda_1, \lambda_2 \in [0,1]$ such that $\boldsymbol{B}\boldsymbol{\beta}^+ = \boldsymbol{A}(\lambda_1 \boldsymbol{\alpha}^+ - \lambda_2(-\boldsymbol{\alpha})^+) = \lambda_1 \boldsymbol{A}\boldsymbol{\alpha} = \lambda_1 \boldsymbol{B}\boldsymbol{\beta}$, we have $\boldsymbol{B}\boldsymbol{\beta}^+ = \boldsymbol{B}_1\delta = \lambda_1 \boldsymbol{A}\boldsymbol{\alpha} = \lambda_1 \delta \boldsymbol{B}_1 + \lambda_1 \gamma \boldsymbol{B}_2$, i.e., $\delta(1-\lambda_1)\boldsymbol{B}_1 = \lambda_1 \gamma \boldsymbol{B}_2$. Since $\delta\gamma \neq 0$, it contradicts with the assumption that $\boldsymbol{B}_1, \boldsymbol{B}_2$ are linearly independent. Q.E.D.

Lemma 8 *Given a set* $\mathcal{A} \subseteq \Re^n$ *and a convex function* $f : \Re^n \to \Re$, *for any* $\boldsymbol{x} \in Conv(\mathcal{A})$, *we have* $f(\boldsymbol{x}) \leq \sup_{\boldsymbol{y} \in \mathcal{A}} f(\boldsymbol{y})$.

Proof. For any $\boldsymbol{x} \in Conv(\mathcal{A})$, we have $\boldsymbol{x} = \sum_{i=1}^m \lambda_i \boldsymbol{y}_i$ for certain $\boldsymbol{y}_i \in \mathcal{A}$, $\lambda_i \geq 0$, $i = 1, 2, \dots, m$ and $\sum_{i=1}^m \lambda_i = 1$. Hence,

$$f(\boldsymbol{x}) = f\left(\sum_{i=1}^m \lambda_i \boldsymbol{y}_i\right) \leq \sum_{i=1}^m \lambda_i f(\boldsymbol{y}_i) \leq \sup_{\boldsymbol{y} \in \mathcal{A}} f(\boldsymbol{y})$$

where the first inequality follows from the convexity of f. Q.E.D.

Lemma 9 *Consider any matrix* $\boldsymbol{A} \in \Re^{m \times n_2}$ *with* $n = Rank(\boldsymbol{A}) < m$. *Suppose that the system of inequalities* $\boldsymbol{A}\boldsymbol{x} \leq \bar{\boldsymbol{b}}$, $-\boldsymbol{A}\boldsymbol{x} \leq -\underline{\boldsymbol{b}}$ *is infeasible. We can choose* $\mathcal{I} \subseteq \{1,\dots,m\}$ *with* $|\mathcal{I}| = n+1$ *and* $Rank(\boldsymbol{A}_{\mathcal{I}}) = n$ *such that the system of inequalities* $\boldsymbol{A}_{\mathcal{I}}\boldsymbol{x} \leq \bar{\boldsymbol{b}}_{\mathcal{I}}$, $-\boldsymbol{A}_{\mathcal{I}}\boldsymbol{x} \leq -\underline{\boldsymbol{b}}_{\mathcal{I}}$ *has no solutions.*

Proof. The lemma is similar to the result in Bertsimas and Tsitsiklis (1997, Exercise 4.29) and can be proved in the same way. Hence, we only sketch the outline of the proof as follows. We can use the given system of inequalities to construct an infeasible linear optimization problem with the trivial objective function $\mathbf{0}^T\boldsymbol{x}$. It has an unbounded dual problem. We can reduce the dimension of the dual problem by keeping only a basis and a corresponding feasible direction, which allow the objective function goes to be unbounded. Getting the dual of the new unbounded problem, we can get the desired infeasible system of inequalities. Q.E.D.

Lemma 10 *If α and β are in series and $\boldsymbol{x}$ is a circulation on $\mathcal{G}$ and $x(\alpha)>0, x(\beta)<0$, then there exist circulations $\boldsymbol{x}^\alpha$ and $\boldsymbol{x}^\beta$ such that*

$$\boldsymbol{x}=\boldsymbol{x}^\alpha+\boldsymbol{x}^\beta$$

$$x^\alpha(\gamma)\cdot x^\beta(\gamma)\geq 0, \qquad \forall\gamma\in\mathcal{A}$$

$$x^\alpha(\beta)=x^\beta(\alpha)=0.$$

Proof. From the Cyclic Decomposition Lemma (see, for instance, Gale and Politof 1981), we can decompose $\boldsymbol{x}$ as $\boldsymbol{x}=\sum_{i=1}^n k_i\boldsymbol{x}_{\Gamma_i}$ for some $k_i>0\ \forall i=1,\dots,n$, and cycles $\boldsymbol{x}_{\Gamma_i}$ (please see the definition of cycle in Gale and Politof (1981)) such that $x_{\Gamma_i}(\gamma)x(\gamma)\geq 0,\ \forall\gamma\in\mathcal{A}$. WLOG, assume that for some r, $x_{\Gamma_i}(\alpha)>0$ if $i\leq r$ and $x_{\Gamma_i}(\alpha)=0$ if $i>r$. Define $\boldsymbol{x}^\alpha=\sum_{i=1}^r k_i\boldsymbol{x}_{\Gamma_i}$, $\boldsymbol{x}^\beta=\sum_{i=r+1}^n k_i\boldsymbol{x}_{\Gamma_i}$. We then have $\boldsymbol{x}=\boldsymbol{x}^\alpha+\boldsymbol{x}^\beta$. In addition, $\forall\gamma\in\mathcal{A}$, we note that $k_ix_{\Gamma_i}(\gamma)$ is with the same sign as $x(\gamma)$, hence,

$$x^\alpha(\gamma)\cdot x^\beta(\gamma)=\left(\sum_{i=1}^r k_ix_{\Gamma_i}(\gamma)\right)\cdot\left(\sum_{i=r+1}^n k_ix_{\Gamma_i}(\gamma)\right)\geq 0.$$

Moreover, by the way we construct r, $x^\beta(\alpha)=\sum_{i=r+1}^n x_{\Gamma_i}(\alpha)=0$. Consider any $i=1,\dots,r$, since $x_{\Gamma_i}(\beta)$ has the same sign as $x(\beta)$, we have $x_{\Gamma_i}(\beta)\leq 0$. Since α,β are in series, $x_{\Gamma_i}(\beta)\geq 0$ is implied from $x_{\Gamma_i}(\alpha)>0$. Therefore, we have $x_{\Gamma_i}(\beta)=0$, and hence $x^\alpha(\beta)=\sum_{i=1}^r k_ix_{\Gamma_{i}(\beta)}=0$. Q.E.D.

Lemma 11 *Consider any matrix $\boldsymbol{Q}\in\Re_+^{(k+1)\times k}$ with rank$(\boldsymbol{Q})=k$, $k\geq 2$. If every 3×2 submatrix of $\boldsymbol{Q}$ contains at least two row vectors which are linearly dependent, then $\boldsymbol{Q}$ has at least two row vectors which are linearly dependent.*

Proof. In this proof, we use "dependent/independent" to represent "linearly dependent/independent", and say a matrix "satisfies 3-2 condition" to represent that its every 3×2 submatrix would contain at least two row vectors which are dependent.

We prove by induction. When $k=2$, the argument in the lemma is true since $\boldsymbol{Q}\in\Re_+^{3\times 2}$ satisfies 3-2 condition. Suppose the argument in the lemma is true when $k=r-1$, $r\in\{3,4,\dots\}$. We now prove the case for $k=r$ by considering such $\boldsymbol{Q}\in\Re_+^{(r+1)\times r}$.

We denote the submatrix consisting of the first $(r-1)$ columns of $\boldsymbol{Q}$ by $\hat{\boldsymbol{Q}} \in \Re_+^{(r+1)\times(r-1)}$. We have $rank(\hat{\boldsymbol{Q}}) = r-1$ since all columns in it are independent. Hence, $\hat{\boldsymbol{Q}}$ has $(r-1)$ independent rows. WLOG, let $\hat{\boldsymbol{q}}_1, \hat{\boldsymbol{q}}_2, \ldots, \hat{\boldsymbol{q}}_{r-1}$ be independent and $\hat{\boldsymbol{Q}}^1 \in \Re_+^{r\times(r-1)}$ be the submatrix of $\hat{\boldsymbol{Q}}$ by deleting the last row, i.e., the $(r+1)$th row. Since we assume the statement in the lemma is true for the case of $k = r-1$, there are two dependent rows in $\hat{\boldsymbol{Q}}^1$, i.e., $\hat{\boldsymbol{q}}_i, \hat{\boldsymbol{q}}_j$ are dependent for some distinct indexes $i, j \in \{1, \ldots, r\}$. Since $\hat{\boldsymbol{q}}_1, \hat{\boldsymbol{q}}_2, \ldots, \hat{\boldsymbol{q}}_{r-1}$ are independent, the two dependent rows cannot be both in the first $(r-1)$ rows. Therefore, we have $\hat{\boldsymbol{q}}_r = \lambda \hat{\boldsymbol{q}}_i$ for some $i \in \{1, \ldots, r-1\}$, $\lambda \geq 0$. WLOG, let $\hat{\boldsymbol{q}}_r = \lambda \hat{\boldsymbol{q}}_1$. Similarly, by letting $\hat{\boldsymbol{Q}}^2 \in \Re_+^{r\times(r-1)}$ be the submatrix of $\hat{\boldsymbol{Q}}$ by deleting only the second to the last row, i.e., the rth row, we can conclude $\hat{\boldsymbol{q}}_{r+1} = \delta \hat{\boldsymbol{q}}_i$ for some $i \in \{1, \ldots, r-1\}$, $\delta \geq 0$. If $\lambda = \delta = 0$, then $\hat{\boldsymbol{q}}_r = \hat{\boldsymbol{q}}_{r+1} = \boldsymbol{0}$, and hence $\boldsymbol{q}_r$, $\boldsymbol{q}_{r+1}$ must be dependent and the argument in the lemma is true for this case. Now, it suffices to consider the case where at least one of λ, δ is nonzero. WLOG, we assume $\lambda > 0$, and normalize it to $\lambda = 1$, i.e., $\hat{\boldsymbol{q}}_r = \hat{\boldsymbol{q}}_1$. If $q_{1r} = q_{rr}$, then $\boldsymbol{q}_1 = \boldsymbol{q}_r$, the statement in the lemma is true for this case since the rows 1 and r are dependent. Hence, it suffices to consider the case where $q_{1r} \neq q_{rr}$. WLOG, let $q_{1r} = 1$, and $q_{rr} \neq 1$. Furthermore, as a nonzero vector, $\hat{\boldsymbol{q}}_1$ contains at least one nonzero element; hence, WLOG, we assume $q_{11} \neq 0$ and normalize it to $q_{11} = 1$. In summary, we have $\hat{\boldsymbol{q}}_r = \hat{\boldsymbol{q}}_1$, $\hat{\boldsymbol{q}}_{r+1} = \delta \hat{\boldsymbol{q}}_i$, $q_{11} = q_{r1} = q_{1r} = 1$, $q_{rr} \neq 1$.

We first consider the case of $i = 1$, which implies $\hat{\boldsymbol{q}}_{r+1} = \delta \hat{\boldsymbol{q}}_1$. We end up with a 3×2 submatrix of $\boldsymbol{Q}$ by deleting all rows except the rows $1, r, r+1$ and deleting all columns except columns $1, r$:

$$\begin{bmatrix} 1 & 1 \\ 1 & q_{rr} \\ \delta & q_{r+1,r} \end{bmatrix}.$$

Since $\boldsymbol{Q}$ satisfies 3-2 condition, the above submatrix contains at least two rows which are dependent. Moreover, as $q_{rr} \neq 1$, we have $q_{r+1,r}$ is either δ or δq_{rr}. While the former results in $\boldsymbol{q}_{r+1} = \delta \boldsymbol{q}_1$, the latter leads to $\boldsymbol{q}_{r+1} = \delta \boldsymbol{q}_r$.

We now consider the case of $i \neq 1$. WLOG, let $i = 2$, i.e., $\hat{\boldsymbol{q}}_{r+1} = \delta \hat{\boldsymbol{q}}_2$. We discuss three possible scenarios.

- If $\delta = 0$, then we end up with a 3×2 submatrix of $\boldsymbol{Q}$ by deleting all rows except rows $1, r, r+1$ and deleting all columns except columns $1, r$:

$$\begin{bmatrix} 1 & 1 \\ 1 & q_{rr} \\ 0 & q_{r+1,r} \end{bmatrix}.$$

As $\boldsymbol{Q}$ satisfies 3-2 condition and $q_{rr} \neq 1$, we have $q_{r+1,r} = 0$. Hence, $\boldsymbol{q}_{r+1} = \mathbf{0}$.

- If $\delta \neq 0$ and $q_{21} = 0$, we end up with a 4×2 submatrix of $\boldsymbol{Q}$ by deleting all rows except rows $1, 2, r, r+1$ and deleting all columns except columns $1, r$:

$$\begin{bmatrix} 1 & 1 \\ 0 & q_{2r} \\ 1 & q_{rr} \\ 0 & q_{r+1,r} \end{bmatrix}.$$

As $\boldsymbol{Q}$ satisfies 3-2 condition and $q_{rr} \neq 1$, we have $q_{2r} = q_{r+1,r} = 0$. Hence, $\boldsymbol{q}_{r+1} = \delta \boldsymbol{q}_2$.

- If $\delta \cdot q_{21} \neq 0$, we normalize $q_{21} = \delta = 1$. Hence, $\hat{\boldsymbol{q}}_{r+1} = \hat{\boldsymbol{q}}_2$. We end up with a 4×2 submatrix of $\boldsymbol{Q}$ by deleting all rows except rows $1, 2, r, r+1$ and deleting all columns except columns $1, r$:

$$\boldsymbol{Q}^o = \begin{bmatrix} 1 & 1 \\ 1 & q_{2r} \\ 1 & q_{rr} \\ 1 & q_{r+1,r} \end{bmatrix}.$$

We now show $q_{r+1,r} = q_{2r}$ by contradiction. Assume to the contrary that $q_{r+1,r} \neq q_{2r}$, i.e., $\boldsymbol{q}_2^o, \boldsymbol{q}_4^o$ are independent. We notice that $\boldsymbol{Q}$ satisfies 3-2 condition and $\boldsymbol{q}_1^o, \boldsymbol{q}_3^o$ are independent since $q_{rr} \neq 1$. Therefore, we have either 1) $q_{2r} = 1$ ($\boldsymbol{q}_1^o$ and $\boldsymbol{q}_2^o$ are dependent), $q_{r+1,r} = q_{rr}$ ($\boldsymbol{q}_3^o$ and $\boldsymbol{q}_4^o$ are dependent), or 2) $q_{r+1,r} = 1$ ($\boldsymbol{q}_1^o$ and $\boldsymbol{q}_4^o$ are dependent), $q_{2r} = q_{rr}$ ($\boldsymbol{q}_2^o$ and $\boldsymbol{q}_3^o$ are dependent). WLOG, we consider the former case. Since $\hat{\boldsymbol{q}}_1, \hat{\boldsymbol{q}}_2$ are independent, we can find $i \neq 1$ such that (q_{11}, q_{21}) and

(q_{1i}, q_{2i}) are independent. WLOG, let $i = 2$. In this case, we end up with a 4×2 submatrix of $\boldsymbol{Q}$ by deleting all rows except rows $1, 2, r, r+1$ and deleting all columns except columns $2, r$:

$$\boldsymbol{Q}^* = \begin{bmatrix} q_{12} & 1 \\ q_{22} & 1 \\ q_{12} & q_{rr} \\ q_{22} & q_{rr} \end{bmatrix}.$$

Let $(\boldsymbol{q}_i^*)^T$ represent the ith row vector of $\boldsymbol{Q}^*$. Since (q_{12}, q_{22}) and $(q_{11}, q_{21}) = (1, 1)$ are independent, we have $q_{12} \neq q_{22}$ and $\boldsymbol{q}_1^*$ and $\boldsymbol{q}_2^*$ are independent. In addition, $\boldsymbol{q}_1^*$ and $\boldsymbol{q}_3^*$ are independent since $q_{rr} \neq 1$. Therefore, as $\boldsymbol{Q}$ satisfies 3-2 condition, $\boldsymbol{q}_2^*$ and $\boldsymbol{q}_3^*$ are dependent, i.e.,

$$q_{12} = q_{22} q_{rr}. \tag{EC.26}$$

Similarly, due to the independence between $\boldsymbol{q}_1^*$ and $\boldsymbol{q}_2^*$ and that between $\boldsymbol{q}_2^*$ and $\boldsymbol{q}_4^*$, we have that $\boldsymbol{q}_1^*$ and $\boldsymbol{q}_4^*$ are dependent, i.e.,

$$q_{12} q_{rr} = q_{22}. \tag{EC.27}$$

If one of q_{12} and q_{22} is 0, the equalities (EC.26) and (EC.27) imply $q_{12} = q_{22} = 0$, which contradicts with $q_{12} \neq q_{22}$. If $q_{12} q_{22} \neq 0$, the equalities (EC.26) and (EC.27) imply $q_{rr}^2 = 1$, which contradicts with $q_{rr} \neq 1$ and $q_{rr} \geq 0$. Hence, we have $q_{r+1,r} = q_{2r}$, which implies $\boldsymbol{q}_{r+1} = \boldsymbol{q}_2$.

Therefore, in all scenarios, the statement in the lemma is true for $k = r$. The proof is complete. Q.E.D.