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, xinchennl@illinois.edu

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

Jin Qi  
Department of Industrial Engineering and Decision Analytics, 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, allowing us to derive complementarity among parameters and monotonic structural properties for optimal policies in many operational models. We introduce the new concepts of mostly-sublattice and additive mostly-sublattice which generalize the commonly imposed sublattice condition significantly, and use them to establish the necessary and sufficient conditions for the feasible set so that supermodularity can be preserved under various assumptions about the objective functions. Further, we identify some classes of polyhedral sets which satisfy these concepts. Finally, we illustrate the use of our results in 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 fields of economics and operations research. It is closely related to the concept of complementarity in economics and has strong economic implications (see Topkis 1998). For example, in a production problem where a firm produces products from available resources, supermodularity of the profit function in available resources implies these resources are complementary in the sense that the marginal profit in one
resource increases as the quantities of the other resources increase. Supermodularity has also proved to be an important tool for deriving monotonic 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). Indeed, when deriving monotonic structural properties of optimal policies in Markovian decision processes, supermodularity is usually the first concept that comes to mind, and it is often necessary to determine whether supermodularity can be preserved within dynamic programming recursions, a series of parametric optimization problems. Several commonly used general results have been established for this purpose (e.g., Topkis 1998, Heyman and Sobel 2003). However, these classic results may not be directly applicable when their underlying assumptions are violated, leading to scattered results due to case-by-case approaches (Chen et al. 2013).

The purpose of this paper is to present a systematic study of parametric optimization problems and provide characterizations of constraint structures which preserve supermodularity with various assumptions on the objective functions. We consider the following parametric optimization problem.

$$g(t) = \max \{ f(x, t) : x \in S_t \}.$$  

Here $f : \mathcal{X} \times \mathcal{T} \to \mathbb{R}$, $\mathcal{X}, \mathcal{T}$ are sublattices of corresponding Euclidean spaces, and $S_t = \{ x \in \mathcal{X} : (x, t) \in S \}$, where $S \subseteq \mathcal{X} \times \mathcal{T}$ is the graph of the constraint set. In Markovian decision process applications (e.g., dynamic inventory control), Equation (1) can be the dynamic recursion where $g(t)$ is the profit-to-go function, with $t$ being the state variables.

To put our research in perspective, we first review some classical results from Topkis (1998) which establish conditions under which supermodularity can be preserved in the optimization operation (1). To present the result, we denote the projection of $S$ on $\mathcal{T}$ by $\Pi_\mathcal{T}S$.

**Theorem 1** (Theorem 2.7.6, Topkis 1998) *If $S$ is a sublattice of $\mathcal{X} \times \mathcal{T}$, $f$ is supermodular on $\mathcal{X} \times \mathcal{T}$, and $g$ is finite on $\Pi_\mathcal{T}S$, then $g$ is supermodular on $\Pi_\mathcal{T}S$.***
Although powerful and widely used in the literature, Topkis’ result imposes the sublattice 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 demonstrate this, consider a common case where $\mathcal{S}$ is defined by linear inequalities:

$$\mathcal{S}_c^\rho = \{(x, t): Ax + Bt \leq c\}$$

for some nonzero matrices $A \in \mathbb{R}^{m \times n_1}, B \in \mathbb{R}^{m \times n_2}$, and a vector $c \in \mathbb{R}^m$.

**Theorem 2 (Page 26, Topkis 1998)** The polyhedron $\mathcal{S}_c^\rho$ defined by Equation (2) is a sublattice of $\mathbb{R}^{n_1+n_2}$ for each $c \in \mathbb{R}^m$ 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 $i$th rows of matrices $A$ and $B$, 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 theorem above is equivalent to saying that the matrix $(A|B)$ concatenating $A$ and $B$ is a lattice-matrix. While many operational problems (for example, the assemble-to-order (ATO) problem which will be discussed in detail in Section 4) have their feasible sets structured as $\mathcal{S}_c^\rho$, the lattice-matrix requirement is often not satisfied. In this case, Theorems 1 and 2 cannot be applied directly to establish the supermodularity of $g$. This renders it necessary to relax the sublattice requirement.

As we show in the next section, this may not always be fruitful if we merely impose supermodularity on $f$. Indeed, to relax the sublattice requirement, we must impose stronger assumptions on $f$, which fortunately are commonly seen in operational applications. For example, $f$ may be linear, concave, or independent of $t$ in its formulation. We want to address questions such as under what conditions on $\mathcal{S}$, for any function $f$ with some given properties (e.g., supermodularity or linearity), $g$ has the corresponding properties (e.g., supermodularity).

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 \( t \);

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

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

Our results illustrate that if (a) supermodularity or (b) concavity and supermodularity are imposed on \( f \), we require \( \mathcal{S} \) to be a mostly-sublattice, which is defined in Definition 1 in Section 2. For (c), (d), (e) and (f), our results require \( \mathcal{S} \) to satisfy certain additive mostly-sublattice conditions, which are defined in Definition 2 in Section 3.

We then focus on \( \mathcal{S} \) with a polyhedral structure (2) and derive conditions for \( A \) and \( B \) such that for any \( c \), \( S_c^{\mathcal{S}} \) satisfies the conditions identified. To avoid trivial cases, we assume that \( A, B \) have no zero columns, \( A \) has no zero rows, and \( n_2 \geq 2 \). Interestingly, we provide a complete characterization of the structures of \( A \) and \( B \) in several settings. Specifically, for (a) and (b), \( A \) and \( B \) are matrices such that the corresponding \( S_c^{\mathcal{S}} \) should be a sublattice. We give a detailed description for (c) in Section 3. For (d), (e) and (f), the key requirement is that given any \( \beta, B \beta^+ \) should be in the column space of \( A \) if \( B \beta \) is in the column space of \( A \).

Table 1 provides an overview of our results for preservation under various settings. The first column lists the properties of \( f \), and the second column lists the properties preserved by \( g \). The last column lists the conditions of the preservation results (under different assumptions on \( \mathcal{S} \)). Throughout the paper, the meet and join operations are defined with respect to Euclidean space,
<table>
<thead>
<tr>
<th>Properties of $f$</th>
<th>Properties of $g$</th>
<th>Set $S$</th>
<th>Convex $S$</th>
<th>Polyhedron $S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supermodular</td>
<td>Supermodular</td>
<td>mostly-sublattice</td>
<td>mostly-sublattice</td>
<td>$(A</td>
</tr>
<tr>
<td>Concave &amp; supermodular</td>
<td>Concave &amp; supermodular</td>
<td>mostly-sublattice and convex</td>
<td>mostly-sublattice</td>
<td>$(A</td>
</tr>
<tr>
<td>Supermodular &amp; independent of $t$</td>
<td>Supermodular</td>
<td>additive mostly-sublattice with $W^1$, $x_\lambda \in Conv(W^2 \cap S_L)$</td>
<td>additive mostly-sublattice with $W^2$, $x_\lambda \in Conv(W^2 \cap S_L)$</td>
<td>Theorem 7</td>
</tr>
<tr>
<td>Concave &amp; supermodular &amp; independent of $t$</td>
<td>Concave &amp; supermodular</td>
<td>additive mostly-sublattice with $W^2$, $x_\lambda \in Conv(W^2 \cap S_L)$</td>
<td>additive mostly-sublattice with $W^3$, $x_\lambda \in Conv(S_L)$</td>
<td>Theorem 9</td>
</tr>
<tr>
<td>Linear</td>
<td>Concave &amp; supermodular</td>
<td>additive mostly-sublattice with $W^3$, $x_\lambda \in Conv(S_L)$</td>
<td>additive mostly-sublattice with $W^3$, $x_\lambda \in Conv(S_L)$</td>
<td>Theorem 11</td>
</tr>
</tbody>
</table>

\[
x_\lambda = \lambda x' + (1-\lambda)x'', \quad t_\lambda = \lambda t' + (1-\lambda)t'', \quad \lambda \in [0,1];
\]

\[
W^1(x',x'') = \{x',x'',x' \wedge x'',x' \vee x''\}, \quad W^2(x',x'') = Conv(x',x'',x' \wedge x'',x' \vee x''), \quad W^3(x',x'') = X, \forall (x',x'') \in X
\]

Table 1  A summary for the preservation result.

Moreover, $Conv(\cdot)$ denotes the convex hull.

As an application, we use our results to analyze an ATO system where the number of outputs (i.e., products) is a linear mapping of the inputs (i.e., components). We derive an efficient method to check any given ATO system whether a subset of components are complementary in the sense that the profit margin of one component’s inventory increases with the other components’ inventories. In particular, we show that for the entire ATO system to have such a property, whenever two products share common components their use of all the components in common must be in the same proportion. This type of system often appears in settings where products mainly differ in size, and the usage of common components is proportional to the size of product.

There are very few papers that derive supermodularity preservation properties when the graph of the constraint set $S$ is not a sublattice. Gale and Politof (1981) and Granot and Veinott (1985) derive supermodularity preservation properties for network flow optimization problems. Zipkin
(2003) focuses on an ATO 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 of 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 of supermodularity preservation without sublattice conditions. Our necessary and sufficient condition for preserving supermodularity covers many of the theoretical tools in Gale and Politof (1981), Granot and Veinott (1985), Zipkin (2003) and Chen et al. (2013) as special cases. A closely related paper is Quah (2007), which shares certain flavor with this paper. In particular, Quah (2007) considers similar parametric optimization problems. However, he focuses on identifying conditions of the feasible sets under which the optimal solutions are monotonic for any concave and supermodular objective functions.

The organization of the paper is as follows. 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 the parameter $t$. We then pay particular attention to the case when $f$ is linear. We apply our results to the ATO system in Section 4. In Section 5, we discuss possible extensions to log-supermodularity and conclude the paper. For the sake of readability, all proofs are relegated to the appendix.

**Notation and convention:** Vectors and matrices are represented by lower- and upper-case boldface characters, respectively. Given any vector $x \in \mathbb{R}^n$, we assume it is a column vector unless otherwise specified and denote its $i^{th}$ element by $x_i$. We use parentheses to construct row vectors where commas separate each element as $x^T = (x_1, \ldots, x_n)$, and deem $x \leq y$ if and only if $x_i \leq y_i$ for all $i$. The $i$th unit vector, whose components are zero except the $i$th component being one, is denoted as $e_i$. For a matrix $A = (a_{ij})_{i=1,\ldots,m;j=1,\ldots,n} \in \mathbb{R}^{m \times n}$, define $a_i^T$ and $A_j$ as its $i$th row vector
and $j$th column vector, respectively. We denote by $A_{\mathcal{I}}$ the submatrix consisting of all rows $a_i^T$, $i \in \mathcal{I} \subseteq \{1, \ldots, m\}$, and denote by $C(A)$ the column space of $A$. Given any two matrices $A$ and $B$ with the same number of rows, their concatenated matrix is denoted by $(A|B)$. The positive part of $x$ is denoted by $x^+ = \max\{x, 0\}$.

2. Main Results for General Problems

In this section, we focus on the general parametric optimization problem (1). We assume that $\mathcal{S}$ is a nonempty closed set, and the maximization problem in Equation (1) is well defined for all $t \in \mathcal{T}$ with nonempty $\mathcal{S}_t$. Furthermore, if $\Pi_T \mathcal{S}$ is a chain (i.e., does not contain any unordered pairs of elements), $g$ is always supermodular. Hence, unless otherwise specified we assume that $\Pi_T \mathcal{S}$ 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 called mostly-sublattice.

Definition 1 A set $\mathcal{S} \subseteq \mathcal{X} \times \mathcal{T}$ is a mostly-sublattice if for any unordered pair $t', t'' \in \Pi_T \mathcal{S}$ and any $x' \in \mathcal{S}_{t'}, x'' \in \mathcal{S}_{t''}$, we have $x' \vee x'' \in \mathcal{S}_{t' \vee t''}$ and $x' \wedge x'' \in \mathcal{S}_{t' \wedge t''}$. Unlike a sublattice, the definition of mostly-sublattice only imposes requirements for unordered pairs $t', t'' \in \Pi_T \mathcal{S}$. It automatically holds if $\Pi_T \mathcal{S}$ does not contain any unordered pair. On the other hand, if $\Pi_T \mathcal{S} = \{t', t'', t' \vee t'', t' \wedge t''\}$ for an unordered pair $t'$ and $t''$, the only condition for $\mathcal{S}$ to be a mostly-sublattice is that $\{x' \vee x'': x' \in \mathcal{S}_{t'}, x'' \in \mathcal{S}_{t''}\} \subseteq \mathcal{S}_{t' \vee t''}$ and $\{x' \wedge x'': x' \in \mathcal{S}_{t'}, x'' \in \mathcal{S}_{t''}\} \subseteq \mathcal{S}_{t' \wedge t''}$, which does not require $\mathcal{S}_{t'}, \mathcal{S}_{t''}, \mathcal{S}_{t' \vee t''}, \mathcal{S}_{t' \wedge t''}$ to be sublattices in $\mathcal{X}$, the very basic requirement for $\mathcal{S}$ to be a sublattice. We next provide two examples to illustrate this definition.

Example 1. Let $\mathcal{X} = \mathcal{T} = \mathbb{R}^2$ and $\mathcal{S} = \{(0, 1, 0, 1), (1, 0, 1, 0), (0, 0, 0, 0), (1, 1, 1, 1), (0.5, 1.5, 1, 1)\}$. $\mathcal{S}$ is not a sublattice since $\mathcal{S}_{(1,1)}$ is not a sublattice. However, $\mathcal{S}$ is a mostly-sublattice.

Example 2. Let $\mathcal{X} = \mathcal{T} = \mathbb{R}^2$ and $\mathcal{S} = \{(1, 0, 0, 0), (0, 1, 1, 1)\}$. Its convex hull can be represented as $\text{Conv}(\mathcal{S}) = \{(x_1, x_2, t_1, t_2) \in \mathbb{R}^4 : x_1 + x_2 = 1, t_1 - t_2 = 0, t_1 - x_2 = 0, x_1 \in [0, 1]\}$. Obviously, neither $\mathcal{S}$ nor $\text{Conv}(\mathcal{S})$ is a sublattice. However, both of them are mostly-sublattices since there is no unordered pair $t', t''$ in their projections on $\mathcal{T}$. 
Interestingly, $S$ being a mostly-sublattice is both necessary and sufficient for the preservation of supermodularity in problem (1).

**Theorem 3** The function $g$ is supermodular on $\Pi_T S$ whenever $f$ is supermodular on $X \times T$ if and only if $S$ is a mostly-sublattice.

Observe that the only difference between a sublattice and a mostly-sublattice lies in the ordered pair in $\Pi_T S$. The mostly-sublattice requirement in Theorem 3 implies that to preserve supermodularity, there is no requirement for the ordered pair $t', t''$, since for the ordered pair $t', t''$, \( \{ t' \wedge t'', t' \vee t'' \} = \{ t', t'' \} \), and hence $g(t') + g(t'') \leq g(t' \wedge t'') + g(t' \vee t'')$ automatically holds. But for the sublattice requirement, even if $t'$ and $t''$ are ordered, we need $(x' \wedge x'', t' \wedge t'')$, $(x' \vee x'', t' \vee t'') \in S$ for any $x' \in S_T$ and $x'' \in S_T$.

By the sufficient condition, $S$ being mostly-sublattice implies the supermodularity of $g$ if $f$ is supermodular. On the other hand, by the necessary condition, $S$ not being mostly-sublattice implies that there exists some supermodular function $f$ such that the corresponding $g$ is not supermodular. That is, if $S$ is not a mostly-sublattice, the supermodularity of $f$ alone is not sufficient for us to claim the supermodularity of $g$, and stronger conditions on $f$ are needed.

**Corollary 1** When the set $S$ is convex, the function $g$ is supermodular on $\Pi_T S$ whenever $f$ is supermodular on $X \times T$ if and only if $S$ is a mostly-sublattice.

To see the above result, note that the same condition in Theorem 3 must be a sufficient condition for the case when $S$ is convex. In addition, the counter example constructed in the proof of the “necessary” direction in Theorem 3, is still a valid counter example when $S$ is convex. Hence, $S$ being a mostly-sublattice is also a necessary condition for the case when $S$ is convex.

We now investigate the conditions on $A \in \mathbb{R}^{m \times n_1}$ and $B \in \mathbb{R}^{m \times n_2}$ under which the corresponding polyhedron $S_c^\nu$ defined in Equation (2) is a mostly-sublattice for any $c \in \mathbb{R}^m$. Recall that Theorem 2 implies the polyhedron $S_c^\nu$ is a sublattice for any $c$ if and only if $(A|B)$ is a lattice-matrix. Interestingly, the lattice-matrix requirement also applies to mostly-sublattices.
Theorem 4 The polyhedron $S_c^p$ defined by Equation (2) is a mostly-sublattice for each $c \in \mathbb{R}^m$ if and only if $(A|B)$ is a lattice-matrix.

Remark 1 We here study the condition for $S_c^p$ to be a mostly-sublattice for every $c$, instead of for a given $c$, for two reasons. First, in many applications (e.g., the ATO problem studied later in Section 4, or dynamic programming recursions), the parametric optimization problem often involves uncertainties. In general, these uncertainties affect the parameter $c$ in the constraints which define the polyhedron $S_c^p$; consequently, $c$ takes values over a wide range rather than being constant. Secondly, it remains open to derive easily checkable conditions for $S_c^p$ to be a sublattice or a mostly-sublattice. In fact, the only general result on the sublattice of $S_c^p$ is Theorem 2, which comes with statement “for every $c$”.

According to Theorems 2 and 4, the requirements from sublattices and mostly-sublattices are equivalent for the polyhedral case when considering every $c$. However, they are not equivalent in general, e.g., $Conv(S)$ in Example 2, and the network flow problem discussed later.

We now present the preservation results when $f$ is both concave and supermodular.

Theorem 5 The function $g$ is concave and supermodular on $\Pi_T S$ whenever $f$ is concave and supermodular on $X \times T$ if and only if the set $S$ is a convex mostly-sublattice.

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

From Theorem 5, we can also easily derive the condition for preserving concavity and supermodularity when $S$ is already convex.
Corollary 2 Consider the case where $\mathcal{S}$ is convex. The function $g$ is concave and supermodular on $\Pi_T \mathcal{S}$ whenever $f$ is concave and supermodular on $\mathcal{X} \times \mathcal{T}$ if and only if $\mathcal{S}$ is a mostly-sublattice.

Since a polyhedral set is convex, we immediately have a sufficient and necessary condition that preserves both concavity and supermodularity under the optimization operation (1) when $\mathcal{S}$ is specified by (2) is $\mathcal{S}_c^{\prime\prime}$ being a mostly-sublattice. The preservation holds for any $c \in \mathbb{R}^m$ if and only if $(A|B)$ 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 parameter $t$, i.e., $f(x, t') = f(x, t'')$, $\forall x \in \mathcal{X}, t', t'' \in \mathcal{T}$. 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 $\mathbb{R}$. Therefore, the problem (1) becomes

$$g(t) = \max \{f(x) : x \in \mathcal{S}_t\}. \quad (3)$$

This special class of parametric optimization problems appears 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-sublattice.

**Definition 2** A set $\mathcal{S} \subseteq \mathcal{X} \times \mathcal{T}$ is an additive mostly-sublattice with a function $\mathcal{W} : \mathcal{X} \times \mathcal{X} \rightarrow 2^\mathcal{X}$, which maps two members of $\mathcal{X}$ to a subset of $\mathcal{X}$, if and only if for any unordered pair $t', t'' \in \mathcal{T}$ and any $x' \in \mathcal{S}_{t'}$, $x'' \in \mathcal{S}_{t''}$, there exist $y \in \text{Conv}(\mathcal{W}(x', x'')) \cap \mathcal{S}_{t' \wedge t''}$ and $z \in \text{Conv}(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''})$ such that $y + z = x' + x''$.

We first state that mostly-sublattices are special cases of additive mostly-sublattices with any $\mathcal{W}$ satisfying $x' \wedge x''$, $x \lor x'' \in \mathcal{W}(x', x'')$. To show that, consider any mostly-sublattice $\mathcal{S}$ and any unordered pair $t', t'' \in \Pi_T \mathcal{S}$, $x' \in \mathcal{S}_{t'}$, $x'' \in \mathcal{S}_{t''}$. Based on the definition, we have $x' \wedge x'' \in \mathcal{S}_{t' \wedge t''}$, $x' \lor x'' \in \mathcal{S}_{t' \vee t''}$. For any mapping $\mathcal{W}$ satisfying the condition $x' \wedge x''$, $x' \lor x'' \in \mathcal{W}(x', x'')$, let $y = x' \wedge x''$ and $z = x' \lor x''$. Clearly, $y \in \text{Conv}(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \wedge t''})$, $z \in \text{Conv}(\mathcal{W}(x', x'') \cap \mathcal{S}_{t' \vee t''})$ and $y + z = x' \wedge x'' + x' \lor x'' = x' + x''$. Hence, $\mathcal{S}$ is an additive mostly-sublattice with $\mathcal{W}$. 

We can understand the extension from mostly-sublattice to additive mostly-sublattice by comparing Definitions 1 and 2. Specifically, given \((x', t'), (x'', t'') \in S\) with unordered \(t', t''\), both mostly-sublattice and additive mostly-sublattice have requirements on \(S_{t' \wedge t''}\) and \(S_{t' \vee t''}\). While mostly-sublattice requires exactly \(y = x' \wedge x'' \in S_{t' \wedge t''}\) and \(z = x' \vee x'' \in S_{t' \vee t''}\), the additive mostly-sublattice allows more flexibility when choosing \(y\) and \(z\) in \(S_{t' \wedge t''}\) and \(S_{t' \vee t''}\). This level of flexibility is controlled by the function \(W\).

To demonstrate better the flexible requirement on \(S_{t' \wedge t''}\) and \(S_{t' \vee t''}\) in additive mostly-sublattice with different mappings, we provide the following examples. In particular, as in Table 1, we define three set-valued functions: \(W^1, W^2\) and \(W^3\), such that at every \(x', x'' \in \mathcal{X}\),

\[
W^1(x', x'') = \{x', x'', x' \wedge x'', x' \vee x''\},
\]

\[
W^2(x', x'') = \text{Conv}(x', x'', x' \wedge x'', x' \vee x''),
\]

\[
W^3(x', x'') = \mathcal{X}.
\]

**Example 3.** Let \(\mathcal{X} = \mathbb{R}, \mathcal{T} = \mathbb{R}^2\) and \(S = \{(0, 0, 1), (1, 1, 0), (\epsilon, 0, 0), (1 - \epsilon, 1, 1)\}\). We now show that 1) \(S\) is a mostly-sublattice if \(\epsilon = 0\), 2) \(S\) is an additive mostly-sublattice with \(W^1\) but not a mostly-sublattice if \(\epsilon = 1\), 3) \(S\) is an additive mostly-sublattice with \(W^2\) but not an additive mostly-sublattice with \(W^3\) if \(\epsilon \notin [0, 1]\). Observe that the only unordered pair in \(\Pi_T S\) is \(t' = (0, 1)\) and \(t'' = (1, 0)\), with the corresponding \(x' = 0, x'' = 1\). Hence, \(t' \wedge t'' = (0, 0)\), \(t' \vee t'' = (1, 1)\), \(x' \wedge x'' = 0, x' \vee x'' = 1\), \(W^1(x', x'') = \{0, 1\}\), \(W^2(x', x'') = [0, 1]\), \(W^3(x', x'') = \mathbb{R}\). We illustrate the set \(S\) under all four cases of \(\epsilon\) in Figure 1.

When \(\epsilon = 0\), we have \(S_{t' \wedge t''} = \{0\}\), \(S_{t' \vee t''} = \{1\}\). Hence, \(x' \wedge x'' \in S_{t' \wedge t''}\) and \(x' \vee x'' \in S_{t' \vee t''}\), \(S\) is a mostly-sublattice.

When \(\epsilon = 1\), we have \(S_{t' \wedge t''} = \{1\}\), \(S_{t' \vee t''} = \{0\}\). Hence, \(x' \wedge x'' \notin S_{t' \wedge t''}\) and \(x' \vee x'' \notin S_{t' \vee t''}\), \(S\) is not a mostly-sublattice. However, we can choose \(y = 1 \in \text{Conv}(W^1(x', x'') \cap S_{t' \wedge t''})\) and \(z = 0 \in \text{Conv}(W^1(x', x'') \cap S_{t' \vee t''})\) such that \(y + z = 1 = x' + x''\). Hence, \(S\) is an additive mostly-sublattice with \(W^1\).
When $\epsilon \in (0, 1)$, we have $S_{U \wedge U'} = \{\epsilon\}$, $S_{U \vee U'} = \{1 - \epsilon\}$. Hence, $\mathcal{W}^1(x', x'') \cap S_{U \wedge U'} = \mathcal{W}^1(x', x'') \cap S_{U \vee U'} = \emptyset$, $\mathcal{S}$ is not an additive mostly-sublattice with $\mathcal{W}^1$. However, we can choose $y = \epsilon \in \text{Conv}(\mathcal{W}^2(x', x'') \cap S_{U \wedge U'})$ and $z = 1 - \epsilon \in \text{Conv}(\mathcal{W}^2(x', x'') \cap S_{U \vee U'})$ such that $y + z = 1 = x' + x''$. Hence, $\mathcal{S}$ is an additive mostly-sublattice with $\mathcal{W}^2$.

When $\epsilon \not\in [0, 1]$, we have $S_{U \wedge U'} = \{\epsilon\}$, $S_{U \vee U'} = \{1 - \epsilon\}$ and $\mathcal{W}^2(x', x'') \cap S_{U \wedge U'} = \emptyset$. Hence, $\mathcal{S}$ is not an additive mostly-sublattice with $\mathcal{W}^2$. Nevertheless, we can choose $y = \epsilon \in (S_{U \wedge U'} \cap \mathcal{W}^3(x', x''))$, $z = 1 - \epsilon \in (S_{U \vee U'} \cap \mathcal{W}^3(x', x''))$ such that $y + z = 1 = x' + x''$. Therefore, $\mathcal{S}$ is an additive mostly-sublattice with $\mathcal{W}^3$.

We now consider the convex hall of $\mathcal{S}$, which is denoted as $\hat{\mathcal{S}} = \text{Conv}(\mathcal{S}) = \{(x, t_1, t_2) \in \mathbb{R}^3 : x - (1 - \epsilon)t_1 + \epsilon t_2 = \epsilon, t_1, t_2 \in [0, 1]\}$. Consider any $(x', t'_1, t'_2), (x'', t''_1, t''_2) \in \hat{\mathcal{S}}$ with $t'_1 > t''_1, t'_2 < t''_2$. We notice that $x' = \epsilon + (1 - \epsilon)t'_1 - \epsilon t'_2, x'' = \epsilon + (1 - \epsilon)t''_1 - \epsilon t''_2$, and $\hat{S}_{U \wedge U'} = \{\epsilon + (1 - \epsilon)t'_1 - \epsilon t'_2\}$, $\hat{S}_{U \vee U'} = \{\epsilon + (1 - \epsilon)t'_1 - \epsilon t''_2\}$. We make similar observations: $\text{Conv}(\mathcal{S})$ is a mostly-sublattice when $\epsilon = 0$, an additive mostly-sublattice with $\mathcal{W}^1$ when $\epsilon = 1$, an additive mostly-sublattice with $\mathcal{W}^2$ when $\epsilon \in (0, 1)$, and an additive mostly-sublattice with $\mathcal{W}^3$ when $\epsilon \not\in [0, 1]$.

![Figure 1](image_url)  
An illustrative example for four definitions with $\mathcal{S} = \{(0, 0, 1), (1, 1, 0), (\epsilon, 0, 0), (1 - \epsilon, 1, 1)\}$.
In Definition 2, when the function $\mathcal{W}$ maps to a larger subset of $\mathcal{X}$, the choice of $\mathbf{y}$ in $\mathcal{S}_{\mathcal{U} \setminus \mathcal{U}'}$ and $\mathbf{z}$ in $\mathcal{S}_{\mathcal{U} \setminus \mathcal{U}''}$ is more flexible and hence the requirement is easier to satisfy. Hence, we have the following relationship,

$$\{\mathcal{S} : \mathcal{S} \text{ is a mostly-sublattice}\} \subseteq \{\mathcal{S} : \mathcal{S} \text{ is an additive mostly-sublattice with } \mathcal{W}^1\} \subseteq \{\mathcal{S} : \mathcal{S} \text{ is an additive mostly-sublattice with } \mathcal{W}^2\} \subseteq \{\mathcal{S} : \mathcal{S} \text{ is an additive mostly-sublattice with } \mathcal{W}^3\}$$

Indeed, the flexibility when choosing $\mathbf{y}$ and $\mathbf{z}$ in the additive mostly-sublattice with $\mathcal{W}^1$, $\mathcal{W}^2$ and $\mathcal{W}^3$ is essential to preserving supermodularity in our settings. We first show the result for the optimization problem (3) with objective functions independent of $\mathbf{t}$. We now present a necessary and sufficient condition that preserves supermodularity.

**Theorem 6** The function $g$ is supermodular on $\Pi_{\mathcal{T}} \mathcal{S}$ whenever $f$ is supermodular on $\mathcal{X}$ if and only if the set $\mathcal{S}$ is an additive mostly-sublattice with $\mathcal{W}^1$.

As discussed before, all mostly-sublattices are additive mostly-sublattices with $\mathcal{W}^1$ but not vice versa. Hence, Theorem 6 implies that when function $f$ is independent of the decision variable $\mathbf{t}$, the condition for preserving supermodularity is less restrictive than that for the general case stated in Theorem 3. To provide intuition into the need for a weaker condition, note that for the general function $f(\mathbf{x}, \mathbf{t})$, we need to find $\mathbf{y} \in \mathcal{S}_{\mathcal{U} \setminus \mathcal{U}'}$ and $\mathbf{z} \in \mathcal{S}_{\mathcal{U} \setminus \mathcal{U}''}$ such that $f(\mathbf{y}, \mathbf{t'} \wedge \mathbf{t''}) + f(\mathbf{z}, \mathbf{t'} \vee \mathbf{t''}) \geq f(\mathbf{x}', \mathbf{t'}) + f(\mathbf{x}'', \mathbf{t''})$, which is only true in general if $\mathbf{y} = \mathbf{x'} \wedge \mathbf{x''}$, $\mathbf{z} = \mathbf{x'} \vee \mathbf{x''}$. However, for the special case of $f(\mathbf{x})$, we just need $f(\mathbf{y}) + f(\mathbf{z}) \geq f(\mathbf{x'}) + f(\mathbf{x''})$. Hence, either $\mathbf{y} = \mathbf{x}', \mathbf{z} = \mathbf{x''}$ or $\mathbf{y} = \mathbf{x'} \wedge \mathbf{x''}, \mathbf{z} = \mathbf{x'} \vee \mathbf{x''}$ can work; moreover, switching the values of $\mathbf{y}$ and $\mathbf{z}$ can also work. Summarizing these four cases, we have the requirement of $\mathbf{y}, \mathbf{z} \in \mathcal{W}^1(\mathbf{x}', \mathbf{x''})$ and $\mathbf{y} + \mathbf{z} = \mathbf{x'} + \mathbf{x''}$, which is the essential concept of the additive mostly-sublattice with $\mathcal{W}^1$.

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

**Corollary 3** When the set $\mathcal{S}$ is convex, the function $g$ is supermodular on $\Pi_{\mathcal{T}} \mathcal{S}$ whenever $f$ is supermodular on $\mathcal{X}$ if and only if the set $\mathcal{S}$ is an additive mostly-sublattice with $\mathcal{W}^1$. 
We now investigate when the polyhedron defined by (2) is an additive mostly-sublattice with $\mathcal{W}^1$.

**Theorem 7** The polyhedron $S^p_c$ defined by Equation (2) is an additive mostly-sublattice with $\mathcal{W}^1$ for each $c \in \mathbb{R}^n$ if and only if one of the following conditions holds.

1. Either $(A|B)$ or $(-A|B)$ is a lattice-matrix.

2. Denote $I = \{ i : b_i \neq 0 \} \subseteq \{1, \ldots, m\}$. There exist $k \in \mathbb{R}^{|I|}$, $d \in \mathbb{R}^n$ such that $A_I = kd^T$, and $(k|B_X)$ is a lattice-matrix.

Recall that the condition on $A, B$ imposed by mostly-sublattice is that $(A|B)$ has to be a lattice-matrix. Here, for $S^p_c$ to be an additive mostly-sublattice with $\mathcal{W}^1$, the first condition in Theorem 7 implies that other than $(A|B)$ being a lattice-matrix, $(-A|B)$ being a lattice-matrix also works. This is due to the greater flexibility of the additive mostly-sublattice over the mostly-sublattice. Consider any unordered pair $t', t''$, and $x', x'' \in S_{t'}$, $x', x'' \in S_{t''}$. Note that mostly-sublattice requires $x' \wedge x'' \in S_{t' \wedge t''}$, $x' \vee x'' \in S_{t' \vee t''}$. The additive mostly-sublattice with $\mathcal{W}^1$, however, is satisfactory with either the same condition (i.e., $x' \wedge x'' \in S_{t' \wedge t''}$ and $x' \vee x'' \in S_{t' \vee t''}$) or its reverse: $x' \vee x'' \in S_{t' \wedge t''}$ and $x' \wedge x'' \in S_{t' \vee t''}$ (e.g., Example 3). Essentially, while the former leads to $(A|B)$ being a lattice-matrix, the latter results in $(-A|B)$ being a lattice-matrix.

We illustrate the condition by two examples: $S^p_{c, 1} = \{ (x, t) : x_1 + t_1 \leq c_1, -x_2 - t_2 \leq c_2 \}$ and $S^p_{c, 2} = \{ (x, t) : -\sum_{i=1}^n x_i + t_1 \leq c_1, -\sum_{i=1}^n x_i + t_2 \leq c_2 \}$. Specifically, $S^p_{c, 1}$ satisfies the first condition while $S^p_{c, 2}$ satisfies the second condition in Theorem 7. Hence, the optimization problem (3) preserves supermodularity in both examples. However, they do not satisfy the condition in Theorem 4. Thus, there exist $c$ and a (concave and) supermodular function $f$ such that $g$ is not (concave and) supermodular under the optimization operation (1) when $S$ is specified by $S^p_{c, 1}$ or $S^p_{c, 2}$.

We now derive a necessary and sufficient condition that preserves both concavity and supermodularity under the optimization operation (3).

**Theorem 8** The function $g$ is concave and supermodular on $\Pi_T S$ whenever $f$ is concave and supermodular on $X$ if and only if the following two conditions hold simultaneously:
1. $\mathcal{S}$ is an additive mostly-sublattice with $\mathcal{W}^2$.

2. For any $t', t'' \in \mathcal{V}$, $x' \in \mathcal{S}_{t'}$, $x'' \in \mathcal{S}_{t''}$ and $\lambda \in [0, 1]$, we have $x_\lambda \in \text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t_\lambda})$.

\[ t_\lambda = \lambda t' + (1 - \lambda)t'' \]
\[ x_\lambda = \lambda x' + (1 - \lambda)x'' \]

When the function $f$ is independent of $t$, comparing the condition that preserves supermodularity (Theorem 6) and that preserves both concavity and supermodularity (Theorem 8), we observe two differences.

First, any additive mostly-sublattice with $\mathcal{W}^1$ in Theorem 6 must be an additive mostly-sublattice with $\mathcal{W}^2$ in Theorem 8. It implies that the additional property of concavity leads to a less restrictive condition. Intuitively, with $y + z = x' + x''$, to have $f(y) + f(z) \geq f(x') + f(x'')$, we need $y, z \in \{x', x'', x' \vee x'', x' \wedge x''\}$ if $f$ is supermodular, but we only need $y, z \in \text{Conv}(x', x'', x' \vee x'', x' \wedge x'')$ if $f$ is both concave and supermodular. Hence, the concavity of $f$ leads to the relaxation from $\mathcal{W}^1$ to $\mathcal{W}^2$.

Second, 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 to preserve concavity under optimization operations. Therefore, the condition that preserves concavity and supermodularity is not a simple combination of the condition which preserves supermodularity and that preserves concavity. We next provide an example which satisfies the conditions in Theorem 8 but not the convexity of $\mathcal{S}$.

**Example 4.** Let $\mathcal{X} = \mathbb{R}, \mathcal{S} = \mathbb{R}^2, \mathcal{S} = \{(x_1, x_2) : x \in [0, 1], t_1, t_2 \in [0, 1]\}$. For any pair $t', t'' \in \Pi_t \mathcal{S}$ 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 $\mathcal{W}^2(x', x'') = \text{Conv}(x', x'', x' \wedge x'', x' \vee x'') = \{x'\} \subseteq [0, 1]$ and $\text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t' \wedge t''}) = \mathcal{W}^2(x', x'') \cap \mathcal{S}_{t_\lambda} = \{x'\}$. Define $y = z = x'$. We have $y \in \mathcal{W}^2(x', x'') \cap \mathcal{S}_{t' \wedge t''}$, $z \in \text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t' \wedge t''})$ and $y + z = x' + x''$. In addition, $x_\lambda = x' \in \text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t_\lambda})$. If $x' \neq x''$, without loss of generality (WLOG), let $x' = 0, x'' = 1$. Then, we have $\mathcal{W}^2(x', x'') = [0, 1]$ and $\text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t' \wedge t''}) = \text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t' \vee t''}) = \text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t_\lambda}) = [0, 1]$.

Let $y = 0 \in \text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t' \wedge t''}), z = 1 \in \text{Conv}(\mathcal{W}^2(x', x'') \cap \mathcal{S}_{t' \vee t''})$.
$S_{t', t''}$. We have $y + z = x' + x''$. In addition, $x_\lambda = \lambda x' + (1 - \lambda)x'' = 1 - \lambda \in \text{Conv}(W^2(x', x'') \cap S_{t, \lambda})$.

Hence, set $S$ satisfies both conditions in Theorem 8 but is not a convex set.

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

**Corollary 4** Consider the case where $S$ is convex. The function $g$ is concave and supermodular on $\Pi_\gamma S$ whenever $f$ is concave and supermodular on $\mathcal{X}$ if and only if the set $S$ is an additive mostly-sublattice with $W^2$.

The condition in Corollary 4 is less restrictive than that in Corollary 2 because $f$ is independent of $t$. We now use two examples to show how this can be applied. First, we use this corollary to cover the result in Lemma 2.6.2 part (b) (Topkis 1998), which shows that $g(t) := f(a_1 t_1 + a_2 t_2)$ is supermodular if $f$ is concave and $a_1 > 0, a_2 < 0$. To see this, we can reformulate $g(t)$ as $g(t) = f(a_1 t_1 + a_2 t_2) = \max_{x = a_1 t_1 + a_2 t_2} f(x)$. Since the function $f(x)$ is concave and supermodular, and we can easily check that the set $S = \{(x, t) : x = a_1 t_1 + a_2 t_2\}$ is an additive mostly-sublattice with $W^2$ when $a_1 > 0, a_2 < 0$ by observing $W^2(x', x'') = \{x', x''\}$, hence the concavity and supermodularity is preserved according to Corollary 4. We also apply this corollary to analyze an inventory management problem with cash flow constraints and show the preservation of supermodularity in the appendix. A similar problem has been considered by Chao et al. (2008). But they define the state differently and do not obtain the property of supermodularity.

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

**Theorem 9** The polyhedron $S_c^P$ defined by Equation (2) is an additive mostly-sublattice with $W^2$ for any $c \in \mathbb{R}^n$ only if for any $\alpha \in \mathbb{R}^m$, $\beta \in \mathbb{R}^{m_2}$, $I \subseteq \{1, \ldots, m\}$, and $B_I \beta = A_I \alpha$, there exist $\lambda_1, \lambda_2 \in [0, 1]$ such that

$$B_I \beta^+ = A_I (\lambda_1 \alpha^+ - \lambda_2 (-\alpha)^+)$$

which in turn implies that $B$ 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 propositions, we show the conjecture is valid for two special cases: 1) $\text{rank}(B) = 1$; 2) $m = 2$ (i.e., $A$ and $B$ have only two rows).
Proposition 1 Suppose \( \text{rank}(B) = 1 \). The polyhedron \( F^c \) defined by Equation (2) is an additive mostly-sublattice with \( W^2 \) for any \( c \in \mathbb{R}^m \) if and only if \( B \) is a lattice-matrix.

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

- **S1.** The polyhedron \( F^c \) defined by Equation (2) is an additive mostly-sublattice with \( W^2 \) for any \( c \in \mathbb{R}^m \).

- **S2.** For any \( \alpha \in \mathbb{R}^{n_1}, \beta \in \mathbb{R}^{n_2}, I \subseteq \{1, 2\}, \) and \( B_k \beta = A_{(I)} \alpha \), there exist \( \lambda_1, \lambda_2 \in [0, 1] \) such that

  \[ B_k \beta^+ = A_{(I)} \left( \lambda_1 \alpha^+ - \lambda_2 (-\alpha)^+ \right). \]

- **S3.** \( B \) is a lattice-matrix. When \( \text{rank}(B) = 2 \), \( B \) 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 \( D = B^{-1} A \) should satisfy 1) \( d_{11}d_{21} \geq 0 \) in case of \( \text{rank}(D) = 1 \), and 2) \( D \geq 0 \) or \( D \leq 0 \) in case of \( \text{rank}(D) = 2 \); When \( n_2 = 3 \), there exists \( i \in \{1, 2, 3\} \) such that \( B_i \) has no zero element, and \( A = B, k^T \) for some \( k \in \mathbb{R}^{n_1} \).

In Proposition 2, Statement **S2** indicates that the necessary condition provided by Theorem 9 is also sufficient in this particular case. Meanwhile, Statement **S3** explicitly characterizes the structure of such matrices \( A, B \). When \( n_2 = 2 \) and \( \text{rank}(D) = 2 \), this condition implies that either all column vectors in \( A \) are conic combinations of \( B_1 \) and \( B_2 \), or all column vectors in \( A \) are conic combination of \( -B_1 \) and \( -B_2 \). Proposition 2 can be slightly extended.

Proposition 3 For \( m = 2 \), given any nonempty convex sublattice of Euclidean space \( D \), both set \( \{(x, t) : Ax + Bt = c, x \in D\} \) and set \( \{(x, t) : Ax + Bt \leq c, x \in D\} \) are additive mostly-sublattices with \( W^2 \) for all \( c \in \mathbb{R}^2 \) if \( A, B \) 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 \( S = \{(x, t) : Ax + Bt = 0, x \in D\} \), with \( B \in \mathbb{R}^{2 \times 2} \), and \( D \) being a nonempty closed convex sublattice. They prove that when \( B^T B \) is an \( L_0 \)-matrix (i.e., with non-negative
diagonal entries and non-positive off-diagonal entries) and $\mathbf{B}^T \mathbf{A} \leq 0$, $g(t) = \max \{ f(x) : (x,t) \in \mathcal{S} \}$ is concave and supermodular if $f$ is. Here we show that their result is an immediate corollary of Proposition 3. In particular, when $\text{rank}(\mathbf{B}) = 1$, $\mathbf{B}^T \mathbf{B}$ being $\mathcal{L}_0$-matrix implies that $\mathbf{B}$ is a lattice-matrix. Hence, the condition in Proposition 3 is satisfied. When $\text{rank}(\mathbf{B}) = 2$, we can equivalently represent $\mathcal{S} = \{(x,t) : (\mathbf{B}^T \mathbf{A})x + (\mathbf{B}^T \mathbf{B})t = 0, x \in \mathcal{D}\}$ and we can check 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-sublattice with $\mathcal{W}^2$. 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 5.** We consider the following sets.

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

\[
\mathcal{S}_2 = \left\{ (x,t) : \begin{bmatrix} 1 & 2 & 3 \\ -2 & -4 & -6 \end{bmatrix} x + \begin{bmatrix} 10 \\ 0 \end{bmatrix} t = 0 \right\},
\]

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

For the first two cases, we do not have $\mathbf{B}^T \mathbf{A} \leq 0$. 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 where the objective function $f(x)$ is a linear function, i.e., $f(x) = \mathbf{a}^T x$ for some vector $\mathbf{a}$. We provide a necessary and sufficient condition for $\mathcal{S}$ under which the function $g$ defined by the optimization problem (3) is supermodular.

**Theorem 10** The function $g$ is supermodular on $\Pi_T \mathcal{S}$ whenever $f$ is linear on $\mathcal{X}$ if and only if the set $\mathcal{S}$ is an additive mostly-sublattice with $\mathcal{W}^3$. 

Not surprisingly, as we have more specific properties for \( f \), the condition in Theorem 10 is less restrictive than those in Theorems 3, 6 and 8. Intuitively, with \( y + z = x' + x'' \), to have \( f(y) + f(z) \geq f(x') + f(x'') \), we do not need any additional requirement when \( f \) is linear; by contrast, we need \( y, z \in \text{Conv}(x', x'', x' \land x'', x' \lor x'') \) if \( f \) is only known to be concave and supermodular. Hence, the requirement is extended from additive mostly-sublattices with \( \mathcal{W}^2 \) when \( f \) is concave and supermodular (Theorem 8) to additive mostly-sublattices with \( \mathcal{W}^3 \) when \( f \) is linear (Theorem 10).

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

**Corollary 5** Consider the case where \( \mathcal{S} \) is convex. The function \( g \) is supermodular on \( \Pi_{\mathcal{T}} \mathcal{S} \) whenever \( f \) is linear on \( \mathcal{X} \) if and only if the set \( \mathcal{S} \) is an additive mostly-sublattice with \( \mathcal{W}^3 \).

We can also derive the following condition for \( \mathcal{S}^P_c \) under which the linearity of the function \( f \) leads to supermodularity of the function \( g \).

**Theorem 11** The polyhedron \( \mathcal{S}^P_c \) defined by Equation (2) is an additive mostly-sublattice with \( \mathcal{W}^3 \) for any \( c \in \mathbb{R}^m \) if and only if either 1) \( n = m \), or 2) for any \( \beta \in \mathbb{R}^{n_2} \) and \( I \subseteq \{1, \ldots, m\} \) with \(|I| = n + 1\), \( \text{Rank}(A_I) = n \), and \( B_I \beta \in C(A_I) \) we have \( B_I \beta^+ \in C(A_I) \). Here \( n = \text{Rank}(A) \).

Given any specific matrices \( A, B \) 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:** \( A \in \mathbb{R}^{m \times n_1}, B \in \mathbb{R}^{m \times n_2}, n = \text{Rank}(A), \) and indicator \( s \).

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

2. Enumerate all index sets \( \hat{I} \subseteq \{1, \ldots, m\} \) satisfying the condition that \( |\hat{I}| = n \) and \( A_{\hat{I}} \) is invertible.

3. For each of index set \( \hat{I} \), calculate the vector \( d_k^T = b_k^T - a_k^T A_{\hat{I}}^{-1} B_{\hat{I}}, \forall k \in \{1, \ldots, m\} \setminus \hat{I} \). If \( \exists k \in \{1, \ldots, m\} \setminus \hat{I} \) such that \( d_k \) 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 enumerations.

**Output:** \( s \).

**Theorem 12** The polyhedron \( S_c^D \) defined by Equation (2) is an additive mostly-sublattice with \( \mathcal{W}^3 \) for any \( c \in \mathbb{R}^m \) if and only if Algorithm 1 returns \( s = 1 \).

We provide two examples as illustrations. 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 6.** \( A \) is any matrix with full row rank, i.e., \( n = m \), and \( B \) is any matrix. More specific examples of \( A \) can be \( A = I \) (it cannot satisfy the condition in Theorem 2 if \( B \) has a strictly positive element), or \( A = [I - I] \).

**Example 7.** \( A = \begin{bmatrix} I_n \\ e^T_i \end{bmatrix} \) with any \( i \in \{1, \ldots, n\} \), \( B \) is any matrix such that \((b_{n+1} - b_i)\) 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 with the rank of \( A \) and thus in \( m \) and \( n_1 \). It remains unclear whether it can be completed in polynomial time.

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

**Theorem 13** The function \( g \) is concave and supermodular on \( \Pi_T \mathcal{S} \) whenever \( f \) is linear on \( \mathcal{X} \) if and only if the following conditions hold:

1. \( \mathcal{S} \) is an additive mostly-sublattice with \( \mathcal{W}^3 \).
2. For any \( t', t'' \in T \), \( x' \in \mathcal{S}_{t'}, x'' \in \mathcal{S}_{t''} \) and \( \lambda \in (0,1) \), we have \( x_\lambda \in \text{Conv}(\mathcal{S}_{t_\lambda}) \).

Here \( t_\lambda = \lambda t' + (1-\lambda)t'' \) and \( x_\lambda = \lambda x' + (1-\lambda)x'' \).
Remark 2 When the set $S$ is convex, the function $g$ is concave when $f$ is linear. Hence, the function $g$ is concave and supermodular on $\Pi_T S$ whenever $f$ is linear on $\mathcal{X}$ if and only if $S$ is an additive mostly-sublattice with $\mathcal{W}^3$. This condition is the same as that in Theorem 10. Therefore, for $S^c\in\Pi_T S$ to be an additive mostly-sublattice with $\mathcal{W}^3$ for every $c$, the requirement for $A, B$ is the same as that in Theorem 11, and can be checked by Algorithm 1.

Remark 3 By the definition of supermodularity, $g$ being supermodular on $\Pi_T S$ requires that the set $\Pi_T S$ is a sublattice. Indeed, this requirement has been incorporated in our proposed condition. More specifically, we can easily show that $\Pi_T S$ is a sublattice as long as $S$ is a mostly-sublattice or an additive mostly-sublattice with $\mathcal{W}^i$, $i \in \{1, 2, 3\}$. Similarly, by the definition of concavity, $g$ being concave on $\Pi_T S$ requires that the set $\Pi_T S$ is convex, and this requirement has also been incorporated in our proposed condition (i.e., the conditions in Theorems 5, 8 and 13).

We next discuss how our results relate to existing literature. Zipkin (2003) analyzes the preservation of supermodularity for a series of problems with linear structure. Specifically, he considers the problem $g(t) = \max \{p^T x : Ax \leq t, Cx \leq 0, x \geq 0, t \in \mathbb{R}^2_+\}$. Corollary 5 implies that it suffices to show that the graph of the constraint mapping is an additive mostly-sublattice with $\mathcal{W}^3$.

**Proposition 4** The set $S = \{(x, t) : Ax \leq t, Cx \leq 0, x \geq 0, t \in \mathbb{R}^2_+\}$ is an additive mostly-sublattice with $\mathcal{W}^3$ for any matrices $A$ and $C$ whose dimensions are consistent with $x$ and $t$.

The set with the form in Proposition 4 has certain business implications. For example, $t$ can be the available resource, $x$ represents intended levels of certain activities, $A$ is the consumption matrix of the activities, and $Cx \leq 0$ represents some homogeneous side constraints.

Some important supermodularity properties in network flow problems developed by Gale and Politof (1981) and Granot and Veinott (1985) can also be studied using Corollary 5. Consider any directed graph $G = (\mathcal{V}, \mathcal{A})$ where $\mathcal{V}$ is the vertex set and $\mathcal{A}$ is the arc set. Let $N \in \{0, +1, -1\}^{\mathcal{V} \times \mathcal{A}}$ be the vertex-arc incidence matrix. Let $q$ be a mapping from $\mathcal{A}$ to $\mathbb{R}^2$ defined as $q(\gamma) = (\xi(\gamma), \tau(\gamma))$, where...
where $c(\gamma)$ and $\bar{c}(\gamma)$ represent the lower and upper bounds of the flow on the arc $\gamma$, respectively. Gale and Politof (1981) and Granot and Veinott (1985) are concerned with the case where $q$ is constant on all arcs except arcs $\alpha$ and $\beta$. 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.,

$$
\mu(q(\alpha), q(\beta)) = \max \left\{ \sum_{\gamma \in \mathcal{A}} w(\gamma) x(\gamma) : \mathbf{N} \mathbf{x} = 0, \ c(\gamma) \leq x(\gamma) \leq \bar{c}(\gamma), \ \forall \gamma \in \mathcal{A} \right\},
$$

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

Gale and Politof (1981) and Granot and Veinott (1985) establish that $\mu(q(\alpha), q(\beta))$ is supermodular if arcs $\alpha$ and $\beta$ are in series, i.e., there is no simple cycle that 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 $\mathcal{S} = \{(x, t) : \mathbf{N} \mathbf{x} = 0, \ (q(\alpha), q(\beta)) = t, \ c(\gamma) \leq x(\gamma) \leq \bar{c}(\gamma), \ \forall \gamma \in \mathcal{A} \}$ is an additive mostly-sublattice with $\mathcal{W}^3$.

It is noteworthy that although sets $\mathcal{S}$ in both Propositions 4 and 5 are polyhedra, Theorem 12 cannot be used to check the preservation of supermodularity directly. The reason is that Theorem 12 explores the condition such that, for any $c$, $\mathcal{S}_{c'}$ is an additive mostly-sublattice with $\mathcal{W}^3$, while in Propositions 4 and 5, we only require the set $\mathcal{S}$ to be an additive mostly-sublattice with $\mathcal{W}^3$ for some specific $c$.

### 4. An Application on ATO Systems

In this section, we use the results of the previous sections on an ATO model. As an important operational problem, it has attracted considerable attention in recent decades (see for instance, Lu et al. (2003), Zhao and Simchi-Levi (2006), Lu et al. (2010)). We refer interested readers to Song and Zipkin (2003) and Atan et al. (2017) for comprehensive literature reviews. To describe the problem formally, we follow the formulation in Song and Zipkin (2003). The system assembles $n$ types of products from $m$ types of components. Let $\hat{g}(\mathbf{x})$ be the maximum expected profit with
the initial inventory levels of components being \( x \in \mathbb{R}_+^m \). Assuming unsatisfied demand is lost, we have

\[
\hat{g}(x) = \max_{y \geq x} \left\{ -c^T (y - x) + \mathbb{E} \left[ g \left( y, \tilde{d} \right) \right] \right\},
\]

where

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

and

\[
\mathcal{P}_A(y, d) = \left\{ (v, u, w) \in \mathbb{R}_+^n \times \mathbb{R}_+^m \times \mathbb{R}_+^n : \begin{array}{l}
Av + u = y, \\
v + w = d
\end{array} \right\}.
\]

Here \( y \) denotes the order-up-to level of components, \( \tilde{d} \) and \( d \) represent the random demands for products and their realizations, \( v, w, u \) denote the sales quantities of products, unmet demands of products, and amounts of leftover components, respectively. Parameters \( c \) and \( h \) represent the components’ per-unit ordering cost and holding cost, and \( r, p \) represent the products’ per-unit revenue from selling and penalties from shortage. In addition, \( A = (a_{ij})_{i=1, \ldots, m; j=1, \ldots, n} \), 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 \( A \) and \( n \) again here but with different meanings from those in previous sections.

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

**Theorem 14** Consider any distinct indexes \( i, j \in \{1, \ldots, m\} \). The function \( g(y, d) \) is supermodular in \((y_l)_{l \in I}\) for any non-negative \( y_l, l \in \{1, \ldots, m\} \setminus \{i, j\} \), \( d, r, p, h \) if and only if for any \( k = 2, \ldots, \min\{m-1, n\} \), any \((k+1) \times k\) submatrix \( Q \) is either 1) \( \text{rank}(Q) < k \), or 2) for any \( \lambda \in \mathbb{R}^{k+1} \) with \( Q^T \lambda = 0 \), we have \( \lambda_1 \lambda_2 \leq 0 \). Here \( Q \) is obtained from \( A \) 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 \( Q \).
The above result allows us to derive a condition under which \( g(\mathbf{y}, \mathbf{d}) \) is supermodular in the whole vector \( \mathbf{y} \).

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

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

- **W system:** \( \mathbf{A} = \begin{bmatrix} \mathbf{D} \\ \mathbf{c}^T \end{bmatrix} \in \mathbb{R}^{(n+1) \times n} \) where \( \mathbf{D} \) is a diagonal matrix and \( D_{ii} > 0 \), \( c_i > 0 \) \( \forall i = 1, \ldots, n \). 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 use Theorem 14 to show that the function \( g(\mathbf{y}, \mathbf{d}) \) is not supermodular for any two distinct components \( i, j \in \{1, \ldots, n\} \) for general parameters \( \mathbf{d}, \mathbf{r}, \mathbf{p} \) and \( \mathbf{h} \). Let the submatrix \( \mathbf{Q} \) be a \( 3 \times 2 \) matrix as

\[
\mathbf{Q} = \begin{bmatrix} D_{ii} & 0 \\ 0 & D_{jj} \\ c_i & c_j \end{bmatrix}.
\]

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

We then show that function \( g(\mathbf{y}, \mathbf{d}) \) 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

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

Hence, for any \( 3 \times 2 \) submatrix \( \mathbf{Q} \) obtained from the above matrix, \( \sum_{i=1}^{3} \lambda_i \mathbf{q}_i = \mathbf{0} \) only if \( \lambda_1 \lambda_2 \leq 0 \).
\textbf{M system:} $A = [D \ c] \in \mathbb{R}^{n \times (n+1)}$ where $D$ 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 $3 \times 2$ submatrix of $A$ contains at least one row dependent on another. Hence, $A$ satisfies Theorem 15 and function $g(y, d)$ is supermodular in $y$. We remark that for the special $W$ system and M system, where all elements in $A$ are either 0 or 1, the same result is derived by Zipkin (2003).

\textbf{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,

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

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 $A$ are 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 $g(y, d)$ in $y$, which can also be derived easily from Theorem 15.

- **Proportional Chained System:** for any two products $i, i'$ with a 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 use of all common components is in the same proportion. This is often the case when the main difference between products is size, so their consumption of common components are proportional to the size of product. Indeed, the propor-
tional 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,

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

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).

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

5. Extension and Conclusion

In this paper, we present a systematic study of 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-sublattice and additive mostly-sublattice, which generalize the concept of sublattice, and illustrate that they provide the needed sufficient and necessary conditions for such preservation. We then characterize some classes of polyhedral sets which satisfy these generalized lattice-like concepts. Finally, these characterizations are used to analyze ATO models.

Our results can be extended to allow for log-supermodular functions, which find interesting applications in the economics literature (see for instance, Topkis (1998), Athey (2002)). In the appendix, we show that the preservation of supermodularity can be extended to the preservation of log-supermodularity. For the general case where \( g(t) = \max_{x \in \mathcal{S}_t} f(x, t) \), to preserve log-supermodularity, we require that \( \mathcal{S} \) is a mostly-sublattice; While for the special case \( g(t) = \max_{x \in \mathcal{S}_t} f(x) \), the requirement for \( \mathcal{S} \) is an additive mostly-sublattice with \( \mathcal{W}^1 \).

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


Appendix

Inventory control with self-financing

We consider a single product inventory management problem over a planning horizon $N$ with the order quantity limited by the cash flow. At the beginning of each period, the manager observes the on-hand inventory $t_1$ and the available cash $s$. With per unit ordering cost $c$, the manager makes the ordering decision $z_1$ which is bounded by $[0, s/c]$. The remaining cash $s - cz_1$ will be deposited to earn an interest rate $(r - 1)$ per period $(r \geq 1)$. After the uncertain demand $\tilde{d}$ is realized as $d$, the revenue $p \min\{t_1 + z_1, d\}$ is collected and will become $r p \min\{t_1 + z_1, d\}$ at the end of the period, where $p$ represents the unit selling price. At the end of this period, the unused inventory is carried over and the unsatisfied demand is lost. The total available cash at the beginning of next period is hence $r (s - cz_1 + p \min\{t_1 + z_1, d\})$. Our objective is to maximize the expected cash at the end.

We let $t_2 = s + pt_1$, which represents the summation of current capital and the revenue if all on-hand inventory is sold, and $V_n(t_1, t_2)$ represents the maximum expected ending cash when the states observed at the beginning of period $n$ are $t_1, t_2$. The dynamic recursion can be written as

$$V_n(t_1, t_2)$$
$$= \max_{0 \leq z_1 \leq \frac{s}{c}} E \left[ V_{n+1}\left( (t_1 + z_1 - \tilde{d})^+, r (s - cz_1 + p \min\{t_1 + z_1, d\}) + p (t_1 + z_1 - \tilde{d})^+ \right) \right]$$
$$= \max_{0 \leq z_1 \leq \frac{s}{c}} E \left[ V_{n+1}\left( (t_1 + z_1 - \tilde{d})^+, r (t_2 - pt_1 - cz_1 + p (t_1 + z_1)) - p (r - 1) (t_1 + z_1 - \tilde{d})^+ \right) \right]$$
$$= \max_{(x, t) \in S} E \left[ V_{n+1}\left( (x_1 - \tilde{d})^+, r x_2 - p (r - 1) (x_1 - \tilde{d})^+ \right) \right]$$

(1)

where we define $x_1 = t_1 + z_1$, $x_2 = t_2 + (p - c) (x_1 - t_1)$ and $S = \{(x, t) : x_1 \geq t_1, \ cx_1 \leq t_2 - (p - c) t_1, \ x_2 - (p - c) x_1 = t_2 - (p - c) t_1\}$. Moreover, $V_{N+1}(t_1, t_2) = t_2 - pt_1$. By showing that $S$ in Equation (1) is an additive mostly-sublattice with $W^2$, we can recursively prove that $V_n$ is concave and supermodular. We formally state the result in the following Proposition.

Proposition 6 The function $V_n(t_1, t_2)$ is concave and supermodular in $(t_1, t_2)$.
Proof. Note that the function $V_n(t_1, t_2)$ is non-increasing in $t_1$ and non-decreasing in $t_2$, which is easy to prove by standard induction. We now use induction again to prove the concavity and supermodularity of the function $V_n$. We assume the statement is true in period $n + 1$, which is straightforward when $n = N$. We let $g_n(x_1, x_2) = \mathbb{E} \left[ h_n(x_1, x_2, \tilde{d}) \right]$ with $h_n(x_1, x_2, d) = V_{n+1}((x_1 - d)^+, rx_2 - p(r - 1)(x_1 - d)^+ )$. We then have $V_n(t_1, t_2) = \max_{(x, t) \in S} g_n(x_1, x_2)$. We now use Corollary 4 to show the concavity and supermodularity of $g_n$ and $V_n$.

First, by the monotonicity of $V_n$, we can have $h_n(x_1, x_2, d) = \max_{(u, x) \in \hat{S}} V_{n+1}(u_1, u_2)$ with $\hat{S} = \{(u, x) : u_1 \geq (x_1 - d)^+, u_2 \leq r x_2 - p(r - 1)(x_1 - d)^+ \}$, which is clearly a convex set. We next show that $\hat{S}$ is an additive mostly-sublattice with $\mathcal{W}^2$. To this end, we consider any $(u', x'), (u'', x'') \in \hat{S}$ with $x_1' < x_1''$, $x_2' > x_2''$, and define $\lambda = \frac{(r-1)p(x_1''-d)+-(x_1'-d)^+}{r(x_2''-x_2')+((r-1)p(x_1''-d)+-(x_1'-d)^+)} \in [0,1)$. We discuss three scenarios. 1) If $u_2' \leq u_2''$, we choose $y = u'$, $z = u''$. 2) If $u_2' > u_2''$ and $u_1' \geq u_1''$, we choose $y = \lambda u' + (1-\lambda) u''$, $z = (1-\lambda) u' + \lambda u''$. 3) If $u_2' > u_2''$ and $u_1' < u_1''$, we choose $y = (u_1', \lambda u_2' + (1-\lambda) u_2'')$, $z = (u_1'', (1-\lambda) u_2' + \lambda u_2'')$. In all three scenarios, we can have $y, z \in \mathcal{W}^2(u', u''), (y, x' \land x''), (z, x' \lor x'') \in \hat{S}$, and $y + z = u' + u''$. Hence, $\hat{S}$ is an additive mostly-sublattice with $\mathcal{W}^2$. Therefore, due to the concavity and supermodularity of $V_{n+1}$, we have the concavity and supermodularity of $h_n(x_1, x_2, d)$ in $(x_1, x_2)$ for any given $d$. Taking expectation, $g_n$ is concave and supermodular.

Secondly, since $V_n(t_1, t_2) = \max_{(x, t) \in S} g_n(x_1, x_2)$ and $g_n$ is concave and supermodular, to prove the concavity and supermodularity of $V_n$, it remains to show that $S$ is also an additive mostly-sublattice with $\mathcal{W}^2$. To this end, we consider any $(x', t'), (x'', t'') \in S$ with $t_1' < t_1'', t_2' > t_2''$. We discuss two scenarios. 1) If $x', x''$ are ordered or $x', x''$ are unordered with $x_1' \geq t_1''$, we define $\lambda = \frac{(p-\rho)^{t_1''}}{(t_2''-t_1'')} \in (0,1)$, and choose $y = \lambda x' + (1-\lambda)x''$, $z = (1-\lambda)x' + \lambda x''$. 2) If $x', x''$ are unordered with $x_1' < t_1''$, we define $\lambda = \frac{t_2''-t_1'}{t_2''-t_1''} \in (0,1)$, and choose $z = \lambda(t_1'', x_2') + (1-\lambda)x''$, $y = x' + x'' - z$. In both scenarios, we can have $y, z \in \mathcal{W}^2(x', x'')$, $(y, t' \land t'')$, $(z, t' \lor t'') \in S$, and $y + z = x' + x''$. Hence, $S$ is an additive mostly-sublattice with $\mathcal{W}^2$. Q.E.D.

Proof for Theorem 3

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 $t', t'' \in \mathcal{T}$, and any $x' \in S_{1'}$ and $x'' \in S_{1''}$,

$$f(x', t') + f(x'', t'') \leq f(x' \lor x'', t' \lor t'') + f(x' \land x'', t' \land t'') \leq g(t' \lor t'') + g(t' \land t''),$$

where $f(x', t')$ and $g(t' \lor t'')$ are defined as

$$f(x', t') = \mathbb{E} \left[ h_n(x_1, x_2, \tilde{d}) \right],$$

$$g(t' \lor t'') = \max_{(x, t) \in S} \mathbb{E} \left[ h_n(x_1, x_2, \tilde{d}) \right].$$

The “only if” part is shown by contradiction. If $f(x', t') + f(x'', t'') > g(t' \lor t'') + g(t' \land t'')$, then there must exist some $(x, t) \in S$ such that $h_n(x_1, x_2, \tilde{d}) > \mathbb{E} \left[ h_n(x_1, x_2, \tilde{d}) \right]$, which contradicts the definition of $g(t' \lor t'')$. Therefore, $f(x', t') + f(x'', t'') \leq g(t' \lor t'') + g(t' \land t'')$ holds.
where the first inequality follows from the supermodularity of \( f \) on \( X \times T \) and the second inequality holds by the definition of \( g \) and the fact that \( S \) is a mostly-sublattice. Taking supremum on the left hand side of the inequalities over \( x' \in S_{t'} \) and \( x'' \in S_{t''} \), we have that

\[
g(t') + g(t'') \leq g(t' \vee t'') + g(t' \wedge t'').
\]

We now prove the “only if” part. Assume to the contrary that \( S \) is not a mostly-sublattice, i.e., there exists an unordered pair \( t', t'' \in \Pi_T S \), and \( x' \in S_{t'} \) and \( x'' \in S_{t''} \) such that \((x', t') \vee (x'', t'') \notin S \) or \((x', t') \wedge (x'', t'') \notin S \). Let \( W = \{(x', t'), (x'', t''), (x', t') \vee (x'', t''), (x', t') \wedge (x'', t'')\} \) and define

\[
f(x, t) = \max_{w \in W} \{-\|w - w\|_1\}. \tag{EC.2}
\]

As the function \(-\|(x, t) - w\|_1\) is supermodular in \((x, t, w)\) (see Example 2.6.2(g) in Topkis (1998)), by Theorem 2.7.6 in Topkis (1998), \( f \) is supermodular on \( X \times T \). Since \( f(w) = 0 \) for any \( w \in W \) and \( f(w) < 0 \) for any \( w \in W \), we have \( g(t) \leq 0 \) for any \( t \in T \), and \( g(t') = g(t'') = 0 \). In addition, \( g(t' \vee t'') < 0 \) if \((x', t') \vee (x'', t'') \notin S \) or \( g(t' \wedge t'') < 0 \) if \((x', t') \wedge (x'', t'') \notin S \) (the strict inequality follows from the closedness assumption of \( S_t \)). In either case,

\[
g(t') + g(t'') > g(t' \vee t'') + g(t' \wedge t''),
\]

which implies that \( g \) is not supermodular. The “only if” part is now completed. Q.E.D.

**Proof for Theorem 4**

The “if” part is straightforward since all sublattices are mostly-sublattices. We now prove the “only if” part by contradiction. Suppose \( \exists i \in \{1, \ldots, m\} \) such that \( (a_{i_1}, \ldots, a_{i_m}, b_{i_1}, \ldots, b_{i_m}) \), \( i \in \{1, \ldots, m\} \) has two nonzero components with the same sign. For all \( j \neq i \), choose \( c_j \) large enough such that the \( j \)th constraint in defining \( S_{e_j} \) can always be satisfied for the vectors to be constructed. Hence, we specifically focus on the \( i \)th constraint. With a slight abuse of notation, denote \( a = a_i, b = b_i, c = c_i \). We discuss the following scenarios and choose \((x', t'), (x'', t'')\) accordingly.

1. \( \exists j, k \ (j \neq k) \) such that \( b_j b_k > 0 \). WLOG, we assume \( b_j, b_k > 0 \). Let \( x' = x'' = 0, t' = \frac{1}{b_j} e_j, t'' = \frac{1}{b_k} e_k, c = 1 \).
2. \( \exists j, k \ (j \neq k) \) such that \( a_j a_k > 0 \). Choose two different indexes \( l_1, l_2 \in \{1, \ldots, n_2\} \). Let \( x' = \frac{1}{a_j} e_j, \ x'' = \frac{1}{a_k} e_k, \ t' = \epsilon e_{l_1}, \ t'' = \epsilon e_{l_2}, \ c = 1.5, \) where \( \epsilon \) is a strictly positive number such that 
\[
\epsilon \max\{b_{i_1}, b_{i_2}, -(b_{i_1} + b_{i_2})\} < 0.5.
\]

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 \( l \in \{1, \ldots, n_2\} \setminus \{k\} \). Let \( x' = \frac{1}{a_j} e_j, \ x'' = 0, \ t' = \epsilon e_{l_1}, \ t'' = \frac{1}{b_k} e_k, \ c = 1.5, \) where \( \epsilon \) is a strictly positive number such that \( \epsilon |b_l| < 0.5 \).

In all of the above scenarios, we can verify that \( (x', t'), (x'', t'') \in S^p_{\mu'} \) and \( t', t'' \) are unordered, but \( (x' \lor x'', t' \lor t'') \not\in S^p_{\mu'} \) (specifically, the \( i \)th constraint is violated). Hence, the polyhedron is not a mostly-sublattice. 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-sublattice. 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-sublattice requirement. Assume to the contrary that \( \mathcal{S} \) is not a mostly-sublattice. It implies that \( \exists x' \in \mathcal{S}_{\mu'}, x'' \in \mathcal{S}_{\mu''} \) with \( t', t'' \) being unordered, \( x' \land x'' \in \mathcal{S}_{\mu' \land \mu''} \) and \( x' \lor x'' \in \mathcal{S}_{\mu' \lor \mu''} \) cannot be true simultaneously. We define a function \( f : \mathcal{X} \times \mathcal{T} \to \mathbb{R} \) as

\[
f(x, t) = \max_{w \in W} \{-\| (x, t) - w \|_1 \},
\]

where

\[
W = \text{Conv} \left( (x', t'), (x'', t''), (x' \land x'', t' \land t''), (x' \lor x'', t' \lor t'') \right).
\]

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

\[
g(t') + g(t'') > g(t' \land t'') + g(t' \lor t''),
\]
which implies that \( g \) is not supermodular.

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

**Scenario 1:** \( t' \) and \( t'' \) 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 \( X \) and \( T \) such that

\[
x' = (x'_1, x'_2), \quad x'' = (x''_1, x''_2), \quad x'_1 \geq x''_1, \quad x'_2 \leq x''_2,
\]

\[
t' = (t'_1, t'_2), \quad t'' = (t''_1, t''_2), \quad t'_1 \geq t''_1, \quad t'_2 \leq t''_2.
\]

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

\[
(x, t_{\lambda}) = \mu_1(x', t') + \mu_2(x'', t'') + \mu_3(x' \land x'', t' \land t'') + \mu_4(x' \lor x'', t' \lor t'')
\]

\[
= \mu_1(x'_1, x'_2, t'_1, t'_2) + \mu_2(x''_1, x''_2, t''_1, t''_2) + \mu_3(x'_1, x'_2, t'_1, t'_2) + \mu_4(x''_1, x''_2, t''_1, t''_2) + \mu_4(x'_1, x''_2, t'_1, t''_2)
\]

\[
= (\mu_1 + \mu_4)x'_1 + (\mu_2 + \mu_3)x'_2 + (\mu_1 + \mu_3)x''_2 + (\mu_2 + \mu_4)x''_2,
\]

\[
= (\mu_1 + \mu_4)t'_1 + (\mu_2 + \mu_3)t'_2 + (\mu_1 + \mu_3)t''_2 + (\mu_2 + \mu_4)t''_2.
\]

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

\[
x = (\lambda x'_1 + (1 - \lambda)x''_1, \lambda x'_2 + (1 - \lambda)x''_2) = \lambda x' + (1 - \lambda) x'' = x_{\lambda}.
\]

That is, \((x, t_{\lambda}) \in W \) if and only if \( x = x_{\lambda} \). If \((x_{\lambda}, t_{\lambda}) \not\in S \), there does not exist \((x, t_{\lambda}) \in W \cap S \), and thus

\[
g(t_{\lambda}) < 0 = \lambda g(t') + (1 - \lambda) g(t''),
\]

which implies that \( g \) is not concave and we have a contradiction. Hence, \((x_{\lambda}, t_{\lambda}) \in S \).

**Scenario 2:** \( t' \) and \( t'' \) are ordered but unequal. WLOG, let \( t' \leq t'' \) and \( t' \neq t'' \).
If $x' \leq x''$, define the function $f$ and set $W$ by Equations (EC.3) and (EC.4). We know that $f$ is concave and supermodular on $X \times T$, $g(t') = g(t'') = 0$. Since $(x', t') \leq (x'', t'')$, $W = \text{Conv}((x', t'), (x'', t''))$. Since $t' \neq t''$, $(x, t_\lambda) \in W$ if and only if $x = x_\lambda$. If $(x_\lambda, t_\lambda) \not\in S$, there does not exist $(x, t_\lambda) \in W \cap S$, and thus

$$g(t_\lambda) < 0 = \lambda g(t') + (1 - \lambda) g(t''),$$

which implies that $g$ is not concave and we have a contradiction. Hence, $(x_\lambda, t_\lambda) \in S$.

Now let us move to the case of $x' \not\leq x''$. Assume to the contrary that $(x_\lambda, t_\lambda) \not\in S$. We obtain the contradiction with the following steps.

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

$$K_1^\alpha = \text{Conv}(\alpha(x' \land x'') + (1 - \alpha)x'', x_\alpha), \quad K_2^\alpha = \text{Conv}(x_\alpha, \alpha x' + (1 - \alpha)(x' \lor x''))$$

by contradiction.

We start by assuming $K_2^\alpha \cap S_{t_\alpha} = \emptyset$. Let $c = x' - (x' \land x'')$, we have $c^T (x' - (x' \land x'')) = \|x' - (x' \land x'')\|_2^2 > 0$, and hence $c^T x' > c^T (x' \land x'')$. Denote $A = \{(x, t) \in S : t = t_\alpha\} \cup \{(\alpha(x' \land x'') + (1 - \alpha)x'', t_\alpha)\}$, $W = \text{Conv}((x', t'), (x'', t''), (x' \land x'', t'), (x' \lor x'', t''))$. We consider any $(x, t) \in A \cap W$ and notice that: 1) $(x, t) = \gamma_1(x', t') + \gamma_2(x' \land x'', t') + \gamma_3(x'', t'') + \gamma_4(x' \lor x'', t'')$ for some $\gamma_i \in [0, 1], i = 1, \ldots, 4$, with $\sum_{i=1}^4 \gamma_i = 1$; 2) $t = t_\alpha$; 3) $t' \neq t''$. Hence, $\gamma_1 + \gamma_2 = \alpha$, $\gamma_3 + \gamma_4 = 1 - \alpha$. WLOG, we can have $\gamma_2 \cdot \gamma_4 = 0$ since $(x' \land x'', t') + (x' \lor x'', t'') = (x', t') + (x'', t'')$. We now prove $\gamma_2 > 0$ by contradiction. Assume the contrary, i.e., $\gamma_2 = 0$. It implies

$$x = \gamma_1 x' + \gamma_3 x'' + \gamma_4 (x' \lor x'') = \frac{\gamma_3}{\gamma_3 + \gamma_4} x_\alpha + \frac{\gamma_4}{\gamma_3 + \gamma_4} (\alpha x' + (1 - \alpha)(x' \lor x'')) \in K_2^\alpha,$$

which contradicts with $K_2^\alpha \cap S_{t_\alpha} = \emptyset$. Hence, $\gamma_2 > 0$, $\gamma_4 = 0$,

$$c^T x = \gamma_1 c^T x' + \gamma_2 c^T (x' \land x'') + \gamma_4 c^T x'' < (\gamma_1 + \gamma_2) c^T x' + \gamma_3 c^T x'' = \alpha c^T x' + (1 - \alpha) c^T x'',$$

where the inequality follows from $\gamma_2 > 0$ and $c^T x' > c^T (x' \land x'')$. Observe that $c^T x' > c^T (x' \land x'')$.
max_{(x,t) \in A \cap W} f(x,t), where f(x,t) = c^T x + K \max_w \{ -\| (x,t) - w \|_1 \}. In addition, by Lemma 1, W is a lattice, and hence the function f is concave and supermodular. We get

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

which implies that g is not concave. Therefore, we have a contradiction. Hence, \( \mathcal{K}_2^o \cap \mathcal{S}_{t_o} \neq \emptyset \).

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

Therefore, \( \mathcal{S}_{t_o} \) has nonempty intersection with both \( \mathcal{K}_1^o \) and \( \mathcal{K}_2^o \).

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

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

\[ y = \beta (x_1 \lor x') + (1-\beta) (x_2 \lor x'') \]

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

Denote the index set \( \mathcal{I} = \{ i : x_i' > x_i'' \} \). Since \( x' \notin x'' \), \( \mathcal{I} \neq \emptyset \). We note that

\[ \forall i \in \mathcal{I} : y_i = \beta (x_1 \lor x_i') + (1-\beta) x_{\lambda_i} = \beta x_i' + (1-\beta) x_{\lambda_i} < x_{\lambda_i}, \]

\[ \forall i \notin \mathcal{I} : y_i = \beta (x_1 \lor x_i') + (1-\beta) x_{\lambda_i} = \beta x_i + (1-\beta) x_{\lambda_i} = x_{\lambda_i}. \]

Within \( (\lambda, \lambda_1) \), we now choose \( \lambda_2 \). By the conclusion in the first step, we can find \( z \in \mathcal{S}_{t_{\lambda_2}} \cap \mathcal{K}_{\lambda_2}^o \).

Similarly, \( \exists \gamma > 0 \) such that

\[ z = \gamma (x_2 \lor (x' \lor x'')) + (1-\gamma) x_{\lambda_2}. \]
We note that

$$
\forall i \notin \mathcal{I} : z_i = \gamma(\lambda_2 x'_i + (1 - \lambda_2)(x'_i \lor x''_i)) + (1 - \gamma)x_{\lambda_2 i} = \gamma x'_{\lambda_2 i} + (1 - \gamma)x_{\lambda_2 i} = x_{\lambda_2 i} \geq x_{\lambda_1 i} = y_i,
$$

(\text{EC.6})

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

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

and the last equality is due to the equality in (EC.5).

Moreover, by the inequality (EC.5), $y_i < x_{\lambda_1 i}$ $\forall i \in \mathcal{I}$, we can choose $\lambda_2$ close to $\lambda_1$ enough such that $x_{\lambda_1}$ and $x_{\lambda_2}$ are also very close, and hence $\forall i \in \mathcal{I}$, $y_i < x_{\lambda_2 i}$. Therefore,

$$
\forall i \in \mathcal{I} : z_i = \gamma(\lambda_2 x'_i + (1 - \lambda_2)(x'_i \lor x''_i)) + (1 - \gamma)x_{\lambda_2 i} = \gamma x'_{\lambda_2 i} + (1 - \gamma)x_{\lambda_2 i} > x_{\lambda_2 i} > y_i,
$$

(\text{EC.7})

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.6) and (EC.7), we get $z \geq y$. Moreover, recall that $y \in S_{\lambda_1} \setminus 1, z \in S_{\lambda_2} \setminus 1, t_{\lambda_1} \leq t_{\lambda_2}$ since

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

Therefore, based on the analysis at the beginning of this \textbf{Scenario 2}, we conclude that $\forall \alpha \in [0, 1],

(\alpha y + (1 - \alpha)z, \alpha t_{\lambda_1} + (1 - \alpha)t_{\lambda_2}) \in S$. Choose $\alpha = \frac{\gamma(1 - \lambda_2)}{\beta \lambda_1 + \gamma(1 - \lambda_2)}$, we can verify that

$$
\alpha y + (1 - \alpha)z = \alpha(\beta (\lambda_1 (x' \land x'')) + (1 - \lambda_1)(x'')) + (1 - \beta)x_{\lambda_1}
$$

$$
+ (1 - \alpha)(\gamma (\lambda_2 x' + (1 - \lambda_2)(x' \lor x'')) + (1 - \gamma)x_{\lambda_2})
$$

$$
= \delta x' + (1 - \delta)x''
$$

$$
= x_{\delta},
$$

$$
\alpha t_{\lambda_1} + (1 - \alpha)t_{\lambda_2} = \alpha(\lambda_1 t' + (1 - \lambda_1)t'') + (1 - \alpha)(\lambda_2 t' + (1 - \lambda_2)t'')
$$

$$
= \delta t' + (1 - \delta)t''
$$

$$
= t_{\delta},
$$

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
\[
\lambda_2 (\beta \lambda_1 + \gamma (1 - \lambda_2)) < \beta \lambda_1 \lambda_2 + \gamma \lambda_1 (1 - \lambda_2) < \lambda_1 (\beta \lambda_1 + \gamma (1 - \lambda_2)).
\]
Therefore, we have $\delta \in (\lambda_2, \lambda_1) \subset [\lambda, \tilde{\lambda}]$ such that $(x_\delta, t_\delta) \in S$, which contradicts with the conclusion in the second step.

Therefore, we conclude that when $x' \not\leq x''$, we also have $(x_\lambda, t_\lambda) \in S$.

**Scenario 3: $t' = t''$.**

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

Therefore, $\forall n = 1, 2, \ldots$, $(x^{o,n}, t^{o,n}) \in S$, where we denote
\[
(x^{o,n}, t^{o,n}) = (x'', t) + \frac{1}{n} ((x^o, t^o) - (x'', t)).
\]

Again, based on the conclusion from **Scenario 1** and **Scenario 2**, $(x^n, t^n) \in S$, where we let
\[
(x^n, t^n) = \lambda(x', t) + (1 - \lambda)(x^{o,n}, t^{o,n}).
\]

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

**Proof for Theorem 6**

We first prove the “if” part. Suppose set $S$ is an additive mostly-sublattice with $W^1$. Consider any unordered pair $t', t'' \in \Pi_T S$ and any $x' \in S_{\nu'}, x'' \in S_{\nu''}$. We abuse the notation by dropping $x', x''$ in $W^1(x', x'')$, i.e., $W^1$ is the set $\{x', x'', x' \wedge x'', x' \vee x''\}$.

If $x', x''$ are ordered, then $W^1 = \{x', x''\}$. Since there exists $y \in \text{Conv}(W^1 \cap S_{\nu' \wedge \nu''})$, we can find $\alpha \in [0, 1]$ such that $y = \alpha x' + (1 - \alpha) x''$. Moreover, there exists $z = x' + x'' - y = (1 - \alpha)x' + \alpha x''$ and $z \in \text{Conv}(W^1 \cap S_{\nu' \vee \nu''})$. For the case of $\alpha = 0$, it is easy to observe that $x'' \in S_{\nu' \wedge \nu''}$ and $x' \in S_{\nu' \vee \nu''}$.

Consider the case where $\alpha \neq 0$. If $x' \not\in S_{\nu' \wedge \nu''}$, we have $W^1 \cap S_{\nu' \wedge \nu''} \subseteq \{x''\}$ which contradicts with the fact $\alpha x' + (1 - \alpha) x'' = y \in \text{Conv}(W^1 \cap S_{\nu' \wedge \nu''})$ and $\alpha \neq 0$. Hence, we have $x' \in S_{\nu' \wedge \nu''}$. Similarly, $x'' \in S_{\nu' \vee \nu''}$. Therefore, no matter whether $\alpha = 0$ or $\alpha \neq 0$, we always have
\[
g(t' \wedge t'') + g(t' \vee t'') = \max_{z \in S_{\nu' \wedge \nu''}} f(z) + \max_{z \in S_{\nu' \vee \nu''}} f(z) \geq f(x') + f(x''). \tag{EC.8}
\]
If \( x', x'' \) are unordered, then \( \mathcal{W}^1 = \{ x', x'', x' \land x'', x' \lor x'' \} \). Since \( y \in \text{Conv}(\mathcal{W}^1 \cap \mathcal{S}_{\nu', \lambda''}) \), there exist \( \alpha, \beta, \gamma, \lambda \geq 0, \alpha + \beta + \gamma + \lambda = 1 \) and \( \lambda \times \gamma = 0 \) such that \( y = \alpha x' + \beta x'' + \gamma (x' \land x'') + \lambda (x' \lor x'') \).

The reason for \( \lambda \times \gamma = 0 \) is that if \( \lambda, \gamma > 0 \), due to \( x' + x'' = (x' \land x'') + (x' \lor x'') \), we can have \( y = \alpha x' + \beta x'' + \gamma (x' \land x'') + \lambda (x' \lor x'') = \alpha x' + \beta x'' + (\gamma \land \lambda)((x' \land x'') + (x' \lor x''))) \). 

In this case, we can have new \( \alpha^0 = \alpha + (\gamma \land \lambda), \beta^0 = \beta + (\gamma \land \lambda), \lambda^0 = (\gamma - \lambda)^+ \) and \( \gamma^0 = (\lambda - \gamma)^+ \) such that \( \alpha^0, \beta^0, \gamma^0, \lambda^0 \geq 0, \alpha^0 + \beta^0 + \gamma^0 + \lambda^0 = 1 \), and \( \lambda^0 \times \gamma^0 = 0 \). WLOG, we assume \( \lambda = 0 \) (the case with \( \lambda \neq 0 \) implies \( \gamma = 0 \) and can be proved similarly). Hence, there exists \( z = x' + x'' - y = (1 - \gamma - \alpha) x' + (1 - \gamma - \beta) x'' + \gamma (x' \land x'') = \beta x' + \alpha x'' + \gamma (x' \land x'') \) and \( z \in \text{Conv}(\mathcal{W}^1 \cap \mathcal{S}_{\nu', \lambda''}) \).

In the case of \( \gamma = 0 \), \( y = \alpha x' + (1 - \alpha) x'' \) and \( z = (1 - \alpha) x' + \alpha x'' \). Similarly to the case with \( x', x'' \) being ordered, we always have either 1) \( x'' \in \mathcal{S}_{\nu' \land \nu''} \) and \( x' \in \mathcal{S}_{\nu' \lor \nu''} \) if \( \alpha = 0 \), or 2) \( x' \in \mathcal{S}_{\nu' \land \nu''} \) and \( x'' \in \mathcal{S}_{\nu' \lor \nu''} \) if \( \alpha \neq 0 \), and thus the inequality (EC.8) holds.

In the case of \( \gamma \neq 0 \), we have \( x' \land x'' \in \mathcal{S}_{\nu' \land \nu''} \), and correspondingly, \( x' \lor x'' \in \mathcal{S}_{\nu' \lor \nu''} \). Hence,

\[
g(t' \land t'') + g(t' \lor t'') = \max_{x \in \mathcal{S}_{\nu' \land \nu''}} f(x) + \max_{x \in \mathcal{S}_{\nu' \lor \nu''}} f(x) \geq f(x' \land x'') + f(x' \lor x'') \geq f(x') + f(x''),
\]

(EC.9)

where the last inequality follows from the supermodularity of \( f \).

For both inequalities (EC.8) and (EC.9), taking the maximum of the right hand side over all \( x' \in \mathcal{S}_{\nu}, x'' \in \mathcal{S}_{\nu''} \), we have \( g(t' \land t'') + g(t' \lor t'') \geq g(t') + g(t'') \), i.e., \( g \) is supermodular.

We next prove the “only if” part. Suppose that there exists an unordered pair \( t', t'' \), and \( x' \in \mathcal{S}_{\nu}, x'' \in \mathcal{S}_{\nu''} \), such that there does not exist \( y \in \text{Conv}(\mathcal{W}^1 \cap \mathcal{S}_{\nu' \land \nu''}) \), and \( z \in \text{Conv}(\mathcal{W}^1 \cap \mathcal{S}_{\nu' \lor \nu''}) \) with \( y + z = x' + x'' \). Similarly to the proof for the “if” part, here we also abuse the notation by dropping \( x', x'' \) in \( \mathcal{W}^1(x', x'') \), i.e., \( \mathcal{W}^1 \) is the set \( \{ x', x'', x' \land x'', x' \lor x'' \} \). Let

\[
\mathcal{H} = \{ y + z : y \in \text{Conv}(\mathcal{W}^1 \cap \mathcal{S}_{\nu' \land \nu''}), z \in \text{Conv}(\mathcal{W}^1 \cap \mathcal{S}_{\nu' \lor \nu''}) \}. 
\]

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

\[
f(x) = a^T x + K \times \max_{w \in \mathcal{W}^1} \{ -\|x - w\|_1 \},
\]
where $K$ is a positive constant chosen based on Lemma 2 in the appendix such that
\[
\max_{x \in S_{t', t''}} f(x) = \max_{x \in W^1 \cap S_{t', t''}} f(x), \quad \max_{x \in S_{t', t''}} f(x) = \max_{x \in W^1 \cap S_{t', t''}} f(x).
\]
Clearly, $f$ is supermodular. However,
\[
g(t') + g(t'') \geq f(x') + f(x'') = a^T(x' + x'') > b,
\]
\[
g(t' \land t'') + g(t' \lor t'') = \max_{x \in W^1 \cap S_{t', t''}} f(x) + \max_{x \in W^1 \cap S_{t', t''}} f(x)
\]
\[
= \max_{x \in W^1 \cap S_{t', t''}} a^T x + \max_{x \in W^1 \cap S_{t', t''}} a^T x
\]
\[
= \max_{x \in \text{Conv}(W^1 \cap S_{t', t''})} a^T x + \max_{x \in \text{Conv}(W^1 \cap S_{t', t''})} a^T x
\]
\[
= \max_{y \in \text{Conv}(W^1 \cap S_{t', t''}), z \in \text{Conv}(W^1 \cap S_{t', t''})} a^T (y + z)
\]
\[
= \max_{h \in H} a^T h
\]
\[
< b,
\]
where the third equality holds since for any $x \in W^1$, $K \times \max_{w \in W^1} \{-\|x - w\|_1\} = 0$ and then $f(x) = a^T x$. Hence, we have $g(t' \land t'') + g(t' \lor t'') < g(t') + g(t'')$, i.e., $g$ is not supermodular.

Q.E.D.

**Proof for Theorem 7**

We first prove the “if” part. Consider any $(x', t'), (x'', t'') \in \mathcal{S}_x^P$ with $t', t''$ unordered.

Assume the first condition holds. If $(A|B)$ is a lattice-matrix, as shown by Theorem 4, $(x' \land x'', t' \land t'')$, $(x' \lor x'', t' \lor t'') \in \mathcal{S}_x^P$. If $(-A|B)$ is a lattice-matrix, with a similar logic, we can show that $(x' \lor x'', t' \land t'')$, $(x' \land x'', t' \lor t'') \in \mathcal{S}_x^P$. Therefore, we can either choose $y = x' \land x''$, $z = x' \lor x''$ in the first case, or $y = x' \lor x''$, $z = x' \land x''$ in the second case. We always have that $y, z$ satisfy the requirement for additive mostly-sublattice with $W^1$.

Assume the second condition holds. Let $L = \{1, \ldots, m\} \setminus I$, and denote scalars $r' = d^T x'$, $r'' = d^T x''$. WLOG, let $r' \leq r''$. By $(x', t'), (x'', t'') \in \mathcal{S}_x^P$, we have
\[
A_{Ix'} x' + B_{Ix'} t' = kr' + B_{Ix'} t' \leq c_I
\]
\[
A_{Ix''} x'' + B_{Ix''} t'' = kr'' + B_{Ix''} t'' \leq c_I
\]

(EC.10)
\[ A_{\mathcal{L}}x' \leq c_{\mathcal{L}}, \]  
\[ A_{\mathcal{L}}x'' \leq c_{\mathcal{L}} \]  
where the inequalities (EC.10) follow from the fact that \( B_{\mathcal{L}} = 0 \). If \( k|B_{\mathcal{L}} \) is a lattice-matrix, by the inequalities (EC.10) and Theorem 4, we have \( k(r' \land r'') + B_{\mathcal{L}}(t' \land t'') \leq c_{\mathcal{L}} \) and \( k(r' \lor r'') + B_{\mathcal{L}}(t' \lor t'') \leq c_{\mathcal{L}} \). Hence,

\[
A_{\mathcal{L}}x' + B_{\mathcal{L}}(t' \land t'') = kr' + B_{\mathcal{L}}(t' \land t'') = k(r' \land r'') + B_{\mathcal{L}}(t' \land t'') \leq c_{\mathcal{L}}.
\]
\[
A_{\mathcal{L}}x'' + B_{\mathcal{L}}(t' \lor t'') = kr'' + B_{\mathcal{L}}(t' \lor t'') = k(r' \lor r'') + B_{\mathcal{L}}(t' \lor t'') \leq c_{\mathcal{L}}.
\]

The above inequalities, together with (EC.11), imply that \((x', t', t'') \), \((x'', t', t'') \) \( \in S_{x}^{p} \), and we can choose \( y = x', z = x'' \) for the requirement of the additive mostly-sublattice with \( W^{1} \).

We now prove the “only if” part. Assuming that \( S_{x}^{p} \) is an additive mostly-sublattice with \( W^{1} \) \( \forall c \in \mathbb{R}^{m} \) but the second condition in Theorem 7 is violated. It suffices to show that the first condition in Theorem 7 is satisfied. We denote \( S_{x,t}^{p} = \{ x : (x, t) \in S_{x}^{p} \} \), and abuse the notation by dropping \( x', x'' \) in \( W^{1}(x', x'') \), i.e., \( W^{1} \) is the set \( \{ x', x'', x' \land x'', x' \lor x'' \} \) for given \( x', x'' \). In addition, WLOG, let \( \mathcal{I} = \{ i : b_{i} \neq 0 \} = \{ 1, \ldots, s \} \) where \( s \geq 2 \) since the second condition in Theorem 7 is violated, we first prove four statements as follows.

**Statement 1**—Each row of \( B \) has at most one nonzero. Assume to the contrary that, WLOG, \( b_{11}b_{12} \neq 0 \). Since \( a_{1} \neq 0 \) (\( A \) 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 \( a_{1}^{T}x + b_{1}^{T}t \leq c_{i} \) would not be violated. Let \( t' = e_{1}, t'' = e_{2} \). Hence, \( t' \land t'' = 0, t' \lor t'' = e_{1} + e_{2} \).

- If \( b_{11} = b_{12} \), let \( c_{1} = b_{11} = b_{12} \). Choosing \( x' = x'' = 0 \), we then have \((x', t'), (x'', t'') \in S_{x}^{p} \) and \( W^{1} = \{ 0 \} \). However, \( 0 \notin S_{x'}^{p} \) if \( b_{11} = 1 \), and \( 0 \notin S_{x''}^{p} \) if \( b_{11} = -1 \).

- If \( b_{11} = -b_{12} \), let \( c_{1} = 1 \). WLOG, let \( b_{11} = 1 \) and \( b_{12} = -1 \). Choosing \( x' = 0, x'' = 2a_{11}e_{1} \), we then have \((x', t'), (x'', t'') \in S_{x}^{p} \) and \( W^{1} = \{ 0, 2a_{11}e_{1} \} \). However, \( W^{1} \cap S_{x', t'}^{p} = W^{1} \cap S_{x'', t''}^{p} = \{ 0 \} \).

In either case, we cannot find \( y, z \) satisfying the requirement of additive mostly-sublattice with \( W^{1} \).
**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 \( b_i \) 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 \( a_i^T x + b_i^T t \leq c_i \) would not be violated. Let \( x' = 0, \ x'' = e_1, \ t' = e_1, \ t'' = e_2 \). We then have \( x' \in S_{c_1 t'}, \ x'' \in S_{c_2 t''} \), and \( \mathcal{W}^1 = \{0, e_1\} \). If \( \exists y, z \) satisfying the requirement of additive mostly-sublattice with \( \mathcal{W}^1 \), by Lemma 3, 
\[
\begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix} y = \begin{bmatrix} a_{11} \\ 0 \end{bmatrix}
\]
However, the structure of \( \mathcal{W}^1 \) implies that \( y_i = 0 \ \forall i \geq 2 \) and hence
\[
\begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix} y = y_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \text{ which is a contradiction as } a_{11} a_{21} \neq 0.
\]

**Statement 3**—\( \exists k \neq j, l_i \neq l_j \) such that \( b_{il_i} b_{jl_j} \neq 0, \ a_i, a_j \) are linearly independent. Suppose that all rows of \( a_1, \ldots, a_s \) 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 \( a_i, a_j \) are linearly independent. Denote the indexes of nonzero element of \( b_i, b_j \) by \( l_i, l_j \). If \( l_i \neq l_j \), then it is done. If \( l_i = l_j \), since \( B \) has at least two nonzero columns, there must exist another row \( k \) such that the index of the nonzero element of \( b_k \) does not equal to \( l_i \). In addition, as \( a_i, a_j \) are linearly independent, \( a_k \) must be linearly independent with at least one of \( a_i \) and \( a_j \). WLOG, let \( a_k, a_j \) 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 \( b_{il_i} b_{jl_j} \neq 0, \ a_i, a_j \) are linearly independent, we have that each row of \( a_i, a_j \) 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 \( a_1, a_2 \) are linearly independent. Since each \( b_i \) has at most one nonzero, \( b_{1q} = 0 \ \forall q \neq 1, \ b_{2q} = 0 \ \forall q \neq 2 \). Since \( a_1, a_2 \) are linearly independent, WLOG, we can assume the following submatrix
\[
\bar{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}
\]
has full rank. Since $\tilde{A}$ is of full rank, WLOG, we let $a_{11}a_{22} \neq 0$. In addition, since $\tilde{A}$ is of full rank, there exists a unique $x', x'' \in \mathbb{R}^2$ with

$$\tilde{A}x' = \begin{bmatrix} -b_{11} \\ 0 \end{bmatrix}, \quad \tilde{A}x'' = \begin{bmatrix} 0 \\ -b_{22} \end{bmatrix}.$$  \hspace{1cm} (EC.12)

Let $x' = (\tilde{x}_1', \tilde{x}_2', 0, \ldots, 0)$, $x'' = (\tilde{x}_1'', \tilde{x}_2'', 0, \ldots, 0)$, $t' = e_1$, $t'' = e_2$, $c_1 = c_2 = 0$, $c_i$ is of large value for all $i \geq 3$ and hence it suffices to focus on the first two constraints. We then have

$$\tilde{A}x' + \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} t' = \tilde{A}x'' + \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} t'' = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix},$$

and hence $x' \in S_{c,t'}^P$, $x'' \in S_{c,t''}^P$. As $S_{c,t'}^P$ is an additive mostly-sublattice with $\mathcal{W}_1$, we can find $y \in \text{Conv}(\mathcal{W}_1 \cap S_{c,t'+t''}^P)$, $z \in \text{Conv}(\mathcal{W}_1 \cap S_{c,t'+t''}^P)$, $y + z = x' + x''$. Since $y, z \in \text{Conv}(\mathcal{W}_1)$, we have $y_i = z_i = 0$, $\forall i \geq 3$. Let $\tilde{y}$ be the column vector of $y_1, y_2$; and denote $\tilde{z}$ similarly. Since $y \in \text{Conv}(S_{c,t'+t''}^P) = S_{c,t'+t''}^P$, $z \in \text{Conv}(S_{c,t'+t''}^P) = S_{c,t'+t''}^P$, by Lemma 3 we have

$$\tilde{A}\tilde{y} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} (t' \land t'') = 0 - \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} 0 = 0,$$  \hspace{1cm} (EC.13)

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

We then demonstrate that $\{y, z\} = \{x' \land x'', x' \lor x''\}$. Note that since $y \in \text{Conv}(\mathcal{W}_1)$, we have

$$y = \lambda_1 x' + \lambda_2 x'' + \lambda_3 (x' \land x'') + \lambda_4 (x' \lor x''),$$

for some $\lambda \geq 0$ with $\sum_{i=1}^3 \lambda_i = 1$. Since $x' + x'' = x' \land x'' + x' \lor x''$, we can find $\lambda$ with $\lambda_1 \lambda_2 = 0$.

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

$$y = \lambda_1 x' + \lambda_3 (x' \land x'') + \lambda_4 (x' \lor x''),$$

$$z = x' + x'' - y = \lambda_1 x'' + (1 - \lambda_1 - \lambda_3)(x' \land x'') + (1 - \lambda_1 - \lambda_4)(x' \lor x'').$$  \hspace{1cm} (EC.14)

Since $y \in \text{Conv}(\mathcal{W}_1 \cap S_{c,t'+t''}^P)$ where $\mathcal{W}_1 = \{x', x'', x' \land x'', x' \lor x''\}$, the expression (EC.14) implies $x' \in S_{c,t'+t''}^P$. Otherwise, $y \in \text{Conv}(\{x'', x' \land x'', x' \lor x''\})$ and cannot be written by expression (EC.14) with $\lambda_1 \neq 0$. Similarly, we can get $x'' \in S_{c,t'+t''}^P$ since $z \in \text{Conv}(\mathcal{W}_1 \cap S_{c,t'+t''}^P)$. 

Hence, by the first constraint of $Ax' + B(t' \land t'') \leq c$, and that of $Ax'' + B(t' \lor t'') \leq c$, we have

$$-b_{11} \leq 0, \quad b_{11} \leq 0,$$

which contradicts with $b_{11} \neq 0$.

- If $\lambda_1 = 0, \lambda_2 \neq 0$, like the previous case ($\lambda_1 \neq 0, \lambda_2 = 0$), we obtain a contradiction.

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

$$y = \lambda_3 (x' \land x'') + \lambda_4 (x' \lor x''), \quad z = \lambda_4 (x' \land x'') + \lambda_3 (x' \lor x'').$$

That implies both $x' \land x''$ and $x' \lor x''$ are elements of $S^P_{c, t', t''}$ and $S^P_{c, t' \lor t''}$. Hence, $y = x' \land x''$ and $z = x' \lor x''$ satisfy the requirement in the additive mostly-sublattice with $\mathcal{W}^1$, that implies $x' \land x'' = y = 0$ based on our analysis after Equation (EC.13). Similarly, $y = x' \lor x''$ and $z = x' \land x''$ also meet the requirement in the additive mostly-sublattice with $\mathcal{W}^1$, which implies $x' \lor x'' = y = 0$.

Therefore, we have $x' = x'' = 0$, which contradicts with Equation (EC.12). Hence, we conclude that $\lambda_3 \lambda_4 = 0$ and $\{y, z\} = \{x' \land x'', x' \lor x''\}$.

Therefore, by $y = 0$ we conclude that either $\bar{x}' = (0, x'_2), \bar{x}'' = (x'_1, 0)$, or $\bar{x}' = (x'_1, 0), \bar{x}'' = (0, x'_2)$ is true. If the first scenario is true, by equation (EC.12) we have

$$\bar{A} = \begin{bmatrix} 0 & -\frac{b_{11}}{x'_2} \\ -\frac{b_{32}}{x'_1} & 0 \end{bmatrix},$$

which contradicts with $a_{11} a_{22} \neq 0$. Therefore, we must have $\bar{x}' = (x'_1, 0), \bar{x}'' = (0, x'_2)$, and

$$\bar{A} = \begin{bmatrix} -\frac{b_{11}}{x'_1} & 0 \\ 0 & -\frac{b_{32}}{x'_2} \end{bmatrix}.$$
Now it remains to show that there is no other nonzero element in the first two rows of $A$. Assume the contrary and let $(a_{13}, a_{23})$ be a nonzero vector. Since $\hat{A}$ 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 $t' = e_1$, $t'' = e_2$, $c_1 = c_2 = 0$, and let $c_i, i \geq 3$ take values large enough such that the constraints $a_i^T x + b_i^T t \leq c_i$ 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 $x' = (-m, 2, -2, 0, \ldots, 0)^T$, $x'' = me_2$, and hence $x' \in S_{c,t'}$, $x'' \in S_{c,t''}$ and we can find $y, z$ satisfying the requirement in the additive mostly-sublattice with $W^1$. Therefore, we can find $\lambda \geq 0$ with $\sum_{i=1}^4 \lambda_i = 1$ such that $y = \lambda_1 x' + \lambda_2 x'' + \lambda_3 (x' \wedge x'') + \lambda_4 (x' \vee x'')$. 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 $x' \in S_{c,t' \wedge t''}$ and $x'' \in S_{c,t' \vee t''}$, 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 $x'' \in S_{c,t' \wedge t''}$ and $x' \in S_{c,t' \vee t''}$, which implies $-b_{22} \leq 0$ and $b_{22} \leq 0$, contradicts $b_{22} \neq 0$. Therefore, $y = \lambda_3 (x' \wedge x'') + \lambda_4 (x' \vee x')$. In addition, by Lemma 3, we have

$$\begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix} y = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} - \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} (t' \wedge t'') = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} b_1^T \\ b_2^T \end{bmatrix} 0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$ (EC.16)

Nevertheless, if $m = 1$, $x' \wedge x'' = (-1, -1, -2, 0, \ldots, 0)^T$, $x' \vee x'' = (0, 2, 0, 0, 0, \ldots, 0)^T$; if $m = -1$, $x' \wedge x'' = (0, -1, -2, 0, \ldots, 0)^T$, $x' \vee x'' = (1, 2, 0, 0, 0, \ldots, 0)^T$. In either case, we cannot have $y$ as a convex combination of $x' \wedge x''$ and $x' \vee x''$ such that Equation (EC.16) holds. Therefore, the assumption is false. There is no other nonzero element in the first two rows of $A$. 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 $a_1, a_2$ be linearly independent and $b_{11}b_{22} \neq 0$. By Statement 4, each vector of $a_1, a_2$ 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 $b_i$ as $l_i$, i.e., $b_{dl_i} \neq 0$. If $a_i$ has more than one nonzero component, then $a_i$ is linearly independent with both $a_1, a_2$, which contradicts with Statement 4. Hence, $a_i$ has only one nonzero component, say $a_{ik_i}$. By Statements 2 and 4, we can conclude $a_{ik_i} b_{dl_i}$ has the same sign as $a_{ik_i} b_{11}$.

We now prove the result for the other rows. Assume the contrary, i.e., $\exists k > s$ such that the row vector $(a_{k1}, \ldots, a_{km})$ 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 $a_i, a_j$ are of different indexes, and the nonzeros in $b_i, b_j$ are also of different indexes. WLOG, let $\{i, j\} = \{1, 2\}$, first two constraints of $Ax + Bt \leq c$ are

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

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 where $a_{11}b_{11} = 1$, the case of $a_{11}b_{11} = -1$ can be proved similarly. In constructing a counter example to the condition of the additive mostly-sublattice with $\mathcal{W}^i$, it suffices to consider the case where the only remaining constraint is the third constraint, where $b_3 = 0$, and the row vector $(a_{31}, \ldots, a_{3n_1})$ has two nonzeros with the same sign. Choose $t' = -e_1$, $t'' = -e_2$, $c_1 = c_2 = 0$. We now discuss three scenarios.

1. $a_{31}a_{32} > 0$. We let $x' = e_1$, $x'' = e_2$, $c_3 = \max\{a_{31}, a_{32}\}$.

2. $a_{31}a_{32} \leq 0$, $\exists i \in \{1, 2\}, j \geq 3$ such that $a_{3i}a_{3j} > 0$. WLOG, let $a_{31}a_{33} > 0$. We let $x' = e_1$, $x'' = e_2 + le_3$, $c_3 = \max\{a_{31}, a_{33}\}$. Since $a_{32}a_{33} \leq 0$, we can choose $l \geq 1$ such that $a_{32} + la_{33} = a_{33}$.

3. $a_{31}a_{32} \leq 0$, $a_{3i}a_{3j} \leq 0$ for all $i \in \{1, 2\}, j \geq 3$. In this case, WLOG, we have $a_{33}a_{34} > 0$. Let $x' = e_1 + l_1 e_3$, $x'' = e_2 + l_2 e_4$, $c_3 = \max\{a_{33}, a_{34}\}$. Since $a_{31}a_{33}, a_{32}a_{34} \leq 0$, we can choose $l_1, l_2 \geq 1$ such that $a_{31} + l_1 a_{33} = a_{33}$, $a_{32} + l_2 a_{34} = a_{34}$.

In all scenarios, we have

$$\begin{bmatrix} a_1^T \\
 a_2^T \end{bmatrix} x' + \begin{bmatrix} b_1^T \\
 b_2^T \end{bmatrix} t' = \begin{bmatrix} a_1^T \\
 a_2^T \end{bmatrix} x'' + \begin{bmatrix} b_1^T \\
 b_2^T \end{bmatrix} t'' = \begin{bmatrix} c_1 \\
 c_2 \end{bmatrix}$$
and \( x' \in S_{c,t', \lambda}^P, x'' \in S_{c,t''}^P \). If the condition for the additive mostly-sublattice with \( W^1 \) is satisfied, we can find \( y \in \text{Conv}(W^1 \cap S_{c,t', \lambda}^P), z \in \text{Conv}(W^1 \cap S_{c,t'', \lambda}^P) \) such that \( y + z = x' + x'' \). By Lemma 3, we have

\[
\begin{bmatrix}
    a_1^T \\
    a_2^T
\end{bmatrix} z =
\begin{bmatrix}
    c_1 \\
    c_2
\end{bmatrix} -
\begin{bmatrix}
    b_1^T \\
    b_2^T
\end{bmatrix} (t' \vee t'') =
\begin{bmatrix}
    0 \\
    0
\end{bmatrix}.
\]  

(EC.17)

Since \( a_1 = a_{11} e_1, a_2 = a_{22} e_2 \), the equality (EC.17) implies \( z_1 = z_2 = 0 \). Recall that \( z \in \text{Conv}(W^1) \). In all the four elements of \( W^1 = \{ x', x'', x' \wedge x'', x' \vee x'' \} \), except \( x' \wedge x'' \), all other elements have the first two components not equal to but no less than \( z_1, z_2 \). Therefore, \( z = x' \wedge x'' \) and hence \( y = x' + x'' - z = x' \vee x'' \). Therefore, by \( y \in \text{Conv}(W^1 \cap S_{c,t', \lambda}^P) \) and \( z \in \text{Conv}(W^1 \cap S_{c,t'', \lambda}^P) \), we have \( x' \vee x'' \in S_{c,t', \lambda}^P, x' \wedge x'' \in S_{c,t'', \lambda}^P \). Nevertheless, if \( c_3 > 0 \), in all scenarios, we cannot get \( x' \vee x'' \in S_{c,t', \lambda}^P, x' \wedge x'' \in S_{c,t'', \lambda}^P \). If \( c_3 < 0 \), in all scenarios, we cannot get \( x' \wedge x'' \in S_{c,t', \lambda}^P \). Therefore, we always have a contradiction. Hence, the first condition in Theorem 7 holds.

Q.E.D.

**Proof for Theorem 8**

We first prove the “sufficient” direction. Suppose \( f \) is supermodular and concave, and both conditions 1 and 2 hold. Consider any \( (x', t'), (x'', t'') \in S \). We abuse the notation by dropping \( x', x'' \) in \( W^2(x', x'') \), i.e., \( W^2 \) is the set \( \{ x', x'', x' \wedge x'', x' \vee x'' \} \). To prove the concavity of \( g(t) \), we need to show \( g(t_\lambda) \geq (1 - \lambda)g(t') + \lambda g(t'') \). Since \( x_\lambda \in \text{Conv}(W^2 \cap S_{t_\lambda}^P) \), following Lemma 4, we have

\[
g(t_\lambda) = \max_{x \in S_{t_\lambda}} f(x) \geq \max_{x \in W^2 \cap S_{t_\lambda}^P} f(x) \geq \lambda f(x') + (1 - \lambda) f(x'').
\]

Taking the maximum on the right hand side over all \( x' \in S_\nu \) and \( x'' \in S_\nu \), we get \( g(t_\lambda) \geq \lambda g(t') + (1 - \lambda) g(t'') \). This is equivalent to the function \( g \) being concave.

To prove the supermodularity of the function \( g \), it suffices to prove \( g(t' \wedge t'') + g(t' \vee t'') \geq g(t') + g(t'') \) for any unordered pair \( t', t'' \). Since set \( S \) is an additive mostly-sublattice with \( W^2 \), we can find \( y \in \text{Conv}(W^2 \cap S_{t_\lambda}^P), z \in \text{Conv}(W^2 \cap S_{t_\lambda}^P) \) such that \( y + z = x' + x'' \).

If \( x', x'' \) are ordered, \( W^2 = \text{Conv}(x', x'') \). In this case, \( y \in \text{Conv}(W^2 \cap S_{t_\lambda}^P) \) implies that \( \exists \beta \in [0, 1] \) such that \( y = \beta x' + (1 - \beta) x'' \). Hence, \( z = x' + x'' - y = (1 - \beta) x' + \beta x'' \). Since \( y \in \text{Conv}(W^2 \cap \)
$S' \cap t''$, there exist $y_1, y_2 \in \mathcal{W}^2 \cap S' \cap t''$ such that $y \in \text{Conv}(y_1, y_2)$. Similarly, there exist $z_1, z_2 \in \mathcal{W}^2 \cap S' \cap t''$ such that $z \in \text{Conv}(z_1, z_2)$. Therefore,

$$g(t' \land t'') + g(t' \lor t'') = \max_{x \in S' \land t''} f(x) + \max_{x \in S' \land t''} f(x) \geq \max\{f(y_1), f(y_2)\} + \max\{f(z_1), f(z_2)\} \geq (\beta f(x') + (1-\beta)f(x'')) + ((1-\beta)f(x') + \beta f(x'')) = f(x') + f(x''),$$

where the second inequality follows from Lemma 5.

If $x', x''$ 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 $y = \lambda_1 x' + \mu_1 x'' + \nu_1 (x' \land x'') + \beta_1 (x' \lor x'')$, $z = \lambda_2 x' + \mu_2 x'' + \nu_2 (x' \land x'') + \beta_2 (x' \lor x'')$. Moreover, WLOG, we can have $\nu_i \times \beta_i = 0, i = 1, 2$. The reason is similar to that for $\lambda \times \gamma = 0$ in the proof of Theorem 6. Since $y \in \text{Conv}(\mathcal{W}^2 \cap S' \land t'')$, $z \in \text{Conv}(\mathcal{W}^2 \cap S' \land t'')$, following Lemma 4, we have

$$g(t' \land t'') = \max_{x \in S' \land t''} f(x) \geq \max_{x \in \mathcal{W}^2 \cap S' \land t''} f(x) \geq \lambda_1 f(x') + \mu_1 f(x'') + \nu_1 f(x' \land x'') + \beta_1 f(x' \lor x'') \geq \lambda_2 f(x') + \mu_2 f(x'') + \nu_2 f(x' \land x'') + \beta_2 f(x' \lor x''),$$

(E.C.18)

$$g(t' \lor t'') = \max_{x \in S' \land t''} f(x) \geq \max_{x \in \mathcal{W}^2 \cap S' \land t''} f(x) \geq \lambda_2 f(x') + \mu_2 f(x'') + \nu_2 f(x' \land x'') + \beta_2 f(x' \lor x'').$$

(E.C.19)

Since $x', x''$ are unordered, based on Lemma 6, 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,

$$g(t' \land t'') + g(t' \lor t'') \geq (\lambda_1 + \lambda_2)f(x') + (\mu_1 + \mu_2)f(x'') + (\nu_1 + \nu_2)(f(x' \land x'') + (\beta_1 + \beta_2)f(x' \lor x'')$$

$$= (\lambda_1 + \lambda_2)f(x') + (\mu_1 + \mu_2)f(x'') + (\nu_1 + \nu_2)f(x' \land x'') + (\nu_1 + \nu_2)f(x' \lor x'') \geq (\lambda_1 + \lambda_2 + \nu_1 + \nu_2)f(x') + (\mu_1 + \mu_2 + \nu_1 + \nu_2)f(x'')$$

$$= f(x') + f(x''),$$

where the first inequality is a result of the inequalities (E.C.18) and (E.C.19) and the second inequality follows from the supermodularity of $f$. Therefore, $g(t' \lor t'') + g(t' \land t'') \geq f(x') + f(x'')$ for all
Taking the maximum of the right hand side over all \( x' \in S_{\nu} \) and \( x'' \in S_{\nu''} \), we have \( g(t' \land t'') + g(t' \lor t'') \geq g(t') + g(t'') \), i.e., \( g \) 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 \( x_\lambda \notin \text{Conv}(\mathcal{W}^2 \cap S_{\lambda}) \) for some \( t', t'' \in \mathcal{T} \), \( x' \in S_{\nu} \), \( x'' \in S_{\nu''} \). Therefore, according to the separating hyperplane theorem, there exist a vector \( a \in \mathbb{R}^{\dim(x')} \) and a scalar \( b \in \mathbb{R} \) such that \( a^T x_\lambda > b > a^T w \), \( \forall w \in \text{Conv}(\mathcal{W}^2 \cap S_{\lambda}) \). We define

\[
f(x) = a^T x + K \times \max_{\mathcal{w} \in \mathcal{W}^2} \{-\| x - w \|_1 \}.
\]

with a large \( K > 0 \) such that \( \max_{x \in S_{\lambda}} f(x) = \max_{x \in \mathcal{W}^2 \cap S_{\lambda}} f(x) \) (see Lemma 2). Since \( \mathcal{W}^2 \) is a lattice and convex set (see Lemma 1), the function \( f \) is concave and supermodular. Observe that \( f(x') = a^T x' \)

\[
f(x''') = a^T x'', \quad \text{with} \quad (1-\lambda)f(x'') = a^T x_\lambda > b,
\]

Therefore, the function \( g \) 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 is \( \mathcal{W}^2 = \text{Conv}(x', x'', x' \lor x'', x' \land x'') \) instead of \( \mathcal{W}^1 = \{x', x'', x' \lor x'', x' \land x''\} \) in Theorem 6. Q.E.D.

### Proof for Theorem 9

Consider any \( \beta \in \mathbb{R}^m \) and \( \alpha \in \mathbb{R}^m \) with \( B_\mathcal{I} \beta = A_\mathcal{I} \alpha \) for some \( \mathcal{I} \subseteq \{1, \ldots, m\} \). Assume that \( \beta \) and \( \mathbf{0} \) are unordered; otherwise, the statement of the theorem is trivially true. Let \( x' = \alpha, t' = 0, x'' = 0, t'' = \beta \). Define the vector \( c_i \) as follows:

\[
c_\mathcal{I} = A_\mathcal{I} x' + B_\mathcal{I} t' = A_\mathcal{I} \alpha = B_\mathcal{I} \beta = A_\mathcal{I} x'' + B_\mathcal{I} t''
\]

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

Notice that \( S^P \) is an additive mostly-sublattice with \( \mathcal{W}^2 \). We abuse the notation by dropping \( x', x'' \) in \( \mathcal{W}^2(x', x'') \), i.e., \( \mathcal{W}^2 \) is the set \( \text{Conv}(x', x'', x' \lor x'', x' \land x'') \). We can find \( y \in \mathcal{W}^2 \cap S_{\mathcal{I} \land \mathcal{I}'} \), \( z \in \mathcal{W}^2 \cap S_{\mathcal{I'} \lor \mathcal{I}'} \) with \( y + z = x' + x'' = \alpha \). Hence, by Lemma 3, we obtain

\[
A_\mathcal{I} y = c_\mathcal{I} - B_\mathcal{I}(0 \land \beta) = B_\mathcal{I} \beta - B_\mathcal{I}(0 \land \beta) = B_\mathcal{I} \beta^{	ext{op}}.
\]
We recall that $y \in W^2$, i.e., $\exists \mu_1, \mu_2, \mu_3, \mu_4 \in [0, 1]$ with $\sum_{i=1}^4 \mu_i = 1$, and

$$y = \mu_1 0 + \mu_2 \alpha + \mu_3 (0 \land \alpha) + \mu_4 (0 \lor \alpha) = \mu_2 \alpha - \mu_3 (-\alpha)^+ + \mu_4 \alpha^+ = (\mu_2 + \mu_4) \alpha^+ - (\mu_2 + \mu_3) (-\alpha)^+.$$  

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

$$B_I \beta^+ = A_I y = A_I (\lambda_1 \alpha^+ - \lambda_2 (-\alpha)^+).$$

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

$$A_I (\lambda_1 \alpha^+ - \lambda_2 (-\alpha)^+) = A_I^T 0 = 0 \neq B_I \beta^+.$$

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 $g(t) = \max \{ f(x) : (x, t) \in S_c^p \}$ is concave and supermodular when $f$ is so. From the lattice-matrix requirement and rank($B$) = 1, $B$ has at most two columns (recall that we assume $B$ does not contain any column with all zeros). Since we assume $n_2 \geq 2$, $B$ has exactly two columns, which are linearly dependent. Let $B_2 = -k B_1$ with $k > 0$. Hence,

$$g(t) = \max \{ f(x) : A x + B t \leq c \} = \max \{ f(x) : A x + B_1 (t_1 - k t_2) \leq c \} = \hat{g}(t_1 - k t_2),$$

where the function $\hat{g} : \mathbb{R} \to \mathbb{R}$ is defined as $\hat{g}(z) = \max \{ f(x) : A x + B_1 z \leq c \}$. Noticing that $\hat{g}$ is concave and supermodular, we can get $g(t) = \hat{g}(t_1 - k t_2)$ 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($B$) = 1 has been analyzed by Proposition 1. Thus, we focus on cases with rank($B$) = 2.

“S1$\rightarrow$S2”. It is the result from Theorem 9.
“S2 $\rightarrow$ S3”. $B$ being a lattice-matrix follows directly from Theorem 9. Assume S3 is false. Since $B$ is a lattice-matrix with two rows, it has at most four columns.

We first consider the case where $B$ has two columns. As in the proposition, we denote $D = B^{-1}A$, which does not have zero column since $A$ has no zero column. If $\text{rank}(D) = 1$, S3 being false implies that $d_{11}d_{21} < 0$. By the definition of $D$, since $A_1 = BD_1 = d_{11}B_1 + d_{21}B_2$, Lemma 7 implies that S2 is false. If $\text{rank}(D) = 2$, S3 being false implies that $D$ has two nonzero elements with opposite sign. If there exists $i \in \{1, \ldots, n_1\}$ such that $d_{1i}d_{2i} < 0$, considering $A_i = BD_i = d_{1i}B_1 + d_{2i}B_2$, again from Lemma 7, S2 is false. If for any $i \in \{1, \ldots, n_1\}$, $d_{1i}d_{2i} \geq 0$, together with $\text{rank}(D) = 2$ and $D$ having elements with opposite signs, we can find $i, j \in \{1, \ldots, n_1\}$ such that $D_i \geq 0$, $D_j \leq 0$, and $D_i, D_j$ are linearly independent. We can choose $\eta, \lambda > 0$ such that $\eta D_i + \lambda D_j \in \mathbb{R}^2$ and 0 are unordered. Let $\alpha = \eta e_i + \lambda e_j \geq 0$. We have $A\alpha = \eta A_i + \lambda A_j = B(\eta D_i + \lambda D_j)$, which implies that S2 is false based on Lemma 7.

We now consider the case where $B$ has three columns. WLOG, let $b_1^T = (b_{11}, b_{12}, 0)$, $b_2^T = (0, b_{22}, b_{23})$ where $b_{11}b_{12} < 0, b_{23} \neq 0$. If S3 is false, we first prove that there exists $\alpha \geq 0$ such that $A\alpha$ is independent from any column of $B$. We show this by contradiction. Suppose for any $\alpha \geq 0$, $A\alpha$ is linearly dependent on one column in $B$, which implies that $A = B_jk^T$ for some $i \in \{1, 2, 3\}$ and $k \in \mathbb{R}^{n_1}$. Since S3 is false, we have either $b_{22} = 0$ or $A = B_jk^T$ or $A = B_3k^T$. In any case, $A$ has a zero row, which contradicts with the original assumption on $A$. Hence, there exists a vector $\alpha \geq 0$ such that $A\alpha$ is independent from any column of $B$. Observing that $b_{11}b_{12} < 0, b_{23} \neq 0, b_{22}b_{23} \leq 0$, by plotting vectors $B_1, B_2, B_3$ and $A\alpha$ on a Cartesian coordinate system, we can get $A\alpha = \delta B_i + \gamma B_j$ for some $i, j \in \{1, 2, 3\}$ and $\delta \gamma < 0$. According to Lemma 7, S2 is false.

We now show that $B$ cannot have four columns. Assume to the contrary that $B$ has four columns. Since we assume it is a lattice-matrix and does not contain zero column, WLOG, we can let $b_1^T = (b_{11}, b_{12}, 0, 0)$, $b_2^T = (0, 0, b_{23}, b_{24})$ 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 $\alpha \geq 0$ such that $A\alpha$ is independent from any column of $B$. Hence, there exist some $i \in \{1, 2\}, j \in \{3, 4\}$ and $\delta \gamma < 0$ such that $A\alpha = \delta B_i + \gamma B_j$. According to Lemma 7, S2 is false.
"S3→S1". The case of $\text{rank}(B) = 1$ has been shown in Proposition 1 and hence we only focus on the case of $\text{rank}(B) = 2$. To show $\mathcal{S}_c^c$ is an additive mostly-sublattice with $\mathcal{W}^2$, by Theorem 8, it suffices to show that $g(t) = \max\{ f(x) : (x, t) \in \mathcal{S}_c^c \}$ is concave supermodular when $f$ is so. We start from the case when $B$ has two columns and denote $D = B^{-1}A$ as in the proposition. Since $B \in \mathbb{R}^{2 \times 2}$ is a lattice-matrix, for each column in $B^{-1}$, the elements in it are with the same sign. Hence, we can denote $B^{-1} = M \times \begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix}$, where $M \geq 0$, $u, v$ are either 1 or $-1$. In that case,

$$g(t) = \max \{ f(x) : Ax + Bt \leq c \}$$

$$= \max \left\{ f(x) : \begin{bmatrix} u & 0 \\ 0 & -v \end{bmatrix} p + Bt = c, \begin{bmatrix} -u & 0 \\ 0 & -v \end{bmatrix} p \in \mathbb{R}^2_+ \right\}$$

$$= \max \left\{ f(x) : -Dx - M \begin{bmatrix} u & 0 \\ 0 & v \end{bmatrix} p - t = -B^{-1}c, \begin{bmatrix} -u & 0 \\ 0 & v \end{bmatrix} p \in \mathbb{R}^2_+ \right\}$$

where $x, p, q$ are decision variables and the set $\mathcal{W}^2 = \{ p \in \mathbb{R}^2_+ : -up_1 \geq 0, -vp_2 \geq 0 \}$ is a sublattice.

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

- Consider the case of $\text{rank}(D) = 1$. We then have $k \in \mathbb{R}^{n_1}$ such that $D = D_1k^T$. Denote the function $\tilde{f}(w) = \max\{ f(x) : -k^T x = w \}$. It is concave since $f$ is concave and the set $\{ x : -k^T x = w \}$ is convex. Moreover, it is supermodular since it is a function of a scalar. Hence,

$$g(t) = \max \left\{ f(x) : -D_1k^T x + Mp + q = t, (x, p, q) \in \mathbb{R}^{n_1} \times \mathbb{R}^2 \times \{ B^{-1}c \} \right\}$$

$$= \max \left\{ \tilde{f}(w) : D_1w + Mp + q = t, (w, p, q) \in \mathbb{R} \times \mathbb{R}^2 \times \{ B^{-1}c \} \right\}.$$
We now move to the case when $B$ has three columns. WLOG, assume $B_3$ does not contain any zero element. We then have $k \in \mathbb{R}^n$ such that $A = B_3 k^T$. Like the discussion in the previous case, let $\bar{f}(w) = \max \{ f(x) : k^T x = w \}$, which is concave and supermodular. In addition,

$$g(t) = \max \{ f(x) : A x + B t \leq c \}$$

$$= \max \{ f(x) : B_3 k^T x + B_1 t_1 + B_2 t_2 + B_3 t_3 \leq c \}$$

$$= \max \{ \bar{f}(z - t_3) : B_1 t_1 + B_2 t_2 + B_3 z \leq c \},$$

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 $B$ is a lattice-matrix, $g$ is concave and supermodular. Q.E.D.

**Proof for Proposition 3**

Assume that $S_3$ of Proposition 2 holds. Here we prove that set $S_c = \{(x, t) : A x + B t = c, x \in \mathcal{D} \}$ is an additive mostly-sublattice with $\mathcal{W}^2$. Consider any $(x', t'), (x'', t'') \in S_c \subseteq S_c^\sigma$, where $S_c^\sigma$ is the polyhedron defined by Equation (2). By Statement $S_1$ of Proposition 2, there exist $y, z \in \mathcal{W}(x', x'')$ such that $y + z = x' + x''$, and $(y, t' \land t''), (z, t' \lor t'') \in S_c^\sigma$. By Lemma 3, we have

$$A y + B (t' \land t'') = c,$$

$$A z + B (t' \lor t'') = c.$$

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

**Proof for Theorem 10**

Note that $S_{r \land t''} \cap \mathcal{W}^3(x', x'') = S_{r \land t''}$ and $S_{r \lor t''} \cap \mathcal{W}^3(x', x'') = S_{r \lor t''}$ for any $x', x'', t', t''$.

We first prove the “if” part. Let $f(x) = a^T x$ for some vector $a$. Given any unordered $t', t'' \in \Pi_{r S}$ and $x' \in S_{r'}, x'' \in S_{r''}$, we can find $y \in \text{Conv}(S_{r \land t''}), z \in \text{Conv}(S_{r \lor t''})$ such that $y + z = x' + x''$.

Hence,

$$f(x') + f(x'') = a^T (x' + x'') = a^T (y + z) = f(y) + f(z) \leq g(t' \land t'') + g(t' \lor t''),$$

(EC.20)
where the second equality holds since \( y + z = x' + x'' \) and the inequality holds from Lemma 8.

Taking supremum on the left hand side over \( x' \in S_{V'} \) and \( x'' \in S_{V''} \), we obtain

\[
g(t') + g(t'') \leq g(t' \land t'') + g(t' \lor t'').
\]

Now we prove the “only if” part. Suppose there exists an unordered pair \( t', t'' \in \Pi T S \) and \( x' \in S_{V'}, x'' \in S_{V''} \), such that \( (x' + x'') \notin \mathcal{H} \), where \( \mathcal{H} = \{ y + z : y \in \text{Conv}(S_{V' \land V''}), z \in \text{Conv}(S_{V' \lor V''}) \} \) is a nonempty closed convex set. By the separating hyperplane theorem, there exists a vector \( \eta \) and a scalar \( \lambda \in \mathbb{R} \) such that \( \eta^T(x' + x'') > \lambda > \eta^T w \) for all \( w \in \mathcal{H} \). Define a linear function \( f(x) = \eta^T x \).

We have

\[
g(t') + g(t'') \geq f(x') + f(x'') = \eta^T(x' + x'') > \lambda.
\]

Moreover, since

\[
g(t' \land t'') + g(t' \lor t'') = \max\{ \eta^T y : y \in S_{V' \land V''} \} + \max\{ \eta^T z : z \in S_{V' \lor V''} \}
\]

\[
= \max\{ \eta^T(y + z) : y \in S_{V' \land V''}, z \in S_{V' \lor V''} \}
\]

\[
\leq \max\{ \eta^T w : w \in \mathcal{H} \}
\]

\[
\leq \lambda,
\]

we have

\[
g(t') + g(t'') > g(t' \land t'') + g(t' \lor t''),
\]

which implies that \( g \) is not supermodular. \( \quad \)Q.E.D.

**Proof for Theorem 11**

We first prove the “only if” part. Assume the contrary, i.e., both conditions in the theorem are not satisfied. In other words, \( n > m \), and we can find \( I \subseteq \{1, \ldots, m\} \) and \( \beta \in \mathbb{R}^{n_2} \) with \( |I| = n + 1 \), \( \text{Rank}(A_I) = n \), \( B_I \beta \in C(A_I) \), but \( B_I \beta' \notin C(A_I) \). In this case, there exists \( \alpha \in \mathbb{R}^{m_1} \) satisfying \( A_I \alpha = B_I \beta \).

Choose any \( x'' \in \mathbb{R}^{n_1}, t'' \in \mathbb{R}^{n_2} \) and denote \( x' = x'' + \alpha, t' = t'' - \beta, c_I = A_I x'' + B_I t'' \). We have

\[
A_I(x' - x'') = A_I \alpha = B_I \beta = B_I(t'' - t'),
\]
and hence \( A_I x' + B_I t' = A_I x'' + B_I t'' = c_I \). Choosing \( c_i, i \notin I \) to be sufficiently large, we have 
\((x', t'), (x'', t'') \in S_c^+ \). Moreover, neither \( \beta \geq 0 \) nor \( \beta \leq 0 \) is true since \( B_I \beta \in C(A_I) \) and \( B_I \beta^+ \notin C(A_I) \), which implies that \( t', t'' \) are unordered.

If we can find \( y, z \) with \((y, t' \land t'') \), \((z, t' \lor t'') \in S_c^p \) and \( y + z = x' + x'' \), by Lemma 3, we have 
\[ A_I y = c_I - B_I (t' \land t'') = A_I x'' + B_I t'' - B_I (t' \land t'') = A_I x'' + B_I (t'' - (t' \land t'')) = A_I x'' + B_I \beta^+. \]

It contradicts \( B_I \beta^+ \notin C(A_I) \). Hence, no such \( y, z \) exist and \( S_c^p \) is not an additive mostly-sublattice with \( W^3 \).

Now we prove the “if” part by contradiction. Suppose there exists \( c \) such that the set \( S_c^p \) is not an additive mostly-sublattice with \( W^3 \). That is, there exist \( x', x'', t', t'' \) such that \( Ax' + Bt' \leq c \), \( Ax'' + Bt'' \leq c \), but there do not exist \( y, z \) satisfying \( Ay + B(t' \land t'') \leq c \), \( Az + B(t' \lor t'') \leq c \), and \( y + z = x' + x'' \). Let \( u' = Ax' + Bt', u'' = Ax'' + Bt'' \). We then have the set 
\[ W = \{ y \in \mathbb{R}^n : (u' \land u'') - B(t' \land t'') \leq Ay \leq (u' \lor u'') - B(t' \land t'') \} = \emptyset; \]
otherwise, we can have \( y \) with \((u' \land u'') - B(t' \land t'') \leq Ay \leq (u' \lor u'') - B(t' \land t'') \), which implies 
\[ Ay + B(t' \land t'') \leq u' \lor u'' \leq c, \]
and contradicts the previous assumption of the nonexistence of \( y, z \). By \( W = \emptyset \), the following problem

\[
\text{max} \quad 0 \\
\text{s.t.} \quad \begin{bmatrix} A \\ -A \end{bmatrix} y \leq \begin{bmatrix} (u' \lor u'') - B(t' \land t'') \\ -((u' \land u'') - B(t' \land t'')) \end{bmatrix} \tag{EC.21}
\]
is infeasible. If \( n = m \), we can solve \( y \) with \( Ay = (u' \lor u'') - B(t' \land t'') \geq (u' \land u'') - B(t' \land t'') \), contradicting the problem (EC.21) being infeasible. Hence, \( n < m \). By Lemma 9, there exists \( I \subseteq \{1, \ldots, m\} \) such that \( |I| = n + 1 \), \( \text{Rank}(A_I) = n \), and

\[
\text{max} \quad 0 \\
\text{s.t.} \quad \begin{bmatrix} A_I \\ -A_I \end{bmatrix} y \leq \begin{bmatrix} (u'_I \lor u''_I) - B_I (t' \land t'') \\ -((u'_I \land u''_I) - B_I (t' \land t'')) \end{bmatrix} \tag{EC.22}
\]
is infeasible.

Therefore, the dual of problem (EC.22)

\[
\begin{align*}
\min \quad & p_1^T ((u'_x \lor u''_x) - B_I(t' \land t'')) - p_2^T ((u'_x \land u''_x) - B_I(t' \land t'')) \\
\text{s.t.} \quad & A^T_I (p_1 - p_2) = 0 \\
& p_1, p_2 \geq 0
\end{align*}
\]

is unbounded. Hence, there exist \( p_1, p_2 \geq 0 \) such that \( A^T_I (p_1 - p_2) = 0 \), and

\[
h = p_1^T ((u'_x \lor u''_x) - B_I(t' \land t'')) - p_2^T ((u'_x \land u''_x) - B_I(t' \land t''))) < 0.
\]

Moreover,

\[
\begin{align*}
0 > h &= h - (p_1 - p_2)^T A_I x' \\
&\geq p_1^T (u'_x - A_I x' - B_I(t' \land t'')) - p_2^T (u'_x - A_I x' - B_I(t' \land t'')) \\
&= (p_1 - p_2)^T (u''_x - A_I x'' - B_I(t' - (t' \land t''))) + B_I(t'' - (t' \land t'')) \\
&= (p_1 - p_2)^T B_I(t'' - (t' \land t'')) + B_I(t'' - (t' \land t'')) \tag{EC.25}
\end{align*}
\]

where in both equations the first equality follows from \( A^T_I (p_1 - p_2) = 0 \), and the second inequality follows from \( p_1, p_2 \geq 0 \). Denote \( q = p_1 - p_2 \), we have \( A^T_I q = 0 \). Let \( \Delta_1 = q^T B_I(t' - (t' \land t'')) < 0 \), \( \Delta_2 = q^T B_I(t'' - (t' \land t'')) < 0 \). We have

\[
0 = \frac{\Delta_1}{\Delta_1} - \frac{\Delta_2}{\Delta_2} = q^T B_I \beta, \tag{EC.26}
\]

where \( \beta = \frac{t' - (t' \land t'')}{\Delta_1} - \frac{t'' - (t' \land t'')}{\Delta_2} \). Note that \( q \neq 0 \); otherwise \( q^T B_I(t' - (t' \land t'')) = 0 \), which contradicts the inequality (EC.24). Moreover, since \( \text{Rank}(A^T_I) = \text{Rank}(A_I) = n \), the solution space of
\( A_T^T v = 0 \) is of dimension 1. Hence, for any \( v \) with \( A_T^T v = 0 \), we have \( v = kq \) for some \( k \in \mathbb{R} \) and hence \( v^T B_{\mathcal{I}} \beta = kq^T B_{\mathcal{I}} \beta = 0 \), which implies that 0 is the optimal value of the problem

\[
\min \ v^T B_{\mathcal{I}} \beta \\
\text{s.t.} \ A_T^T v = 0,
\]

and its dual problem

\[
\max \ 0 \\
\text{s.t.} \ A_T^T \alpha = B_{\mathcal{I}} \beta
\]

is feasible. Hence, we can find \( \alpha \) satisfying \( A_T^T \alpha = B_{\mathcal{I}} \beta \), i.e., \( B_{\mathcal{I}} \beta \in C(A_T) \). Nevertheless, since the inequality (EC.25) still holds when multiplying \( q \) by any positive value and \( \beta^+ = -\frac{t'' - (t' \wedge t'')}{\Delta_2} \), the problem

\[
\min \ v^T B_{\mathcal{I}} \beta^+ \\
\text{s.t.} \ A_T^T v = 0.
\]

is unbounded. Therefore, its dual problem

\[
\max \ 0 \\
\text{s.t.} \ A_T^T \alpha = B_{\mathcal{I}} \beta^+,
\]

is infeasible, i.e., \( B_{\mathcal{I}} \beta^+ \not\in C(A_T) \). Hence, we show that if there is no \( y, z \) such that \( Ay + B(t' \wedge t'') \leq c, Az + B(t' \vee t'') \leq c \), and \( y + z = x' + x'' \), then \( n < m \), and there exists \( \mathcal{I} \subseteq \{1, \ldots, m\} \) with \( |\mathcal{I}| = n + 1 \), \( \beta \in \mathbb{R}^m \) such that \( B_{\mathcal{I}} \beta \in C(A_T) \) and \( B_{\mathcal{I}} \beta^+ \not\in C(A_T) \). 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 \( B \) and \( C(A) \). Therefore, removing the dependent columns of \( A \) will 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 \( k \in \{1, \ldots, m\} \setminus \hat{\mathcal{I}} \) and \( d_{ki}d_{kj} > 0 \). Let

\[
\beta = \frac{1}{d_{ki}} e_i - \frac{1}{d_{kj}} e_j, \quad \alpha = A_{\mathcal{I}}^{-1} \left( \frac{1}{d_{ki}} B_{\hat{\mathcal{I}},i} - \frac{1}{d_{kj}} B_{\hat{\mathcal{I}},j} \right),
\]
where $B_{\hat{I},i}, B_{\hat{I},j}$ are the $i, j$th columns of the submatrix $B_{\hat{I}}$, respectively. Denote $\hat{I} = \hat{I} \cup \{k\}$. We have
\[
\begin{bmatrix}
B_{\hat{I}} \\
b_k^T
\end{bmatrix}
\begin{bmatrix}
\beta \\
a_k^T
\end{bmatrix} = \begin{bmatrix}
\frac{B_{\hat{I},i}}{d_{ki}} - \frac{B_{\hat{I},j}}{d_{kj}} \\
\frac{b_{ki}}{d_{ki}} - \frac{b_{kj}}{d_{kj}}
\end{bmatrix} = \begin{bmatrix}
\frac{B_{\hat{I},i}}{d_{ki}} - \frac{B_{\hat{I},j}}{d_{kj}} \\
\frac{d_{ki} + a_k^T A_{\hat{I}}^{-1} B_{\hat{I},i}}{d_{ki}} - \frac{d_{kj} + a_k^T A_{\hat{I}}^{-1} B_{\hat{I},j}}{d_{kj}}
\end{bmatrix} = \begin{bmatrix}
A_{\hat{I}} \\
a_k^T
\end{bmatrix} \begin{bmatrix}
\alpha \\
a_k^T
\end{bmatrix},
\]
where the last equality follows from the definition of $\alpha$. Hence, $B_{\hat{I}} \beta = A_{\hat{I}} \alpha \in C(A_{\hat{I}})$. However, if $d_{ki} > 0$, we have $(-d_{kj}) < 0$, and
\[
B_{\hat{I}} \beta^+ = \begin{bmatrix}
B_{\hat{I}} \\
b_k^T
\end{bmatrix} \begin{bmatrix}
e_i \\
a_k^T
\end{bmatrix} = \begin{bmatrix}
\frac{B_{\hat{I},i}}{d_{ki}} \\
\frac{b_{ki}}{d_{ki}}
\end{bmatrix} = \begin{bmatrix}
\frac{B_{\hat{I},i}}{d_{ki}} \\
\frac{a_k^T A_{\hat{I}}^{-1} B_{\hat{I},i}}{d_{ki}} + 1
\end{bmatrix} = \begin{bmatrix}
A_{\hat{I}} \\
a_k^T
\end{bmatrix} \begin{bmatrix}
\alpha \\
a_k^T
\end{bmatrix} \begin{bmatrix}
\alpha \\
1
\end{bmatrix}.
\]
On the right-hand side, the first term is in $C(A_{\hat{I}})$ but the second term is not (as $A_{\hat{I}}$ is invertible, there is no $x$ satisfying $\begin{bmatrix}
A_{\hat{I}} \\
a_k^T
\end{bmatrix} x = \begin{bmatrix}
0 \\
1
\end{bmatrix}$). Similarly, we can also prove $B_{\hat{I}} \beta^+ \notin C(A_{\hat{I}})$ in the case of $d_{ki} < 0$. Therefore, the condition of Theorem 11 cannot be satisfied.

We now prove the other direction. Assume that the condition in Theorem 11 is violated, i.e., there exist $\beta \in \mathbb{R}^{n^2}$ and $\hat{I} \subseteq \{1, \ldots, m\}$ with $|\hat{I}| = n + 1$, $\text{Rank}(A_{\hat{I}}) = n$, $B_{\hat{I}} \beta \in C(A_{\hat{I}})$ and $B_{\hat{I}} \beta^+ \notin C(A_{\hat{I}})$. Let set $\hat{I} \subseteq \hat{I}$ be the set such that $A_{\hat{I}}$ is invertible, then $|\hat{I}| = n$ and $\{k\} = \hat{I} \setminus \hat{I}$. We have
\[
B_{\hat{I}} = \begin{bmatrix}
B_{\hat{I}} \\
b_k^T
\end{bmatrix} = \begin{bmatrix}
0 \\
b_k^T - a_k^T A_{\hat{I}}^{-1} B_{\hat{I}}
\end{bmatrix} + \begin{bmatrix}
A_{\hat{I}} \\
a_k^T
\end{bmatrix} A_{\hat{I}}^{-1} B_{\hat{I}} = \begin{bmatrix}
0 \\
d_k^T
\end{bmatrix} + A_{\hat{I}} A_{\hat{I}}^{-1} B_{\hat{I}},
\]
where we denote $d_k^T = b_k^T - a_k^T A_{\hat{I}}^{-1} B_{\hat{I}}$ as in the algorithm. Therefore, for any $\gamma$, we have
\[
B_{\hat{I}} \gamma = \begin{bmatrix}
0 \\
d_k^T
\end{bmatrix} \gamma + A_{\hat{I}} A_{\hat{I}}^{-1} B_{\hat{I}} \gamma.
\]
Since the second term is always in $C(A_{\hat{I}})$, $B_{\hat{I}} \gamma \in C(A_{\hat{I}})$ if and only if $\begin{bmatrix}
0 \\
d_k^T
\end{bmatrix} \gamma \in C(A_{\hat{I}}) = C\left(\begin{bmatrix}
A_{\hat{I}} \\
a_k^T
\end{bmatrix}\right)$. It is equivalent to $d_k^T \gamma = 0$, since $A_{\hat{I}}$ is invertible. Therefore, we have $d_k^T \beta = 0$, and $d_k^T \beta^+ \neq 0$. It can only happen when $d_k$ 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. Like the proof in Theorem 10, we can show the preservation of supermodularity. Furthermore, we consider any $t', t'' \in T, x' \in S_{t'}, x'' \in S_{t''}$ and any $\lambda \in (0, 1)$. Since $x_\lambda \in \text{Conv}(S_{t_\lambda})$, there exists an integer $k$ such that $x_\lambda = \sum_{i=1}^{k} \theta_i x_i$, where $\sum_{i=1}^{k} \theta_i = 1, \theta_i \geq 0$ and $x_i \in S_{t_\lambda}, i = 1, \ldots, k$. Hence, we have

$$g(S_{t_\lambda}) = \max_{x \in S_{t'}} a^T x \geq \sum_{i=1}^{k} \theta_i a^T x_i = a^T x_\lambda = \lambda a^T (\lambda x' + (1-\lambda)x'') = \lambda a^T x' + (1-\lambda)a^T x''.$$ 

Since this inequality holds for every $x' \in S_{t'}, x'' \in S_{t''}$, we have that

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

Equivalently, the function $g$ is concave on $t \in T$.

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

$$g(t_\lambda) = \max_{x \in S_{t'}} a^T x \leq \max_{x \in \text{Conv}(S_{t_\lambda})} a^T x < \gamma < \eta^T x_\lambda = \lambda \eta^T x' + (1-\lambda)\eta^T x'' \leq \lambda g(t') + (1-\lambda)g(t'').$$

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

Proof for Proposition 4

We consider any $(x', t'), (x'', t'') \in \mathbb{R}^n \times \mathbb{R}^2_+$ such that $Ax' \leq t', Cx' \leq 0$, $Ax'' \leq t''$, $Cx'' \leq 0$, $x', x'' \geq 0$. Since $t', t'' \in \mathbb{R}^2_+$, there exist $\alpha, \beta \in [0, 1]$ such that

$$t' \wedge t'' = \alpha t' + \beta t''.$$ 

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

$$Cy = \alpha Cx' + \beta Cx'' \leq 0, \ Cz = (1-\alpha)Cx' + (1-\beta)Cx'' \leq 0,$$
and $y + z = x' + x''$. Moreover,

$$Ay = \alpha Ax' + \beta Ax'' \leq \alpha t' + \beta t'' = t' \land t''\),

and $Az = (1 - \alpha) Ax' + (1 - \beta) Ax'' \leq (1 - \alpha) t' + (1 - \beta) t'' = t' \lor t''$.

Hence, $y \in S_{t' \land t''}$ and $z \in S_{t' \lor t''}$, set $S$ is an additive mostly-sublattice with $W^3$. Q.E.D.

**Proof for Proposition 5**

Let function $q'(\gamma) \in \mathbb{R}^3$ and $q'(\alpha) \geq q(\alpha), q'(\beta) \geq q(\beta)$, and $G$ be the graph with parameters $q(\alpha), q(\beta)$ on the arcs $\alpha, \beta$, respectively. We denote the graphs $G^\alpha$ with $q(\alpha)$ replaced by $q'(\alpha)$, $G^\beta$ with $q(\beta)$ replaced by $q'(\beta)$ and $G^\alpha_\beta$ with both replacements. The quantities $\xi^\alpha, \xi^\beta, \xi_\alpha^\beta, \xi_\beta^\alpha$ are defined correspondingly. To prove supermodularity of $\mu$, by Theorem 10, it suffices to prove that for any $x^\alpha, x^\beta$ feasible on $G^\alpha, G^\beta$, respectively, we can find $y, z$ feasible on $G, G^\alpha_\beta$, respectively, such that $y + z = x^\alpha + x^\beta$.

**Case 1.** $x^\alpha(\beta) \geq \xi^\beta_\alpha(\beta)$ and $x^\beta(\beta) \leq \bar{c}(\beta)$. In this case, we have

$$\xi^\alpha_\beta(\alpha) = \xi^\alpha(\alpha) \leq x^\alpha(\alpha) \leq \bar{c}^\alpha(\alpha) = \bar{c}^{\alpha\beta}(\alpha), \quad \xi^\beta_\beta(\beta) = \xi^\beta(\beta) \leq x^\alpha(\beta) \leq \bar{c}(\beta) \leq \bar{c}^{\alpha\beta}(\beta),$$

$$\xi(\alpha) = \xi^\beta(\alpha) \leq x^\beta(\alpha) \leq \bar{c}^\beta(\alpha) = \bar{c}(\alpha), \quad \xi(\beta) \leq \xi^\beta(\beta) \leq x^\beta(\beta) \leq \bar{c}(\beta).$$

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

**Case 2.** $x^\beta(\alpha) \geq \xi^\alpha(\alpha)$ and $x^\alpha(\alpha) \leq \bar{c}(\alpha)$. It can be proved symmetrically with **Case 1**.

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

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

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

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

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

$$v = v^\alpha + v^\beta, \quad \text{(EC.27)}$$
$$v^\alpha(\gamma) \cdot v^\beta(\gamma) \geq 0, \quad \forall \gamma \in \mathcal{A},$$  

(EC.28)

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

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

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

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

For any arc $\gamma \in \mathcal{A} \setminus \{\alpha, \beta\}$, following equations (EC.27) and (EC.28), $v^\alpha(\gamma)$ 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

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

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

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

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

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

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

Moreover, we observe that

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

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

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

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

Therefore, $y$ is feasible on $G$, and $z$ is feasible on $G^{\alpha\beta}$. 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, \( Q \) is the \((k + 1) \times k\) submatrix obtained from \( A \) 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.

\[
g_{12}(y_1, y_2) = -\sum_{i=3}^{m} h_i y_i + \max \left( r + p + A^T h \right)^T v - \sum_{i=1}^{2} h_i y_i - p^T d
\]

\[
A^o v + B^o \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq c^o,
\]

where

\[
A^o = \begin{bmatrix} A \\ I_n \\ -I_n \end{bmatrix}, \quad B^o = \begin{bmatrix} -I_2 \\ 0_{(2n+m-2)\times 2} \end{bmatrix}, \quad c^o = \begin{bmatrix} 0_{2\times 1} \\ y_3 \\ \vdots \\ y_m \\ d \\ 0_{n\times 1} \end{bmatrix},
\]

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

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

The case where \( \{1, 2\} \not\subset \mathcal{I} \) is trivial, since in that case, \( B^o_{\mathcal{I}} \beta^+ \) is either \( B^o_{\mathcal{I}} \beta \) or \( 0 \), both of which are in \( C(A^o_{\mathcal{I}}) \). 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 \( \mathcal{I}_1 \subset \{1, \ldots, m\} \), \( \mathcal{I}_2 \subset \{m+1, \ldots, 2n+m\} \). That is, \( \mathcal{I}_1 \) contains the indexes of rows from \( A \), and \( \mathcal{I}_2 \) contains the indexes of other rows. In addition, let \( k + 1 = |\mathcal{I}_1| \leq \min\{m, n+1\} \), and \( |\mathcal{I}_2| = |\mathcal{I}| - |\mathcal{I}_1| = n - k \).
• Consider the case where rank($A^o_{x_2}$) ≤ $n - k - 1$, i.e., there exists a row in $A^o_{x_2}$ such that it is linearly dependent on other rows. By the structure of $A^o_{x_2}$, we can find $i, j \in \mathcal{I}_2$ such that $a^o_i = -a^o_j$. Together with rank($A^o_{x_2}$) = $n$, we have rank($A^o_{\mathcal{I}(j)}$) = $n$. Hence, we can always find $\alpha \in \mathbb{R}^n$ such that $A^o_{\mathcal{I}(j)} \alpha = B^o_{\mathcal{I}(j)} \beta^+$. Furthermore, since $b^o_i = b^o_j = 0$, we get $(b^o_i)^T \beta^+ = 0 = (b^o_j)^T \beta^+ = (a^o_i)^T \alpha = (a^o_j)^T \alpha$, where the last equality follows from $a^o_i = -a^o_j$ and $(a^o_i)^T \alpha = 0$. Therefore, we have $B^o_{x_2} \beta^+ = A^o_{x} \alpha \in C(A^o_{x})$.

• Consider the case where rank($A^o_{x_2}$) = $n - k$. By $B^o_{x} \beta \in C(A^o_{x})$ we can find $\alpha \in \mathbb{R}^n$ such that $A^o_{x} \alpha = B^o_{x} \beta$. For all $j \in \mathcal{I}_2$, we define the index $s(j) \in \{1, \ldots, n\}$ such that $a^o_j$ is either $e_{s(j)}$ or $-e_{s(j)}$. Since rank($A^o_{x_2}$) = $n - k$, $s(j), j \in \mathcal{I}_2$ take distinct values. WLOG, let $\{s(j) : j \in \mathcal{I}_2\} = \{k+1, \ldots, n\}$.

In addition, for all $j \in \mathcal{I}_2$, we have $\alpha_{s(j)} = 0$ since $b^o_j = 0$ and hence $0 = |(b^o_j)^T \beta| = |(a^o_j)^T \alpha| = |\alpha_{s(j)}|$. Therefore, $\alpha_j = 0$ for all $j \geq k + 1$, and $A^o_{x_1} \alpha = \tilde{A}^o_{x_1} \tilde{\alpha}$ where $\tilde{A}^o$ refers to the matrix consisting of the first $k$ columns of $A^o$, and $\tilde{\alpha}$ refers to the vector consisting of the first $k$ elements of $\alpha$.

Furthermore, since rank($A^o_{x_2}$) = $n$, all columns in $\tilde{A}^o_{x_1}$ are linearly independent. It implies that all columns in $\tilde{A}^o_{x_1}$ are linearly independent since $\forall j \in \mathcal{I}_2$, $a^o_j$ is either $e_s$ or $-e_s$ for some $s \geq k + 1$ and hence $\tilde{A}^o_{x_2} = 0$. Therefore, rank($\tilde{A}^o_{x_1}$) = $k$. Recall that $|\mathcal{I}_1| = k + 1$. Hence, we can find $\lambda_i, i \in \mathcal{I}_1$ such that $\sum_{i \in \mathcal{I}_1} \lambda_i \tilde{a}^o_i = 0$ and $\lambda_i, i \in \mathcal{I}_1$ are not all zero.

If at least one of $\lambda_1$ and $\lambda_2$ is nonzero, WLOG, let $\lambda_1 \neq 0$ and normalize it to $\lambda_1 = -1$. We then have $\tilde{a}^o_1 = \sum_{i \in \mathcal{I}_1 \setminus \{1\}} \lambda_i \tilde{a}^o_i$. Notice that $\forall i \in \mathcal{I}_1 \setminus \{1, 2\}$, $(\tilde{a}^o_i)^T \tilde{\alpha} = (a^o_i)^T \alpha = (b^o_i)^T \beta = 0$ since $b^o_i = 0$, we have

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

Note that the matrix $A$ satisfies the condition stated in Theorem 14. Consider its submatrix $\tilde{A}^o_{x_1} \in \mathbb{R}^{(k+1) \times k}$ whose rank is $k$. By the condition stated in Theorem 14, we have $\lambda_1 \lambda_2 \leq 0$. As $\lambda_1 = -1$, we have $\lambda_2 \geq 0$. Therefore, $\beta_1$ and $\beta_2$ have the same sign, which implies that $B^o_{x} \beta^+$ is either $0$ or $B^o_{x} \beta$, and always in $C(A^o_{x})$. 
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 \( \text{rank}(\tilde{A}_{\mathcal{I}_1 \setminus \{i\}}) = k \) and hence we must have \( \tilde{\gamma} \in \mathbb{R}^k \) with \( \tilde{A}_{\mathcal{I}_1 \setminus \{i\}}^o \tilde{\gamma} = B_{\mathcal{I}_1 \setminus \{i\}}^o \beta^+ \). Normalizing \( \lambda_i = -1 \), we have \( \tilde{a}_i^o = \sum_{j \in \mathcal{I}_1 \setminus \{i, 2\}} \lambda_j \tilde{a}_j^o = \sum_{j \in \mathcal{I}_1 \setminus \{i, 2\}} \lambda_j \tilde{a}_j^o \) since \( \lambda_1 = 2 = 0 \), and hence

\[
(\tilde{a}_i^o)^T \tilde{\gamma} = \sum_{j \in \mathcal{I}_1 \setminus \{i, 2\}} \lambda_j (\tilde{a}_j^o)^T \tilde{\gamma} = \sum_{j \in \mathcal{I}_1 \setminus \{i, 2\}} \lambda_j (b_j^o)^T \beta^+ = 0 \Rightarrow (b_i^0)^T \beta^+,
\]

where the last two equalities follow from \( b_j^o = 0 \) \( \forall j \in \mathcal{I}_1 \setminus \{1, 2\} \). Hence, we have \( \tilde{A}_{\mathcal{I}_1}^o \tilde{\gamma} = B_{\mathcal{I}_1}^o \beta^+ \).

Choosing \( \gamma \in \mathbb{R}^n \) such that \( \gamma^T = (\gamma_1, \ldots, \gamma_k, 0, \ldots, 0) \), we have \( A_{\mathcal{I}_1}^o \gamma = \tilde{A}_{\mathcal{I}_1}^o \tilde{\gamma} = B_{\mathcal{I}_1}^o \beta^+ \). In addition, since \( \tilde{A}_{\mathcal{I}_2}^o = 0 \) and \( B_{\mathcal{I}_2}^o = 0 \), \( A_{\mathcal{I}_2}^o \gamma = \tilde{A}_{\mathcal{I}_2}^o \tilde{\gamma} = 0 = B_{\mathcal{I}_2}^o \beta^+ \). Therefore, we have \( B_{\mathcal{I}}^o \beta^+ = A_{\mathcal{I}}^o \gamma \in C(A_{\mathcal{I}}^o) \).

We now prove the “only if” part by contradiction. Suppose \( A \) does not satisfy the condition in the theorem, i.e., there exists a submatrix \( Q \) with rank(\( Q \)) = \( k \) and \( \lambda \in \mathbb{R}^{k+1} \) with \( \lambda_1 \lambda_2 > 0 \) such that \( \sum_{i=1}^{k+1} \lambda_i q_i = 0 \). We will show that the function \( g \) is not supermodular. WLOG, normalize \( \lambda_1 = -1 \). Hence \( q_1 = \sum_{i=2}^{k+1} \lambda_i q_i \), where \( \lambda_2 < 0 \). WLOG, let \( Q \) be the submatrix in the upper left of \( A \), i.e., it is obtained from \( A \) by deleting the last \( (m - (k + 1)) \) rows and last \( (n - k) \) columns.

Since rank(\( Q \)) = \( k \), \( q_1 \) and \( q_2 \) cannot be both linearly dependent on \( q_3, \ldots, q_{k+1} \). WLOG, let \( q_2 \) be linearly independent from \( q_3, \ldots, q_{k+1} \). Therefore, we can find \( v^o, v^p \in \mathbb{R}^k \) such that it is perpendicular to \( q_3, \ldots, q_{k+1} \) but not \( q_2 \). WLOG, let \( \eta_2^o T v^o > 0 \). We can find \( v^o, v^p \in \mathbb{R}^k_+ \) such that both \( v^o + v^p \geq 0 \) and \( v^o + 2v^p \geq 0 \). Define \( v', v'' \in \mathbb{R}^n_+ \) 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 \( \forall i = 3, \ldots, k + 1, a_i T v' = a_i T v'' = q_i T v' \). Denote \( C \) as the lower left submatrix of \( A \) obtained by deleting the first \( (k + 1) \) rows and the last \( (n - k) \) columns. We can choose \( s' \in \mathbb{R}^{m-(k+1)}_+ \) appropriately such that

\[
s'' = s' - C v^p \in \mathbb{R}^{m-(k+1)}_+.
\]

Define \( y', y'' \in \mathbb{R}^n \) as

\[
y' = A v' + \begin{bmatrix} 0 \\ s' \end{bmatrix} = \begin{bmatrix} Q & A_{k+1} & A_{k+2} & \cdots & A_n \\ C \end{bmatrix} \begin{bmatrix} v^o + v^p \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ s' \end{bmatrix} = \begin{bmatrix} Q(v^o + v^p) \\ C(v^o + v^p) + s' \end{bmatrix},
\]

\[
y'' = A v'' + \begin{bmatrix} 0 \\ s'' \end{bmatrix} = \begin{bmatrix} Q & A_{k+1} & A_{k+2} & \cdots & A_n \\ C \end{bmatrix} \begin{bmatrix} v^o + 2v^p \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ s'' \end{bmatrix} = \begin{bmatrix} Q(v^o + 2v^p) \\ C(v^o + 2v^p) + s'' \end{bmatrix}.
\]
Then \( y' \) and \( y'' \) have the same elements except the first two. Here we have

\[
y'_1 = q_1^T(v^o + v^P) = \sum_{i=2}^{k+1} \lambda_i q_i^T(v^o + v^P) = \sum_{i=2}^{k+1} \lambda_i q_i^T v^o + \lambda_2 q_2^T v^P > \sum_{i=2}^{k+1} \lambda_i q_i^T v^o + 2\lambda_2 q_2^T v^P = y''_1,
\]

\[
y'_2 = q_2^T v^o + q_2^T v^P < q_2^T v^o + 2q_2^T v^P = y''_2,
\]

where the inequalities follow from \( q_2^T v^P > 0, \lambda_2 < 0 \).

Choose \( r = p = 0, h = \sum_{i=1}^{k+1} e_i \), and \( d = v' \vee v'' \). We then have \( g(y, d) \leq 0 \) for all \( y \). Moreover,

\[
g(y', d) = 0 \text{ since we can choose } v = v', u = \begin{bmatrix} 0 \\ s' \end{bmatrix}. \text{ Similarly, } g(y'', d) = 0.
\]

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

\[
y_i' \wedge y_i'' = a_i^T v = q_i^T \hat{v}, \quad i = 1, \ldots, k+1.
\]

(EC.29)

Since \( q_2, \ldots, q_{k+1} \) are linearly independent, the last \( k \) equations of (EC.29) has a unique solution, which is exactly \( \hat{v} = v^o + v^P \). This solution, unfortunately, does not satisfy the first equation. Hence, (EC.29) is infeasible. Therefore, for every \( (v, u, w) \in \mathcal{P}_A(y' \wedge y'', d) \), the first \( (k+1) \) elements of \( u \) cannot be all 0. As the set \( \mathcal{P}_A(y' \wedge y'', d) \) is closed, the minimal value of \( h^Tu = \sum_{i=1}^{k+1} u_i \) subject to \( (v, u, w) \in \mathcal{P}_A(y' \wedge y'', d) \) is strictly positive. Hence, \( g(y' \wedge y'', d) < 0 \),

\[
g(y', d) + g(y'', d) = 0 \geq g(y' \wedge y'', d) \geq g(y' \wedge y'', d) + g(y' \vee y'', d),
\]

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

**Proof for Theorem 15**

We first prove the “only if” direction. Suppose \( A \) has a \( 3 \times 2 \) submatrix \( \hat{A} \) which contains no pair of rows that are linearly dependent. In this case, \( \exists \lambda \in \mathbb{R}^3 \) with \( \hat{A}^T \lambda = 0, \lambda_i \neq 0 \ \forall i \in \{1, 2, 3\} \).

Hence, there are two elements in \( \lambda \) with the same sign. WLOG, let \( \lambda_1 \lambda_2 > 0 \). We also let \( \hat{A} \) be the submatrix in the upper left of \( A \), i.e., it is obtained from \( A \) by deleting all rows except the
first three and deleting all columns except the first two. However, \( g(y, d) \) is not supermodular in \((y_1, y_2)\) since the condition in Theorem 14 is violated when we choose \( k = 2 \) and \( Q = \hat{A} \).

We now prove the “if” direction by showing that for all pairs of distinct indexes \( i, j \in \{1, \ldots, n\} \), the function \( g(y, d) \) is supermodular in \((y_i, y_j)\). To this end, following Theorem 14, we consider any \((k + 1) \times k\) submatrix of \( Q \). With \( \text{rank}(Q) = k \), it suffices to show that for any \( \lambda \in \mathbb{R}^{k+1} \) satisfying \( Q^T \lambda = 0 \), we have \( \lambda_i \lambda_j \leq 0 \) for all distinct indexes \( i, j \in \{1, \ldots, k + 1\} \). By Lemma 11, for such \( Q \) (i.e., it is with rank \( k \) and its every \( 3 \times 2 \) submatrix contains at least two row vectors which are linearly dependent), we always have \( \gamma q_s = q_t \) for some \( \gamma \in \mathbb{R}^+ \) and distinct indexes \( s, t \in \{1, \ldots, k + 1\} \), where \( q_s^T \) (\( q_t^T \)) is the \( s \)th (\( t \)th) row vector in \( Q \). WLOG, let \( s = 1, t = k + 1 \), i.e., \( \gamma q_1 = q_{k+1} \). We then have

\[
0 = Q^T \lambda = \sum_{i=1}^{k+1} \lambda_i q_i^T = \sum_{i=1}^k \hat{\lambda}_i q_i^T, \tag{EC.30}
\]

where \( \hat{\lambda}_1 = \lambda_1 + \gamma \lambda_{k+1}, \hat{\lambda}_j = \lambda_j, j = 2, \ldots, k \). Since \( \text{rank}(Q) = k \) and \( \gamma q_1 = q_{k+1} \), vectors \( q_1, \ldots, q_k \) are linearly independent. Hence, the equality (EC.30) implies \( \hat{\lambda} = 0 \), 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 \( i, j \in \{1, \ldots, k + 1\} \) such that \( \{i, j\} \neq \{1, k + 1\} \). Therefore, we have \( \lambda_i \lambda_j \leq 0 \) for all pairs of distinct indexes \( i, j \in \{1, \ldots, k + 1\} \). The proof is complete. Q.E.D.

**Theorem 16** For the general case where \( g(t) = \max_{x \in S_t} f(x, t) \), the function \( g \) is log-supermodular on \( \Pi_T S \) whenever \( f \) is log-supermodular on \( \mathcal{X} \times \mathcal{T} \) if and only if \( S \) is a mostly-sublattice.

**Proof.** The proof is like that for Theorem 3. For completeness, we briefly describe the proof here.

For the “if” direction, we consider any unordered pair \( t', t'' \in \Pi_T S \), and any \( x' \in S_{t'} \) and \( x'' \in S_{t''} \). We then have

\[
f(x', t') \times f(x'', t'') \leq f(x' \vee x'', t' \vee t'') \times f(x' \wedge x'', t' \wedge t''), \leq g(t' \vee t'') \times g(t' \wedge t''),
\]

where the first inequality follows from the log-supermodularity of \( f \) on \( \mathcal{X} \times \mathcal{T} \) and the second inequality holds by the definition of \( g \) and the fact that \( S \) is a mostly-sublattice. Taking supremum on the left-hand side of the inequalities over \( x' \in S_{t'} \) and \( x'' \in S_{t''} \), we have that

\[
g(t') \times g(t'') \leq g(t' \vee t'') \times g(t' \wedge t'').
\]
Taking log on both sides we can have the log-supermodularity of $g$.

For the “only if” direction, we assume to the contrary that $S$ is not a mostly-sublattice, i.e., there exists an unordered pair $t', t'' \in \Pi_T S$, and $x' \in S_{t'}$ and $x'' \in S_{t''}$ such that $(x', t') \vee (x'', t'') \not\in S$ or $(x', t') \land (x'', t'') \not\in S$. Let $W = \{ (x', t'), (x'', t''), (x', t') \vee (x'', t''), (x', t') \land (x'', t'') \}$ and define

$$f(x, t) = \exp \left( \max_{w \in W} \{-\| (x, t) - w \|_1 \} \right),$$

which is log-supermodular since $\log(f)$ is supermodular by the proof for Theorem 3. Since $f(w) = 1$ for any $w \in W$ and $f(w) < 1$ for any $w \not\in W$, we have $g(t) \leq 1$ for any $t \in T$, and $g(t') = g(t'') = 1$.

In addition, $g(t' \lor t'') < 1$ if $(x', t') \lor (x'', t'') \not\in S$ or $g(t' \land t'') < 1$ if $(x', t') \land (x'', t'') \not\in S$ (the strict inequality follows from the closedness assumption of $S_\ell$). In either case,

$$g(t') \times g(t'') = 1 > g(t' \lor t'') \times g(t' \land t''),$$

which implies that $g$ is not log-supermodular. The “only if” part is now completed. Q.E.D.

**Lemma 1** $\text{Conv}(x', x'', x' \land x'', x' \lor x'')$ is a lattice.

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

$$x' \land x'' = (x''_1, x'_2), \quad x' \lor x'' = (x'_1, x''_2).$$

Denote $W = \text{Conv}(x', x'', x' \land x'', x' \lor x'') = \text{Conv}((x'_1, x'_2), (x''_1, x''_2), (x''_1, x'_2), (x'_1, x''_2))$.

If $y', y'' \in W$, 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

$$y' = \lambda_1 x' + \lambda_2 x'' + \lambda_3 (x' \land x'') + \lambda_4 (x' \lor x'') = ((\lambda_1 + \lambda_4) x'_1 + (\lambda_2 + \lambda_3) x''_1, (\lambda_1 + \lambda_3) x'_2 + (\lambda_2 + \lambda_4) x''_2),$$

$$y'' = \mu_1 x' + \mu_2 x'' + \mu_3 (x' \land x'') + \mu_4 (x' \lor x'') = ((\mu_1 + \mu_4) x'_1 + (\mu_2 + \mu_3) x''_1, (\mu_1 + \mu_3) x'_2 + (\mu_2 + \mu_4) x''_2).$$
Take the meet and get

\[ y' \land y'' = \left( (\lambda_1 + \lambda_4)x'_1 + (\lambda_2 + \lambda_3)x''_1 \right) \land \left( (\mu_1 + \mu_4)x'_1 + (\mu_2 + \mu_3)x''_1 \right), \]

\[ (\lambda_1 + \lambda_3)x'_2 + (\lambda_2 + \lambda_4)x''_2 \land (\mu_1 + \mu_3)x'_2 + (\mu_2 + \mu_4)x''_2. \]

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 \( x'_1 > x''_1 \), we have

\[ ((\lambda_1 + \lambda_3)x'_1 + (\lambda_2 + \lambda_3)x''_1) \land ((\mu_1 + \mu_4)x'_1 + (\mu_2 + \mu_3)x''_1) = (\mu_1 + \mu_4)x'_1 + (\mu_2 + \mu_3)x''_1. \]

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

\[ ((\lambda_1 + \lambda_3)x'_2 + (\lambda_2 + \lambda_4)x''_2) \land ((\mu_1 + \mu_3)x'_2 + (\mu_2 + \mu_4)x''_2) = (\mu_1 + \mu_3)x'_2 + (\mu_2 + \mu_4)x''_2, \]

and \( y' \land y'' = y'' \in \mathcal{W} \). If \( \lambda_1 + \lambda_3 > \mu_1 + \mu_3 \), we have

\[ ((\lambda_1 + \lambda_3)x'_2 + (\lambda_2 + \lambda_4)x''_2) \land ((\mu_1 + \mu_3)x'_2 + (\mu_2 + \mu_4)x''_2) = (\lambda_1 + \lambda_3)x'_2 + (\lambda_2 + \lambda_4)x''_2, \]

then

\[ y' \land y'' = ((\mu_1 + \mu_4)x'_1 + (\mu_2 + \mu_3)x''_1, (\lambda_1 + \lambda_3)x'_2 + (\lambda_2 + \lambda_4)x''_2). \]

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

\[ m_2 = m_1 - (\lambda_1 + \lambda_3) + \mu_2 + \mu_3, \]

\[ m_3 = -m_1 + \lambda_1 + \lambda_3, \]

\[ m_4 = -m_1 + \mu_1 + \mu_4. \]

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

\[ y' \land y'' = m_1(x'_1, x''_2) + m_2(x''_1, x'_2) + m_3(x''_1, x'_2) + m_4(x'_1, x''_2) \]

and we get \( y' \land y'' \in \mathcal{W} \).

Similarly, we can show that \( y' \lor y'' \in \text{Conv}(x', x'', x' \land x'', x' \lor x'') \).

Q.E.D.
Lemma 2. For any closed sets \( \mathcal{A}, \mathcal{H} \subseteq \mathbb{R}^n \) with \( \mathcal{A} \cap \mathcal{H} \neq \emptyset \) and \( c \in \mathbb{R}^n \), there exists a \( K \geq 0 \) such that

\[
\max_{x \in \mathcal{A}} f(x) = \max_{x \in \mathcal{A} \cap \mathcal{H}} f(x),
\]

where the function \( f : \mathcal{A} \rightarrow \mathbb{R} \) is defined as

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

Proof. By optimality, we observe that for any \( K > 0 \),

\[
\max_{x \in \mathbb{R}} \left\{ c^T x - K \min_{h \in \mathcal{H}} \|x - h\|_1 - K \min_{h \in \mathcal{A}} \|x - h\|_1 \right\} \geq \max_{x \in \mathcal{A}} \left\{ c^T x - K \min_{h \in \mathcal{H}} \|x - h\|_1 \right\} \geq \max_{x \in \mathcal{A} \cap \mathcal{H}} \left\{ c^T x \right\}.
\]

(EC.31)

Note that both sets \( \mathcal{A}, \mathcal{H} \subseteq \mathbb{R}^n \) 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 \( K \geq \bar{K} \),

\[
\max_{x \in \mathbb{R}} \left\{ c^T x - K \min_{h \in \mathcal{H}} \|x - h\|_1 - K \min_{h \in \mathcal{A}} \|x - h\|_1 \right\} = \max_{x \in \mathcal{A} \cap \mathcal{H}} c^T x.
\]

(EC.32)

Based on relationships in (EC.31) and (EC.32), the lemma is proved. Q.E.D.

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

Proof. We note that

\[
a^T y = a^T (x' + x'' - z) + b^T (t' + t'' - (t' \wedge t'')) - b^T (t' \wedge t'')
\]

\[
\geq 2c - (a^T z + b^T (t' \vee t'')) - b^T (t' \wedge t'')
\]

\[
= c - b^T (t' \wedge t''),
\]

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

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

$$\max_{x \in \mathcal{N}} f(x) \geq af(x') + bf(x'') + cf(x' \land x'') + df(x' \lor x'').$$

**Proof.** The case where \(x', x''\) are ordered can be derived directly from Lemma 5. Now we just consider the case where \(x', x''\) are unordered. Since \(x^o \in Conv(\mathcal{N})\), we have there exists \(x_1, \ldots, x_n \in \mathcal{N}\) and \(\beta \in \mathbb{R}_+^n\) such that

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

Moreover, since \(x_1, \ldots, x_n \in \mathcal{N} \subseteq \mathcal{W} = Conv(x', x'', x' \land x'', x' \lor x'')\), we have

$$\begin{align*}
x_1 &= \alpha_1^1 x' + \alpha_2^2 x'' + \alpha_3^3 (x' \land x'') + \alpha_4^4 (x' \lor x''), \\
\vdots \\
x_n &= \alpha_1^n x' + \alpha_2^n x'' + \alpha_3^n (x' \land x'') + \alpha_4^n (x' \lor x''), \\
\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 &= 1, \\
\alpha_1, \alpha_2, \alpha_3, \alpha_4 &\geq 0,
\end{align*}$$

where \(\alpha_i\) is the column vector of \((\alpha_1^i, \ldots, \alpha_n^i)^T\). Therefore,

$$x^o = \beta^T \alpha_1 x' + \beta^T \alpha_2 x'' + \beta^T \alpha_3 (x' \land x'') + \beta^T \alpha_4 (x' \lor x'') = ax' + bx'' + c(x' \land x'') + d(x' \lor x'').$$

WLOG, let \(x' = (x'_1, x'_2)\), \(x'' = (x''_1, x''_2)\) such that \(x'_1 < x''_1\), \(x'_2 \geq x''_2\) and \(x'_2 \neq x''_2\). Then we have \(x' \land x'' = (x'_1, x'_2)\), \(x' \lor x'' = (x''_1, x''_2)\), and

$$0 = (\beta^T \alpha_1 - a)x' + (\beta^T \alpha_2 - b)x'' + (\beta^T \alpha_3 - c)(x' \land x'') + (\beta^T \alpha_4 - d)(x' \lor x'')$$

$$= ((\beta^T \alpha_1 + \beta^T \alpha_3 - a - c)x'_1 + (\beta^T \alpha_2 + \beta^T \alpha_4 - b - d)x'_2)$$

$$= (\beta^T \alpha_1 + \beta^T \alpha_3 - a - c)(x'_1 - x''_1), (\beta^T \alpha_1 + \beta^T \alpha_3 - a - d)(x'_2 - x''_2),$$

where the last equality holds since \(\beta^T (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) = \beta^T 1 = 1 = a + b + c + d\). Hence, we have

$$\begin{align*}
\beta^T (\alpha_1 + \alpha_3) &= a + c, \\
\beta^T (\alpha_2 + \alpha_4) &= a + d,
\end{align*}$$

$$\begin{align*}
\beta^T \alpha_1 &= a + c - \beta^T \alpha_3, \\
\beta^T \alpha_2 &= b + c - \beta^T \alpha_3, \\
\beta^T \alpha_4 &= d - c + \beta^T \alpha_3.
\end{align*}$$
Therefore, based on the concavity of function $f$, we can get

$$
\max \{ f(x_1), \ldots, f(x_n) \} 
\geq
\max \{ \alpha_1 f(x') + \alpha_2 f(x''), \alpha_3 f(x' \wedge x'') + \alpha_4 f(x' \vee x''), \ldots, \\
\alpha_1 f(x') + \alpha_2 f(x'') + \alpha_3 f(x' \wedge x'') + \alpha_4 f(x' \vee x'') \}
$$

$$
\geq
\beta_1 (\alpha_1 f(x') + \alpha_2 f(x'')) + \alpha_3 f(x' \wedge x'') + \alpha_4 f(x' \vee x'') + \ldots
$$

$$
+ \beta_n (\alpha_1 f(x') + \alpha_2 f(x'') + \alpha_3 f(x' \wedge x'') + \alpha_4 f(x' \vee x''))
$$

$$
= \beta^T \alpha_1 f(x') + \beta^T \alpha_2 f(x'') + \beta^T \alpha_3 f(x' \wedge x'') + \beta^T \alpha_4 f(x' \vee x'').
$$

Hence, we have

$$
\max \{ f(x_1), \ldots, f(x_n) \} - (af(x') + bf(x'') + cf(x' \wedge x'') + df(x' \vee x'')) 
\geq (\beta^T \alpha_1 - a) f(x') + (\beta^T \alpha_2 - b) f(x'') + (\beta^T \alpha_3 - c) f(x' \wedge x'') + (\beta^T \alpha_4 - d) f(x' \vee x'').
$$

$$
= (\beta^T \alpha_3 - c) (f(x' \wedge x'') + f(x' \vee x'') - f(x') - f(x'')).
$$

(EC.33)

Recall that $c \times d = 0$ and $\beta^T \alpha_4 = d - c + \beta^T \alpha_3$. Therefore, we have $\beta^T \alpha_3 - c \geq 0$. Together with the supermodularity of $f$, we know that the right hand side of Equation (EC.33) is non-negative.

Hence,

$$
\max_{x \in \mathcal{X}} f(x) \geq \max \{ f(x_1), \ldots, f(x_n) \} \geq af(x') + bf(x'') + cf(x' \wedge x'') + df(x' \vee x'').
$$

Q.E.D.

**Lemma 5** Given $x_1, x_2$, we consider any concave function $f(x)$ defined on $\text{Conv}(x_1, x_2)$. Given $x_\lambda = \lambda x_1 + (1 - \lambda)x_2$, where $\lambda \in [0, 1]$ and $y, z \in \text{Conv}(x_1, x_2)$, if $x_\lambda \in \text{Conv}(y, z)$, then we must have $\max \{ f(y), f(z) \} \geq \lambda f(x_1) + (1 - \lambda)f(x_2)$.

**Proof.** Since $y, z \in \text{Conv}(x_1, x_2)$, $\exists \beta, \mu \in [0, 1]$ such that $y = \beta x_1 + (1 - \beta)x_2$, $z = \mu x_1 + (1 - \mu)x_2$.

Moreover, as $x_\lambda \in \text{Conv}(y, z)$, we get $\lambda \in \text{Conv}(\beta, \mu)$. Therefore,

$$
\max \{ f(y), f(z) \} = \max \{ f(\beta x_1 + (1 - \beta)x_2), f(\mu x_1 + (1 - \mu)x_2) \}
$$

$$
\geq \max \{ \beta f(x_1) + (1 - \beta)f(x_2), \mu f(x_1) + (1 - \mu)f(x_2) \}
$$

$$
\geq \lambda f(x_1) + (1 - \lambda)f(x_2).
$$
Lemma 6 For any unordered \( x', x'' \), if \( y + z = x' + x'' \) and \( y, z \in \text{Conv}(x', x'', x' \land x'', x' \lor x'') \) 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 \( y = \lambda_1 x' + \mu_1 x'' + \nu_1(x' \land x'') + \beta_1(x' \lor x'') \), \( z = \lambda_2 x' + \mu_2 x'' + \nu_2(x' \land x'') + \beta_2(x' \lor x'') \). We have

\[
\begin{align*}
\mu_1 + \nu_1 + \mu_2 + \nu_2 &= \lambda_1 + \nu_1 + \lambda_2 + \nu_2 = 1, \\
\beta_1 + \beta_2 &= \nu_1 + \nu_2.
\end{align*}
\]

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

\[
y + z = (\lambda_1 + \lambda_2)x' + (\mu_1 + \mu_2)x'' + (\nu_1 + \nu_2)(x' \land x'') + (\beta_1 + \beta_2)(x' \lor x'') \\
= (\lambda_1 + \lambda_2)x' + (\mu_1 + \mu_2 + \nu_1 + \nu_2)(x' + x'' - x' \land x'') + (\beta_1 + \beta_2)(x' \lor x'') \\
= (\lambda_1 + \nu_1 + \lambda_2 + \nu_2)(x'_1, x'_2) + (\mu_1 + \nu_1 + \mu_2 + \nu_2)(x''_1, x''_2) + (\beta_1 - \nu_1 + \beta_2 - \nu_2)(x'_1, x''_2) \\
= \left( (\lambda_1 + \nu_1 + \lambda_2 + \nu_2)x'_1 + (\mu_1 + \nu_1 + \mu_2 + \nu_2)x''_1, \\
(1 + \nu_1 + \lambda_2 + \nu_2)x'_2 + (\mu_1 + \nu_1 + \mu_2 + \nu_2)x''_2 \right)
\]

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

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

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

Q.E.D.

Lemma 7 Statement S2 in Proposition 2 cannot hold if there exist \( j, k \in \{1, \ldots, n_2\}, \alpha \geq 0 \) such that \( B_j, B_k \) are linearly independent, and \( A\alpha = \delta B_j + \gamma B_k \) for some \( \delta, \gamma < 0 \).

Proof. We prove by contradiction. WLOG, assume \( A\alpha = \delta B_1 + \gamma B_2 \) with \( B_1 \) and \( B_2 \) linearly independent, and \( \delta > 0, \gamma < 0, \alpha \geq 0 \). To show that S2 is false, we choose \( \beta = \delta e_1 + \gamma e_2 \). Then
\( A \alpha = B \beta \). If there exist \( \lambda_1, \lambda_2 \in [0, 1] \) such that \( B \beta^+ = A(\lambda_1 \alpha^+ - \lambda_2 (\alpha)^+) \) = \( \lambda_1 A \alpha = \lambda_1 B \beta \), we have \( B \beta^+ = B_1 \delta = \lambda_1 A \alpha = \lambda_1 \delta B_1 + \lambda_1 \gamma B_2 \), i.e., \( \delta (1 - \lambda_1) B_1 = \lambda_1 \gamma B_2 \). Since \( \delta \gamma \neq 0 \), it contradicts with the assumption that \( B_1, B_2 \) are linearly independent.

Q.E.D.

**Lemma 8** Given a set \( \mathcal{A} \subseteq \mathbb{R}^n \) and a convex function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \), for any \( x \in \text{Conv}(\mathcal{A}) \), we have

\[
\forall y \in \mathcal{A}, \quad f(x) \leq \sup_{y \in \mathcal{A}} f(y).
\]

**Proof.** For any \( x \in \text{Conv}(\mathcal{A}) \), we have \( x = \sum_{i=1}^{m} \lambda_i y_i \) for certain \( y_i \in \mathcal{A}, \lambda_i \geq 0, i = 1, 2, \ldots, m \) and \( \sum_{i=1}^{m} \lambda_i = 1 \). Hence,

\[
f(x) = f \left( \sum_{i=1}^{m} \lambda_i y_i \right) \leq \sum_{i=1}^{m} \lambda_i f(y_i) \leq \sup_{y \in \mathcal{A}} f(y)
\]

where the first inequality follows from the convexity of \( f \).

Q.E.D.

**Lemma 9** Consider any matrix \( A \in \mathbb{R}^{m \times n_2} \) with \( n = \text{Rank}(A) < m \). Suppose that the system of inequalities \( Ax \leq \bar{b}, -Ax \leq -\bar{b} \) is infeasible. We can choose \( I \subseteq \{1, \ldots, m\} \) with \( |I| = n + 1 \) and \( \text{Rank}(A_I) = n \) such that the system of inequalities \( A_I x \leq \bar{b}_I, -A_I x \leq -\bar{b}_I \) 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 \( 0^T x \). 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 obtain the desired infeasible system of inequalities.

Q.E.D.

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

\[
x = x^\alpha + x^\beta
\]

\[
x^\alpha(\gamma) \cdot x^\beta(\gamma) \geq 0, \quad \forall \gamma \in \mathcal{A}
\]

\[
x^\alpha(\beta) = x^\beta(\alpha) = 0.
\]
Proof. From the Cyclic Decomposition Lemma (see, for instance, Gale and Politof 1981), we can decompose $x = \sum_{i=1}^{n} k_i x_{\Gamma_i}$ for some $k_i > 0 \forall i = 1, \ldots, n$, and cycles $x_{\Gamma_i}$ (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 $x^\alpha = \sum_{i=1}^{r} k_i x_{\Gamma_i}$, $x^\beta = \sum_{i=r+1}^{n} k_i x_{\Gamma_i}$. We then have $x = x^\alpha + x^\beta$. In addition, $\forall \gamma \in \mathcal{A}$, we note that $k_i x_{\Gamma_i}(\gamma)$ is with the same sign as $x(\gamma)$, hence,

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

Moreover, by the way we construct $r$, $x^\beta(\alpha) = \sum_{i=r+1}^{n} x_{\Gamma_i}(\alpha) = 0$. Consider any $i = 1, \ldots, r$, since $x_{\Gamma_i}(\beta)$ has the same sign as $x(\beta)$, we have $x_{\Gamma_i}(\beta) \leq 0$. Since $\alpha, \beta$ 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 $x^\alpha(\beta) = \sum_{i=1}^{r} k_i x_{\Gamma_i}(\beta) = 0$. Q.E.D.

**Lemma 11** Consider any matrix $Q \in \mathbb{R}^{(k+1) \times k}$ with rank$(Q) = k$, $k \geq 2$. If every $3 \times 2$ submatrix of $Q$ contains at least two row vectors which are linearly dependent, then $Q$ 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 $3 \times 2$ 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 $Q \in \mathbb{R}^{3 \times 2}$ satisfies 3-2 condition. Suppose the argument in the lemma is true when $k = r - 1$, $r \in \{3, 4, \ldots \}$. We now prove the case for $k = r$ by considering such $Q \in \mathbb{R}^{(r+1) \times r}$.

We denote the submatrix consisting of the first $(r - 1)$ columns of $Q$ by $\hat{Q} \in \mathbb{R}^{(r+1) \times (r-1)}$. We have rank$(\hat{Q}) = r - 1$ since all columns in it are independent. Hence, $\hat{Q}$ has $(r - 1)$ independent rows. WLOG, let $\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_{r-1}$ be independent and $Q^1 \in \mathbb{R}^{r \times (r-1)}$ be the submatrix of $\hat{Q}$ 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 $Q^1$, i.e., $\hat{q}_i, \hat{q}_j$ are dependent for some distinct indexes $i, j \in \{1, \ldots, r\}$. Since $\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_{r-1}$ are independent, the two dependent rows cannot be
both in the first \((r - 1)\) rows. Therefore, we have \(\hat{q}_r = \lambda \hat{q}_i\) for some \(i \in \{1, \ldots, r - 1\}\), \(\lambda \geq 0\). WLOG, let \(\hat{q}_r = \lambda \hat{q}_1\). Similarly, by letting \(\hat{Q}^2 \in \mathbb{R}^{r \times (r-1)}\) be the submatrix of \(\hat{Q}\) by deleting only the second to the last row, i.e., the \(r\)th row, we can conclude \(\hat{q}_{r+1} = \delta \hat{q}_i\) for some \(i \in \{1, \ldots, r - 1\}\), \(\delta \geq 0\). If \(\lambda = \delta = 0\), then \(\hat{q}_r = \hat{q}_{r+1} = \mathbf{0}\), and hence \(q_r, q_{r+1}\) 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 \(\lambda, \delta\) is nonzero. WLOG, we assume \(\lambda > 0\), and normalize it to \(\lambda = 1\), i.e., \(\hat{q}_r = \hat{q}_1\). If \(q_{1r} = q_{rr}\), then \(q_1 = q_r\). 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, \(\hat{q}_1\) 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 \(\hat{q}_r = \hat{q}_1\), \(\hat{q}_{r+1} = \delta \hat{q}_1\), \(q_{11} = q_{r1} = q_{1r} = 1\), \(q_{rr} \neq 1\).

We first consider the case of \(i = 1\), which implies \(\hat{q}_{r+1} = \delta \hat{q}_1\). We end up with a \(3 \times 2\) submatrix of \(Q\) 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 \(Q\) 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 \(\delta\) or \(\delta q_{rr}\). While the former results in \(q_{r+1} = \delta q_1\), the latter leads to \(q_{r+1} = \delta q_r\).

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

- If \(\delta = 0\), then we end up with a \(3 \times 2\) submatrix of \(Q\) 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 \(Q\) satisfies 3-2 condition and \(q_{rr} \neq 1\), we have \(q_{r+1,r} = 0\). Hence, \(q_{r+1} = \mathbf{0}\).
- If $\delta \neq 0$ and $q_{21} = 0$, we end up with a $4 \times 2$ submatrix of $Q$ 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 $Q$ satisfies 3-2 condition and $q_{rr} \neq 1$, we have $q_{2r} = q_{r+1,r} = 0$. Hence, $q_{r+1} = \delta q_2$.

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

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

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., $\mathbf{q}_2^o, \mathbf{q}_4^o$ are independent. We notice that $Q$ satisfies 3-2 condition and $\mathbf{q}_1^o, \mathbf{q}_3^o$ are independent since $q_{rr} \neq 1$. Therefore, we have either 1) $q_{2r} = 1$ ($\mathbf{q}_2^o$ and $\mathbf{q}_2^o$ are dependent), $q_{r+1,r} = q_{rr}$ ($\mathbf{q}_3^o$ and $\mathbf{q}_4^o$ are dependent), or 2) $q_{r+1,r} = 1$ ($\mathbf{q}_1^o$ and $\mathbf{q}_2^o$ are dependent). $q_{2r} = q_{rr}$ ($\mathbf{q}_2^o$ and $\mathbf{q}_3^o$ are dependent). WLOG, we consider the former case. Since $\mathbf{q}_1, \mathbf{q}_2$ are independent, we can find $i \neq 1$ such that $(q_{1i}, q_{2i})$ and $(q_{1i}, q_{2i})$ are independent. WLOG, let $i = 2$. In this case, we end up with a $4 \times 2$ submatrix of $Q$ by deleting all rows except rows 1, 2, $r, r+1$ and deleting all columns except columns 2, $r$:

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

\[
q_{12} = q_{22}q_{rr},
\]

(EC.34)

Similarly, due to the independence between \(q^*_1\) and \(q^*_2\) and that between \(q^*_2\) and \(q^*_4\), we have that \(q^*_1\) and \(q^*_4\) are dependent, i.e.,

\[
q_{12}q_{rr} = q_{22}.
\]

(EC.35)

If one of \(q_{12}\) and \(q_{22}\) is 0, the equalities (EC.34) and (EC.35) imply \(q_{12} = q_{22} = 0\), which contradicts with \(q_{12} \neq q_{22}\). If \(q_{12}q_{22} \neq 0\), the equalities (EC.34) and (EC.35) 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 \(q_{r+1} = q_2\).

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