Operations Research

Preservation of Structural Properties in Optimization with Decisions Truncated by Random Variables and Its Applications

Xin Chen, Xiangyu Gao, Zhan Pang

Please scroll down for article—it is on subsequent pages

With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

CONTEXTUAL AREAS

Preservation of Structural Properties in Optimization with Decisions Truncated by Random Variables and Its Applications

Xin Chen,[a] Xiangyu Gao,[b] Zhan Pang[c]

[a] Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801;
[b] Department of Decision Sciences and Managerial Economics, CUHK Business School, The Chinese University of Hong Kong, Hong Kong;
[c] College of Business, City University of Hong Kong, Kowloon Tong, Hong Kong
Contact: xinchen@illinois.edu, http://orcid.org/0000-0002-5168-4823 (XC); xiangyu@cuhk.edu.hk,
http://orcid.org/0000-0002-7126-7330 (XG); zhan.pang@cityu.edu.hk, http://orcid.org/0000-0001-6534-7820 (ZP)

Abstract. A common technical challenge encountered in many operations management models is that decision variables are truncated by some random variables and the decisions are made before the values of these random variables are realized, leading to nonconvex minimization problems. To address this challenge, we develop a powerful transformation technique that converts a nonconvex minimization problem to an equivalent convex minimization problem. We show that such a transformation enables us to prove the preservation of some desired structural properties, such as convexity, submodularity, and $L^\natural$-convexity, under optimization operations, that are critical for identifying the structures of optimal policies and developing efficient algorithms. We then demonstrate the applications of our approach to several important models in inventory control and revenue management: dual sourcing with random supply capacity, assemble-to-order systems with random supply capacity, and capacity allocation in network revenue management.

1. Introduction

In operations management literature, a common technical challenge encountered in many models is that decision variables are truncated by some random variables and the decisions are made before the values of these random variables are realized. A notable example is inventory control problems with supply capacity uncertainty in which the replenishment decision is truncated by the random supply capacity (see, e.g., Ciarallo et al. 1994, Wang and Gerchak 1996, Bollapragada et al. 2004, Hu et al. 2008, Feng 2010, and Feng and Shi 2012). Another example is capacity allocation problems in revenue management where the booking limit of each demand class is truncated by the random demand (see, e.g., Brumellem and Mcgill 1993, Robinson 1995, and Chen and Homem-de-Mello 2010). This type of variable truncation often leads to stochastic optimization problems in the following form:

$$g(x, z) = \inf_{u: (x, z, u) \in \mathcal{A}} E[f(x, u \wedge (z + \Xi))], \qquad (1)$$

where f is a function in decision variables u and state variables (x, z), $\mathcal{A}$ is the constraint set, Ξ is a random vector, and $\wedge$ denotes componentwise minimum.

For these applications, it is natural to ask how to solve problem (1) efficiently and whether the optimization operation can preserve some desired structural properties of f such as convexity or submodularity. However, solving and analyzing such a problem can be very difficult. An intrinsic challenge arises from the fact that the truncation by random variables may destroy convexity: the objective function may not be convex in the decision variables even if the function f is convex. Without the regular properties such as convexity, the problem could be both analytically and computationally intractable, in particular when facing multidimensional state and decision variables.

Our paper aims at addressing this challenge when the random variables are independently distributed by developing a novel transformation technique that converts the nonconvex minimization problem (1) to an equivalent convex minimization problem. As we mentioned earlier, the original problem formulation may be nonconvex for a convex function f because in the objective function there are terms involving the minimum of decision variables and random variables. The key idea is to relax the original problem by replacing $u \wedge (z + \xi)$ by new variables $v(\xi) = (v_1(\xi_1), v_2(\xi_2), \ldots, v_n(\xi_n))$ and

imposing $v(\xi) \leqslant z + \xi$ in the constraints. We prove that the optimal objective values of the original and transformed problems are the same when f is convex and certain regularity conditions are imposed on $\mathscr{A}$. Furthermore, our transformation technique allows us to show that the optimization operation in problem (1) can preserve convexity, submodularity, or $L^\natural$-convexity, which then enables us to perform comparative statics analysis in multidimensional state and decision spaces and characterize the monotone structure of optimal policies.

Our approach has a wide range of applications. In this paper, we focus on the applications of the transformation technique to three models with multidimensional state spaces. Our first application is an inventory system with two capacitated suppliers, a regular one with a longer lead time and an expedited one with a shorter lead time. The two suppliers have independent supply capacity uncertainties. The objective of the firm is to find a dual-sourcing strategy to minimize the total expected cost. The second application is an assemble-to-order (ATO) inventory system with multiple components and products. The order quantity of each component cannot exceed a random capacity. The firm decides the ordering quantities of all components and then the number of products assembled to minimize the expected cost. The third application is the capacity allocation in network revenue management where fixed capacities of resources are allocated dynamically to different products with random demands. In the airline industry, this corresponds to setting booking limits for each itinerary-fare class combination. The booking limits are truncated by the random demand. The firm aims to maximize the expected total revenue. In all the above applications, we employ the transformation technique to prove that the apparently nonconvex minimization problems (or nonconcave maximization problems) can be converted to equivalent convex minimization problems (or concave maximization problems), and under some conditions, the optimal decisions are monotone in terms of the state variables with limited sensitivities. Without the transformation technique, the structural analyses would have been much more complicated, if not impossible, to carry out.

1.1. Related Literature

We next review two streams of related literature: (1) inventory management with special emphasis on supply capacity uncertainty, and (2) capacity allocation in network revenue management.

Supply uncertainty of inventory/production systems can be driven by a variety of factors. Most studies in this literature focus on random yield problems where the supply is a random proportion of the order quantity; see Henig and Gerchak (1990), Federgruen and Yang (2008, 2011), Chen et al. (2013), and the references therein. Such an issue usually arises from the

quantity uncertainty of items produced in a batch. Another important supply uncertainty is the supply capacity uncertainty due to the unreliability of the supply processes (e.g., partial delivery or cancellation of an order by the supplier). In such an environment, the firm has to place orders before knowing the actual supply capacity. There are relatively few papers addressing the random capacity problems.

Ciarallo et al. (1994) consider an inventory control problem, assuming that the replenishment decisions are made before the capacity uncertainty is realized and the replenishment lead time is zero. They show that the presence of capacity uncertainty does not affect the optimality of a base-stock policy. Wang and Gerchak (1996) extend the analysis to systems with both random supply capacity and random yield. Feng (2010) addresses a joint pricing and inventory control problem with supply capacity uncertainty and zero lead time and shows that the optimal policy is characterized by two critical values: a reorder point and a target safety stock. The common technical challenge of these models is that with random supply capacity, the corresponding dynamic programming recursions, though all involving one-dimensional state spaces, are not convex minimization (concave maximization) problems anymore, and delicate analyses are needed to characterize the structures of optimal policies.

Our transformation technique can be readily applied to the aforementioned models to simplify the structural analysis. More importantly, such an approach allows us to address more general inventory models under supply capacity uncertainty with multidimensional state spaces using the concept of $L^\natural$-convexity. This paper demonstrates two applications in the area of inventory management with supply uncertainty, i.e., the dual-sourcing problem and the assemble-to-order problem. In the following, we introduce the literature related to these two applications.

There is an extensive literature on the dual-sourcing problem. It was first studied by Barankin (1961) in a one-period setting and then extended by Daniel (1963), Fukuda (1964), and Whittmore and Saunders (1977) to various settings with multiperiod horizons. Feng and Shi (2012) consider a joint inventory control and pricing problem with multiple suppliers whose replenishment lead times are zero and supply capacities are uncertain. They show that with deterministic capacities a multilevel base-stock list-price policy plus a cost-based supplier selection (i.e., ordering from a cheaper source first) is optimal. However, with general random supply capacities, such a policy is no longer optimal. They show that the optimal policy can be characterized by a *near* reorder point such that a positive order is placed (almost everywhere) if and only if the inventory level is below this point. They also identify a condition under which a strict reorder-point policy and a

cost-based supplier-selection criterion become optimal. More recently, Zhou and Chao (2014) address the dual-sourcing problem with price sensitive demand, a regular supplier with one-period lead time, and an expedited supplier with zero lead time, and characterize the structure of the optimal policy. Gong et al. (2014) further generalize the structural analysis to a dual-sourcing problem with price sensitive demand and Markovian supply interruptions. In both models, there are no capacity limits on the supplies. To the best of our knowledge, our paper is the first addressing the dual-sourcing system with arbitrary deterministic lead time discrepancies and supply capacity uncertainties.

The assemble-to-order system is one of the most important production/inventory systems; see Song and Zipkin (2003) for a review of the research literature and applications of assemble-to-order systems up to the early 2000s. Lu and Song (2005) study a continuous-review assemble-to-order system with random demands and lead times with an order-based approach. Nadar et al. (2014) develop the optimal structural results for a continuous-review assemble-to-order generalized M-system with lost sales. Bollapragada et al. (2004) study multiechelon assembly systems under installation base-stock policies where the component suppliers have various lead times and random supply capacities. They propose a decomposition approach and their numerical study shows that their heuristic performs well in comparison with the optimal base-stock policy. In this paper, we show that our approach applies to the assemble-to-order system with random component capacity. Moreover, for the generalized M-system, we show that the cost-to-go functions are $L^{\natural}$-convex, which allows us to characterize the monotone structure of the optimal policy.

Revenue management (RM), also known as yield management, has been widely adopted in various industries such as airlines, hotels, car rentals, and cruise lines. Driven by its prevalence in the service industry, the research interest in RM has been growing rapidly over the last two decades; see Talluri and van Ryzin (2005) for a comprehensive introduction to the practice and theoretical developments of RM.

The network revenue management problem, which involves managing multiple resources (such as airline seats in different leg-cabin combinations), is notoriously challenging. Indeed, as mentioned by Talluri and van Ryzin (2005, p. 83), "in the network case, exact optimization is for all practical purposes impossible," and thus the literature focuses predominantly on various approximations. One approximation is to formulate a stochastic programming problem (see Cooper and Homem-de-Mello 2007, Möller et al. 2008, Chen and Homem-de-Mello 2010, and the references therein). For example, one can formulate a two-stage stochastic linear programming problem (SLP) by aggregating the

demand over the planning horizon and determining the booking limits at the beginning (see Talluri and van Ryzin 2005, Section 3.3.1). To improve upon the SLP, one can consider a multistage stochastic programming (MSSP), in which the policy of booking limits is revised from time to time in order to take into account the information about demand learned so far. The MSSP is challenging, evidenced by Chen and Homem-de-Mello (2010, p. 93): "even the continuous relaxation of that problem does not have a concave expected recourse function," as its objective function and constraints involve booking limits truncated by realized demands. As a compromise, they propose an approximation based on re-solving a sequence of two-stage stochastic programs.

We consider the MSSP with continuous relaxation. In each time period, the firm decides the booking limits allocated to each demand class before the demand is realized. Interestingly, our transformation technique preserves concavity in the dynamic programming recursions, and hence overcomes the difficulty stated by Chen and Homem-de-Mello (2010). Under certain network structures, we further show that $L^{\natural}$-concavity can be preserved and use it to derive some monotonicity properties of the optimal booking limits. Our approach opens the door to the development of effective algorithms to solve MSSP directly.

1.2. Our Contribution

As evidenced by the literature review, structural analyses for many important models (such as the inventory control problem under supply capacity uncertainty and the capacity allocation problem in revenue management) involve solving challenging stochastic optimization problems with a form similar to problem (1). Our transformation technique provides a unified technical tool to facilitate the structural analysis of this type of problems, which is our primary contribution to the literature. The power of this technique is demonstrated by its applications to several important inventory control and revenue management models that generalize the corresponding ones in the literature. The preservation results for structural properties such as convexity, submodularity, and $L^{\natural}$-convixity enabled by our transformation technique can also be potentially exploited to develop efficient algorithms.

Recently, Feng and Shanthikumar (2018) use the notion of *stochastic linearity in midpoint* to develop a different technique to show that a class of nonlinear supply and demand functions (in the almost sure sense) are in fact linear in the stochastic sense. Like ours, their approach allows them to convert some nonconvex minimization problems, including those in Ciarallo et al. (1994), Wang and Gerchak (1996), Feng (2010), and Feng and Shi (2012), into convex minimization problems. Treating the means of the supply and demand

functions as decision variables instead of the original decisions (ordering quantity and price), they show that supply and demand functions are stochastically linear in midpoint with respect to their means and the objective functions are concave in the means of supply and demand. Note that they focus on the concavity property but do not touch upon supermodularity or $L^\natural$-concavity. Different from their approach, our approach works on the original decision variables and transforms the original optimization problem into an equivalent constrained optimization problem, which allows us to readily show the preservation of convexity, submodularity, and $L^\natural$-convexity. Hence, our approach is more suitable for problems with high-dimensional state spaces like the applications we present in this paper. In the appendix we provide a detailed comparison between our transformation technique and their approach. In particular, we show that although their approach can also preserve convexity and submodularity, it does not preserve $L^\natural$-convexity.

The remainder of this paper is organized as follows. Section 2 develops the transformation technique and the relevant preservation results. Sections 3–5 focus on the applications of our approach to the dual-sourcing problem, the assemble-to-order system, and the capacity allocation problem, respectively. The paper is concluded in Section 6.

Throughout this paper, we use decreasing, increasing, and monotonicity in a weak sense. We use $\mathbb{R}$ and $\mathbb{R}_+$ to denote the real space and the set with nonnegative reals, $\mathbb{Z}$, and $\mathbb{Z}_+$ to denote the set of integers and the set of nonnegative integers, respectively. For convenience, let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{Z}$. Define $\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, $e \in \mathbb{F}^n$ a vector whose components are all ones, e_j a unit vector whose jth component is one, and for $x, y \in \mathbb{F}^n$, $x \leqslant y$ if and only if $x_i \leqslant y_i$ for any $i = 1, \ldots, n$, $x^+ = \max(x, 0)$, $x \wedge y = \min(x, y)$, and $x \vee y = \max(x, y)$ (the componentwise minimum and maximum operations). The indicator function of any set $\mathcal{V} \subseteq \mathbb{F}^n$, denoted by $\delta_{\mathcal{V}}$, is defined as $\delta_{\mathcal{V}}(x) = 0$ for $x \in \mathcal{V}$ and $+\infty$ otherwise. We use the superscript T to denote the transpose of a vector or a matrix. We use uppercase letters (e.g., Ξ) to denote random vectors and lowercase letters (e.g., ξ) for their realizations. Given a random vector $\Xi = (\Xi_1, \ldots, \Xi_n)^T$, we use $\mathcal{X} = \mathrm{Supp}(\Xi)$ to denote the support of this random vector. In addition, we define $\bar{\xi}_j = \mathrm{ess\,sup}\{\xi_j \mid \xi_j \in \mathcal{X}_j\}$, $\underline{\xi}_j = \mathrm{ess\,inf}\{\xi_j \mid \xi_j \in \mathcal{X}_j\}$ for $j = 1, \ldots, n$, where $\mathcal{X}_j$ is $\mathcal{X}$'s projection into the jth coordinate. Let $\bar{\xi} = (\bar{\xi}_1, \ldots, \bar{\xi}_n)^T$, $\underline{\xi} = (\underline{\xi}_1, \ldots, \underline{\xi}_n)^T$, and almost surely is abbreviated as a.s.

2. Transformation Technique and Preservation Properties

In this section, we first develop the transformation technique for a class of stochastic optimization problems and then show several preservation results that are useful in structural analysis.

2.1. Transformation

Given a function $f: \mathbb{F}^n \to \bar{\mathbb{R}}$ and a random vector Ξ with $\mathrm{Supp}(\Xi) = \mathcal{X} \subseteq \mathbb{F}^n$, consider the following optimization problem:

$$\tau^* = \inf_{u \in \mathbb{F}^n} E[f(u \wedge \Xi)]. \tag{2}$$

In general, the above problem may not be a convex minimization problem even if the function f is convex. For instance, let $f(u) = u^2$ and Ξ be Bernoulli distributed with success probability 0.5. One can easily see $E[f(u \wedge \Xi)]$ is not convex in u. Interestingly, we show that under certain conditions, we can convert it into an equivalent convex minimization problem.

For this purpose, note that the optimization problem (2) can be rewritten as follows:

$$\begin{aligned}
\inf \quad & E[f(v(\Xi))] \\
\text{s.t.} \quad & v(\xi) = u \wedge \xi \quad \forall \xi \in \mathcal{X}, \\
& u \in \mathbb{F}^n, v(\cdot) \in \mathcal{M},
\end{aligned} \tag{3}$$

where $\mathcal{M}$ is the set of measurable functions. The feasible region of (3) is $\mathbb{F}^n \times (\mathbb{F}^n)^{\mathcal{X}}$, while the feasible region of (2) is $\mathbb{F}^n$. In the following theorem, we show that the equality constraint $v(\xi) = u \wedge \xi$ can be relaxed by the inequality constraint $v(\xi) \leqslant \xi$, $\forall \xi \in \mathcal{X}$, with $v(\xi) = (v_1(\xi_1), \ldots, v_n(\xi_n)) \in \mathbb{F}^n$. For the rest of the paper, we require that $v(\cdot)$ is measurable in all of our formulations and therefore omit $v(\cdot) \in \mathcal{M}$ for brevity. The following lemma will be useful for the proof of the theorem.

Lemma 1. *Suppose that the function $f: \mathbb{F} \to \bar{\mathbb{R}}$ is quasi-convex. If x^* is a minimizer of $f(x)$ over $\mathbb{F}$, we have $f(x^* \wedge b) \leqslant f(a)$ for any $a, b \in \mathbb{F}$ with $a \leqslant b$.*

Proof. The quasi-convexity of f implies that $f(x)$ decreases in x as $x \leqslant x^*$ and increases in x as $x \geqslant x^*$. If $a \geqslant x^*$, we have $b \geqslant a \geqslant x^*$, which implies that $a \geqslant x^* \wedge b = x^*$. If $a \leqslant x^*$, since $a \leqslant b$, we have $a \leqslant x^* \wedge b \leqslant x^*$. In either case, $f(x^* \wedge b) \leqslant f(a)$.   Q.E.D.

Theorem 1 (Equivalent Transformation). *Suppose that (a) the function $f: \mathbb{F}^n \to \bar{\mathbb{R}}$ is lower semicontinuous with $f(x) \to +\infty$ for $|x| \to \infty$; (b) f is componentwise convex (componentwise discrete convex if $\mathbb{F} = \mathbb{Z}$); (c) the random vector Ξ has independent components and it has realizations $\xi \in \mathcal{X} = \mathrm{Supp}(\Xi)$. Then, τ^* defined in (2) is also the optimal objective value of the following optimization problem:*

$$\begin{aligned}
\inf \quad & E[f(v(\Xi))] \\
\text{s.t.} \quad & v(\xi) \leqslant \xi \quad \forall \xi \in \mathcal{X}, \\
& v(\xi) = (v_1(\xi_1), \ldots, v_n(\xi_n)) \in \mathbb{F}^n \quad \forall \xi \in \mathcal{X}.
\end{aligned} \tag{4}$$

Proof. Let π^* be the optimal objective value of problem (4). Since for any $u \in \mathbb{F}^n$, $v(\xi) = u \wedge \xi$ is feasible for problem (4), $\pi^* \leqslant \tau^*$.

It remains to show that $\tau^* \leqslant \pi^*$. Clearly, it holds when $\pi^* = \infty$. Thus, in the following, we assume that $\pi^* < \infty$, which together with assumption (a) implies

that all optimization problems involved below, as well as problems (2) and (4), admit finite optimal solutions. Given any optimal solution of (4) denoted by $v^* = (v^*(\xi) \mid \xi \in \mathscr{X})$, we will show that we can find a solution $\hat{u} \in \mathbb{F}^n$ such that $E[f(\hat{u} \wedge \Xi)] = E[f(v^*(\Xi))]$.

We first show that it is true for $n = 1$. Let $\hat{u} = \arg\min_{u \in \mathbb{F}} f(u)$ (when there are multiple optimal solutions, we choose the smallest one). Consider any feasible solution $v = (v(\xi) \mid \xi \in \mathscr{X})$ of problem (4). We have $f(\hat{u} \wedge \xi) \leqslant f(v(\xi))$ for any $\xi \in \mathscr{X}$ according to Lemma 1. Hence, $E[f(\hat{u} \wedge \Xi)] \leqslant \pi^*$. Note that $\hat{u}$ is a feasible solution for problem (2), which implies that $\tau^* = E[f(\hat{u} \wedge \Xi)] \leqslant \pi^*$. Combined with the fact that $\pi^* \leqslant \tau^*$, we have $\tau^* = \pi^*$.

We now consider the general case with $n \geqslant 1$. Use v_i^* to represent the ith component of v^* for $i = 1, \ldots, n$. Starting from the first component, define

$$\pi_1(u_1) = E[f(u_1, v_2^*(\Xi_2), \ldots, v_n^*(\Xi_n))].$$

The componentwise convexity of f implies that $\pi_1(u_1)$ is convex in u_1. Since the components of the vector Ξ are independently distributed, $E_{\Xi_1}[\pi_1(v_1(\Xi_1))] = E_\Xi[f(v_1(\Xi_1), v_2^*(\Xi_2), \ldots, v_n^*(\Xi_n))]$ for any measurable function $v_1(\cdot)$, and the preceding analysis for $n = 1$ implies that there exists a $\hat{u}_1$ such that

$$\begin{aligned}\pi^* &= \min\{E[\pi_1(v_1(\Xi_1))] \mid v_1(\xi_1) \leqslant \xi_1, v_1(\xi_1) \in \mathbb{F}, \forall \xi_1 \in \mathscr{X}_1\} \\ &= \min_{u_1 \in \mathbb{F}} E[\pi_1(u_1 \wedge \Xi_1)] = E[\pi_1(\hat{u}_1 \wedge \Xi_1)].\end{aligned}$$

Next define $\pi_2(u_2) = E[f(\hat{u}_1 \wedge \Xi_1, u_2, v_3^*(\Xi_3), \ldots, v_n^*(\Xi_n))]$. Clearly, π_2 is convex. Following the preceding analysis, there exists a $\hat{u}_2$ such that

$$\begin{aligned}\pi^* &= \min\{E[\pi_2(v_2(\Xi_2))] \mid v_2(\xi_2) \leqslant \xi_2, v_2(\xi_2) \in \mathbb{F}, \forall \xi_2 \in \mathscr{X}_2\} \\ &= \min_{u_2 \in \mathbb{F}} E[\pi_2(u_2 \wedge \Xi_2)] = E[\pi_2(\hat{u}_2 \wedge \Xi_2)].\end{aligned}$$

Continue this process and define $\pi_i(u_i) = E[f(\hat{u}_1 \wedge \Xi_1, \ldots, \hat{u}_{i-1} \wedge \Xi_{i-1}, u_i, v_{i+1}^*(\Xi_{i+1}), \ldots, v_n^*(\Xi_n))]$. Applying the same approach, we can find $\hat{u}_i, i = 3, \ldots, n$, such that

$$\begin{aligned}\pi^* &= \min\{E[\pi_i(v_i(\Xi_i))] \mid v_i(\xi_i) \leqslant \xi_i, v_i(\xi_i) \in \mathbb{F}, \forall \xi_i \in \mathscr{X}_i\} \\ &= \min_{u_i \in \mathbb{F}} E[\pi_i(u_i \wedge \Xi_i)] = E[\pi_i(\hat{u}_i \wedge \Xi_i)].\end{aligned}$$

Therefore,

$$\pi^* = E[\pi_n(\hat{u}_n \wedge \Xi_n)] = E[f(\hat{u}_1 \wedge \Xi_1, \ldots, \hat{u}_n \wedge \Xi_n)].$$

Since $\hat{u}$ is a feasible solution to problem (2), we have $\tau^* \leqslant E[f(\hat{u} \wedge \Xi)] = \pi^*$. Combined with the fact that $\pi^* \leqslant \tau^*$, we have $\pi^* = \tau^*$. Q.E.D.

Remark 1. In the proof of the above theorem, we illustrate that when $n = 1$,

$$\min_{u \in \mathbb{F}} E[f(u \wedge \Xi)] = E[f(\hat{u} \wedge \Xi)],$$

where $\hat{u}$ is any minimizer of the function f. In fact, this observation is still valid when f is quasi-convex.

However, when $n > 1$, such a result no longer holds, i.e., $\min_{u \in \mathbb{F}} E[f(u \wedge \Xi)] \neq E[f(\hat{u} \wedge \Xi)]$, even if f is jointly convex. We now present an example. Specifically, let

$$n = 2, \mathbb{F} = \mathbb{R},$$
$$f(u_1, u_2) = (u_1 + u_2 - 2)^2 + (u_1 - 1)^2 + (u_2 - 1)^2,$$

and Ξ_1 and Ξ_2 be independent and identically distributed and take values 0 and 2 with equal probabilities. In this case, $\hat{u} = (1, 1)$. However, one can easily verify that $\arg\min_{u \in \mathbb{F}^n} E[f(u \wedge \Xi)] = (1.2, 1.2) \neq \hat{u}$.

Remark 2. In the above theorem, we require that $v(\xi) = (v_1(\xi_1), \ldots, v_n(\xi_n))$. This cannot be relaxed to allow $v(\xi) = (v_1(\xi), \ldots, v_n(\xi))$. To illustrate this, we use the above example again. Note that for problem (4), the optimal objective value is 2.4 and an optimal solution is given by

$$v_1^*(0) = v_2^*(0) = 0, \quad v_1^*(2) = v_2^*(2) = 1.2.$$

However, if one replaces $v(\xi) = (v_1(\xi_1), \ldots, v_n(\xi_n))$ by $v(\xi) = (v_1(\xi), \ldots, v_n(\xi))$ in problem (4), the optimal objective value becomes 2.25 and an optimal solution is given by

$$\begin{aligned}v^*(0, 0) &= (0, 0), & v^*(0, 2) &= (0, 1.5), \\ v^*(2, 0) &= (1.5, 0), & v^*(2, 2) &= (1, 1).\end{aligned}$$

Remark 3. It is interesting to observe that u does not appear in problem (4). Our proof implies that given an optimal solution u^* of problem (2), $v^* = (v^*(\xi) = u^* \wedge \xi \mid \xi \in \mathscr{X})$ is optimal for problem (4). On the other hand, given an optimal solution v^* of problem (4), we can directly construct an optimal solution of problem (2) without solving any additional optimization problem. To see this, we start with $n = 1$ and define $S = \{\xi \mid v^*(\xi) < \xi, \xi \in \mathscr{X}\}$ (for simplicity, we drop the subscript 1 in the presentation when $n = 1$). We consider two cases depending on whether the probability of event S, denoted by $P(S)$, is zero or not. In the first case, $P(S) > 0$. Randomly pick $\hat{\xi} \in S$ according to the probability distribution of Ξ conditional on S and define $\hat{u} = v^*(\hat{\xi})$. It suffices to show that $\hat{u}$ is optimal for the optimization problem $\min_{u \in \mathbb{F}} f(u)$ with probability 1. Suppose this is not true and $P(S') > 0$, where S' is the event such that $\hat{\xi} \in S$ and $v^*(\hat{\xi})$ is not optimal for $\min_{u \in \mathbb{F}} f(u)$. We define a new feasible solution of problem (4):

$$\hat{v}(\xi) = \begin{cases} v^*(\xi), & \text{if } \xi \notin S', \\ u^0 \wedge \xi, & \text{if } \xi \in S', \end{cases}$$

where u^0 is an optimal solution of $\min_{u \in \mathbb{F}} f(u)$. If $\xi \notin S'$, then $\hat{v}(\xi) = v^*(\xi)$ and $f(\hat{v}(\xi)) = f(v^*(\xi))$. If $\xi \in S'$

and $\xi > u^0$, $f(\hat{v}(\xi)) = f(u^0) < f(v^*(\xi))$. If $\xi \in S'$ and $\xi \leqslant u^0$, $v^*(\xi) < \xi \leqslant u^0$ and the convexity of f implies that $f(\hat{v}(\xi)) = f(\xi) < f(v^*(\xi))$. Since $P(S') > 0$, we have $E[f(\hat{v}(\xi))] < E[f(v^*(\xi))]$, which is a contradiction. Therefore, with probability 1, $\hat{u}$ is optimal for the optimization problem $\min_{u \in \mathbb{F}} f(u)$. In the second case, $P(S) = 0$. Note that f must be decreasing over $\mathscr{X}$, otherwise we can easily construct a feasible solution of problem (4) with a lower cost. Hence, assumption (a) implies that $\bar{\xi} < \infty$, and $\hat{u} = \bar{\xi}$ is a minimizer of the function f on $\mathbb{F}$. For $n > 1$, define, for $i = 1, \dots, n$, event $S_i = \{\xi_i \mid v_i^*(\xi_i) < \xi_i\}$. If the probability of S_i is positive, randomly pick $\hat{\xi}_i \in S_i$ according to the probability distribution of Ξ conditional on S_i and define $\hat{u}_i = v_i^*(\hat{\xi}_i)$; otherwise, define $\hat{u}_i = \bar{\xi}_i$ (again $\bar{\xi}_i < \infty$). Since the components of the random vector Ξ are independent, we can extend the above analysis to show that, with probability 1, $\hat{u} = (\hat{u}_1, \dots, \hat{u}_n)$ is an optimal solution of problem (2).

We can explicitly incorporate constraints on u in Theorem 1 and consider a more general optimization model. To simplify notations, we define an operator $\diamondsuit_k$ as $u \diamondsuit_k \xi \triangleq (u_1 \wedge \xi_1, \dots, u_k \wedge \xi_k, u_{k+1} \vee \xi_{k+1}, \dots, u_n \vee \xi_n)$. The problem of interest is

$$\inf_{u \in \mathscr{U}} E[f(u \diamondsuit_k \Xi)], \tag{5}$$

where $f: \mathbb{F}^n \to \bar{\mathbb{R}}$ and $\mathscr{U} \subseteq \mathbb{F}^n$. Define a set

$$\mathscr{V} = \{u \diamondsuit_k \xi \mid u \in \mathscr{U}, \xi \in \mathscr{X}\}. \tag{6}$$

We impose the following assumption:

Assumption 1. (a) *For any $u \in \mathbb{F}^n$ such that $u \diamondsuit_k \xi \in \mathscr{V}$, $\forall \xi \in \mathscr{X}$, there exists $u' \in \mathscr{U}$ such that $u' \diamondsuit_k \xi = u \diamondsuit_k \xi$, $\forall \xi \in \mathscr{X}$.*

(b) *The indicator function of the set $\mathscr{V}$ is componentwise convex (componentwise discrete convex if $\mathbb{F} = \mathbb{Z}$).*

Notice that part (a) implies that if $u \diamondsuit_k \xi \in \mathscr{V}$, $\forall \xi \in \mathscr{X}$, we do not necessarily need $u \in \mathscr{U}$. Instead, we only require that there exists $u' \in \mathscr{U}$ such that $u' \diamondsuit_k \xi = u \diamondsuit_k \xi$, $\forall \xi \in \mathscr{X}$. As can be seen from the proof of Theorem 2, Assumption 1 allows us to convert the constrained optimization problem (5) to an equivalent unconstrained optimization problem so that Theorem 1 can be applied. We provide a nontrivial example under which Assumption 1 holds in Lemma 2.

Theorem 2. *Consider the optimization problem (5), where $f: \mathbb{F}^n \to \bar{\mathbb{R}}$ and the random vector Ξ in $\mathbb{F}^n$ satisfy the assumptions in Theorem 1. Suppose that Assumption 1 is satisfied. Problem (5) and the following optimization problem have the same optimal objective value:*

$$\begin{aligned} \inf \quad & E[f(v_1(\Xi_1), \dots, v_n(\Xi_n))] \\ \text{s.t.} \quad & v_j(\xi_j) \leqslant \xi_j \quad \forall \xi_j \in \mathscr{X}_j, \ j = 1, \dots, k, \\ & v_j(\xi_j) \geqslant \xi_j \quad \forall \xi_j \in \mathscr{X}_j, \ j = k+1, \dots, n, \\ & (v_1(\xi_1), \dots, v_n(\xi_n)) \in \mathscr{V} \quad \forall \xi \in \mathscr{X}. \end{aligned} \tag{7}$$

Proof. Problem (5) is equivalent to the following unconstrained optimization problem.

$$\inf_{u \in \mathbb{F}^n} \{E[f(u \diamondsuit_k \Xi)] + \delta_{\mathscr{U}}(u)\}. \tag{8}$$

Define for any $v \in \mathbb{F}^n$,

$$\hat{f}(v) = f(v) + \delta_{\mathscr{V}}(v),$$

where $\mathscr{V}$ is defined in (6). Then by Assumption 1 the optimal objective value of problem (8) is equivalent to that of the following problem:

$$\inf_{u \in \mathbb{F}^n} E[\hat{f}(u \diamondsuit_k \Xi)]. \tag{9}$$

To see this, note that for any $u \in \mathscr{U}$, we have $u \diamondsuit_k \xi \in \mathscr{V}$, $\forall \xi \in \mathscr{X}$; hence, $\inf_{u \in \mathbb{F}^n} E[\hat{f}(u \diamondsuit_k \Xi)] \leqslant \inf_{u \in \mathscr{U}} E[f(u \diamondsuit_k \Xi)]$. On the other hand, because of Assumption 1, we have $\inf_{u \in \mathbb{F}^n} E[\hat{f}(u \diamondsuit_k \Xi)] \geqslant \inf_{u \in \mathscr{U}} E[f(u \diamondsuit_k \Xi)]$.

Define a new random vector $\tilde{\Xi}$ with $(\tilde{\Xi}_1, \dots, \tilde{\Xi}_k, \tilde{\Xi}_{k+1}, \dots, \tilde{\Xi}_n) = (\Xi_1, \dots, \Xi_k, -\Xi_{k+1}, \dots, -\Xi_n)$ and a new function $\tilde{f}: \mathbb{F}^n \to \bar{\mathbb{R}}$ by

$$\tilde{f}(u_1, \dots, u_k, u_{k+1}, \dots, u_n) = \hat{f}(u_1, \dots, u_k, -u_{k+1}, \dots, -u_n).$$

Then problem (9) is equivalent to the problem

$$\inf_{\tilde{u} \in \mathbb{F}^n} E[\tilde{f}(\tilde{u} \wedge \tilde{\Xi})].$$

By Theorem 1, it has the same optimal objective value with the following problem:

$$\begin{aligned} \inf \quad & E[\tilde{f}(\tilde{v}(\tilde{\xi}))] \\ \text{s.t.} \quad & \tilde{v}(\tilde{\xi}) \leqslant \tilde{\xi} \quad \forall \tilde{\xi} \in \mathrm{Supp}(\tilde{\Xi}), \\ & \tilde{v}(\tilde{\xi}) = (\tilde{v}_1(\tilde{\xi}_1), \dots, \tilde{v}_n(\tilde{\xi}_n)) \in \mathbb{F}^n, \end{aligned}$$

which is clearly equivalent to problem (7) from the definition of $\tilde{f}$. Notice that the indicator function of the set $\mathscr{V}$ needs to be componentwise convex to ensure that $\tilde{f}$ is componentwise convex. Q.E.D.

The following lemma provides an example that satisfies Assumption 1.

Lemma 2. *Assume that $\mathscr{U} = \{u \in \mathbb{F}^n \mid Au \leqslant b, u_1 \geqslant \underline{u}_1, \dots, u_k \geqslant \underline{u}_k, u_{k+1} \leqslant \bar{u}_{k+1}, \dots, u_n \leqslant \bar{u}_n\}$, where $b, \underline{u}_1, \dots, \underline{u}_k, \bar{u}_{k+1}, \dots, \bar{u}_n$ are given constants, $A = (a_{ij})$ with entries $a_{ij} \geqslant 0$ for any i and $j = 1, \dots, k$, and $a_{ij} \leqslant 0$ for any i and $j = k+1, \dots, n$. In addition $\mathscr{X}_j$ is contained in $[\underline{u}_j, +\infty)$ for $j = 1, \dots, k$, and $\mathscr{X}_j$ is contained in $(-\infty, \bar{u}_j]$ for $j = k+1, \dots, n$. Then Assumption 1 is satisfied.*

Proof. For notational convenience, we only prove the case where $k = n$, i.e., there is only the $\wedge$ operation. This is because we can apply the same technique used in the proof of Theorem 2 to convert a problem with the $\vee$ operation to a new one only with the $\wedge$ operation. In this case the set $\mathscr{U} = \{u \mid Au \leqslant b, u_j \geqslant \underline{u}_j, j = 1, \dots, n\}$,

where $a_{ij} \geq 0$ for all $i = 1, \ldots, m$, $j = 1, \ldots, n$, and $\xi_j \geq \underline{u}_j$ $\forall \xi_j \in \mathscr{X}_j$ for $j = 1, \ldots, n$.

Recall that we define $\bar{\xi}_j = \mathrm{ess\,sup}\{\xi_j \mid \xi \in \mathscr{X}\}$. We first consider the case where $\bar{\xi}_j < \infty$ for all j. Note that $\mathscr{V} = \{u \wedge \xi \mid u \in \mathscr{U}, \xi \in \mathscr{X}\}$ is equivalent to the following set $\{w \mid Aw \leq b, \underline{u}_j \leq w_j \leq \bar{\xi}_j, j = 1, \ldots, n\}$, denoted by $\mathscr{V}_w$. For any $w = u \wedge \xi \in \mathscr{V}$, we have $Aw = A(u \wedge \xi) \leq b$ since $a_{ij} \geq 0$ for all i, j and $Au \leq b$; $\underline{u}_j \leq u_j \wedge \xi_j = w_j \leq \bar{\xi}_j$ since $\underline{u}_j \leq \xi_j \ \forall \xi_j \in \mathscr{X}_j$. For any $w \in \mathscr{V}_w$, let $u = w$, $\xi_j = \bar{\xi}_j$ for all j. Then $w = u \wedge \xi$ since $w_j \leq \bar{\xi}_j$ for all j, and $u \in \mathscr{U}$. Hence, $\mathscr{V} = \mathscr{V}_w$. Clearly $\mathscr{V}$ is a convex set. Given any u satisfying $u \wedge \xi \in \mathscr{V} \ \forall \xi \in \mathscr{X}$, we define u' such that for $j = 1, \ldots, n$,

$$u'_j = \begin{cases} u_j, & \text{if } u_j \leq \bar{\xi}_j, \\ \bar{\xi}_j, & \text{if } u_j > \bar{\xi}_j. \end{cases}$$

One can easily check that $u' \wedge \xi = u \wedge \xi \ \forall \xi \in \mathscr{X}$. We only need to show $u' \in \mathscr{U}$. Since $\bar{\xi}_j \geq \underline{u}_j$ and $u_j \geq \underline{u}_j$, we have $u'_j \geq \underline{u}_j$ for $j = 1, \ldots, n$. Because $A(u \wedge \xi) \leq b \ \forall \xi \in \mathscr{X}$ and Ξ has independent components, we obtain $A(u \wedge \bar{\xi}) \leq b$, which is the same as $Au' \leq b$.

If $\bar{\xi}_j = \infty$ for any j, then $u'_j = u_j$ and following similar arguments we can obtain the desired results.   Q.E.D.

2.2. Preservation of Structural Properties

One advantage of our transformation technique is that it can be used to establish the preservation of not only convexity and submodularity but also $L^\natural$-convexity under optimization operations, which plays a critical role in characterizing the structure of the optimal policies for many dynamic decision-making problems and facilitates their efficient computations. To see this, we first provide a brief review of the concept of $L^\natural$-convexity and some structural properties. The concept of $L^\natural$-convexity was introduced by Murota (1998) to extend convex analysis from real space to spaces with integers (see Murota 2009 for a survey of the recent developments in discrete convex analysis). It was first introduced into the inventory management literature by Lu and Song (2005) and used by Zipkin (2008) to characterize the optimal structural policy of lost-sales inventory models with positive lead times. Since then, $L^\natural$-convexity was found to be powerful enough to establish the structures of optimal policies in various other inventory models: serial inventory systems (Huh and Janakiraman 2010), inventory-pricing models with positive lead times (Pang et al. 2012), perishable inventory models (Chen et al. 2014), etc.

In the transformed problem the decisions are $v = (v(\xi) \mid \xi \in \mathscr{X}) \in (\mathbb{F}^n)^{\mathscr{X}}$. Note that the direct product of lattices is still a lattice under the componentwise partial order (see Example 2.2.3 (d) of Topkis 1998). Therefore, if X_α is a lattice for each $\alpha \in \mathbb{A}$, where $\mathbb{A}$ is an index set, then the direct product of sets X_α, is also a lattice. In the following, we present the definition of

$L^\natural$-convexity with domain $\mathscr{Y} \triangleq (\mathbb{F}^n)^{\mathscr{X}}$, where $\mathbb{A}$ is any index set.

Definition 1. A function $f \colon \mathscr{Y} \to \bar{\mathbb{R}}$ is $L^\natural$-convex if for any $x, x' \in \mathscr{Y}$, $\lambda \in \mathbb{F}_+$,

$$f(x) + f(x') \geq f((x + \lambda e) \wedge x') + f(x \vee (x' - \lambda e)),$$

where e is the all-ones vector in $\mathscr{Y}$. A set $\mathscr{V} \subseteq \mathscr{Y}$ is said to be $L^\natural$-convex if its indicator function $\delta_{\mathscr{V}}$ is $L^\natural$-convex.

For an $L^\natural$-convex function f, its effective domain $\mathrm{dom}(f) = \{x \in \mathscr{Y} \mid f(x) < +\infty\}$ is an $L^\natural$-convex set. We sometimes say a function f is $L^\natural$-convex on a set $\mathscr{V}$ with the understanding that $\mathscr{V}$ is an $L^\natural$-convex set and the extension of f to the whole space by defining $f(\mathbf{v}) = +\infty$ for $v \notin \mathscr{V}$ is $L^\natural$-convex. One can also show that an $L^\natural$-convex function restricted to an $L^\natural$-convex set is also $L^\natural$-convex. Following a similar proof in Simchi-Levi et al. (2014), we can show that an equivalent definition of $L^\natural$-convexity is given as follows: A function $f \colon \mathscr{Y} \to \bar{\mathbb{R}}$ is $L^\natural$-convex if and only if $g(x, \xi) \triangleq f(x - \xi e)$ is submodular in $(x, \xi) \in \mathscr{Y} \times \mathscr{S}$, where $\mathscr{S}$ is the intersection of $\mathbb{F}$ and any unbounded interval in $\mathbb{R}$, and e is the all-ones vector in $\mathscr{Y}$.

We now list some of the commonly used properties of $L^\natural$-convexity. To describe the monotonicity of optimal solution sets, we use the *induced set ordering* $\sqsubseteq$, which defines $X' \sqsubseteq X''$ for two nonempty sets X' and X'' if $x' \in X'$ and $x'' \in X''$ imply that $x' \wedge x'' \in X'$ and $x' \vee x'' \in X''$ (see Topkis 1998, p. 32). For a nonempty set X_t that depends on the parameter t in a partial order set T, we say that X_t is increasing in t on T if $\{X_t, t \in T\}$ has the induced set ordering $\sqsubseteq$. The proofs of these properties are relegated to the appendix.

Proposition 1 ($L^\natural$-Convexity). (a) *Any nonnegative linear combination of $L^\natural$-convex functions is $L^\natural$-convex. That is, if $f_i \colon \mathscr{Y} \to \bar{\mathbb{R}} (i = 1, 2, \ldots, n)$ are $L^\natural$-convex, then for any scalar $\alpha_i \geq 0$, $\sum_{i=1}^{m} \alpha_i f_i$ is also $L^\natural$-convex.*

(b) *If f_k is $L^\natural$-convex for $k = 1, 2, \ldots$ and $\lim_{k \to \infty} f_k(x) = f(x)$ for any $x \in \mathscr{Y}$, then f is $L^\natural$-convex.*

(c) *Assume that a function $f(\cdot, \cdot)$ is defined on the product space $\mathscr{Y} \times \mathbb{F}^m$. If $f(\cdot, y)$ is $L^\natural$-convex for any given $y \in \mathbb{F}^m$, then for a random vector ζ defined on $\mathbb{F}^m$, $E_\zeta[f(x, \zeta)]$ is $L^\natural$-convex, provided it is well defined.*

(d) *If $f \colon \mathscr{Y} \to \bar{\mathbb{R}}$ is an $L^\natural$-convex function, then $g \colon \mathscr{Y} \times \mathbb{F} \to \bar{\mathbb{R}}$ defined by $g(x, \lambda) = f(x - \lambda e)$ is also $L^\natural$-convex.*

(e) *Assume that $\mathscr{A}$ is an $L^\natural$-convex set of $\mathbb{F}^n \times \mathscr{Y}$ and $f(\cdot, \cdot) \colon \mathbb{F}^n \times \mathscr{Y} \to \bar{\mathbb{R}}$ is an $L^\natural$-convex function. Then the function*

$$g(x) = \inf_{y \colon (x, y) \in \mathscr{A}} f(x, y) \tag{10}$$

is $L^\natural$-convex over $\mathbb{F}^n$ if $g(x) \neq -\infty$ for any $x \in \mathbb{F}^n$.

(f) *Let e and $\tilde{e}$ be the all-ones vectors corresponding to the state space of x and the decision space of y, respectively, in (10). Then $\arg\min_{y \colon (x, y) \in \mathscr{A}} f(x, y)$ is increasing in x and*

$$\operatorname*{arg\,min}_{y \colon (x + \omega e, y) \in \mathscr{A}} f(x + \omega e, y) \sqsubseteq \omega \tilde{e} + \operatorname*{arg\,min}_{y \colon (x, y) \in \mathscr{A}} f(x, y).$$

(g) *Denote x_i a component of $x \in \mathcal{Y}$. A set with a representation $\{x \in \mathcal{Y}: l \leqslant x \leqslant u, x_i - x_j \leqslant v_{ij}, \forall i \neq j\}$, is $L^\natural$-convex in the space $\mathcal{Y}$, where $l, u \in \mathcal{Y}$ and $v_{ij} \in \mathbb{F}$.*

We now show how our transformation technique can be used to establish preservation properties of convexity, submodularity, and $L^\natural$-convexity under optimization operations.

Consider the following optimization problem

$$g(x, z) = \inf_{u:(x,z,u) \in \mathcal{A}} E[f(x, u \lozenge_k (z + \Xi))], \quad (11)$$

where $f(\cdot, \cdot): \mathbb{F}^m \times \mathbb{F}^n \to \bar{\mathbb{R}}$, $x \in \mathbb{F}^m, z \in \mathbb{F}^n$ and set $\mathcal{A} \subseteq \mathbb{F}^m \times \mathbb{F}^n \times \mathbb{F}^n$ is nonempty.

Define a set

$$\mathcal{A}^\Xi = \{(x, z, w) \mid w = u \lozenge_k (z + \xi), (x, z, u) \in \mathcal{A}, \xi \in \mathcal{X}\}.$$

Similar to Assumption 1, we specify the following condition:

Assumption 2. (a) *For any (x, z, u) such that $(x, z, u \lozenge_k (z + \xi)) \in \mathcal{A}^\Xi \; \forall \xi \in \mathcal{X}$, there exists $(x, z, u') \in \mathcal{A}$ such that $u' \lozenge_k (z + \xi) = u \lozenge_k (z + \xi) \; \forall \xi \in \mathcal{X}$.*

(b) *The indicator function of the set $\mathcal{A}^\Xi$ is componentwise convex in w (componentwise discrete convex if $\mathbb{F} = \mathbb{Z}$).*

Similar to Lemma 2, we provide an example with linear constraints, which satisfies Assumption 2. The proof is similar and thus omitted for brevity.

Lemma 3. *Assume that $\mathcal{A} = \{(x, z, u) \mid Au \leqslant b, u_1 \geqslant \underline{u}_1, \ldots, u_k \geqslant \underline{u}_k, u_{k+1} \leqslant \bar{u}_{k+1}, \ldots, u_n \leqslant \bar{u}_n\}$, where $b, \underline{u}_1, \ldots, \underline{u}_k, \bar{u}_{k+1}, \ldots, \bar{u}_n$ are parameters that may depend on x and z, $A = (a_{ij})$ with entries $a_{ij} \geqslant 0$ for any i and $j = 1, \ldots, k$, and $a_{ij} \leqslant 0$ for any i and $j = k+1, \ldots, n$. In addition $\mathcal{X}_j$ is contained in $[\underline{u}_j - z_j, +\infty)$ for $j = 1, \ldots, k$, and $\mathcal{X}_j$ is contained in $(-\infty, \bar{u}_j - z_j]$ for $j = k+1, \ldots, n$. Then Assumption 2 is satisfied.*

Now we are ready to present our main result in this section.

Theorem 3 (Preservation). *Consider the optimization problem (11), where $f(x, \cdot)$ and Ξ satisfy the assumptions in Theorem 1 for any given x. If Assumption 2 is satisfied, then we have the following results:*

(a) *If f and $\mathcal{A}^\Xi$ are convex, then g is also convex.*

(b) *If f is submodular and $\mathcal{A}^\Xi$ is a lattice, then g is also submodular.*

(c) *If f and $\mathcal{A}^\Xi$ are $L^\natural$-convex, then g is also $L^\natural$-convex.*

Proof. Theorem 2 implies that problem (11) can be equivalently converted to the following one:

$$\begin{aligned}
\inf \quad & E[f(x, v_1(\Xi_1), \ldots, v_n(\Xi_n))] \\
\text{s.t.} \quad & v_j(\xi_j) \leqslant z_j + \xi_j \quad \forall \xi_j \in \mathcal{X}_j, \forall j = 1, \ldots, k, \\
& v_j(\xi_j) \geqslant z_j + \xi_j \quad \forall \xi_j \in \mathcal{X}_j, \forall j = k+1, \ldots, n, \\
& (x, z, v_1(\xi_1), \ldots, v_n(\xi_n)) \in \mathcal{A}^\Xi \quad \forall \xi \in \mathcal{X}. \quad (12)
\end{aligned}$$

To see this, given fixed (x, z), let $\mathcal{U}(x, z)$ denote the constraint set $\{u: (x, z, u) \in \mathcal{A}\}$, $f_x(u) = f(x, u)$, $\tilde{\Xi}_z = z + \Xi$, and $\tilde{\mathcal{X}}_z = \mathrm{Supp}(\tilde{\Xi}_z)$. Then (11) is equivalent to

$$\inf_{u \in \mathcal{U}(x,z)} E[f_x(u \lozenge_k \tilde{\Xi}_z)]. \quad (13)$$

Let $\mathcal{V}(x, z) = \{u \lozenge_k \tilde{\xi}: u \in \mathcal{U}(x, z), \tilde{\xi} \in \tilde{\mathcal{X}}_z\}$. Given any $u \lozenge_k \tilde{\xi} \in \mathcal{V}(x, z) \; \forall \tilde{\xi} \in \tilde{\mathcal{X}}_z$, we have (x, z, u) satisfying $(x, z, u \lozenge_k (z + \xi)) \in \mathcal{A}^\Xi \; \forall \xi \in \mathcal{X}$. According to Assumption 2, there exists $(x, z, u') \in \mathcal{A}$ such that $u' \lozenge_k (z + \xi) = u \lozenge_k (z + \xi) \; \forall \xi \in \mathcal{X}$. Thus we have $u' \in \mathcal{U}(x, z)$ and $u' \lozenge_k \tilde{\xi} = u \lozenge_k \tilde{\xi} \; \forall \tilde{\xi} \in \tilde{\mathcal{X}}_z$. If the indicator function of $\mathcal{A}^\Xi$ is componentwise convex in w, it is clear that the indicator function of $\mathcal{V}(x, z)$ is also componentwise convex. Therefore, if Assumption 2 is satisfied, then Assumption 1 is also satisfied. According to Theorem 2, we can transform (13) into

$$\begin{aligned}
\inf \quad & E[f_x(v_1(\tilde{\Xi}_{z1}), \ldots, v_n(\tilde{\Xi}_{zn}))] \\
\text{s.t.} \quad & v_j(\tilde{\xi}_j) \leqslant \tilde{\xi}_j \quad \forall \tilde{\xi}_j \in \tilde{\mathcal{X}}_{zj}, \forall j = 1, \ldots, k, \\
& v_j(\tilde{\xi}_j) \geqslant \tilde{\xi}_j \quad \forall \tilde{\xi}_j \in \tilde{\mathcal{X}}_{zj}, \quad \forall j = k+1, \ldots, n, \\
& (v_1(\tilde{\xi}_1), \ldots, v_n(\tilde{\xi}_n)) \in \mathcal{V}(x, z) \quad \forall \tilde{\xi} \in \tilde{\mathcal{X}}_z,
\end{aligned}$$

which is equivalent to (12).

It is straightforward to check that the constraint set involving $(x, z, (v_1(\xi_1), \ldots, v_n(\xi_n))_{\xi \in \mathcal{X}}$ is a convex set, a lattice, and an $L^\natural$-convex set (Proposition 1 part (g)) on the product set $\mathbb{F}^m \times \mathbb{F}^n \times (\mathbb{F}^n)^{\mathcal{X}}$ for cases (a), (b), and (c), respectively.

In the following, we show that the objective function $E[f(x, v_1(\Xi_1), \ldots, v_n(\Xi_n))]$ is convex, submodular, $L^\natural$-convex in $(x, (v_1(\xi_1), \ldots, v_n(\xi_n))_{\xi \in \mathcal{X}} \in \mathbb{F}^m \times (\mathbb{F}^n)^{\mathcal{X}}$ for cases (a), (b), and (c), respectively. Define $\tilde{f}: \mathbb{F}^m \times (\mathbb{F}^n)^{\mathcal{X}} \times \mathcal{X} \to \bar{\mathbb{R}}$ such that $\tilde{f}(x, v, \xi) \triangleq f(x, v(\xi)) \; \forall \xi \in \mathcal{X}$. Clearly, $E[\tilde{f}(x, v, \Xi)] = E[f(x, v(\Xi))]$. Given any realization ξ, if $f(\cdot, \cdot)$ is convex/submodular/$L^\natural$-convex, then one can easily prove by definition that $\tilde{f}(\cdot, \cdot, \xi)$ is also convex/submodular/$L^\natural$-convex. We show the proof for convexity; the proofs for submodularity and $L^\natural$-convex are similar and simply follow their definitions, respectively. Given ξ, for any (x, v), (x', v'), and $\lambda \in [0, 1]$, we have

$$\begin{aligned}
& \tilde{f}(\lambda x + (1-\lambda)x', \lambda v + (1-\lambda)v', \xi) \\
& \quad = f(\lambda x + (1-\lambda)x', \lambda v(\xi) + (1-\lambda)v'(\xi)) \\
& \quad \leqslant \lambda f(x, v(\xi)) + (1-\lambda)f(x', v'(\xi)) \\
& \quad = \lambda \tilde{f}(x, v, \xi) + (1-\lambda)\tilde{f}(x', v', \xi).
\end{aligned}$$

Since $\tilde{f}(\cdot, \cdot, \xi)$ is convex/submodular/$L^\natural$-convex for any given ξ, we have that the objective function $E[f(x, v(\Xi))] = E[\tilde{f}(x, v, \Xi)]$ is also convex/submodular/$L^\natural$-convex because of Proposition 1 part (c).

Part (a) follows immediately from the theorem of convexity preservation under minimization (see Simchi-Levi et al. 2014, Proposition 2.1.15, for the case with

finite-dimensional spaces, and Zălinescu 2002, Theorem 2.1.3(v), for the case with general vector spaces). Part (b) follows from Topkis (1998, Theorem 2.7.6). Part (c) follows from Proposition 1 part (e).   Q.E.D.

The following theorem characterizes the monotonicity properties of the solution set to the optimization problem (11). The proof is relegated to the appendix.

Theorem 4. *Consider the optimization problem* (11), *where* $f(x, \cdot)$ *and* Ξ *satisfy the assumptions in Theorem 1 for any given* x. *Let* $\mathcal{U}^*(x, z)$ *denote the optimal solution set of* (11) *and let* $\hat{e}$ *and* $\tilde{e}$ *be the all-ones vectors in* $\mathbb{F}^m$ *and* $\mathbb{F}^n$, *respectively, and* $e = (\hat{e}, \tilde{e})$. *If Assumption 2 is satisfied,* $\mathscr{A}, \mathscr{A}^{\Xi}$ *are closed, and in addition* $u_j \leqslant z_j + \bar{\xi}_j, j = 1, \ldots, k, u_j \geqslant z_j + \underline{\xi}_j,$ $j = k+1, \ldots, n$, *then we have the following results:*

(a) *If* f *is a submodular function, and* $\mathscr{A}, \mathscr{A}^{\Xi}$ *are lattices, then* $\mathcal{U}^*(x, z)$ *is increasing in* (x, z). *There exist a greatest element and a least element in* $\mathcal{U}^*(x, z)$, *which are increasing in* (x, z).

(b) *If* f *is an* $L^{\natural}$-*convex function, and* $\mathscr{A}, \mathscr{A}^{\Xi}$ *are* $L^{\natural}$-*convex sets, then* $\mathcal{U}^*(x, z)$ *is increasing in* (x, z) *and* $\mathcal{U}^*((x, z) + \omega e) \sqsubseteq \mathcal{U}^*(x, z) + \omega \tilde{e}$ *for any* $\omega > 0$. *Within* $\mathcal{U}^*(x, z)$, *there exist a greatest element and a least element that satisfy the above monotonicity properties with limited sensitivity.*

In the following, we provide an example to show that the assumption $u_j \leqslant z_j + \bar{\xi}_j, j = 1, \ldots, k, u_j \geqslant z_j + \underline{\xi}_j,$ $j = k+1, \ldots, n$ is needed.

Example 1. Suppose that $f(u) = u^2$ and the support of Ξ is $[-3, -1]$. Let $\mathcal{U}^*(z) = \arg\min_{u \in \mathcal{U}} E[f(u \wedge (z + \Xi))]$, and $z = 0, \omega = 2$. When $\mathcal{U} = \mathbb{R}$, we have $\mathcal{U}^*(z) = [-1, \infty)$, $\mathcal{U}^*(z + \omega) = \{0\}$. Notice that $\mathcal{U}^*(z) \sqsubseteq \mathcal{U}^*(z + \omega)$ does not hold. However, when $\mathcal{U}(z) = \{u \in \mathbb{R}: u \leqslant z + \bar{\xi}\}$, we have $\mathcal{U}(z) = (-\infty, -1], \mathcal{U}(z + \omega) = (-\infty, 1]$. Then $\mathcal{U}^*(z) = \{-1\}, \mathcal{U}^*(z + \omega) = \{0\}$. Clearly $\mathcal{U}^*(z) \sqsubseteq \mathcal{U}^*(z + \omega)$.

Notice that if the conditions in Lemma 3 are satisfied, then the assumptions $u_j \leqslant z_j + \bar{\xi}_j, j = 1, \ldots, k, u_j \geqslant z_j + \underline{\xi}_j, j = k+1, \ldots, n$ in Theorem 4 are without loss of generality. To see this, given any (x, z, u) that is feasible for problem (11), $(x, z, u_1 \wedge (z_1 + \bar{\xi}_1), \ldots, u_k \wedge (z_k + \bar{\xi}_k),$ $u_{k+1} \vee (z_{k+1} + \underline{\xi}_{k+1}), \ldots, u_n \vee (z_n + \underline{\xi}_n))$ is also feasible and yields the same objective value. In all of our applications, the constraint set satisfies the conditions in Lemma 3.

In the following three sections, we apply the transformation technique and the relevant preservation results to three fundamental models from inventory and revenue management literature and demonstrate how these results facilitate the structural analyses.

3. Dual Sourcing Under Supply Capacity Uncertainty

Consider a firm managing a T-period periodic-review inventory system in the presence of two capacitated suppliers (or delivery modes): a regular supplier with a longer replenishment lead time of l_R periods and a unit ordering cost c_R, and an expedited (emergency) supplier with a shorter replenishment lead time of l_E periods and a unit ordering cost c_E, where l_R and l_E are nonnegative integers and $l_R > l_E$. There are no fixed ordering costs. Both suppliers offer limited and uncertain capacities, denoted by $K_{R,t}$ and $K_{E,t}, t \in \{1, \ldots, T\}$, for regular and expedited suppliers, respectively. The processes $\{K_{R,t}\}_{t=1}^T$ and $\{K_{E,t}\}_{t=1}^T$ are both independent over time and independent of each other. Note that the independence assumption on the supply capacity distributions can be justified by the dual-sourcing practice with two geographically distant locations, such as China and Mexico in the case study of Van Mieghem (2008), where the production processes are typically independent of each other. Demands of successive periods, denoted by D_t for period t, are stochastic, independent over time, and independent of the supply capacities. For convenience, let $D_{[t,t+l]}$ be the total demand from period t to period $t + l$, i.e., $D_{[t,t+l]} = D_t + \cdots + D_{t+l}$. We use d_t and $d_{[t,t+l]}$ to denote the realization of D_t and $D_{[t,t+l]}$.

It is notable that a typical assumption in the dual-sourcing literature without capacity limits is that the expedited ordering cost c_E is greater than the regular ordering cost c_R, because otherwise it is trivial for the firm to procure exclusively from the expedited supplier (see, e.g., Veeraraghavan and Scheller-Wolf 2006, Sheopuri et al. 2010). We do not make this assumption here. In fact, if the expedited capacity is limited, even when the regular ordering cost is higher, it may still be beneficial to order from the regular supplier.

The sequence of events is as follows. At the beginning of period t, orders from the regular supplier l_R periods ago and the expedited supplier l_E periods ago (if $l_E \geqslant 1$) are received. (Note that if $l_E = 0$, we assume that an order from the expedited supplier is received right away.) The firm then reviews the inventory level and the orders outstanding, and determines how much to order from the two suppliers before observing the suppliers' capacities $K_{R,t}$ and $K_{E,t}$. Let q_R and q_E be the (target) order quantities from the regular and expedited channels, respectively. After the orders are placed, the suppliers' capacities $K_{R,t}$ and $K_{E,t}$ are realized. We use $k_{R,t}$ and $k_{E,t}$ to denote realizations of $K_{R,t}$ and $K_{E,t}$, respectively. Then the amounts of inventories shipped from the regular and expedited suppliers are $q_R \wedge k_{R,t}$ and $q_E \wedge k_{E,t}$, respectively. Note that here we assume that the supply capacity uncertainties are resolved in the same period when the orders are placed (see Federgruen and Yang 2011 for a similar treatment for the random yield problem). This is reasonable when the capacity uncertainties are mainly driven by the unreliability of the production process and the production time is no more the ordering

costs are given by $c_R(q_R \wedge k_{R,t})$ and $c_E(q_E \wedge k_{E,t})$. Here we assume that the ordering cost is proportional to the quantity actually delivered, which is a common assumption in the literature of inventory control with random capacities (see Ciarallo et al. 1994, Wang and Gerchak 1996, Feng 2010, and so on). This assumption is appropriate when the payment is made upon the receipt of the shipments and the firms only pay the actual delivered amount. At the end of this period, the demand is realized and met with on-hand inventory (if any). Unmet demand is fully backlogged with a unit shortage cost h^-. Excess inventory is carried over to the next period with a unit holding cost h^+.

The objective of the firm is to find a dual-sourcing strategy so as to minimize the total expected discounted cost, including ordering cost, holding cost, and backorder cost, over the planning horizon. To present the dynamic programming model for deriving the optimal strategy, one can naturally describe the system state right before the firm places orders by a vector $s = (s_0, \ldots, s_{l_R-1})$, where s_i denotes the amount of on-hand net inventory plus outstanding orders that will arrive within i periods, $i = 1, \ldots, l_R - 1$. However, in a backlogging model, since the orders of each period will have an influence only l_E periods later, and the on-hand net inventory level l_E periods later solely depends on s_{l_E}, it suffices to use the now standard accounting technique to discount the future inventory cost to the current period and focus on the pipeline inventory levels $s_{l_E}, \ldots, s_{l_R-1}$. Specifically, we can reduce the state space to $k = l_R - l_E$ dimensions by defining the system state as $z = (z_1, \ldots, z_k)$, where $z_i = s_{i+l_E-1}, i = 1, \ldots, k$. The state space is given by

$$\mathscr{S} = \{(z_1, \ldots, z_k): z_1 \leqslant z_2 \leqslant \cdots \leqslant z_k\}.$$

Given the system state z, the system state of the next period is given by

$$\tilde{z} = (z_2 + q_E \wedge k_{E,t} - d_t, \ldots, z_k + q_E \wedge k_{E,t} - d_t,$$
$$y \wedge (z_k + k_{R,t}) + q_E \wedge k_{E,t} - d_t),$$

where $y = z_k + q_R$ is the (target) order-up-to level from the regular channel. For reasons that will become clear later, we denote $u = -q_E$ and $\tilde{k}_{E,t} = -k_{E,t}$. The dynamics of the system state can be rewritten as

$$\tilde{z} = [(z_2, \ldots, z_k, y \wedge (z_k + k_{R,t})) - (u \vee \tilde{k}_{E,t} + d_t)e],$$

where e is the k-dimensional all-ones vector.

We are now ready to present the dynamic program to derive the firm's optimal strategy. Let $\alpha \in (0, 1]$ be the discount factor. The optimality equations can be written as follows. For $t = 1, \ldots, T$,

$$v_t(z) = \min_{y \geqslant z_k, u \leqslant 0} \{E[g_t(z, y \wedge (z_k + K_{R,t}), u \vee \tilde{K}_{E,t})]\}$$
$$\forall z \in \mathscr{S}, \quad (14)$$

where

$$g_t(z, y, u) = c_R(y - z_k) - c_E u + B_t(z_1 - u)$$
$$+ \alpha E[v_{t+1}((z_2, \ldots, z_k, y) - (D_t + u)e)], \quad (15)$$

and

$$B_t(x) = \alpha^{l_E} E[h^+(x - D_{[t,t+l_E]})^+ + h^-(D_{[t,t+l_E]} - x)^+].$$

Note that the expectation of the right-hand side of Equation (14) is taken over the random capacities. The function g_t represents the expected total discounted cost after the capacities are realized but before the demand is realized. The first term of the right-hand side of Equation (15) is the ordering cost from the regular supplier, the second term is the ordering cost from the expedited supplier, the third term is the expected discounted holding and shortage cost of period $t + l_E$, and the last term is the expected total discounted future costs. For simplicity, we assume the terminal value function $v_{T+1}(z) = 0$ for any z, which implies that there is no salvage value for leftover inventory and no backlogging cost for unfilled demand after period $T + l_E$. That is, the firm makes decisions in the first T periods but takes into account the inventory cost up to period $T + l_E$. Our structural results and analysis still hold if $v_{T+1}(z)$ is assumed to be $L^\natural$-convex.

Problem (14) admits optimal solutions under rather general and standard conditions. Nevertheless, it is a challenging problem. First, the state space is multidimensional. A more severe issue is that the objective function of problem (14) is not convex. Note that for the last period with $v_{T+1} = 0$, the objective function has a structure similar to that in (1), which may not be convex. Thus, it is far from being clear whether the cost-to-go functions v_t are convex, and even if they are, the objective function of problem (14) is not. However, with the transformation technique developed in Section 2 we can convert the nonconvex minimization problem (14) into an equivalent convex minimization problem and show that v_t is actually $L^\natural$-convex.

In the following analysis, we assume that both c_E and c_R are smaller than $h^-/(1 - \alpha)$, which ensures that it is not optimal to never order anything and merely accumulate penalty costs. Let $(y_t(z), u_t(z))$ denote the optimal solution for problem (14). When there are multiple optimal solutions, we assume it is the greatest one, which will be shown to be well defined later.

Theorem 5. *For all t, $v_t(z)$ is $L^\natural$-convex in $z \in \mathscr{S}$. The optimal solution $(y_t(z), u_t(z))$ is increasing in z with limited sensitivity. (When there are multiple optimal solutions, we assume it is the greatest one.) That is, for any $\omega > 0$,*

$$y_t(z) \leqslant y_t(z + \omega e) \leqslant y_t(z) + \omega,$$
$$u_t(z) \leqslant u_t(z + \omega e) \leqslant u_t(z) + \omega. \quad (16)$$

Proof. The proof is by induction. Suppose that v_{t+1} is $L^{\natural}$-convex. By Proposition 1(d), for any d_t, $v_{t+1}[(z_2, \ldots, z_k, y) - (d_t + u)e]$ is $L^{\natural}$-convex in (z, y, u) and so is $\alpha v_{t+1}[(z_2, \ldots, z_k, y) - (d_t + u)e]$. Clearly all the other terms of g_t are $L^{\natural}$-convex in (z, y, u) (That is why we define $u = -q_E$ and $\tilde{k}_{e,t} = -k_{E,t}$). Thus, g_t is $L^{\natural}$-convex in (z, y, u). Let $\mathscr{A} = \{(z, y, u) \mid y \geqslant z_k, u \leqslant 0\}$ and

$$\mathscr{A}^{\Xi} = \{(z, y \wedge (z_k + k_{R,t}), u \vee \tilde{k}_{e,t}) \mid y \geqslant z_k, u \leqslant 0,$$
$$k_{R,t} \in \mathrm{Supp}(K_{R,t}), \tilde{k}_{e,t} \in \mathrm{Supp}(\tilde{K}_{E,t})\}.$$

Since $K_{R,t} \geqslant 0$ and $\tilde{K}_{E,t} \leqslant 0$ almost surely, it is easy to see that

$$\mathscr{A}^{\Xi} = \{(z, w_1, w_2) \mid z_k + k^u_{R,t} \geqslant w_1 \geqslant z_k, \tilde{k}^l_{E,t} \leqslant w_2 \leqslant 0\},$$

where we set $k^u_{R,t} = \mathrm{ess\,sup}(\mathrm{Supp}(K_{R,t}))$ and $\tilde{k}^l_{E,t} = \mathrm{ess\,inf}(\mathrm{Supp}(\tilde{K}_{E,t}))$.

The constraint set $\mathscr{A}^{\Xi}$ forms an $L^{\natural}$-convex set because of Proposition 1(g). It is straightforward to see that the set $\mathscr{A} = \{(z, y, u) \mid y \geqslant z_k, u \leqslant 0\}$ is of the form in Lemma 3. Applying Theorem 3, we know $v_t(z)$ is $L^{\natural}$-convex in $z \in \mathscr{S}$. According to Theorem 4, the greatest optimal solution $(y_t(z), u_t(z))$ is well defined and has the desired monotonicity property with limited sensitivity. Q.E.D.

The monotonicity and limited sensitivity of $y_t(z)$ imply that the optimal regular order quantity $q_{R,t}(z)$, which is equal to $y_t(z) - z_k$, increases in $z_1, \ldots, z_{k-1}$, but decreases in z_k and satisfies $-\omega \leqslant q_{R,t}(z + \omega e) - q_{R,t}(z) \leqslant 0$.

To gain more insights, we can transform the state vector to $x = (x_1, \ldots, x_k)$, where $x_1 = z_1$ and $x_i = z_i - z_{i-1}$, $i = 2, \ldots, k$. Note that $x_1 = z_1$ represents the amount of on-hand net inventory plus outstanding orders that will arrive within l_E periods, and x_i represents the size of the outstanding order that will arrive $l_E + i - 1$ periods later. Denote the corresponding optimal order quantities by $\hat{q}_{R,t}(x) = q_{R,t}(z)$ and $\hat{q}_{E,t}(x) = q_{E,t}(z)$. The monotonicity and limited sensitivity of $y_t(z)$ imply the following inequalities:

$$-\omega \leqslant \hat{q}_{R,t}(x + \omega e_k) - \hat{q}_{R,t}(x) \leqslant \hat{q}_{R,t}(x + \omega e_{k-1}) - \hat{q}_{R,t}(x)$$
$$\leqslant \cdots \leqslant \hat{q}_{R,t}(x + \omega e_1) - \hat{q}_{R,t}(x) \leqslant 0. \tag{17}$$

Compare states $z + \omega e_i$ and z. For $i = 1$, the former has ω more units of on-hand inventory or outstanding orders that will arrive within l_E periods. For $i = 2, \ldots, k$, the former has ω more units of outstanding order that will arrive $l_E + i - 1$ periods later. Thus, inequalities (17) imply that the regular order quantity decreases in on-hand inventory level and the sizes of the outstanding orders. The sensitivity decreases in the age of the outstanding order, where the age refers to the number of periods passed since the order was placed. In other

words, the regular order quantity is most sensitive to the size of the most recently placed order.

Similarly, for the expedited order quantity $\hat{q}_{E,t}(x) = -u_t(z)$, we have

$$-\omega \leqslant \hat{q}_{E,t}(x + \omega e_1) - \hat{q}_{E,t}(x) \leqslant \hat{q}_{E,t}(x + \omega e_2) - \hat{q}_{E,t}(x)$$
$$\leqslant \cdots \leqslant \hat{q}_{E,t}(x + \omega e_k) - \hat{q}_{E,t}(x) \leqslant 0. \tag{18}$$

That is, the expedited order quantity decreases in the sizes of outstanding orders in the pipeline, but the sensitivity increases in the age of the outstanding order. In other words, the expedited order quantity is least sensitive to the most recently placed order, which is opposite to the sensitivity of the regular order quantity.

Such monotone properties with limited sensitivity are also observed in Hua et al. (2015) who consider an uncapacitated dual-sourcing problem, and in the joint inventory-pricing control problems with positive lead time where the replenishment decision has a decreasing sensitivity in the age of the outstanding order, whereas the pricing decision has an increasing sensitivity in the age of the outstanding order (see, e.g., Chen et al. 2014). The implication is that the decisions whose immediate impacts are closer to the on-hand stock (e.g., pricing or expedited order) are more sensitive to the on-hand inventory level and older outstanding orders while the decisions whose immediate impacts are further away from the on-hand stock (e.g., regular order) is more sensitive to the younger outstanding orders.

4. Assemble-to-Order Systems with Random Capacity

Consider an ATO system over a planning horizon with T periods. The ATO system consists of m components indexed by $i \in \{1, 2, \ldots, m\}$ and n products indexed by $j \in \{1, 2, \ldots, n\}$. At the beginning of each period, the firm observes on-hand inventory levels of the m components $x = (x_1, \ldots, x_m)^T$, and then decides the order-up-to inventory levels of components $y = (y_1, \ldots, y_m)^T$. The delivered quantity of each component i cannot exceed a random capacity, denoted by $\Xi_{t,i}$, which is realized after the order is placed. The capacities are independent of each other and over time. Inventory replenishment lead time is assumed to be zero. The demand for product j in period t is $D_{t,j}$ and we assume that they are independent over time and independent of capacities. Let $D_t = (D_{t,1}, \ldots, D_{t,n})^T$. The bill of materials is specified by an $m \times n$ matrix A, whose component a_{ij} denotes the units of component i required to make one unit of product j. Unmet demands are assumed to be lost. Let c_i and h_i represent the ordering cost and holding cost of component i per unit, respectively, and b_j denote the per unit shortage cost of product j. We use c, h, b to denote the vectors $(c_1, \ldots, c_m)^T$, $(h_1, \ldots, h_m)^T$, $(b_1, \ldots, b_n)^T$, respectively. The one-period

discount factor is $\alpha \in (0,1]$. The objective of the firm is to minimize the total expected discounted cost.

Let $f_t(x)$ be the cost-to-go function with initial inventory levels x at the beginning of period t. We omit the subscript t for notational brevity when no ambiguity occurs. The optimality equation is

$$f_t(x) = \min_{y \geqslant x} \{E[c^T(y \wedge (x+\Xi) - x)] + E[g_t(y \wedge (x+\Xi)|D)]\},$$
(19)

where

$$g_t(z \mid d) = \min_{u:(z,u) \in \mathcal{U}(d)} \{\mathcal{L}(z, u \mid d) + \alpha f_{t+1}(z - Au)\}, \quad (20)$$

and

$$\mathcal{L}(z, u \mid d) = h^T(z - Au) + b^T(d - u). \quad (21)$$

The boundary condition is assumed to be $f_{T+1}(x) = 0$ without loss of generality. The first term in the objective function of (19) is the ordering cost. Similar to the dual-sourcing model, we assume that the ordering cost is proportional to the quantity actually delivered. The feasible set in (20) is given by $\mathcal{U}(d) = \{(z, u) \mid Au \leqslant z, 0 \leqslant u_j \leqslant d_j, j = 1, 2, \ldots, n\}$, where z is the on-hand inventory level after the inventory ordered in the current period arrives, and u is the vector of assembled-product quantities. The inventory holding and shortage costs are given in $\mathcal{L}(z, u \mid d)$.

Because of the complexity of general ATO systems, some important special systems are studied in the literature, one of which is a generalized M-system (see Nadar et al. 2014). A generalized M-system has m components and $m + 1$ products, where each product i requires a single unit of component i for $i \leqslant m$ and product $m + 1$ consumes one unit of each component. This ATO system reduces to an M-system when $m = 2$. The bill of materials matrix has the following form:

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 1 \end{pmatrix}. \quad (22)$$

We summarize the structural results of this section in the following theorem.

Theorem 6. (a) *For a general ATO system, the optimal cost function $f_t(x)$ is convex in x for all t.*

(b) *For a generalized M-system, the optimal cost function $f_t(x)$ is $L^\natural$-convex in x for all t. The optimal order-up-to level $y_t(x)$ is increasing in x with limited sensitivity. That is, for any $\omega > 0$, $y_t(x) \leqslant y_t(x + \omega e) \leqslant y_t(x) + \omega e$. (When there are multiple optimal solutions, we assume it is the greatest one.)*

Proof. (a) We prove by induction. Assume that f_{t+1} is convex. It is easy to see that $g_t(z \mid d)$ in (20) is convex in z

for any demand realization d since the objective function is jointly convex in (z, u) and the constraints form a convex set. Define $G_t(z) \triangleq c^T z + E[g_t(z \mid D)]$, which is convex in z. Then $f_t(x) = \min_{y \geqslant x} E[G_t(y \wedge (x+\Xi))] - c^T x$. The constraint set is $\mathscr{A} = \{(x, y) \mid y \geqslant x\}$. By definition, $\mathscr{A}^\Xi = \{(x, y \wedge (x+\xi)) \mid y \geqslant x, \xi \in \text{Supp}(\Xi)\}$. This is equivalent to the set $\{(x, w) \mid x_i \leqslant w_i \leqslant x_i + \bar{\xi}_i, \forall i = 1, \ldots, m\}$, which is convex. In addition, Assumption 2 is satisfied since $\mathscr{A}$ is of the form given in Lemma 3. Therefore, following Theorem 3 we know that $f_t(x)$ is convex in x.

(b) For $j = 1, \ldots, m$, define $\hat{u}_j = z_j - u_j$. Let $\hat{u}_{m+1} = u_{m+1}$ and

$$\hat{A} = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & -1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & -1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}. \quad (23)$$

Then $g_t(y \mid d)$ can be written as

$$g_t(z \mid d) = \min_{\hat{u}:(z,\hat{u}) \in \hat{\mathcal{U}}(d)} \{\hat{\mathcal{L}}(z, \hat{u} \mid d)$$
$$+ \alpha f_{t+1}(\hat{u}_1 - \hat{u}_{m+1}, \ldots, \hat{u}_m - \hat{u}_{m+1})\}, \quad (24)$$

where

$$\hat{\mathcal{L}}(z, \hat{u} \mid d) = \sum_{i=1}^{m} h_i(\hat{u}_i - \hat{u}_{m+1}) + \sum_{j=1}^{m} b_j(d_j - z_j + \hat{u}_j)$$
$$+ b_{m+1}(d_{m+1} - \hat{u}_{m+1}), \quad (25)$$

and

$$\hat{\mathcal{U}}(d) = \{(z, \hat{u}) \mid \hat{A}\hat{u} \leqslant 0, 0 \leqslant z_j - \hat{u}_j \leqslant d_j,$$
$$j = 1, 2, \ldots, m, 0 \leqslant \hat{u}_{m+1} \leqslant d_{m+1}\}. \quad (26)$$

We then prove by induction. Clearly $f_{T+1}(x)$ is $L^\natural$-convex. If f_{t+1} is $L^\natural$-convex then the objective function of (24) is also $L^\natural$-convex in $(z, \hat{u})$ because of Proposition 1(a) and (d). The constraint set $\hat{\mathcal{U}}(d)$ forms a $L^\natural$-convex set by Proposition 1(g). Therefore, $g_t(z \mid d)$ is $L^\natural$-convex in z for any d according to Proposition 1(e). Similar to part (a) we have $\mathscr{A}^\Xi = \{(x, w) \mid x_i \leqslant w_i \leqslant x_i + \bar{\xi}_i, \forall i = 1, \ldots, m\}$, which is $L^\natural$-convex following Proposition 1(g). One can easily check that $\mathscr{A}$ is of the form given in Lemma 3. Therefore, applying Theorem 3 and 4, we know that $f_t(x)$ is also $L^\natural$-convex and the sensitivity results hold. Q.E.D.

Theorem 6 summarizes the sensitivity results for the stage when components are ordered. For a generalized M-system, the order-up-to level y_i of any component i increases in its own inventory level x_i as well as the inventory level of any other component $j \neq i$. The limited sensitivity implies that for any component i the ordering quantity $y_i - x_i$ decreases in the inventory level of any component. One can easily check that the following sensitivity results hold during the stage when

products are assembled. For a generalized M-system, the quantity of product $m + 1$ increases in the quantity of each component, while the quantity of product $j (\neq m + 1)$ increases in the quantity of component j but decreases in the quantity of component $k (\neq j)$.

5. Capacity Allocation in Network Revenue Management

We consider a network system consisting of m resources (airline seats in different legs), indexed by $i \in \{1, \dots, m\}$, with initial capacity levels $C = (C_1, \dots, C_m)^T$, and n products (itinerary-class combinations), indexed by $j \in \{1, \dots, n\}$. The corresponding prices, denoted by $p = (p_1, \dots, p_n)^T$, are exogenously given. Each product needs at most one unit of each resource. Let $A = (a_{ij})$ be the resource coefficient matrix, where $a_{ij} = 1$ if product j uses one unit of resource i and $a_{ij} = 0$ otherwise. Define $D_t = (D_{t,1}, \dots, D_{t,n})^T$, where $D_{t,j}$ is demand of product j in period t. Assume that the demands of different products are independent and the demands are independent over time. The objective of the firm is to decide the booking limits for all demand classes dynamically so as to maximize the total expected profit over the planning horizon.

As mentioned in Section 1, the model we consider here is MSSP in Chen and Homem-de-Mello (2010) with continuous relaxations. Chen and Homem-de-Mello (2010) point out that the major difficulty of the above model is that it is not a concave maximization problem, since the decisions are truncated by random demands. Therefore, they re-solve a sequence of two-stage stochastic programs for approximation. Interestingly, as we show in this section, our transformation technique can overcome this difficulty and allows us to preserve concavity in the dynamic programming recursions. Under a certain network structure, we further demonstrate that $L^\natural$-concavity can be preserved and use it to derive monotone properties of the optimal booking limits. Note that the model considered here is different from the one in Talluri and van Ryzin (2005, section 3.2.1). Their model assumes there is at most one demand request in any period. We do not impose this assumption. Since in our model each time period corresponds to the time when the firm needs to revise its capacity allocation policy, it may not be practical to divide the planning horizon so much so that there is at most one demand in any period because of the increased computational complexity.

In the following, we omit the subscript t for notational brevity when no ambiguity occurs. The state variable is denoted by the vector $x = (x_1, \dots, x_m)^T$ in which x_i is the capacity level of the resource i in the current period. At the beginning of the planning horizon, we have $x = C$. In each period, the firm observes the current capacity level x and decides the booking limits for different demand classes. The decision variable

is denoted by vector $u = (u_1, \dots, u_n)^T$, where u_j is the booking limit for class j demand in the current period. The action space can be defined as $\mathscr{A} = \{(x, u) \mid Au \leqslant x, u \geqslant 0\}$. Let $f_t(x)$ be the optimal value. The optimality equations can be expressed as

$$f_t(x) = \max_{u: (x, u) \in \mathscr{A}} E[p^T(u \wedge D) + f_{t+1}(x - A(u \wedge D))],$$
$$t = 1, \dots, T, \quad (27)$$

where $f_{T+1}(x) = 0$. For $\xi \in \mathbb{F}_+^m$, define the function $g_t: \mathbb{F}_+^{m+1} \to \mathbb{R}$ such that

$$g_t(x, \xi) = p^T \xi + f_{t+1}(x - A\xi).$$

Then the optimality equation can be expressed as

$$f_t(x) = \max_{u: (x, u) \in \mathscr{A}} E[g_t(x, u \wedge D)], \quad t = 1, \dots, T. \quad (28)$$

We also consider a special case where the resource coefficient matrix has the same format as the bill of materials matrix in the assemble-to-order generalized M-system, i.e., the resource coefficient matrix is given by (22). When the number of resources $m = 2$, one can relate this type of resource coefficient matrix to the following setting. There are two legs in the network: A to B and B to C. There are three types of consumers. Type one consumers travel from A to B, type two consumers travel from B to C, and type three consumers travel from A to C with a transition at B.

We summarize the structural results in the following theorem.

Theorem 7. (a) *For the network revenue management problem* (27), *the optimal value function $f_t(x)$ is concave in x for all t.*

(b) *If, in addition, the resource coefficient matrix is given by* (22), *then for all t, $f_t(x)$ is $L^\natural$-concave. The optimal booking limit $u_{m+1}^*(x)$ is increasing in x with limited sensitivity, i.e., for any $\omega > 0$, $u_{m+1}^*(x) \leqslant u_{m+1}^*(x + \omega e) \leqslant u_{m+1}^*(x) + \omega$. For $j = 1, \dots, m$, $u_j^*(x)$ is increasing in x_j and decreasing in $x_k, k \neq j$, with limited sensitivity, i.e., $u_j^*(x) \leqslant u_j^*(x + \omega e_j) \leqslant u_j^*(x) + \omega$ and $u_j^*(x) - \omega \leqslant u_j^*(x + \omega e_k) \leqslant u_j^*(x)$ for any $\omega > 0, k \neq j$. (When there are multiple optimal solutions, we choose the one such that $(-u_1^*(x), \dots, -u_m^*(x), u_{m+1}^*(x))$ is the greatest.)*

Proof. (a) We prove by induction. Assume f_{t+1} is concave. In the objective function of (28), $g_t(\cdot, \cdot)$ is clearly concave. Since $\mathscr{A} = \{(x, u) \mid Au \leqslant x, u \geqslant 0\}$, $\mathscr{A}^\Xi = \{(x, u \wedge d) \mid (x, u) \in \mathscr{A}, d \in \text{Supp}(D)\}$, which is equivalent to the convex set $\{(x, w): Aw \leqslant x, 0 \leqslant w_j \leqslant \bar{d}_j, \forall j = 1, \dots, n\}$. In addition, Assumption 2 is satisfied since $\mathscr{A}$ is of the form given in Lemma 3. Then it follows from Theorem 3 that $f_t(x)$ is concave.

(b) For $j = 1, \ldots, m$, define $\hat{u}_j = x_j - u_j$, $\hat{u}_{m+1} = u_{m+1}$ and $\hat{A}$ is given in (23). The optimality equations can be rewritten as

$$f_t(x) = \max_{\hat{u}:(x,\hat{u})\in\mathscr{A}} E\left[p_{m+1}(\hat{u}_{m+1} \wedge D_{m+1}) + \sum_{j=1}^{m} p_j x_j \right.$$
$$\left. - \sum_{j=1}^{m} p_j(\hat{u}_j \vee (x_j - D_j)) + f_{t+1}(\tilde{x}) \right],$$

where $\hat{\mathscr{A}} = \{(x,\hat{u}) \mid \hat{u} \geqslant 0, \hat{A}\hat{u} \leqslant 0, x_j - \hat{u}_j \geqslant 0, j = 1, \ldots, m\}$ and

$$\tilde{x} = [\hat{u}_1 \vee (x_1 - D_1), \ldots, \hat{u}_m \vee (x_m - D_m)] - (\hat{u}_{m+1} \wedge D_{m+1})e.$$

Define $h_t(x,\xi) = p_{m+1}\xi_{m+1} + \sum_{j=1}^{m} p_j x_j - \sum_j^m p_j \xi_j + f_{t+1}((\xi_1, \ldots, \xi_m) - \xi_{m+1}e)$. Then for $t = 1, \ldots, T$,

$$f_t(x) = \max_{\hat{u}:(x,\hat{u})\in\mathscr{A}} E[h_t(x, \hat{u}_1 \vee (x_1 - D_1), \ldots,$$
$$\hat{u}_m \vee (x_m - D_m), \hat{u}_{m+1} \wedge D_{m+1})].$$

Clearly $f_{T+1}(x)$ is $L^\natural$-concave. If $f_{t+1}(x)$ is $L^\natural$-concave, then $h_t(x,\xi)$ is also $L^\natural$-concave by Proposition 1(a) and (d). We have

$$\mathscr{A}^\Xi = \{(x, \hat{u}_1 \vee (x_1 - d_1), \ldots, \hat{u}_m \vee (x_m - d_m),$$
$$\hat{u}_{m+1} \wedge d_{m+1}) \mid (x,\hat{u}) \in \hat{\mathscr{A}}, d \in \text{Supp}(D)\}.$$

Notice that $\mathscr{A}^\Xi$ is equivalent to the following set:

$$\{(x,w) \mid x_j - \bar{d}_j \leqslant w_j \leqslant x_j, w_{m+1} \leqslant w_j, w_j \geqslant 0,$$
$$j = 1, \ldots, m, 0 \leqslant w_{m+1} \leqslant \bar{d}_{m+1}\},$$

where $\bar{d}_j = \text{ess sup}\{d_j \mid d \in \text{Supp}(D)\}$. It follows from Proposition 1(g) that $\mathscr{A}^\Xi$ is $L^\natural$-convex. One can easily check that $\mathscr{A}$ is of the form in Lemma 3. Therefore, Theorem 3 can be applied to show that the $L^\natural$-concavity of $f_t(x)$ is preserved. It follows from Theorem 4 that there exists a greatest solution $\hat{u}^*(x)$ such that $\hat{u}_j^*(x)$ is increasing in x for all j with limited sensitivity, which implies that $u_{m+1}^*(x) = \hat{u}_{m+1}^*(x)$ is increasing in x with limited sensitivity while $u_j^*(x) = x_j - \hat{u}_j^*(x)$ is increasing in x_j, and decreasing in $x_k, k \neq j$ with limited sensitivity. Q.E.D.

The sensitivity result from Theorem 7 implies that if the current capacity level of any resource i increases by ω, then the allocated capacity of product i and $m+1$ should also increase, but the allocated capacity of product $j, j \neq i, j \neq m+1$ will decrease. All the above changes are bounded by ω because of the limited sensitivity.

Remark 4. Even though the resource coefficient matrix here is the same as the bill of materials matrix in Section 4, the analyses of the two models have a significant difference. For the ATO model, the decision variable is truncated by random capacity and the bill of materials matrix does not enter the constraints when we apply the transformation technique. However, for the revenue management model the decision variable is truncated by random demand and the resource coefficient matrix affects the constraints when applying the transformation.

6. Conclusion

In this paper, we develop a transformation technique for a class of stochastic optimization problems. This transformation technique allows us to convert a non-convex minimization problem to an equivalent convex minimization problem, and to prove the preservation of some desirable structural properties (e.g., convexity, submodularity, and $L^\natural$-convexity). We apply these results to several important applications: dual sourcing with random supply capacity, ATO systems with random supply capacity, and network revenue management. Our transformation technique is not limited to the aforementioned applications. For instance, Chen et al. (2015) applied our results, together with a preservation property of concavity and supermodularity with a nonlattice constraint structure developed in Chen et al. (2013), to provide a significantly simplified analysis to the two-facility joint inventory and transshipment problem with uncertain capacities analyzed in Hu et al. (2008). Recently, Demirel et al. (2015) analyze a calibrate-to-order system where a firm produces two products on dedicated production lines that are then calibrated according to the specifications of customer orders on a shared resource. Both the dedicated product lines and the shared resource face random capacities. It can be readily shown that our analysis also applies to their model when the uncertain capacities are independent of each other and over time, and can significantly simplify the analysis. We believe our transformation technique can find many more applications in inventory control, revenue management, and beyond.

It is notable that our transformation technique requires the assumption that the random components are independently distributed. It is likely that new approaches are needed to extend the analysis to cases with correlated random components. Another future research direction is to design efficient algorithms for the applications considered here employing properties of convexity or $L^\natural$-convexity enabled by our transformation technique.

Appendix A. Proofs
Proof of Proposition 1
Parts (a)–(c) are from Murota (2003). The proofs follow directly from the definition.

(d) We need to show that $g[(x, \lambda) - \xi(e, 1)]$ is submodular. Notice that $g[(x, \lambda) - \xi(e, 1)] = g(v - \xi e, \lambda - \xi) = f((v - \xi e) - (\lambda - \xi)e) = f(v - \lambda e)$, which is submodular.

(e) We assume without loss of generality that $\mathscr{A} = \mathbb{F}^n \times \mathscr{Y}$; otherwise we can focus on the restriction of f on $\mathscr{A}$ and let f be infinity outside of $\mathscr{A}$. We know that $r(x, y, \xi) = f[(x, y) - \xi e]$ is submodular, and we want to show that $g(x - \xi e)$ is submodular in (x, ξ). We have $g(x - \xi e) = \inf_{y \in \mathscr{Y}} f(x - \xi e, y) = \inf_{y \in \mathscr{Y}} f[(x, y + \xi \tilde{e}) - \xi e] = \inf_{y \in \mathscr{Y}} r(x, y + \xi \tilde{e}, \xi) = \inf_{z - \xi \tilde{e} \in \mathscr{Y}} r(x, z, \xi)$. Notice that $\{(z, \xi): z - \xi \tilde{e} \in \mathscr{Y}\}$ is a lattice and $r(x, z, \xi)$ is submodular; it follows from Theorem 2.7.6 of Topkis (1998) that $g(x - \xi e)$ is submodular in (x, ξ).

(f) It follows from Theorem 2.8.2 of Topkis (1998) that $\arg\min_{y: (x, y) \in \mathscr{A}} f(x, y)$ is increasing in x. For any $\omega > 0$ and any $x \in dom(g)$, define $\omega \tilde{e} + \arg\min_{y:(x,y) \in \mathscr{A}} f(x, y)$ as the set $\{u + \omega \tilde{e}: u \in \arg\min_{y:(x,y) \in \mathscr{A}} f(x, y)\}$. Pick y' in $\arg\min_{y:(x,y) \in \mathscr{A}} f(x, y)$ and y'' in $\arg\min_{y:(x+\omega e, y) \in \mathscr{A}} f(x + \omega e, y)$. Then for any $\omega > 0$ such that $(x + \omega e, y' + \omega \tilde{e}) \in \mathscr{A}$ and $(x, y'' - \omega \tilde{e}) \in \mathscr{A}$ we have $(x + \omega e, y'' \wedge (y' + \omega \tilde{e})) = (x + \omega e, y'') \wedge (x + \omega e, y' + \omega \tilde{e}) \in \mathscr{A}$, $(x, (y'' - \omega \tilde{e}) \vee y') = (x, y'' - \omega \tilde{e}) \vee (x, y') \in \mathscr{A}$ and

$$0 \geq f(x + \omega e, y'') - f(x + \omega e, y'' \wedge (y' + \omega \tilde{e}))$$
$$= f[(x, y'' - \omega \tilde{e}) + \omega(e, \tilde{e})] - f[(x, (y'' - \omega \tilde{e}) \wedge y') + \omega(e, \tilde{e})]$$
$$\geq f(x, y'' - \omega \tilde{e}) - f(x, (y'' - \omega \tilde{e}) \wedge y')$$
$$\geq f(x, (y'' - \omega \tilde{e}) \vee y') - f(x, y') \geq 0,$$

where the first and the last inequalities are due to the optimality of y'' and y' for $x + \omega e$ and x, respectively, the second inequality is due to the $L^\natural$-convexity of f, which implies that $f(x - \omega e, y - \omega \tilde{e})$ is submodular in (x, y, ω), and the third inequality is due to the submodularity of $f(x, y)$ in y. The first and the last inequalities then imply that equality holds throughout the above inequalities and so $y'' \wedge (y' + \omega \tilde{e}) \in \arg\min_{y:(x+\omega e, y) \in \mathscr{A}} f(x + \omega e, y)$, and $(y'' - \omega \tilde{e}) \vee y' \in \arg\min_{y:(x, y) \in \mathscr{A}} f(x, y)$, which then implies that $y'' \vee (y' + \omega \tilde{e}) \in \omega \tilde{e} + \arg\min_{y:(x, y) \in \mathscr{A}} f(x, y)$. Therefore,

$$\arg\min_{y: (x+\omega e, y) \in \mathscr{A}} f(x + \omega e, y) \sqsubseteq \omega \tilde{e} + \arg\min_{y:(x, y) \in \mathscr{A}} f(x, y).$$

(g) Let $\mathscr{A} = \{x \in \mathscr{Y}: l \leq x \leq u, x_i - x_j \leq v_{ij}, \forall i \neq j\}$. For any $x, x' \in \mathscr{A}, \lambda \in \mathbb{F}_+$, we only need to show that $(x + \lambda e) \wedge x'$, $x \vee (x' - \lambda e) \in \mathscr{A}$. Firstly we have $(x + \lambda e) \wedge x' \leq x' \leq u$, and $l \leq x \wedge x' \leq (x + \lambda e) \wedge x'$. For any $i \neq j$, if $x_i' \leq x_i + \lambda, x_j' \leq x_j + \lambda$, then $(x_i + \lambda) \wedge x_i' - (x_j + \lambda) \wedge x_j' = x_i' - x_j' \leq v_{ij}$. If $x_i' \geq x_i + \lambda, x_j' \geq x_j + \lambda$, then $(x_i + \lambda) \wedge x_i' - (x_j + \lambda) \wedge x_j' = x_i - x_j \leq v_{ij}$. If $x_i' \leq x_i + \lambda, x_j + \lambda \leq x_j'$, then $(x_i + \lambda) \wedge x_i' - (x_j + \lambda) \wedge x_j' = x_i' - (x_j + \lambda) \leq x_i + \lambda - (x_j + \lambda) \leq v_{ij}$. If $x_i' \geq x_i + \lambda, x_j + \lambda \geq x_j'$, then $(x_i + \lambda) \wedge x_i' - (x_j + \lambda) \wedge x_j' = (x_i + \lambda) - x_j' \leq x_i' - x_j' \leq v_{ij}$. Thus we have $(x + \lambda e) \wedge x' \in \mathscr{A}$. Similarly we can show that $x \vee (x' - \lambda e) \in \mathscr{A}$.

Proof of Theorem 4
We only provide the proof of part (b). Since part (a) can be proved using almost the same arguments (as $L^\natural$-convexity implies submodularity), its proof is omitted for brevity.

We first show that the optimal solution set for the optimization problem (11), $\mathscr{U}^*(x, z)$, is increasing in x for any given z. The $L^\natural$-convexity of f implies that $f(x, z)$ is submodular in (x, z). Note that for the submodular function $f(x, z)$ it is straightforward to show that $f(g_1(x_1), \ldots, g_2(x_n), h_1(x_1), \ldots, h_m(x_m))$ is also submodular in (x, z) for any increasing univariate functions $g_i, i = 1, \ldots, n$ and $h_j, j = 1, \ldots, m$. Then $f(x, u \Diamond_k (z + \Xi))$ is submodular in (x, u) for any given z and realization of Ξ. Since the expected value of a submodular function is submodular (Corollary 2.6.2 of Topkis 1998), $E[f(x, u \Diamond_k (z + \Xi))]$ is submodular in (x, u) for any given z. Since $\mathscr{A}$ is an $L^\natural$-convex set, we know that for any given z the constraint set of the optimization problem (11) is a sublattice. It follows from Theorem 2.8.2 of Topkis (1998) that $\mathscr{U}^*(x, z)$ is increasing in x for any given z.

We then show the optimal solution set satisfies the limited sensitivity, i.e., $\mathscr{U}^*((x, z) + \omega e) \sqsubseteq \mathscr{U}^*(x, z) + \omega \tilde{e}$ for any $\omega > 0$. Note that for any $\omega \geq 0$, letting $\tilde{\omega} = -\omega$, we have

$$g((x, z) + \omega e) = \inf_{u: ((x, z) + \omega e, u) \in \mathscr{A}} E[f(x + \omega \hat{e}, u \Diamond_k (z + \omega \tilde{e} + \Xi))]$$
$$= \inf_{\tilde{u}: ((x, z) - \tilde{\omega} e, \tilde{u} - \tilde{\omega} \tilde{e}) \in \mathscr{A}} E[f((x, \tilde{u} \Diamond_k (z + \Xi)) - \tilde{\omega} e)]. \tag{A.1}$$

The $L^\natural$-convexity of f implies that $f((x, \tilde{u}) - \tilde{\omega} e)$ is submodular in $(x, \tilde{u}, \tilde{\omega})$, which implies that $f((x, \tilde{u} \Diamond_k (z + \Xi)) - \tilde{\omega} e)$ is submodular in $(\tilde{u}, \tilde{\omega})$ for any given (x, z) and any realization of Ξ. Since the expected value of a submodular function is submodular (Corollary 2.6.2 of Topkis 1998), $E[f((x, \tilde{u} \Diamond_k (z + \Xi)) - \tilde{\omega} e)]$ is submodular in $(\tilde{u}, \tilde{\omega})$ for any given (x, z). Since $\mathscr{A}$ is an $L^\natural$-convex set, $\{(x, z, \tilde{u}, \tilde{\omega}): ((x, z) - \tilde{\omega} e, \tilde{u} - \tilde{\omega} \tilde{e}) \in \mathscr{A}, \tilde{\omega} \leq 0\}$ is also an $L^\natural$-convex set, which implies that $\{(\tilde{u}, \tilde{\omega}): ((x, z) - \tilde{\omega} e, \tilde{u} - \tilde{\omega} \tilde{e}) \in \mathscr{A}, \tilde{\omega} \leq 0\}$ is a sublattice for any given (x, z). It follows from Theorem 2.8.2 of Topkis (1998) that the solution set of $\tilde{u}$ for optimization problem (A.1), denoted by $\tilde{\mathscr{U}}^*(\tilde{\omega}, x, z)$, is increasing in $\tilde{\omega}$, i.e., $\tilde{\mathscr{U}}^*(\tilde{\omega}, x, z) \sqsubseteq \tilde{\mathscr{U}}^*(0, x, z)$. Note that $\tilde{\mathscr{U}}^*(\tilde{\omega}, x, z) = \mathscr{U}^*((x, z) + \omega e) - \omega \tilde{e}$. Hence we have $\mathscr{U}^*((x, z) + \omega e) - \omega \tilde{e} \sqsubseteq \mathscr{U}^*(x, z)$.

We next proceed to show that $\mathscr{U}^*(x, z)$ is increasing in z for any given x, which turns out to be more challenging. Note that it is difficult to show this monotonicity property directly for $\mathscr{U}^*(x, z)$ based on the original optimization problem (11). In what follows, we show that the projection of the solution set of the transformed problem (12) on the constraint set $\mathscr{U}(x, z)$ of the original problem (11) is exactly $\mathscr{U}^*(x, z)$ and that the monotonicity properties of the solution set of the transformed problem (12) can be preserved by the projection.

Let $\tilde{\mathscr{V}}(x, z)$ denote the constraint set of the transformed problem (12). Define the projection of the solution set of problem (12), $\mathscr{S}^*(x, z) = \arg\min_{(v(\xi), \xi \in \mathscr{X}) \in \tilde{\mathscr{V}}(x, z)} E[f(x, v(\Xi))]$, on the constraint set $\mathscr{U}(x, z)$ as

$$\Pi_{\mathscr{U}} \mathscr{S}^*(x, z) = \{u \in \mathscr{U}(x, z) \mid (u \Diamond_k (z + \xi), \xi \in \mathscr{X}) \in \mathscr{S}^*(x, z)\}.$$

We claim that $\Pi_{\mathscr{U}} \mathscr{S}^*(x, z)$ is the solution set of the original problem for any given (x, z), i.e., $\Pi_{\mathscr{U}} \mathscr{S}^*(x, z) = \mathscr{U}^*(x, z)$. In fact, if u^* is an optimal solution to the original problem, $(u^* \Diamond_k (z + \xi), \xi \in \mathscr{X})$ is a minimizer of the transformed problem, i.e., $(u^* \Diamond_k (z + \xi), \xi \in \mathscr{X}) \in \mathscr{S}^*(x, z)$. On the other hand, if $u \in \Pi_{\mathscr{U}} \mathscr{S}^*(x, z)$, then $(v(\xi) \mid v(\xi) = u \Diamond_k (z + \xi), \xi \in \mathscr{X})$ is an optimal solution of the transformed problem because of the definition

of $\Pi_{\mathcal{U}}\mathcal{S}^*$. Since $E[f(x,v(\Xi))] = E[f(x,u\Diamond_k(z+\Xi))] = \tau^*$, u is optimal for the original problem. Hence the claim is true.

It remains to show $\Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ is increasing in z for any given x, i.e.,

$$\Pi_{\mathcal{U}}\mathcal{S}^*(x,z) \sqsubseteq \Pi_{\mathcal{U}}\mathcal{S}^*(x,z+\omega\tilde{e}_j), \quad j=1,\ldots,n. \qquad (A.2)$$

First note from Proposition 1 that $\mathcal{S}^*(x,z)$ is increasing in (x,z) and satisfies the limited sensitivity property with respect to (x,z), i.e., $\mathcal{S}^*((x,z)+\omega e) \sqsubseteq \omega e + \mathcal{S}^*(x,z)$.

For any $j \in \{1,\ldots,n\}$, it suffices to show that $u' \wedge u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ and $u' \vee u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z+\omega\tilde{e}_j)$ for any $u' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ and $u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z+\omega\tilde{e}_j)$. Since the proofs for $u' \wedge u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ and $u' \vee u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z+\omega\tilde{e}_j)$ are similar, we only provide the detailed arguments of the former and omit that of the latter for brevity.

It is notable that since $\mathcal{U}$ is a lattice and $u',u'' \in \mathcal{U}$ (by the definition of $\Pi_{\mathcal{U}}S^*(x,z)$) we know that $u' \wedge u'' \in \mathcal{U}$ and $u' \vee u'' \in \mathcal{U}$. Hence, to show $u' \wedge u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$ it suffices to show $((u' \wedge u'')\Diamond_k(z+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)$. Since $\mathcal{S}^*(x,z) \sqsubseteq \mathcal{S}^*(x,z+\omega\tilde{e}_j)$, we have $(u'\Diamond_k(z+\xi), \xi \in \mathcal{X}) \wedge (u''\Diamond_k(z+\omega\tilde{e}_j+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)$.

For any component z_l, $l \in \{1,\ldots,j-1,j+1,\ldots,n\}$, we have

$$(u'_l \diamond_l (z_l+\xi_l), \xi \in \mathcal{X}) \wedge (u''_l \diamond_l (z_l+\xi_l), \xi \in \mathcal{X})$$
$$= ((u'_l \wedge u''_l) \diamond_l (z_l+\xi_l), \xi \in \mathcal{X}),$$

where $\diamond_l$ represents the corresponding operation in $\Diamond_k$ for component l.

For component z_j, if $j \in \{1,\ldots,k\}$, the operation in $\Diamond_k$ for component j is $\diamond_j = \wedge$ and we have

$$(u'_j \diamond_j (z_j+\xi_j), \xi \in \mathcal{X}) \wedge (u''_j \diamond_j (z_j+\omega+\xi_j), \xi \in \mathcal{X})$$
$$= ((u'_j \wedge u''_j) \diamond_j (z_j+\xi_j), \xi \in \mathcal{X}).$$

If $j \in \{k+1,\ldots,n\}$, we have $\diamond_j = \vee$. If, in addition, $u'_j \leqslant u''_j$, then

$$(u'_j \diamond_j (z_j+\xi_j), \xi \in \mathcal{X}) \wedge (u''_j \diamond_j (z_j+\omega+\xi_j), \xi \in \mathcal{X})$$
$$= ((u'_j \wedge u''_j) \diamond_j (z_j+\xi_j), \xi \in \mathcal{X}),$$

which, together with the above analysis for other components, implies that

$$((u' \wedge u'')\diamond_j(z+\xi), \xi \in \mathcal{X})$$
$$= (u'\diamond_j(z+\xi), \xi \in \mathcal{X}) \wedge (u''\diamond_j(z+\omega\tilde{e}_j+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z).$$

If, otherwise, $u'_j < u''_j$, defining $\tilde{v}(\xi) \triangleq (u' \Diamond_k (z+\xi)) \wedge (u'' \Diamond_k (z+\omega\tilde{e}_j+\xi))$, we have

$$\tilde{v}_j(\xi_j) = (u'_j \vee (z_j+\xi_j)) \wedge (u''_j \vee (z_j+\omega+\xi_j))$$

$$= \begin{cases} z_j+\xi_j, & \text{if } \xi_j \geqslant u'_j-z_j, \\ u'_j, & \text{if } u'_j-z_j-\omega \leqslant \xi_j < u'_j-z_j, \\ z_j+\omega+\xi_j, & \text{if } u''_j-z_j-\omega \leqslant \xi_j < u'_j-z_j-\omega, \\ u''_j, & \text{if } \xi_j < u''_j-z_j-\omega. \end{cases}$$

We use $\tilde{v}_{-j}(\xi_{-j})$ to denote $(\tilde{v}_1(\xi_1),\ldots,\tilde{v}_{j-1}(\xi_{j-1}),\tilde{v}_{j+1}(\xi_{j+1}),\ldots,\tilde{v}_n(\xi_n))$. Let π^* denote the optimal objective value of the transformed problem with parameters (x,z). Similar to the arguments in Theorem 2, let $\hat{f}(x,z,v) = f(x,v) + \delta_{\mathcal{V}(x,z)}(v)$, where $\mathcal{V}(x,z)$ denotes the constraint set $\{v(\xi) \mid (x,z,v(\xi)) \in \mathcal{A}^\Xi, \xi \in \mathcal{X}\}$. Following the proof in Theorem 1, define

$\hat{g}(\cdot) = E[\hat{f}(x,z,\tilde{v}_{-j}(\Xi_{-j}),\cdot)]$. Since the function f is componentwise convex, lower semicontinuous with $f(x,u) \to +\infty$ for $|u| \to \infty$ and any given x, and the constraint set is componentwise convex and closed, we know that $\hat{g}(\cdot)$ is convex, lower semicontinuous, and $\hat{g}(u) \to +\infty$ for $|u| \to \infty$. Then $\min_{u_j \in \mathbb{F}} \hat{g}(u_j)$ contains a smallest minimizer and a greatest minimizer, denoted by $\underline{u}_i$ and $\bar{u}_i$, respectively, and given any $\tilde{u}_j \in \arg\min_{u_j \in \mathbb{F}} \hat{g}(\cdot)$, we have

$$\pi^* = \min\{E[\hat{g}(v_j(\Xi_j))] \mid v_j(\xi_j) \geqslant z_j+\xi_j, \forall \xi_j \in \mathcal{X}_j\}$$
$$= \min_{u_j \in \mathbb{F}} E[\hat{g}(u_j \vee (z_j+\Xi_j))] = E[\hat{g}(\tilde{u}_j \vee (z_j+\Xi_j))].$$

We next prove by contradiction that $u''_j \in [\underline{u}_j,\bar{u}_j]$, which implies that $u''_j \in \arg\min_{u_j \in \mathbb{F}} \hat{g}(u_j)$.

Suppose $\bar{u}_j < u''_j$. For $\xi_j < u'_j-z_j$ we have $\bar{u}_j \leqslant \bar{u}_j \vee (z_j+\xi_j) < \tilde{v}_j(\xi_j)$ and thus $\hat{g}(\bar{u}_j \vee (z_j+\xi_j)) < \hat{g}(\tilde{v}_j(\xi_j))$. By the assumption of this theorem, we have $\underline{\xi}_j + z_j + \omega \leqslant u''_j < u'_j$, which implies $\Pr(\xi_j < u'_j-z_j) > 0$. For $\xi_j \geqslant u'_j-z_j$, we have $\hat{g}(\bar{u}_j \vee (z_j+\xi_j)) = \hat{g}(z_j+\xi_j) = \hat{g}(\tilde{v}_j(\xi_j))$. Then we must have $\pi^* = E[\hat{g}(\bar{u}_j \vee (z_j+\Xi_j))] < E[\hat{g}(\tilde{v}_i(\Xi_j))]$, which contradicts the optimality of $(\tilde{v}(\xi), \xi \in \mathcal{X})$ for the transformed problem.

Suppose $\underline{u}_j > u''_j$. For $\xi_j < \underline{u}_j-z_j-\omega$, we have $\underline{u}_j \vee (z_j+\xi_j) > \tilde{v}_j(\xi_j)$ and thus $\hat{g}(\underline{u}_j \vee (z_j+\xi_j)) < \hat{g}(\tilde{v}_j(\xi_j))$. Since $\underline{\xi}_j+\omega+z_j \leqslant u''_j < \underline{u}_j$, we have $\Pr(\xi_j < \underline{u}_j-z_j-\omega) > 0$. For $\underline{u}_j-z_j-\omega \leqslant \xi_j < \underline{u}_j-z_j$, we have $\hat{g}(\underline{u}_j \vee (z_j+\xi_j)) = \hat{g}(\underline{u}_j) \leqslant \hat{g}(\tilde{v}_j(\xi_j))$. For $\xi_j \geqslant \underline{u}_j-z_j$, we have $\hat{g}(\underline{u}_j \vee (z_j+\xi_j)) = \hat{g}(z_j+\xi_j) \leqslant \hat{g}((u'_j \vee (z_j+\xi_j)) \wedge (z_j+\omega+\xi_j)) = \hat{g}(\tilde{v}_j(\xi_j))$. Then we must have $\pi^* = E[\hat{g}(\underline{u}_j \vee (z_j+\Xi_j))] < E[\hat{g}(\tilde{v}_i(\Xi_j))]$, which contradicts the optimality of $(\tilde{v}(\xi), \xi \in \mathcal{X})$ for the transformed problem.

The above contradictions imply $u''_j \in [\underline{u}_j,\bar{u}_j]$ and $\pi^* = \min_{u_j \in \mathbb{F}} E[\hat{g}(u''_j \vee (z_j+\Xi_j))]$. Clearly, for any $l \neq j$, $\tilde{v}_l(\xi_l) = (u'_l \wedge u''_l)\diamond_l(z_l+\xi_l), \forall \xi \in \mathcal{X}$. Then $(u' \wedge u'')\Diamond_k(z+\omega\xi) = ((u'_j \wedge u''_j) \vee (z_i+\xi_j), \tilde{v}_{-j}(\Xi_{-j}), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)$. In summary, we have $((u' \wedge u'')\Diamond_k(z+\xi), \xi \in \mathcal{X}) \in \mathcal{S}^*(x,z)$ and hence $u' \wedge u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z)$. Following the similar arguments, we can show that $u' \vee u'' \in \Pi_{\mathcal{U}}\mathcal{S}^*(x,z+\omega\tilde{e}_j)$ and then (A.2) holds.

Finally, we show that the optimal solution set $\mathcal{U}^*(x,z)$ is a lattice, and it has a greatest element and a least element. Given fixed (x,z), let $h(u) = E[f(x,u\Diamond_k(z+\Xi))]$. For any realization of Ξ, denoted by ξ, given any u' and u'', we have

$$f(x,u' \Diamond_k (z+\xi)) + f(x,u'' \Diamond_k (z+\xi))$$
$$= f(x,u'_1 \wedge (z_1+\xi_1),\ldots,u'_n \vee (z_n+\xi_n))$$
$$\quad + f(x,u''_1 \wedge (z_1+\xi_1),\ldots,u''_n \vee (z_n+\xi_n))$$
$$\geqslant f(x,(u'_1 \wedge (z_1+\xi_1)) \wedge (u''_1 \wedge (z_1+\xi_1)),\ldots,$$
$$\quad (u'_n \vee (z_n+\xi_n)) \wedge (u''_n \vee (z_n+\xi_n)))$$
$$\quad + f(x,(u'_1 \wedge (z_1+\xi_1)) \vee (u''_1 \wedge (z_1+\xi_1)),\ldots,$$
$$\quad (u'_n \vee (z_n+\xi_n)) \vee (u''_n \vee (z_n+\xi_n)))$$
$$= f(x,(u'_1 \wedge u''_1) \wedge (z_1+\xi_1),\ldots,(u'_n \wedge u''_n) \vee (z_n+\xi_n))$$
$$\quad + f((u'_1 \vee u''_1) \wedge (z_1+\xi_1),\ldots,(u'_n \vee u''_n) \vee (z_n+\xi_n)).$$

The inequality is due to the submodularity of f. Then $h(u') + h(u'') \geqslant h(u' \wedge u'') + h(u' \vee u'')$. Therefore, the objective function of (11) is submodular in u. By our assumptions the function f satisfies $f(x,z) \to +\infty$ for $|z| \to \infty$ for any given x, and the constraint set is closed and is a lattice, it is equivalent to restricting our constraint set to a compact sublattice of $\mathbb{R}^n$. By Topkis (1998, Corollary 2.7.1) the solution set $\mathcal{U}^*(x,z)$ is a

compact sublattice of $\mathbb{R}^n$, and there exist a greatest element and a least element in the solution set and they satisfy the desired monotonicity properties.

Appendix B. A Comparison to the Stochastic Linearity Approach

Feng and Shanthikumar, (hereafter referred to as FS, 2014) also consider optimization problems with objective function $E[f(u \wedge \Xi)]$ and convert them to convex minimization problems using stochastic linearity in mid-point (SL(mp)). Specifically, given a stochastic function $Y(u) \triangleq \psi(u, \Xi)$, let $\mu(u) = E[\psi(u, \Xi)]$, and $u(\mu)$ be the inverse of $\mu(u)$, i.e., $u(\mu) = \inf\{u \mid E[\psi(u, \Xi)] \geqslant \mu\}$. Then $g(\mu) \triangleq E[f(Y(u(\mu)))]$ is convex in μ as long as $Y(u(\mu))$ is SL(mp). FS prove that, along with several other supply functions, if $\psi(u, \Xi) = u \wedge \Xi$, then $Y(u(\mu))$ is SL(mp). This allows them to convert nonconvex minimization problems to equivalent convex minimization problems by a variable transformation.

Our transformation technique can preserve convexity as well as submodularity and $L^\natural$-convexity. FS do not mention whether their approach can preserve submodularity or $L^\natural$-convexity. It turns out that their approach can preserve submodularity of the objective function, but can *not* preserve $L^\natural$-convexity. These are shown in the following Proposition 2 and Example 2, respectively.

Proposition 2. *Suppose that $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is a submodular function, and Ξ is a random vector with support $\mathcal{X} \in \mathbb{R}^n$, in which any component Ξ_i is independent of each other. Let $\mu(u) = E[u \wedge \Xi]$ and $u(\mu)$ be the inverse of $\mu(u)$. Then*
   (a) *$h(u) = E[f(u \wedge \Xi)]$ is submodular in u.*
   (b) *$g(\mu) = E[f(u(\mu) \wedge \Xi)]$ is submodular in μ.*

Proof. (a) For any realization ξ, given any u and u', we have $f(u \wedge \xi) + f(u' \wedge \xi) \geqslant f((u \wedge \xi) \wedge (u' \wedge \xi)) + f((u \wedge \xi) \vee (u' \wedge \xi)) = f((u \wedge u') \wedge \xi) + f((u \vee u') \wedge \xi)$. It follows from Corollary 2.6.2 of Topkis 1998 that the submodularity property is preserved under expectation, which implies that $h(u) = E[f(u \wedge \Xi)]$ is submodular.

(b) Notice that for any component $i = 1, \ldots, n$, $u_i(\mu_i)$ is an increasing function of μ_i. Hence, for any realization ξ, $u_i(\mu_i) \wedge \xi_i$ is also an increasing function of μ_i for component $i = 1, \ldots, n$. It follows immediately that $f(u(\mu) \wedge \xi) = f(u_1(\mu_1) \wedge \xi_1, \ldots, u_n(\mu_n) \wedge \xi_n)$ is submodular in μ for any realization ξ. Taking expectation, we know that $g(\mu) = E[f(u(\mu) \wedge \Xi)]$ is submodular in μ. Q.E.D.

Example 2. Consider $E[f(u_1 \wedge \Xi_1, u_2 \wedge \Xi_2)]$, where $f(u_1, u_2) = e^{u_1 - u_2}$ is an $L^\natural$-convex function. Suppose both Ξ_1 and Ξ_2 follow exponential distribution with mean 1, and they are independent of each other. Then $\forall i = 1, 2, \mu_i(u_i) = E[u_i \wedge \Xi_i] = \int_0^{u_i} \xi_i e^{-\xi_i} d\xi_i + \int_{u_i}^{\infty} u_i e^{-\xi_i} d\xi_i = 1 - e^{-u_i}$. We have $u_i(\mu_i) = -\ln(1 - \mu_i)$ and

$$E[f(u_1 \wedge \Xi_1, u_2 \wedge \Xi_2)] = \int_0^{u_2} \int_0^{u_1} e^{\xi_1 - \xi_2} e^{-\xi_1} e^{-\xi_2} \, d\xi_1 \, d\xi_2$$
$$+ \int_0^{u_2} \int_{u_1}^{\infty} e^{u_1 - \xi_2} e^{-\xi_1} e^{-\xi_2} \, d\xi_1 \, d\xi_2$$
$$+ \int_{u_2}^{\infty} \int_0^{u_1} e^{\xi_1 - u_2} e^{-\xi_1} e^{-\xi_2} \, d\xi_1 \, d\xi_2$$
$$+ \int_{u_2}^{\infty} \int_{u_1}^{\infty} e^{u_1 - u_2} e^{-\xi_1} e^{-\xi_2} \, d\xi_1 \, d\xi_2$$

$$= \frac{1}{2}(1 + u_1)(1 + e^{-2u_2})$$
$$= \frac{1}{2}(1 - \ln(1 - \mu_1))(1 + (1 - \mu_2)^2).$$

Define $g(\mu_1, \mu_2) = \frac{1}{2}(1 - \ln(1 - \mu_1))(1 + (1 - \mu_2)^2)$. Let $\mu = [0.7, 0.2]$, $\mu' = [0.8, 0.4]$, $\alpha = 0.1$. We have $g(\mu) + g(\mu') \approx 3.5817$ while $g((\mu + \alpha e) \wedge \mu') + g(\mu \vee (\mu' - \alpha e)) \approx 3.5860$. Therefore,

$$g(\mu) + g(\mu') < g((\mu + \alpha e) \wedge \mu') + g(\mu \vee (\mu' - \alpha e)),$$

which means that $g(\mu_1, \mu_2)$ is not $L^\natural$-convex.

Note that the approach introduced by FS requires computing the inverse of $\mu(u)$, which may not have a closed-form solution. When considering a constrained optimization problem, even if all the constraints in the original problem are linear, their approach will very likely add nonlinear constraints explicitly, whereas our transformation technique only adds linear constraints, although a potentially infinite number of them. More importantly, under the conditions in Lemma 2, the constraint set can also preserve $L^\natural$-convexity with our transformation technique, but this may not hold under their approach, which is illustrated in Example 3.

Example 3. Consider $\inf_{u \in \mathcal{U}} E[f(u_1 \wedge \Xi_1, u_2 \vee \Xi_2)]$, where $f(\cdot, \cdot)$ is an $L^\natural$-convex function and $\mathcal{U} = \{(u_1, u_2) \mid u_1 - u_2 \leqslant \frac{1}{2}, 0 \leqslant u_1 \leqslant 1, 0 \leqslant u_2 \leqslant 1\}$. Suppose Ξ_1 and Ξ_2 are both uniformly distributed between 0 and 1, and they are independent of each other. Applying our transformation technique, we have

$$\begin{aligned} \inf \quad & E[f(v_1(\Xi_1), v_2(\Xi_2))] \\ \text{s.t.} \quad & v(\xi_1) \leqslant \xi_1 \quad \forall, \xi_1 \in [0, 1], \\ & v_2(\xi_2) \geqslant \xi_2 \quad \forall \xi_2 \in [0, 1], \\ & (v_1(\xi_1), v_2(\xi_2)) \in \mathcal{V} \quad \forall \xi \in [0, 1] \times [0, 1], \end{aligned} \tag{B.1}$$

where $\mathcal{V} = \{(v_1, v_2) \mid v_1 - v_2 \leqslant \frac{1}{2}, 0 \leqslant v_1 \leqslant 1, 0 \leqslant v_2 \leqslant 1\}$. All constraints in the transformed problem are linear, and they form an $L^\natural$-convex set.

Next we apply the transformation of FS. We have $\mu_1(u_1) = E[u_1 \wedge \Xi_1] = u_1 - \frac{1}{2}u_1^2$, $\mu_2(u_2) = E[u_2 \vee \Xi_2] = \frac{1}{2}(1 + u_2^2)$. Then we have $u_1(\mu_1) = 1 - \sqrt{1 - 2\mu_1}$, $u_2(\mu_2) = \sqrt{2\mu_2 - 1}$. Hence, the constraint set after the transformation becomes $\tilde{\mathcal{U}} = \{(\mu_1, \mu_2) \mid \sqrt{1 - 2\mu_1} + \sqrt{2\mu_2 - 1} \geqslant \frac{1}{2}, 0 \leqslant \mu_1 \leqslant \frac{1}{2}, \frac{1}{2} \leqslant \mu_2 \leqslant 1\}$, which consists of nonlinear constraints. One can also check that $\tilde{\mathcal{U}}$ is not an $L^\natural$-convex set. To see this, notice that $\mu = [0.3, 0.5] \in \tilde{\mathcal{U}}$ and $\mu' = [0.41, 0.51] \in \tilde{\mathcal{U}}$, but $\mu \vee (\mu' - \alpha e) \notin \tilde{\mathcal{U}}$ with $\alpha = 0.01$.

References

Barankin EW (1961) A delivery-lag inventory model with an emergency provision (the one-period case). *Naval Res. Logist. Quart.* 8(3):258–311.

Bollapragada R, Rao US, Zhang J (2004) Managing inventory and supply performance in assembly systems with random capacity and demand. *Management Sci.* 50(12):1729–1743.

Brumelle SL, Mcgill JI (1993) Airline seat allocation with multiple nested fare classes. *Oper. Res.* 41(1):127–137.

Chen L, Homem-de-Mello T (2010) Re-solving stochastic programming models for airline revenue managment. *Ann Oper. Res.* 177(1):91–114.

Chen W, Feng Q, Seshadri S (2013) Sourcing from suppliers with random yield for price-dependent demand. *Ann. Oper. Res.* 208(1):557–579.

Chen X, Gao X, Hu Z (2015) A new approach to two-location joint inventory and transshipment control via $L^\natural$-convexity. *Oper. Res. Lett.* 43(1):65–68.

Chen X, Hu P, He S (2013) Technical note—Preservation of super-modularity in parametric optimization problems with nonlattice structures. *Oper. Res.* 61(5):1166–1173.

Chen X, Pang Z, Pan L (2014) Coordinating inventory control and pricing strategies for perishable products. *Oper. Res.* 62(2):284–300.

Ciarallo FW, Akella R, Morton TE (1994) A periodic review, production planning model with uncertain capacity and uncertain demand: Optimality of extended myopic policies. *Management Sci.* 40(3):320–332.

Cooper W, Homem-de-Mello T (2007) Some decomposition methods for revenue management. *Transportation Sci.* 41(3):332–353.

Daniel K (1963) A delivery-lag inventory model with emergency. Scarf H, Gilford D, Shelly M, eds. *Multistage Inventory Models and Techniques* (Stanford University Press, Stanford, CA), 32–46.

Demirel S, Duenyas I, Kapuscinski R (2015) Production and inventory control for a make-to-stock/calibrate-to-stock system with dediated shared resources. *Oper. Res.* 63(4):823–839.

Federgruen A, Yang N (2008) Selecting a portfolio of suppliers under demand and supply risks. *Oper. Res.* 56(4):916–936.

Federgruen A, Yang N (2011) Procurement strategies with unreliable suppliers. *Oper. Res.* 59(4):1033–1039.

Feng Q (2010) Integrating dynamic pricing and replenishment decisions under supply capacity uncertainty. *Management Sci.* 56(12):2154–2172.

Feng Q, Shanthikumar G (2018) Supply and demand functions in inventory models. *Oper. Res.* 66(1):77–91.

Feng Q, Shi R (2012) Sourcing from multiple suppliers for price-dependent demands. *Production Oper. Management* 21(3):547–563.

Fukuda Y (1964) Optimal policies for the inventory problem with negotiable leadtime. *Management Sci.* 10(4):690–708.

Gong X, Chao X, Zheng S (2014) Dynamic pricing and inventory management with dual suppliers of different leadtimes and disruption risks. *Production Oper. Management* 23(12):2058–2074.

Henig M, Gerchak Y (1990) The structure of periodic review policies in the presence of random yield. *Oper. Res.* 38(4):634–643.

Hu X, Duenyas I, Kapuscinski R (2008) Optimal joint inventory and transshipment control under uncertain capacity. *Oper. Res.* 56(4):881–897.

Hua Z, Yu Y, Zhang W, Xu X (2015) Structural properties of the optimal policy for dual-sourcing systems with general lead times. *IIE Trans.* 47(8):841–850.

Huh W, Janakiraman G (2010) On the optimal policy structure in serial inventory systems with lost sales. *Oper. Res.* 58(2):481–491.

Lu Y, Song JS (2005) Order-based cost optimization in assemble-to-order systems. *Oper. Res.* 53(1):151–169.

Möller A, Römisch W, Weber K (2008) Airline network revenue management by multistage stochastic programming. *Computational Management Sci.* 5(4):355–377.

Murota K (1998) Discrete convex analysis. *Math. Programming* 83(1–3):313–371.

Murota K (2003) *Discrete Convex Analysis*, SIAM Monographs on Discrete Mathematics and Applications (SIAM, Philadelphia).

Murota K (2009) Recent developments in discrete convex analysis. Cook W, Lovasz L, Vygen J, eds. *Research Trends in Combinatorial Optimization, Bonn* 2008, Chap. 11 (Springer, Berlin), 219–260.

Nadar E, Akan M, Sheller-Wolf A (2014) Optimal structural results for assemble-to-order generalized *M*-systems. *Oper. Res.* 62(3):571–579.

Pang Z, Chen F, Feng Y (2012) A note on the structure of joint inventory-pricing control with leadtimes. *Oper. Res.* 60(3):581–587.

Robinson L (1995) Optimal and approximate control policies for airline booking with sequential nonmonotomic fare classes. *Oper. Res.* 43(2):252–263.

Sheopuri A, Janakiraman G, Seshadri S (2010) New policies for the stochastic inventory control problem with two supply sources. *Oper. Res.* 58(3):734–745.

Simchi-Levi D, Chen X, Bramel J (2014) *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics Management*, 3rd ed. (Springer, New York).

Song JS, Zipkin P (2003) Supply chain operations: Assemble-to-order systems. de Kok T, Graves S, eds. *Supply Chain Management*, Handbooks in Operations Research and Management Science, Vol. 30 (North-Holland, Amsterdam), 561–596.

Talluri KT, van Ryzin GJ (2005) *The Theory and Practice of Revenue Management*, International Series in Operations Research and Management Science, Vol. 68 (Springer, New York).

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

Van Mieghem JA (2008) *Operations Strategy: Principles and Practice* (Dynamic Ideas, Belmont, MA), 230–232.

Veeraraghavan S, Scheller-Wolf AA (2006) Now or later: A simple policy for effective dual sourcing in capacitated systems. *Oper. Res.* 56(4):850–864.

Wang YZ, Gerchak Y (1996) Periodic review production models with variable capacity, random yield, and uncertain demand. *Management Sci.* 42(1):130–137.

Whittmore AS, Saunders SC (1977) Optimal inventory under stochastic demand with two supply options. *SIAM J. Appl. Math.* 32(2):293–305.

Zălinescu C (2002) *Convex Analysis in General Vector Spaces* (World Scientsfic, Singapore).

Zhou X, Chao X (2014) Dynamic pricing and inventory management with regular and expedited supplies. *Production Oper. Management* 23(1):65–80.

Zipkin P (2008) On the structure of lost-sales inventory models. *Oper. Res.* 56(4):937–944.

Xin Chen is a professor, an Abel Bliss Faculty Scholar, and the Jerry S. Dobrovolny Faculty Scholar at the University of Illinois at Urbana-Champaign. His research interest lies in optimization, data analytics, revenue management, and supply chain management. He received the INFORMS revenue management and pricing section prize in 2009, and is the coauthor of the book *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management*.

Xiangyu Gao is an assistant professor in the department of Decision Sciences and Managerial Economics at the CUHK business school, the Chinese University of Hong Kong. His research interests include inventory and supply chain management, dynamic pricing and revenue management, and collaborations in supply chains.

Zhan Pang is an associate professor in management science at the College of Business, City University of Hong Kong. His current research interests include supply chain and operations management, pricing and revenue management, and corporate risk management.