MANAGEMENT SCIENCE
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0025-1909 | EISSN 1526-5501 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Sustainable Inventory with Robust Periodic-Affine Policies

Chaithanya Bandi
Kellogg School of Management, Northwestern University, Evanston, IL 60208
c-bandi@kellogg.northwestern.edu

Eojin Han and Omid Nohadani
Industrial Engineering & Management Sciences, Northwestern University, Evanston, IL 60208
eojinHan2020@u.northwestern.edu, nohadani@northwestern.edu

Abstract: We introduce a new class of adaptive policies called *periodic-affine policies*, that allows a decision maker to optimally manage and control large-scale newsvendor networks in the presence of uncertain but non-stochastic demand. These policies are data-driven and model many features of the demand such as correlation, and remain robust to parameter mis-specification. We present a model that can be generalized to multi-product settings and extended to multi-period problems. This is accomplished by modeling the uncertain demand via sets. In this way, it offers a natural framework to study competing policies such as base-stock, affine, and approximative approaches with respect to their profit, sensitivity to parameters and assumptions, and computational scalability. We show that the periodic-affine policies are sustainable, i.e. time consistent, because they warrant optimality both within subperiods and over the entire planning horizon. This approach is tractable and free of distributional assumptions, and hence, suited for real-world applications. We provide efficient algorithms to obtain the optimal periodic-affine policies and demonstrate their advantages on the sales data from one of India's largest pharmacy retailers.

Key words: Newsvendor Network, Robust Optimization, Demand Uncertainty, Correlation, Affine Policies, Healthcare: Pharmaceutical Retailer.

1. Introduction

Despite the physicians' diagnostic matching of patients to drugs, the heterogeneity in patients' illness, drug's efficacy, potential side effects, and varying length of treatment lead to sizable uncertainty in drug's demand (Crawford and Shum 2005). Retailers are mandated to service level guarantees, and overstocking drugs is neither economical nor practical since they are perishable. Such healthcare problems affect a wide section of the population and has large societal implications. In this context, newsvendor models offer a natural framework and are used for decision making.

Practical solutions to such problems are critical to a broad range of industries. In particular, pharmaceutical companies with a large turnover are interested in optimal inventory management.

GlaxoSmithKline (GSK) spends over \$4.5 billion each year on manufacturing and supplying products. Johnson & Johnson (JNJ) spends approximately \$30 billion annually in leveraging its purchasing power to set sustainability expectations beyond its operations. Similarly, companies like Teva Pharmaceuticals (TEVA), Pfizer (PFE), and Merck (MRK) spend millions of dollars to ensure the safety and supply of their products, even though they have manufacturing units in multiple locations. Therefore, any variation in inventories can lead to multiple disturbances in the system. A pharmacy's inventory represents its single, largest investment. In a common pharmacy, cost of goods sold accounts for approximately 68% of total expenditures. For every 1% change in costs of goods, profits may increase or decrease by more than 20%. Thus, the sheer magnitude of dollars involved makes seemingly minor inefficiencies in purchasing and inventory control matter of great importance to both cash flow and profitability.

The challenges of such networks are multifold. Real-world settings are typically high-dimensional with multiple products and multiple stages of decision-making. These settings also suffer from substantial uncertainties in demand. Modeling such demand uncertainty is challenging because demand is often not stationary.

In this work, we consider a newsvendor network with uncertain and correlated demand. Using the paradigm of robust optimization, we model such demand to reside in uncertainty sets and provide tractable formulations and associated algorithms for sustainable policies. To gain insight from a real-world setting, we apply the results to a major online pharmacy retailer in India, where a prohibitively large penalty occurs when customers' demand is not satisfied. This company carries over 163 different brands, and the sales grow at about 23% per year. Their distribution network spans the entire country through fixed retail locations and online platforms. The decision makers of this company observe a sizable uncertainty in demand over the course of the year (in addition to seasonality) and significant correlations amongst various product categories. In close collaboration with this company's managers, we seek to design optimal implementable policies to control their inventory levels in their network.

Our contributions are:

- *Modeling*: We provide a distribution-free description of uncertainty in demand using two types of sets. Independent demands are modeled via budget constraints. We also incorporate correlated demands using a factor model approach. The inventory control problem is then cast as a multi-stage robust optimization problem. As a result, a novel solution concept of *periodic-affine policy* is provided for newsvendor networks with time-dependent and potentially correlated demand uncertainty.
- *Algorithms*: We provide a tractable algorithm that provides periodic-affine policies. These policies decompose the overall problem into a more tractable formulation than affine policies.

- *Application*: We analyze the sales data of a pharmacy retailer in India for the fourth quarter of 2016. This entails 1.5 million transactions for 228 different products. We construct the demand uncertainty set for the 20 most-popular products, comprising 80% of all transactions. Our numerical experiments show that even for the single-station case, the computational burden for the optimal periodic-affine policies is significantly reduced over affine policies (by 100× for a 15-period problem), making the proposed approach practical for real-world and large-sized problems. Moreover, the periodic-affine policy improved the cost effectiveness of the operation by 19% over a base-stock policy for realistic penalty costs.

1.1. Literature review

The seminal work of Arrow et al. (1951) introduced the multistage periodic review inventory model, where the inventory is reviewed once every period and a decision is made to place an order, if a replenishment is necessary. The (s, S) inventory policy establishes a lower (minimum) stock point s and an upper (maximum) stock point S. When the inventory level drops below s, an order is placed "up to S." The (s, S) ordering policy has been proven optimal for simple stochastic inventory systems. Scarf (1960) proved that base-stock policies are optimal for a single installation model. Clark and Scarf (1960) extended the result to serial supply chains without capacity constraints and showed that the optimal ordering policy for the multi-echelon system can be decomposed into decisions based on the echelon inventories. Karlin (1960) and Morton (1978) showed that base-stock policies are optimal for single-state systems with non-stationary demands. Federgruen and Zipkin (1986) generalized the analysis to a single-stage capacitated system, and Rosling (1989) extended the analysis of serial systems to assembly systems. For more work, refer to Langenhoff and Zijm (1990), Sethi and Cheng (1997), Muharremoglu and Tsitsiklis (2008), Huh and Janakiraman (2008).

Simulation optimization has attempted to take advantage of the availability of computational resources and the power of simulation for evaluating functions. For a comprehensive overview of commonly used simulation optimization techniques, we refer the reader to the survey by Fu et al. (2005). Fu (1994), Glasserman and Tayur (1995), Fu and Healy (1997) and Kapuscinski and Tayur (1999) have developed various gradient-based algorithms to study inventory systems. These methods are practical whenever the input variables are continuous and their success depends on the quality of the gradient estimator.

On the other hand, Scarf (1958), Kasugai and Kasegai (1961), Gallego and Moon (1993), Graves and Willems (2000) developed distribution-free approaches to inventory theory. Bertsimas and Thiele (2006) took a robust optimization approach to inventory theory and showed that base-stock policies are optimal in the case of serial supply chain networks. Bienstock and Özbay (2008) presented a family of decomposition algorithms aimed at solving for the optimal base-stock policies using a robust optimization approach. Rikun (2011) extended the robust framework introduced by Bienstock and Özbay (2008) to compute optimal (s, S) policies in supply chain networks and compared their performance to optimal policies obtained via stochastic optimization. Ben-Tal et al. (2004) extended the robust optimization framework to dynamic settings and explored the use of disturbance-affine policies by allowing the decision maker to adjust their strategy leveraging the information revealed over time. Bertsimas and Thiele (2006) and Bienstock and Özbay (2008) studied the performance of base-stock policies, and Ben-Tal et al. (2005), Kuhn et al. (2011), and Bertsimas et al. (2010) investigated polices that are affine in prior demands under a robust optimization lens. Within the robust optimization framework, affine policies have gained much attention due to their tractability; depending on the class of the nominal problem, the optimal policy can be solved via linear, quadratic, conic or semidefinite programs (see Löfberg (2003), Kerrigan and Maciejowski (2004)). Empirically, Ben-Tal et al. (2005) and Kuhn et al. (2011) have reported that affine policies have excellent performance and in many instances optimal.

In the context of pharmaceutical systems, Guerrero et al. (2013) provided a near-optimal base-stock policy for two-echelon distribution networks with multiple products, where every sink node is replenished by a single supplier. They provided a Markov chain formulation and a heuristic algorithm for Poisson distributed and independent demands. For a combined setting of a pharmaceutical compony and a hospital, Uthayakumar and Priyan (2013) developed a two-echelon supply chain model to determine the optimal lot size, lead time, and total number of deliveries between the pharmaceutical compony and a hospital. Using Lagrange multipliers, they provided decision tools for optimal costs while ensuring required service levels. In a two-level pharmaceutical supply chain, Baboli et al. (2011) studied a specific product with a constant demand rate and numerically showed that the overall cost is improved when pharmacies and hospitals are centralized.

Notation. Lowercase italic is used to denote scalars; lowercase bold is used to denote vectors, and uppercase bold is used to denote matrices. Sets are in calligraphic. Section specific notation is introduced where needed. All proofs are relegated to the appendix.

2. Model

We consider a newsvendor network in which inventories are reviewed periodically and unfulfilled orders are backlogged. For simplicity, we assume zero lead times throughout the network; however, our framework can be adapted to systems with non-zero lead times. We consider a T-period time horizon and, within each period, events occur in the following order: (1) the ordering decision is made at the beginning of the period, (2) demands for the period occur and are filled or backlogged depending on the available inventory, (3) the stock availability is updated for the next period.

- $\mathcal{N}$: Set of all installations where ordering decisions are made (source nodes) with $|\mathcal{N}| = m$
- $\mathcal{S}$: Set of all installations with external demand (sink nodes) with $|\mathcal{S}| = n$
- $\mathcal{L}$: Set of all links (edges) within the inventory network with $|\mathcal{L}| = p$
- $\mathcal{N}_j$: Set of source nodes supplying stock to a sink node $j \in \mathcal{S}$
- $\mathcal{S}_i$: Set of sink installations that are fed from a source node $i \in \mathcal{N}$
- s_t^i : Amount of order at the beginning of period t at a source $i \in \mathcal{N}$
- d_t^j : Demand observed at a sink $j \in \mathcal{S}$ throughout time period
- x_t^ℓ : Stock delivered along a link $\ell \in \mathcal{L}$ at time t
- $u_t^{s,i}$: Stock available after the period t at a source node $i \in \mathcal{N}$
- $u_t^{d,j}$: Backorders after the period t at a sink node $j \in \mathcal{S}$.

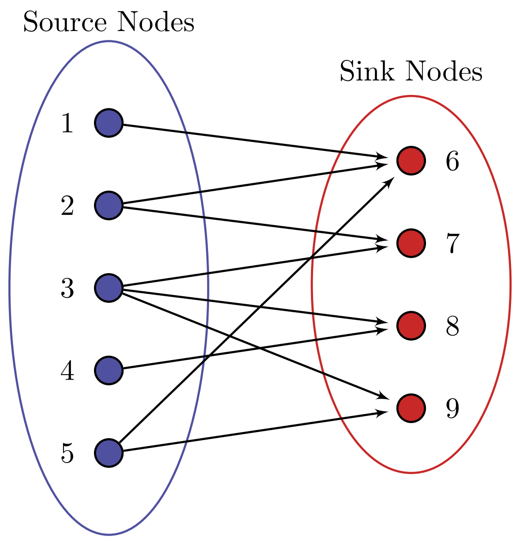

Figure 1 Example of a nine-installation network with $n = 4$ sink nodes and $m = 5$ source nodes.

To track the system's operation, we capture information about the stock available and the stock ordered at source installations at the beginning of each time period as well as the demand at each sink installation throughout each time period. Specifically, assuming zero initial inputs and demands, we can express the dynamics of inventory levels and backlogged demands for $t = 1, \dots, T$ as

$$\begin{aligned} u_t^{s,i} &= u_{t-1}^{s,i} + s_t^i - \sum_{\ell=(i,j), j\in\mathcal{S}_i} x_t^\ell = \sum_{\tau=1}^t s_\tau^i - \sum_{\ell=(i,j), j\in\mathcal{S}_i} \sum_{\tau=1}^t x_\tau^\ell \qquad \forall\, i \in \mathcal{N} \\ u_t^{d,j} &= u_{t-1}^{d,j} + d_t^j - \sum_{\ell=(i,j), i\in\mathcal{N}_j} x_t^\ell = \sum_{\tau=1}^t d_\tau^j - \sum_{\ell=(i,j), i\in\mathcal{N}_j} \sum_{\tau=1}^t x_\tau^\ell \qquad \forall\, j \in \mathcal{S}. \end{aligned} \tag{1}$$

Note that the ordering quantities $s_t^i = s_t^i(\pi, \mathbf{d})$, and therefore the amount of available stock $u_t^{s,i} = u_t^{s,i}(\pi, \mathbf{d})$ and backorders $u_t^{d,j} = u_t^{d,j}(\pi, \mathbf{d})$, are functions of the ordering policy π and the demand $\mathbf{d}$.

The high-dimensional nature of modeling demand uncertainty probabilistically and the complex dependence on random variables underscore the difficulty of analyzing and optimizing the expected total cost. Instead, we propose a framework that builds upon the robust optimization paradigm.

2.1. Robust Newsvendor Network Formulation

To describe our framework, we first introduce a robust approach to single-period models. Our models are based on assumption that we have the following cost and revenue structure:

- c_S^i: Purchasing cost per unit at the source node $i \in \mathcal{N}$
- c_H^i: Holding cost per unit for the leftover stock at the source node $i \in \mathcal{N}$
- c_P^j: Penalty cost per unit for the unsatisfied demand at the sink node $j \in \mathcal{S}$
- r^ℓ: Revenue by satisfying a unit demand occurred at the sink node j via $\ell = (i,j) \in \mathcal{L}$.

The goal of our decision maker is to order a proper amount of products $\{s_i : i \in \mathcal{N}\}$ and to process network activities $\{x_\ell : \ell \in \mathcal{L}\}$ to satisfy the customer demand at the sink nodes, so that the firm maximizes an overall profit. If we denote $\mathcal{U}$ as a demand uncertainty set, then a single-period problem is formulated as a two-stage robust optimization problem

$$\begin{aligned} \max_{s_i \geq 0} \quad & \left[-\sum_{i \in \mathcal{N}} c_S^i s_i + \min_{\mathbf{d} \in \mathcal{U}} \max_{x_\ell \geq 0} \left[\sum_{\ell \in \mathcal{L}} r_\ell x_\ell - \sum_{j \in \mathcal{S}} c_P^j \Big(d_j - \sum_{\ell=(i,j), i \in \mathcal{N}_j} x_\ell \Big) - \sum_{i \in \mathcal{N}} c_H^i \Big(s_i - \sum_{\ell=(i,j), j \in \mathcal{S}_i} x_\ell \Big) \right] \right] \\ \text{s.t.} \quad & \sum_{\ell=(i,j), j \in \mathcal{S}_i} x_\ell \leq s_i \quad \forall\, i \in \mathcal{N}, \sum_{\ell=(i,j), i \in \mathcal{N}_j} x_\ell \leq d_j \quad \forall\, j \in \mathcal{S}, \end{aligned} \tag{2}$$

where the constraints are network constraints and affect the inner maximization problem. Note that the order quantities $\{s_i : i \in \mathcal{N}\}$ are "here-and-now" decisions; it must be placed before demands are realized, while the network activities $\{x_\ell : \ell \in \mathcal{L}\}$ are "wait-and-see" solutions and assigned after demands are observed.

Notation. To simplify (2), we define $\mathbf{c}_S \in \mathbb{R}_+^m$, $\mathbf{c}_H \in \mathbb{R}_+^m$, $\mathbf{c}_P \in \mathbb{R}_+^n$ and $\mathbf{r} \in \mathbb{R}_+^p$ as cost and revenue vectors, and define $\mathbf{R}_S \in \mathbb{R}_+^{m \times p}$ and $\mathbf{R}_D \in \mathbb{R}_+^{n \times p}$ as matrices that describe the two constraints, respectively. Decision variables and uncertain demands are $\mathbf{s} \in \mathbb{R}_+^m$, $\mathbf{x} \in \mathbb{R}_+^p$, and $\mathbf{d} \in \mathbb{R}_+^n$. We obtain

$$\begin{aligned} \Pi(\mathcal{U}) := \max_{\mathbf{s} \geq \mathbf{0}} \quad & \left[-(\mathbf{c}_S + \mathbf{c}_H)^\top \mathbf{s} \; + \; \min_{\mathbf{d} \in \mathcal{U}} \max_{\mathbf{x} \geq \mathbf{0}} \left[\mathbf{v}^\top \mathbf{x} - \mathbf{c}_P^\top \mathbf{d} \right] \right] \\ \text{s.t.} \quad & \mathbf{R}_S \mathbf{x} \leq \mathbf{s}, \quad \mathbf{R}_D \mathbf{x} \leq \mathbf{d}, \end{aligned}$$

with $\mathbf{v} = \mathbf{r} + \mathbf{R}_S^\top \mathbf{c}_H + \mathbf{R}_D^\top \mathbf{c}_P$.

2.2. Modeling Demand Uncertainty

For the sake of simplicity, we assume that there is no demand seasonality and that the demand realizations are light-tailed in nature (i.e., the demand variance is finite). For each sink installation $j \in \mathcal{S}$, we denote the demand mean by μ_j and the demand standard deviation by σ_j. Our framework also captures correlation among the demand, where we denote $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ as its nominal value. Note that all these values can be inferred from historical data. Instead of describing the demand as a random variable, we describe the demand and its correlation by using budget uncertainty sets (Bertsimas and Sim 2004) and a factor-based approach (Bandi and Bertsimas 2012). Such sets do not require any distributional assumption other than first two moments, and consequently, they are robust to the distribution choice.

We capture the correlations via the covariance matrix $\boldsymbol{\Sigma}$ with rank $l \leq n$. This means, there exist $\mathbf{A}$ and $\lambda_1, \ldots, \lambda_l > 0$ that satisfy $\boldsymbol{\Sigma} = \mathbf{A} \cdot \mathbf{diag}(\lambda_1^2, \ldots, \lambda_l^2) \cdot \mathbf{A}^\top$.

Definition 1 (Single-period Uncertainty set). The uncertainty set for correlated demands at sink nodes $\mathbf{d} = (d_1, \ldots, d_n)$ with variability parameters $\Gamma,\ \Gamma_B \geq 0$ is

$$\mathcal{U} = \left\{ \mathbf{d} \in \mathbb{R}_+^n \;\middle|\; \mathbf{d} = \boldsymbol{\mu} + \mathbf{A} \cdot \mathbf{z},\ \sum_{i=1}^{l} \left| \frac{z_i}{\lambda_i} \right| \leq \Gamma,\ \left| \frac{z_i}{\lambda_i} \right| \leq \Gamma_B \ \ \forall\, i = 1, \ldots, l \right\}. \tag{3}$$

Note that in this definition, $\Gamma,\ \Gamma_B \geq 0$ can control the degree of conservatism. The first constraint in $\mathcal{U}$ captures correlation, and the others are budget constraints which limit the absolute deviation from its nominal value. While $\mathcal{U}$ is data driven, it also captures previous results on the effect of mean and standard deviation on the profit in newsvendor networks. In particular, $\mathcal{U}$ recovers the insightful properties in Van Mieghem and Rudi (2002), as proposed in the following.

Proposition 1. *For a single-period robust newsvendor network with the uncertainty set $\mathcal{U}$, the worst-case profit increases in μ_i and decreases in λ_i.*

This proposition shows that our framework generalizes the structural properties from stochastic networks without distributional assumptions. We extend our model to multi-period cases in the next section.

3. Multi-period Robust Newsvendor Networks

To extend the single-period models into dynamic cases, we consider a decision maker who has multiple processing points of T periods. We assume that all parameters $\mathbf{c}_S, \mathbf{c}_H, \mathbf{c}_P, \mathbf{r}$, and matrices $\mathbf{R}_S, \mathbf{R}_D$ remain constant over the time horizon. As in Section 2, on-hand input stocks at source nodes and unsatisfied demand at sink nodes are backlogged to the next periods. We also assume that the demands are correlated over sink nodes, but independent over time, with nominal mean vector $\boldsymbol{\mu}_t$ and covariance matrix $\boldsymbol{\Sigma}_t$ for each time period $t = 1, \ldots, T$.

Notation. Order quantities at time t are denoted by $\mathbf{s}_t$, customer demands by $\mathbf{d}_t$, and network activities by $\mathbf{x}_t$. single-station quantities are denoted by s_t, d_t, x_t. Aggregated amount of orders up to time t are denoted by uppercases $\mathbf{S}_t$, customer demands by $\mathbf{D}_t$, and network activities by $\mathbf{X}_t$ (S_t, D_t, X_t for single-station). Inventory levels and backlogged demands after time t are denoted by $\mathbf{u}_t^{\mathbf{s}}$ and $\mathbf{u}_t^{\mathbf{d}}$. Finally, $\mathbf{d}_{[t_1:t_2]} = (\mathbf{d}_{t_1}, \dots, \mathbf{d}_{t_2}) \in \mathbb{R}_+^{n\times(t_2-t_1)}$ contains every realized demand from time t_1 to t_2. Other quantities such as $\mathbf{S}_{[t_1:t_2]}$ and $\mathbf{X}_{[t_1:t_2]}$ are defined similarly. We define $\mathbf{A}_t$ and $\lambda_{t,i}, \dots, \lambda_{t,l_t}$ for each t, with $rank(\boldsymbol{\Sigma}_t) = l_t$ and $\boldsymbol{\Sigma}_t = \mathbf{A}_t \cdot \operatorname{diag}(\lambda_{t,i}, \dots, \lambda_{t,l_t})\mathbf{A}_t^\top$.

In the following, we generalize Definition 1 for multi-period demand.

Definition 2 (Multi-period Uncertainty Set). The uncertainty set for the demand at sink nodes $(\mathbf{d}_1, \dots, \mathbf{d}_T) \in \mathbb{R}^{n\times T}$ over T periods is

$$\mathcal{U}^T = \left\{ (\mathbf{d}_1, \dots, \mathbf{d}_T) \;\middle|\; \begin{array}{l} \mathbf{d}_t = \boldsymbol{\mu}_t + \mathbf{A}_t \mathbf{z}_t \quad \forall t = 1, \dots, T \\ \displaystyle\sum_{t=1}^{T} \sum_{i=1}^{l_t} \left| \frac{z_{t,i}}{\lambda_{t,i}} \right| \le \Gamma, \ \sum_{i=1}^{l_t} \left| \frac{z_{t,i}}{\lambda_{t,i}} \right| \le \Gamma_t, \ \left| \frac{z_{t,i}}{\lambda_{t,i}} \right| \le \Gamma_B \quad \forall i = 1, \dots, l_t, \ t = 1, \dots, T \end{array} \right\}.$$

In this set, the additional constraint controls the absolute deviation over nodes and time periods. It prevents the demand to take extreme values in every period t, which reduces the conservatism over time. This definition can also describe seasonality of demands, which applies to many areas. When there is an explicit time-dependence between the periods, $\mathcal{U}^T$ can be expressed as a conic set (Nohadani and Roy 2017).

For the multi-period newsvendor networks, we can express the dynamics of inventories and backlogged demands in (1) with vectors and matrices as

$$\mathbf{u}_t^{\mathbf{s}} = \mathbf{u}_{t-1}^{\mathbf{s}} + \mathbf{s}_t - \mathbf{R}_S \mathbf{x}_t = \sum_{\tau=1}^{t} \Big(\mathbf{s}_\tau - \mathbf{R}_S \mathbf{x}_\tau \Big)$$
$$\mathbf{u}_t^{\mathbf{d}} = \mathbf{u}_{t-1}^{\mathbf{d}} + \mathbf{d}_t - \mathbf{R}_D \mathbf{x}_t = \sum_{\tau=1}^{t} \Big(\mathbf{d}_\tau - \mathbf{R}_D \mathbf{x}_\tau \Big),$$

and model a multi-stage robust optimization problem as

$$\max_{\mathbf{S}_t(\mathbf{d}_{[1:t-1]}) \ge \mathbf{0}} \ \min_{\mathbf{d}_{[1:T]} \in \mathcal{U}^T} \ \max_{\mathbf{X}_t \in \mathcal{P}_t(\mathbf{S}_t, \mathbf{D}_t, \mathbf{X}_{t-1})} \Bigg[-\mathbf{c}_S^\top \mathbf{S}_T(\mathbf{d}_{[1:T-1]}) - \mathbf{c}_H^\top \sum_{t=1}^{T} \Big[\mathbf{S}_t(\mathbf{d}_{[1:t-1]}) - \mathbf{R}_S \mathbf{X}_t(\mathbf{d}_{[1:t]}) \Big] - \mathbf{c}_P^\top \sum_{t=1}^{T} \Big[\mathbf{D}_t - \mathbf{R}_D \mathbf{X}_t(\mathbf{d}_{[1:t]}) \Big] + \mathbf{r}^\top \mathbf{X}_T \Bigg], \tag{4}$$

where $\mathbf{d}_{[1:0]} = \mathbf{0}$, $\mathbf{X}_0 = \mathbf{0}$. Note that $\mathbf{X}_t$ is determined after $\mathbf{S}_t$ and $\mathbf{D}_t$, within a set

$$\mathcal{P}_t(\mathbf{S}_t, \mathbf{D}_t, \mathbf{X}_{t-1}) = \left\{ \mathbf{X}_t \in \mathbb{R}_+^p \;\middle|\; \mathbf{R}_S \mathbf{X}_t \le \mathbf{S}_t, \ \mathbf{R}_D \mathbf{X}_t \le \mathbf{D}_t, \ \mathbf{X}_t \ge \mathbf{X}_{t-1} \right\},$$

which is defined for $\mathbf{X}_t$ to maximize profit, where the last constraint requires non-negative network activities. The main difference between single-period and multi-period models is that the order quantities are not *static*. That means, in order to obtain an optimal solution, one should find $\mathbf{S}_t$ as a function of $\mathbf{d}_{[1:t-1]}$ so that they are *fully-adjustable* to all previous demands. Such policies also need to be *non-anticipative*, i.e., adjustable decisions should only be based on realized uncertainties.

Even for $T=1$, the problem (4) is a two-stage robust optimization problem and shown to be NP-hard (Ben-Tal et al. 2004) For multi-period setting, the complexity only worsens and, to our knowledge, no tractable algorithm has been proposed to exactly solve the general problem in (4). Because of this restrictions to specific policies have been considered. In particular, *affine policies* have been proposed, where adaptive decisions are assumed to be an affine function of realized uncertainties. These policies have exhibited excellent performance in many real-world applications. With such policies, the multi-period problem (4) converts to determining the parameters $\{\mathbf{w}_t \in \mathbb{R}^m : 1 \le t \le T\}$ and $\{\mathbf{W}_{\tau,t} \in \mathbb{R}^{m\times n} : 1 \le \tau \le t-1,\ 1 \le t \le T\}$, where

$$\mathbf{s}_1 = \mathbf{w}_1, \quad \mathbf{s}_t = \mathbf{w}_t + \sum_{\tau=1}^{t-1} \mathbf{W}_{\tau,t}\mathbf{d}_\tau \qquad t = 2,\ldots,T. \tag{5}$$

These policies force non-anticipativity of $\mathbf{s}_t$ and one can reformulate (4) as a two-stage adaptive linear optimization problem

$$\max_{\mathbf{w}_t,\mathbf{W}_{\tau,t}} \min_{\mathbf{d}_{[1:T]}\in\mathcal{U}^T} \max_{\mathbf{S}_{[1:T]},\mathbf{X}_{[1:T]}} \left[-\mathbf{c}_S^\top \mathbf{S}_T - \mathbf{c}_H^\top \sum_{t=1}^T \Big(\mathbf{S}_t - \mathbf{R}_S\mathbf{X}_t\Big) - \mathbf{c}_P^\top \sum_{t=1}^T \Big(\mathbf{D}_t - \mathbf{R}_D\mathbf{X}_t\Big) + \mathbf{r}^\top \mathbf{X}_T \right] \tag{6}$$

$$\text{s.t.} \quad \left. \mathbf{w}_t + \sum_{\tau=1}^{t-1}\mathbf{W}_{\tau,t}\mathbf{d}_\tau \ \ge\ \mathbf{0}, \quad \mathbf{w}_1 \ \ge\ \mathbf{0} \right\} \quad \forall\ t = 2,\ldots,T,\ \forall\ \mathbf{d}_{[1:T]} \in \mathcal{U}^T \tag{7}$$

$$\left. \begin{aligned} &\mathbf{S}_1 = \mathbf{w}_1 \\ &\mathbf{S}_t = \mathbf{w}_1 + \sum_{j=2}^t \Big(\mathbf{w}_j + \sum_{\tau=1}^{j-1}\mathbf{W}_{\tau,j}\mathbf{d}_\tau\Big) \\ &\mathbf{R}_S\mathbf{X}_t \ \le\ \mathbf{S}_t \\ &\mathbf{R}_D\mathbf{X}_t \ \le\ \mathbf{D}_t \\ &\mathbf{X}_T \ge \mathbf{X}_{T-1} \ge \cdots \ge \mathbf{X}_1 \ge \mathbf{0} \end{aligned} \right\} \quad \forall t = 1,\ldots,T. \tag{8}$$

Constraint (7) implies that the order quantities are non-negative for any realizations of past demands, and constraint (8) affects the inner maximization problem, which determines the processing activities.

PROPOSITION 2. *Finding an optimal affine policy for a multi-period newsvendor network in (6–8) is a convex optimization problem.*

REMARK 1. The network activities $\mathbf{x}_{[1:T]}$ maximize the net profit over the entire horizon, not just at time t, i.e., we relax non-anticipativity of $\mathbf{x}_{[1:T]}$ in the optimization problem (6–8). However, we claim that this relaxation will not be loose, because penalty cost and holding cost force $\mathbf{x}_t$ to maximize profit in the corresponding period. As a special case, one can show that in single-station models, $\mathbf{x}_t$ maximizes the overall profit if and only if it maximizes the profit at time t. This relaxation facilitates generality, as problem (6–8) is defined for any polyhedral uncertainty sets, whereas in the stochastic case optimal strategies are only available for restricted cases (demands are i.i.d. over time as in Van Mieghem and Rudi (2002)).

Remark 2. The inner maximization problem in (6) is concave, making the minimization problem non-convex. The overall problem can be solved with cut generation. Proposition 2 guarantees that if $\mathcal{U}^T$ is a polyhedron, then a (global) solution is obtained within finite iterations. However, the problem is NP-hard and the computation grows as T increases. Our main contribution is motivated by taking an alternative approach to (6–8), as presented next.

4. Periodic-affine policies for single-station models

As discussed, affine policies face computational difficulties when a decision maker has a larger number of resources and products over an extended period of time. We propose a new solution concept, denoted as *periodic-affine policies* (PA), where the overall time horizon is separated into subperiods, that are interconnected by the preceding surplus to become the proceeding demand. In this approach, the order quantities are determined as an affine function of past demands realized only within its subperiod, as opposed to affine policies where all previous demands are considered. This scheme reduces the number of decision variables and consequently the computation time. Our framework constructs this policy by first formulating a dynamic programming (DP) problem, where each stage corresponds to a subperiod. We also propose an algorithm to compute such periodic-affine policies and show that they are computationally more tractable than affine policies. In addition, we present a sufficient condition for the optimal solution to the original affine problem. We first consider T-period single-station models in this section. However, our framework is naturally extended to multi-station networks which we discuss in the subsequent section.

Notation. We use same notations for all cost parameters c_S, c_H, c_P with revenue per item, r, and we may assume that $R_S = R_D = 1$ without loss of generality in single-station models. In this section $\mathbf{d} = d_{[1:T]}$, $\mathbf{s} = s_{[1:T]}$, and $\mathbf{x} = x_{[1:T]}$. We denote $\pi(w_t, W_{\tau,t})$ as an affine policy with affine parameters $\{w_t, W_{\tau,t} : 1 \leq \tau \leq t-1, 1 \leq t \leq T\}$. Furthermore, the problem of a T-period single-station newsvendor model is denoted as $\Phi(s_0, d_0)$ for an uncertainty set $\mathcal{U}^T$ with initial input $s_0 \geq 0$ and demands $d_0 \geq 0$.

Analysis of initial input and demand. We first study the role of initial input and demand for the optimal affine policy in the multi-period model, given by

$$\Phi(s_0, d_0) := \max_{\pi} \min_{\mathbf{d} \in \mathcal{U}^T} \max_{\mathbf{x}, \mathbf{s} \in \mathcal{X}(\pi, \mathbf{d}, s_0, d_0)} P\Big(\pi(w_t, W_{\tau,t}), \mathbf{d}, \mathbf{x}; s_0, d_0\Big),$$

where the profit during the period is

$$P\Big(\pi(w_t, W_{\tau,t}), \mathbf{d}, \mathbf{x}; s_0, d_0\Big) = -\,c_S\left(\sum_{t=1}^{T} s_t\right) - c_H \sum_{t=1}^{T}\left(s_0 + \sum_{\tau=1}^{t}(s_\tau - x_\tau)\right) - c_P \sum_{t=1}^{T}\left(d_0 + \sum_{\tau=1}^{t}(d_\tau - x_\tau)\right) + r\left(\sum_{t=1}^{T} x_t\right)$$

and the feasible set $\mathcal{X}(\pi, \mathbf{d}, s_0, d_0)$ is given by

$$\mathcal{X}(\pi, \mathbf{d}, s_0, d_0) = \left\{ \mathbf{s}, \mathbf{x} \geq \mathbf{0} \;\middle|\; \begin{array}{ll} s_1 = w_1,\ s_t = w_t + \sum_{\tau=1}^{t-1} W_{\tau,t} d_\tau & \forall\, t = 2, \ldots, T \\ \sum_{\tau=1}^{t} x_\tau \leq s_0 + \sum_{\tau=1}^{t} s_\tau & \forall\, t = 1, \ldots, T \\ \sum_{\tau=1}^{t} x_\tau \leq d_0 + \sum_{\tau=1}^{t} d_\tau & \forall\, t = 1, \ldots, T \end{array} \right\}. \tag{9}$$

The result is intuitive and plays a key role in establishing periodic-affine policies.

Proposition 3. *For an optimal affine policy $\pi^*(w_t^*, W_{\tau,t}^*)$ of $\Phi(0,0)$ with no initial input and demand, if $s_0 \leq w_1^*$, then:*

(i) An optimal affine policy $\overline{\pi} = \overline{\pi}(\overline{w}_t^*, \overline{W}_{\tau,t}^*)$ of $\Phi(s_0, d_0)$ is characterized as

$$\begin{cases} \overline{w}_1^* = w_1^* - s_0 + d_0 & \\ \overline{w}_t^* = w_t^* & \forall\, t = 2, \ldots, T \\ \overline{W}_{\tau,t}^* = W_{\tau,t}^* & \forall\, \tau = 1, \ldots, t-1,\ \forall\, t = 1, \ldots, T. \end{cases} \tag{10}$$

(ii) There exists a single worst-case demand $\mathbf{d}^* \in \mathcal{U}^T$ for both $\Phi(0,0)$ and $\Phi(s_0, d_0)$.

This result implies that for small enough connecting inventories, the subperiods become effectively decoupled.

4.1. Model Formulation

We now introduce the dynamic programming formulation for a multi-period newsvendor network.

Notation. For a T-period single-station model, we partition the time period into N subperiods sorted as $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$. In interval $I_j = \{t_{j-1}+1, \ldots, t_j\}$, the uncertainty set $\mathcal{U}^j \in \mathbb{R}_+^{|I_j|}$ for every $j = 1, \ldots, N$. The amount of on-hand input stock and backlogged demands after time t are u_t^s and u_t^d. A class of affine policies for j^{th} subperiod is denoted by $\Pi_{\text{aff}}(\mathcal{U}^j, \Xi_{j-1})$ on the uncertainty set $\mathcal{U}^j$, where the state Ξ_{j-1} contains all past information at the beginning of j^{th} period with $\Xi_0 = 0$.

We formulate an N-stage robust dynamic programming problem as

$$\max_{\pi_1\in\Pi_{\text{aff}}(\mathcal{U}^1,\Xi_0)}\left[\min_{d_{I_1}\in\mathcal{U}^1}\max_{x_{I_1},s_{I_1}\in\mathcal{X}_1}\left[P_1\Big(\pi_1,d_{I_1},x_{I_1};0,0\Big)\right.\right.$$
$$+\max_{\pi_2\in\Pi_{\text{aff}}(\mathcal{U}^2,\Xi_1)}\left[\min_{d_{I_2}\in\mathcal{U}^2}\max_{x_{I_2},s_{I_2}\in\mathcal{X}_2}\left[P_2\Big(\pi_2,d_{I_2},x_{I_2};u^s_{t_1},u^d_{t_1}\Big)+\cdots\right.\right. \tag{11}$$
$$\left.\left.\left.\left.+\max_{\pi_N\in\Pi_{\text{aff}}(\mathcal{U}^N,\Xi_{N-1})}\left[\min_{d_{I_N}\in\mathcal{U}^N}\max_{x_{I_N},s_{I_N}\in\mathcal{X}_N}P_N\Big(\pi_N,d_{I_N},x_{I_N};u^s_{t_{N-1}},u^d_{t_{N-1}}\Big)\right]\cdots\right]\right]\right]\right],$$

where $P_j\Big(\pi_j,d_{I_j},x_{I_j};u^s_{t_{j-1}},u^d_{t_{j-1}}\Big)$ is a profit generated during the j^{th} subperiod with an initial input and demand

$$P_j\Big(\pi_j,d_{I_j},x_{I_j};u^s_{t_{j-1}},u^d_{t_{j-1}}\Big)=-c_S\left(\sum_{t\in I_j}s_t\right)-c_H\sum_{t\in I_j}\left(u^s_{t_{j-1}}+\sum_{\tau=t_{j-1}+1}^{t}(s_\tau-x_\tau)\right)$$
$$-c_P\sum_{t\in I_j}\left(u^d_{t_{j-1}}+\sum_{\tau=t_{j-1}+1}^{t}(d_\tau-x_\tau)\right)+r\left(\sum_{t\in I_j}x_t\right),$$

and s_{I_j} and x_{I_j} are determined within a feasible set $\mathcal{X}_j=\mathcal{X}(\pi_j,d_{I_j},u^s_{t_{j-1}},u^d_{t_{j-1}})$ from (9). In (11), s_t denotes the order quantity at time t, if ordering decision is made by π_j and d_{I_j} is realized.

Periodic-affine policy formulation. With initial input u^s_0 bounded above with s^*_0, we define *affine-IBS* policies by modifying the initial period of an affine policy $\pi(w_t,W_{\tau,t})$.

DEFINITION 3 (AFFINE-IBS). For j^{th} subperiod, the *affine Initial Base-Stock policy* $\overline{\pi}_j(w^{(j)}_t,W^{(j)}_{\tau,t})$ associated with an affine policy $\pi_j(w^{(j)}_t,W^{(j)}_{\tau,t})$ determines order quantity by

$$s_t(u^s,u^d,d_{I_j})=\begin{cases}w^{(j)}_1-u^s+u^d & t=t_{j-1}+1\\ w^{(j)}_t+\displaystyle\sum_{\tau=1}^{i-1}W^{(j)}_{\tau,i}d_{t_{j-1}+\tau} & t=t_{j-1}+i,\ i\ge 2,\ t\in I_j.\end{cases}$$

Note that affine-IBS policies adapts to initial input and demand by adjusting the order quantity at the first period. From the second period, affine-IBS and its associated affine policies are equivalent.

We now consider a sequence of affine-IBS policies $\overline{\pi}=(\overline{\pi}_1,\ldots,\overline{\pi}_N)$, where each $\overline{\pi}_j=\overline{\pi}_j(w^{(j)}_t,W^{(j)}_{\tau,t})$ is for j^{th} subperiod. Note that this policy may not be well-defined for each subperiod because it does not guarantee that every order quantity is non-negative. That means, if an input stock after t_j is greater than $w^{(j+1)}_1$, then the policy would not be feasible. To account for this, we impose

$$w^{(j+1)}_1\ \ge\ u^s_{t_j}=u^s_{t_j}\Big(\overline{\pi}_j,d_{I_j}\Big)\qquad\forall d_{I_j}\in\mathcal{U}^j\ \forall j=1,\ldots,N-1.$$

By Definition 3, the right-hand side is equivalent to

$$u^s_{t_j}\Big(\overline{\pi}_j,d_{I_j}\Big)=\max\left(0,\ \left(u^s_{t_{j-1}}+\sum_{t\in I_j}s_t(u^s_{t_{j-1}},u^d_{t_{j-1}},d_{I_j})\right)-\left(u^d_{t_{j-1}}+\sum_{t\in I_j}d_t\right)\right)$$
$$=\max\left(0,\ \left(\sum_{t=1}^{t_j-t_{j-1}}w^{(j)}_t+\sum_{t=2}^{t_j-t_{j-1}}\sum_{\tau=1}^{t-1}W^{(j)}_{\tau,t}d_{t_{j-1}+\tau}\right)-\sum_{t=1}^{t_j-t_{j-1}}d_{t_{j-1}+t}\right).$$

Since $w_1^{(j+1)}$ has to be non-negative for any demand realization, the periodic-affine policy is well-defined if

$$w_1^{(j+1)} \;\geq\; S_j^* := \max_{d_{I_j}\in\mathcal{U}^j} \left[\sum_{t=1}^{t_j - t_{j-1}} w_t^{(j)} + \sum_{t=2}^{t_j - t_{j-1}} \sum_{\tau=1}^{t-1} W_{\tau,t}^{(j)} d_{t_{j-1}+\tau} - \sum_{t=1}^{t_j - t_{j-1}} d_{t_{j-1}+t} \right] \tag{12}$$

for every $j = 0, \ldots, N-1$, where S_j^* denotes the maximum leftover input after j^{th} subperiod with $S_0^* = 0$. Now we can define periodic-affine policies.

Definition 4 (Periodic-affine Policy). A *periodic-affine policy* $\overline{\pi}_{\text{PA}} := (\overline{\pi}_1, \ldots, \overline{\pi}_N)$ is an affine-IBS policy $\overline{\pi}_i$ satisfying (12) for affine policies π_i.

For a periodic-affine policy $\overline{\pi}_{\text{PA}} = (\overline{\pi}_1, \ldots, \overline{\pi}_N)$, where $\overline{\pi}_j = \overline{\pi}_j(w_t^{(j)}, W_{\tau,t}^{(j)})$, order quantities at time t are determined as follows:

$$s_t = s_t(\overline{\pi}_{\text{PA}}) = \begin{cases} w_1^{(j)} - u^s_{t_{j-1}} + u^d_{t_{j-1}} & \forall\ t = t_{j-1}+1 \\ w_i^{(j)} + \displaystyle\sum_{\tau=1}^{i-1} W_{\tau,i}^{(j)} d_{t_{j-1}+\tau} & \forall\ t = t_{j-1}+i,\ i \geq 2,\ t \in I_j, \end{cases} \tag{PA}$$

for every $j = 1, \ldots, N$. Since $\overline{\pi}_j \in \mathit{\Pi}_{\text{aff}}(\mathcal{U}^j, \mathit{\Xi}_{j-1})$ and $\overline{\pi}_{\text{PA}}$ satisfies (12), every periodic-affine policy is a feasible solution to the DP in (11).

In the next section, we present our algorithm to compute periodic-affine policies.

4.2. Periodic-affine algorithm

Our algorithm obtains affine-IBS policies for each subperiod by solving smaller subproblems. However, since affine-IBS policies takes initial input and demands into account, we construct the objective function to account for leftover resources and demands. We identify such objective functions from the DP problem (11). We first show that if initial input is small, an affine-IBS policy will be optimal among $\mathit{\Pi}_{\text{aff}}(\mathcal{U}^N, \mathit{\Xi}_{N-1})$. The proof is similar to the Proposition 3 and is omitted.

Corollary 1. *Let $\pi_N(w_t^{(N)}, W_{\tau,t}^{(N)})$ be an optimal affine policy with zero initial input and demands. If $u^s_{t_{N-1}} \leq w_1^{(N)}$ for any realization of $u^s_{t_{N-1}}$, then its associated affine-IBS policy $\overline{\pi}_N$ is an optimal solution among $\mathit{\Pi}_{\text{aff}}(\mathcal{U}^N, \mathit{\Xi}_{N-1})$. Moreover,*

$$\begin{aligned} &\max_{\pi\in\mathit{\Pi}_{\text{aff}}(\mathcal{U}^N, \mathit{\Xi}_{N-1})} \min_{\mathbf{d}\in\mathcal{U}^N} \max_{\mathbf{x}\in\mathcal{X}} \left[P\Big(\pi, \mathbf{d}, \mathbf{x}; u_0^s, u_0^d\Big) \right] \\ &\qquad = c_S u_0^s + (r - c_S) u_0^d + \max_{\pi\in\mathit{\Pi}_{\text{aff}}(\mathcal{U},0)} \min_{\mathbf{d}\in\mathcal{U}} \max_{\mathbf{x}\in\mathcal{X}} \left[P\Big(\pi, \mathbf{d}, \mathbf{x}; 0, 0\Big) \right]. \end{aligned} \tag{13}$$

Using this Corollary, we reformulate an optimality condition for the last stage as

$$\begin{aligned} &\max_{\pi_N\in\mathit{\Pi}_{\text{aff}}(\mathcal{U}^N, \mathit{\Xi}_{N-1})} \min_{d_{I_N}\in\mathcal{U}^N} \max_{x_{I_N}\in\mathcal{X}_N} \left[P_N\Big(\pi_N, d_{I_N}, x_{I_N}; u^s_{t_{N-1}}, u^d_{t_{N-1}}\Big) \right] \\ &\qquad = c_S \cdot u^s_{t_{N-1}} + (r - c_S)\cdot u^d_{t_{N-1}} + \max_{\pi_N\in\mathit{\Pi}_{\text{aff}}(\mathcal{U}^N,0)} \min_{d_{I_N}\in\mathcal{U}^N} \max_{x_{I_N}\in\mathcal{X}_N} \left[P_N\Big(\pi_N, d_{I_N}, x_{I_N}; 0, 0\Big) \right]. \end{aligned}$$

In single-station cases, $u^s_{t_{N-1}}$ and $u^d_{t_{N-1}}$ can be rewritten as

$$u^s_{t_{N-1}} = u^s_{t_{N-2}} + \sum_{t\in I_{N-1}} \Big(s_t - x_t\Big), \quad u^d_{t_{N-1}} = u^d_{t_{N-2}} + \sum_{t\in I_{N-1}} \Big(d_t - x_t\Big)$$

This can be incorporated with $P_{N-1}\Big(\pi_{N-1}, d_{I_{N-1}}, x_{I_{N-1}}; u^s_{t_{N-2}}, u^d_{t_{N-2}}\Big)$ as

$$\begin{aligned}
&P_{N-1}\Big(\pi_{N-1}, d_{I_{N-1}}, x_{I_{N-1}}; u^s_{t_{N-2}}, u^d_{t_{N-2}}\Big) + c_S\cdot u^s_{t_{N-1}} + (r-c_S)\cdot u^d_{t_{N-1}} \\
&= -c_S\left(\sum_{t\in I_{N-1}} s_t\right) - c_H \sum_{t\in I_{N-1}}\left(u^s_{t_{N-2}} + \sum_{\tau=t_{N-2}+1}^{t}(s_\tau - x_\tau)\right) - c_P \sum_{t\in I_{N-1}}\left(u^d_{t_{N-2}} + \sum_{\tau=t_{N-2}+1}^{t}(d_\tau - x_\tau)\right) \\
&\qquad + r\left(\sum_{t\in I_{N-1}} x_t\right) + c_S\left(u^s_{t_{N-2}} + \sum_{t\in I_{N-1}}(s_t - x_t)\right) + (r-c_S)\left(u^d_{t_{N-2}} + \sum_{t\in I_{N-1}}(d_t - x_t)\right) \\
&= c_S\cdot u^s_{t_{N-2}} + (r-c_S)\cdot u^d_{t_{N-2}} - c_H \sum_{t\in I_{N-1}}\left(u^s_{t_{N-2}} + \sum_{\tau=t_{N-2}+1}^{t}(s_\tau - x_\tau)\right) \\
&\qquad - c_P \sum_{t\in I_{N-1}}\left(u^d_{t_{N-2}} + \sum_{\tau=t_{N-2}+1}^{t}(d_\tau - x_\tau)\right) + (r-c_S)\left(\sum_{t\in I_{N-1}} d_t\right).
\end{aligned} \tag{14}$$

We define a modified objective function $\widetilde{P}_{N-1}$ as

$$\begin{aligned}
\widetilde{P}_{N-1}\Big(\pi_{N-1}, d_{I_{N-1}}, x_{I_{N-1}}; u^s_{t_{N-2}}, u^d_{t_{N-2}}\Big) &= -c_H \sum_{t\in I_{N-1}}\left(u^s_{t_{N-2}} + \sum_{\tau=t_{N-2}+1}^{t}(s_\tau - x_\tau)\right) \\
&\quad - c_P \sum_{t\in I_{N-1}}\left(u^d_{t_{N-2}} + \sum_{\tau=t_{N-2}+1}^{t}(d_\tau - x_\tau)\right) + (r-c_S)\left(\sum_{t\in I_{N-1}} d_t\right).
\end{aligned} \tag{15}$$

If we assume that both $u^s_{t_{N-2}}$ and $u^s_{t_{N-1}}$ are small, we can rewrite

$$\begin{aligned}
&\max_{\pi_{N-1}\in \Pi_{\text{aff}}(\mathcal{U}^{N-1}, \Xi_{N-2})} \min_{d_{I_{N-1}}\in\mathcal{U}^{N-1}} \max_{x_{I_{N-1}}\in\mathcal{X}_{N-1}} \Bigg[P_{N-1}\Big(\pi_{N-1}, d_{I_{N-1}}, x_{I_{N-1}}; u^s_{t_{N-2}}, u^d_{t_{N-2}}\Big) \\
&\qquad + \max_{\pi_N\in \Pi_{\text{aff}}(\mathcal{U}^{N}, \Xi_{N-1})} \min_{d_{I_N}\in\mathcal{U}^{N}} \max_{x_{I_N}\in\mathcal{X}_{N}} \Big[P_N\Big(\pi_N, d_{I_N}, x_{I_N}; u^s_{t_{N-1}}, u^d_{t_{N-1}}\Big)\Big]\Bigg] \\
&= \max_{\pi_{N-1}\in \Pi_{\text{aff}}(\mathcal{U}^{N-1}, \Xi_{N-2})} \min_{d_{I_{N-1}}\in\mathcal{U}^{N-1}} \max_{x_{I_{N-1}}\in\mathcal{X}_{N-1}} \Bigg[P_{N-1}\Big(\pi_{N-1}, d_{I_{N-1}}, x_{I_{N-1}}, u^s_{t_{N-2}}, u^d_{t_{N-2}}\Big) \\
&\qquad + c_S\cdot u^s_{t_{N-1}} + (r-c_S)\cdot u^d_{t_{N-1}} + \max_{\pi_N\in \Pi_{\text{aff}}(\mathcal{U}^{N}, 0)} \min_{d_{I_N}\in\mathcal{U}^{N}} \max_{x_{I_N}\in\mathcal{X}_{N}} \Big[P_N\Big(\pi_N, d_{I_N} x_{I_N}; 0, 0\Big)\Big]\Bigg] \\
&= \max_{\pi_{N-1}\in \Pi_{\text{aff}}(\mathcal{U}^{N-1}, \Xi_{N-2})} \min_{d_{I_{N-1}}\in\mathcal{U}^{N-1}} \max_{x_{I_{N-1}}\in\mathcal{X}_{N-1}} \Bigg[c_S\cdot u^s_{t_{N-2}} + (r-c_S)\cdot u^d_{t_{N-2}} \\
&\qquad + \widetilde{P}_{N-1}\Big(\pi_{N-1}, d_{I_{N-1}}, x_{I_{N-1}}; u^s_{t_{N-2}}, u^d_{t_{N-2}}\Big)\Bigg] + \max_{\pi_N\in \Pi_{\text{aff}}(\mathcal{U}^{N}, 0)} \min_{d_{I_N}\in\mathcal{U}^{N}} \max_{x_{I_N}\in\mathcal{X}_{N}} \Big[P_N\Big(\pi_N, d_{I_N} x_{I_N}; 0, 0\Big)\Big]
\end{aligned}$$

$$= c_S \cdot u^s_{t_{N-2}} + (r - c_S) \cdot u^d_{t_{N-2}} + \max_{\pi_{N-1} \in \Pi_{\text{aff}}(\mathcal{U}^{N-1},0)} \min_{d_{I_{N-1}} \in \mathcal{U}^{N-1}} \max_{x_{I_{N-1}} \in \mathcal{X}_{N-1}} \left[\widetilde{P}_{N-1}\Big(\pi_{N-1}, d_{I_{N-1}}, x_{I_{N-1}}; 0, 0\Big) \right]$$
$$+ \max_{\pi_N \in \Pi_{\text{aff}}(\mathcal{U}^N,0)} \min_{d_{I_N} \in \mathcal{U}^N} \max_{x_{I_N} \in \mathcal{X}_N} \left[P_N\Big(\pi_N, d_{I_N} x_{I_N}; 0, 0\Big) \right]. \tag{16}$$

Note that the second equality comes from (14), and one can verify similarly from Corollary 1 that the last equality holds. This reformulation shows that if leftover input after every subperiod is small enough, we can solve the DP problem in (11) by solving smaller subproblems. These subproblems are defined with modified objective function $\widetilde{P}_j$ with no backlogged input and demand, hence we can solve them independently. Proceeding iteratively, we define an objective $P_j^{\text{PA}}\Big(\pi_j, d_{I_j}, x_{I_j}\Big)$ as

$$P_j^{\text{PA}}\Big(\pi_j, d_{I_j}, x_{I_j}\Big) = \begin{cases} -c_H \sum_{t \in I_j} \Big(\sum_{\tau = t_{j-1}+1}^{t} (s_\tau - x_\tau) \Big) - c_P \sum_{t \in I_j} \Big(\sum_{\tau = t_{j-1}+1}^{t} (d_\tau - x_\tau) \Big) & \\ \qquad + (r - c_S) \sum_{t \in I_j} d_t, & j = 1, \ldots, N-1 \\ -c_S \sum_{t \in I_j} s_t - c_H \sum_{t \in I_j} \Big(\sum_{\tau = t_{j-1}+1}^{t} (s_\tau - x_\tau) \Big) & \\ \qquad - c_P \sum_{t \in I_j} \Big(\sum_{\tau = t_{j-1}+1}^{t} (d_\tau - x_\tau) \Big) + r \sum_{t \in I_j} x_t, & j = N. \end{cases} \tag{17}$$

We now propose the periodic-affine algorithm. This algorithm (i) ensures that the solution is well-defined, and (ii) exploits the modified objective functions P_j^{PA}. The j^{th} subproblem can be solved by the following optimization problem

$$\begin{aligned} \max_{\pi_j \in \Pi_{\text{aff}}(\mathcal{U}^j,0)} \min_{d_{I_j} \in \mathcal{U}^j} \max_{x_{I_j}, s_{I_j}} \quad & P_j^{\text{PA}}\Big(\pi_j, d_{I_j}, x_{I_j}\Big) \\ \text{s.t.} \quad & (x_{I_j}, s_{I_j}) \in \mathcal{X}(\pi_j, d_{I_j}, 0, 0) \\ & w_1^{(j)} \ge S^*_{j-1}. \end{aligned} \tag{18}$$

The last constraint ensures that periodic-affine policy is well-defined, where S^*_{j-1} is the maximum amount of on-hand input after $(j-1)^{\text{th}}$ subperiod, computed by (12). The overall procedure solves (18) and (12) iteratively, as summarized in Algorithm 1.

Algorithm 1 Periodic-affine algorithm for single-station problems

---

Given. time indices $0 = t_0 < t_1 < \cdots < t_N = T$, uncertainty set $\mathcal{U} = \mathcal{U}^1 \times \cdots \times \mathcal{U}^N$, $i = 1$, and $S_0^* = 0$.

Step 1. Solve (18) to obtain π_j for the j^{th} subperiod.

Step 2. Using π_j, compute the maximum leftover input S_j^* by (12).

Step 3. If $j = N$, return $\overline{\pi}_{\text{PA}} = (\overline{\pi}_1, \ldots, \overline{\pi}_N)$ and STOP. Otherwise, $j \leftarrow j + 1$ and go to *Step 1.*

We present an additional useful property of the periodic-affine algorithm, namely that the worst-case scenario can be obtained from each iteration.

PROPOSITION 4. *Let $\overline{\pi}_{PA}$ be a solution of periodic-affine algorithm and $d^*_{I_j}$ be a worst-case scenario from the j^{th} subproblem. Then $\overline{\pi}_{PA}$ has a worst-case scenario $(d^*_{I_1}, \ldots, d^*_{I_N})$.*

So far, we discussed a single-station, multi-period robust newsvendor model, where the uncertainty set over time periods is defined as a Cartesian product of pre-specified uncertainty sets for each subperiod. We formulated a dynamic programming problem, where each stage corresponds to each subperiod. Motivated from this formulation, we developed an algorithm to find a periodic-affine policy, by defining the modified objective functions P^{PA}_j's. In the next section, we present theoretical properties of periodic-affine policies, where we provide a sufficient condition for the algorithm to have an optimal solution.

4.3. Optimality of periodic-affine policies

In this section, we present theoretical properties of PA by analyzing the effect of base-stock levels on the worst-case performance. Specifically, we provide a sufficient condition under which the periodic-affine algorithm solves the DP problem (11). To compare the worst-case performance of PA with affine policies, we consider affine policies under the rectangular uncertainty set $\mathcal{U} = \mathcal{U}^1 \times \cdots \times \mathcal{U}^N$ so that both policies are defined equivalently. Moreover, we present an analytical approximation for the suboptimality of PA. Note that this is a *posterior* approximation, i.e., it is computed during the algorithm.

Let the worst-case profit of the two policies be V^*_{PA} and V^*_{Aff} and V^*_{DP} as an optimal value of the DP problem (11). The following assumption guarantees the optimality of PA policies.

ASSUMPTION 1. *For a solution of the PA algorithm $\overline{\pi}_{\mathrm{PA}}$, assume that the maximum leftover input level after each subperiod S^*_j satisfies the last constraint in (18).*

For a solution of the PA algorithm, the last constraint in (18) is not active at every iteration. Since an optimal periodic-affine policy maximizes the overall profit, it tends to have lower leftovers at each time period and is likely to satisfy Assumption 1 (See Section 6 for empirical validity).

THEOREM 1. *For a single-station network, if Assumption 1 holds, then $V^*_{\mathrm{Aff}} \leq V^*_{\mathrm{PA}} = V^*_{\mathrm{DP}}$.*

Since affine policies are defined on a subset of every feasible solution of the DP in (11) by switching the order of max and min operators, the worst-case profit of PA is guaranteed to be greater than or equal to that of affine policies under Assumption 1. We show in the following proposition that in single-station problems, their worst-case performance is indeed equal.

PROPOSITION 5. *For single-station models under Assumption 1, $V^*_{\mathrm{Aff}} = V^*_{\mathrm{PA}} = V^*_{\mathrm{DP}}$.*

Remark. All previous theoretical properties of PA rely on Assumption 1. In other words, the worst-case performance of PA may not match that of affine policies without the assumption.

We now relax Assumption 1 and provide a suboptimality bound for PA. For this, we compare the solutions of the following optimization problems

$$\begin{aligned}\mathcal{P}_j := \max_{\pi_j \in \Pi_{\text{aff}}(\mathcal{U}^j,0)} \min_{d_{I_j} \in \mathcal{U}^j} \max_{x_{I_j}, s_{I_j}} \quad & P_j^{\text{PA}}\Big(\pi_j, d_{I_j}, x_{I_j}\Big) \\ \text{s.t.} \quad & (x_{I_j}, s_{I_j}) \in \mathcal{X}(\pi_j, d_{I_j}, 0, 0) \\ & w_1^{(j)} \geq S_{j-1}^* \end{aligned} \qquad \begin{aligned}\widetilde{\mathcal{P}}_j := \max_{\pi_j \in \Pi_{\text{aff}}(\mathcal{U}^j,0)} \min_{d_{I_j} \in \mathcal{U}^j} \max_{x_{I_j}, s_{I_j}} \quad & P_j^{\text{PA}}\Big(\pi_j, d_{I_j}, x_{I_j}\Big) \\ \text{s.t.} \quad & (x_{I_j}, s_{I_j}) \in \mathcal{X}(\pi_j, d_{I_j}, 0, 0), \end{aligned}$$

where $\mathcal{P}_j$ is the j^{th} subproblem during the PA algorithm and $\widetilde{\mathcal{P}}_j$ solves a subproblem with assuming no leftover input from the previous subperiods.

Theorem 2. *For any single-station newsvendor networks with an objective value $\widetilde{f}_j^*$ of $\widetilde{\mathcal{P}}_j$,*

$$V_{\text{PA}}^* \;\leq\; V_{\text{Aff}}^* \;\leq\; V_{\text{DP}}^* \;\leq\; \widetilde{V}_{PA}^* := \sum_{j=1}^{N} \widetilde{f}_j^*.$$

Theorem 2 provides a tight bound. All the inequalities hold with equalities if Assumption 1 holds. Note that one may not need to resolve a problem $\widetilde{\mathcal{P}}_j$; recall that every subproblem in the periodic-affine algorithm is solved by generating cuts. Once a solution of $\mathcal{P}_j$ is obtained, one can relax the last constraint in $\mathcal{P}_j$ and continue cut generation in order to solve $\widetilde{\mathcal{P}}_j$. This requires less iterations, since (i) an optimal solution of $\mathcal{P}_j$ serves as a warm-start initial point for additional cuts, and (ii) the previously generated cuts are still valid without any modifications to $\widetilde{\mathcal{P}}_j$.

So far, we introduced theoretical properties of periodic-affine policies. We showed that under mild condition, the algorithm finds an optimal solution to the DP problem, and thus achieve an equal worst-case performance as an optimal affine policy for single-station problems. If this assumption does not hold, we provided a tight bound that can measure the gap between periodic and affine policies. Moreover, this gap can be computed by minimally modifying the algorithm with similar computational requirements. In the proceeding section, we extend this framework to general multi-station newsvendor networks and infinite-horizon problems.

5. Extensions of Periodic-affine Policies

We extend our approach in Section 4 to general multi-station networks, where a decision-maker intends to satisfy customers' demand at multiple locations. First, we extend Algorithm 1 for multi-station networks. Then, we develop periodic-affine policies for infinite horizon problems.

5.1. Multi-station networks

Here, we follow the flow of Section 4. To set this up, we define a matrix that plays a key role in implementing periodic-affine policies.

Definition 5 (Basic Matrix). Let the demand at each sink node $j \in \mathcal{S}$ is satisfied most profitably through an arc $\ell_j \in \mathcal{L}$. A *basis matrix* $\mathbf{R}_B \in \mathbb{R}^{p\times n}$ is given by

$$\mathbf{R}_B(\ell, j) = 1 \text{ if } \ell = \ell_j \ \forall j, \quad \text{and} \quad 0 \text{ otherwise.}$$

Using such a basis matrix, we obtain a closed-form expressions of ordering quantities and network activities if the demand is deterministic. In particular, for any $\mathbf{d} \in \mathbb{R}^n$, an optimal decision is given by $\mathbf{s} = \mathbf{R}_S\mathbf{R}_B\mathbf{d}$ and $\mathbf{x} = \mathbf{R}_B\mathbf{d}$.

Recall that we have defined a modified objective functions for each stage of the DP problem (11) to separate the overall problem into subproblems. In single-station models, the values of on-hand products and backlogged demands at the beginning of the $(j+1)^{\text{th}}$ subperiod are expressed as

$$\begin{aligned} c_S \cdot u^s_{t_j} &= c_S \cdot \sum_{t\in I_j} \Big(s_t - x_t \Big) \\ (r - c_S)\cdot u^d_{t_j} &= (r - c_S)\cdot \sum_{t\in I_j} \Big(d_t - x_t \Big), \end{aligned}$$

which are taken into j^{th} subproblem. After the j^{th} subperiod, $c_S \cdot u^s_{t_j}$ and $(r-c_S)\cdot u^d_{t_j}$ are deterministic and hence, their values can be expressed by using the basis matrix $\mathbf{R}_B$ as

$$\begin{aligned} \mathbf{c}_S^\top \mathbf{u}^{\mathbf{s}}_{t_j} &= \mathbf{c}_S^\top \sum_{t\in I_j} \Big(\mathbf{s}_t - \mathbf{R}_S\mathbf{x}_t \Big) \\ \Big(\mathbf{R}_B^\top \mathbf{r} - \mathbf{R}_B^\top \mathbf{R}_S^\top \mathbf{c}_S \Big)^\top \mathbf{u}^{\mathbf{d}}_{t_j} &= \Big(\mathbf{R}_B^\top \mathbf{r} - \mathbf{R}_B^\top \mathbf{R}_S^\top \mathbf{c}_S \Big)^\top \sum_{t\in I_j} \Big(\mathbf{d}_t - \mathbf{R}_D\mathbf{x}_t \Big). \end{aligned}$$

Note that the value of $\mathbf{u}^{\mathbf{d}}_{t_j}$ is determined by ordering $\mathbf{R}_B\mathbf{u}^{\mathbf{d}}_{t_j}$ and processing $\mathbf{R}_S\mathbf{R}_B\mathbf{u}^{\mathbf{d}}_{t_j}$. This allows us to extend the definition of affine-IBS policies as follows.

Definition 6 (Affine-IBS for Multi-station). For j^{th} subperiod, the affine-IBS policy $\overline{\boldsymbol{\pi}}_j(\mathbf{w}^{(j)}_t, \mathbf{W}^{(j)}_{\tau,t})$ associated with an affine policy $\boldsymbol{\pi}_j(\mathbf{w}^{(j)}_t, \mathbf{W}^{(j)}_{\tau,t})$ determines order quantity by

$$\mathbf{s}_t(\mathbf{u}^{\mathbf{s}}, \mathbf{u}^{\mathbf{d}}, \mathbf{d}_{I_j}) = \begin{cases} \mathbf{w}^{(j)}_1 - \mathbf{u}^{\mathbf{s}} + \mathbf{R}_S\mathbf{R}_B\mathbf{u}^{\mathbf{d}} & t = t_{j-1}+1 \\ \mathbf{w}^{(j)}_i + \displaystyle\sum_{\tau=1}^{i-1} \mathbf{W}^{(j)}_{\tau,i}\mathbf{d}_{t_{j-1}+\tau} & t = t_{j-1}+i, \ i \geq 2, \ t \in I_j. \end{cases}$$

Period-affine policy for multi-station networks. As in Definition 4, periodic-affine policies are defined as a sequence of affine-IBS policies. Eq. (12), which is required for periodic-affine policies to be well-defined, is readily extended by replacing with a vector inequality. With this generalization, all the arguments in Section 4.2 can be repeated in multi-station network setting.

As a result, the objective function for each subproblem is given by

$$\mathbf{P}_j^{\mathrm{PA}}\Big(\boldsymbol{\pi}_j, \mathbf{d}_{I_j}, \mathbf{x}_{I_j}\Big) = \begin{cases} -\mathbf{c}_H^\top \sum_{t\in I_j}\Big(\sum_{\tau=t_{j-1}+1}^{t}(\mathbf{s}_\tau - \mathbf{R}_S\mathbf{x}_\tau)\Big) \\ \quad -\mathbf{c}_P^\top \sum_{t\in I_j}\Big(\sum_{\tau=t_{j-1}+1}^{t}(\mathbf{d}_\tau - \mathbf{R}_D\mathbf{x}_\tau)\Big) + \mathbf{v}_{\mathbf{d}}^\top \sum_{t\in I_j}\mathbf{d}_t + \mathbf{v}_{\mathbf{x}}^\top \sum_{t\in I_j}\mathbf{x}_t, & j \le N-1 \\ -\mathbf{c}_S^\top \sum_{t\in I_j}\mathbf{s}_t - \mathbf{c}_H^\top \sum_{t\in I_j}\Big(\sum_{\tau=t_{j-1}+1}^{t}(\mathbf{s}_\tau - \mathbf{R}_S\mathbf{x}_\tau)\Big) \\ \quad -\mathbf{c}_P^\top \sum_{t\in I_j}\Big(\sum_{\tau=t_{j-1}+1}^{t}(\mathbf{d}_\tau - \mathbf{R}_D\mathbf{x}_\tau)\Big) + \mathbf{r}^\top \sum_{t\in I_j}\mathbf{x}_t, & j = N \end{cases} \tag{19}$$

where $\mathbf{v}_{\mathbf{d}} = \mathbf{R}_B^\top\mathbf{r} - \mathbf{R}_B^\top\mathbf{R}_S^\top\mathbf{c}_S$ and $\mathbf{v}_{\mathbf{x}} = \mathbf{r} - \mathbf{R}_S^\top\mathbf{c}_S - \mathbf{R}_D^\top\mathbf{R}_B^\top\mathbf{r} + \mathbf{R}_D^\top\mathbf{R}_B^\top\mathbf{R}_S^\top\mathbf{c}_S$.

Period-affine algorithm for multi-station networks. As in Section 4, Problem (18) can be readily converted into multidimensional form by replacing the last inequality with a vector inequality. However, it is challenging to obtain the multi-station version of (12), which computes the maximum amount of leftover resources. Therefore, we incorporate these into a single robust two-stage optimization problem, as follows:

$$\begin{aligned} \max_{\mathbf{S}_j, \boldsymbol{\pi}_j \in \boldsymbol{\Pi}_{\mathrm{aff}}(\mathcal{U}^j, \mathbf{0})} \quad & \min_{\mathbf{d}_{I_j}\in\mathcal{U}^j} \max_{\mathbf{s}_{I_j}, \mathbf{x}_{I_j}} \; \mathbf{P}_j^{\mathrm{PA}}\Big(\boldsymbol{\pi}_j, \mathbf{d}_{I_j}, \mathbf{x}_{I_j}\Big) - \delta\cdot\mathbf{1}^\top\mathbf{S}_j \\ \text{s.t.} \quad & (\mathbf{s}_{I_j}, \mathbf{x}_{I_j}) \in \mathcal{X}(\boldsymbol{\pi}_j, \mathbf{d}_{I_j}, \mathbf{0}, \mathbf{0}) \\ & \mathbf{w}_1^{(j)} \;\ge\; \mathbf{S}_{j-1}^* \\ & \sum_{t\in I_j}(\mathbf{s}_t - \mathbf{R}_S\mathbf{x}_t) \;\le\; \mathbf{S}_j \end{aligned}$$

where $\delta > 0$ is a small real number. In this way, the periodic-affine algorithm for multi-station networks proceeds by iteratively solving subproblems, similar to Algorithm 1.

Properties of periodic-affine for multi-station networks. We now generalize the theoretical properties for the multi-station networks. We consider the DP problem (11) by replacing every single-dimensional quantity by multi-dimensional quantities. Assumption 1 is extended with vector inequalities, each of which is for each source node. We use V^*_{Aff}, V^*_{PA}, and V^*_{DP} for the worst-case objective values for affine, PA, and the DP problem, and define $\widetilde{V}^*_{PA}$ similar to single-station problems.

Theorem 3. *For multi-station networks, if Assumption 1 holds, then* $V^*_{\mathrm{Aff}} \le V^*_{\mathrm{PA}} = V^*_{\mathrm{DP}}$. *Otherwise,* $V^*_{\mathrm{PA}} \;\le\; V^*_{\mathrm{DP}} \;\le\; \widetilde{V}^*_{PA}$.

Note that Proposition 5 cannot be extended to multi-station networks. In other words, PA policies for multi-station networks are not necessarily be an affine policy. Theorem 3 implies that an optimal PA policy has a worst-case performance not less than an optimal affine policy. However, for multi-station networks, we cannot compare the two policies without Assumption 1.

5.2. Infinite horizon problems

So far, PA policies are based on multi-period problems of finite horizon. In this section, we extend these PA policies to infinite horizon problems with a discount factor of $\beta < 1$. For this, we assume that nominal means and covariances of demands have periodicity with the period $k \geq 1$. We then define an uncertainty set $\mathcal{U}^k \in \mathbb{R}^{n \times k}$ to describe demand uncertainties for each period. This framework models settings where demand has stationary mean and covariance along the periods.

Definition 7 (infinite-horizon uncertainty set). The *infinite-horizon demand uncertainty set* is a Cartesian product of $\mathcal{U}^k$ via

$$\mathcal{U}^\infty = \mathcal{U}^k \times \mathcal{U}^k \times \mathcal{U}^k \times \cdots .$$

We implement PA policies for infinite horizon problems by *replicating* policies over the periods. We construct a PA policy of period k by solving a single problem of duration k. As in previous sections, we define the objective function $\mathbf{P}_\infty^{\mathrm{PA}}(\boldsymbol{\pi}, \mathbf{d}, \mathbf{x})$, by taking leftover inventories, unsatisfied demands, and the discount factor into account as

$$\begin{aligned}
&\mathbf{P}_\infty^{\mathrm{PA}}\Big(\boldsymbol{\pi}, \mathbf{d}_{[1:k]}, \mathbf{x}_{[1:k]}\Big) \\
&\quad = -\mathbf{c}_H^\top \sum_{t=1}^{k} \beta^t \Big(\sum_{\tau=1}^{t} (\mathbf{s}_\tau - \mathbf{R}_S \mathbf{x}_\tau) \Big) - \mathbf{c}_P^\top \sum_{t=1}^{k} \beta^t \Big(\sum_{\tau=1}^{t} (\mathbf{d}_\tau - \mathbf{R}_D \mathbf{x}_\tau) \Big) + \mathbf{r}^\top \sum_{t=1}^{k} \beta^t \mathbf{x}_t \\
&\qquad - \mathbf{c}_S^\top \sum_{t=1}^{k} \beta^t \mathbf{s}_t + \beta^{k+1} \mathbf{c}_S^\top \sum_{t=1}^{k} (\mathbf{s}_t - \mathbf{R}_S \mathbf{x}_t) + \beta^{k+1} \Big(\mathbf{R}_B^\top \mathbf{r} - \mathbf{R}_B^\top \mathbf{R}_S^\top \mathbf{c}_S \Big)^\top \sum_{t=1}^{k} (\mathbf{d}_t - \mathbf{R}_D \mathbf{x}_t) \\
&\quad = -\sum_{t=1}^{k} \beta^t \ \mathbf{c}_{S,t}^{*\top} \mathbf{s}_t - \sum_{t=1}^{k} \beta^t \ \mathbf{c}_{H,t}^{*\top} \Big(\sum_{\tau=1}^{t} (\mathbf{s}_\tau - \mathbf{R}_S \mathbf{x}_\tau) \Big) - \sum_{t=1}^{k} \beta^t \ \mathbf{c}_{P,t}^{*\top} \Big(\sum_{\tau=1}^{t} (\mathbf{d}_\tau - \mathbf{R}_D \mathbf{x}_\tau) \Big) + \sum_{t=1}^{k} \beta^t \ \mathbf{r}_t^{*\top} \mathbf{x}_t,
\end{aligned}$$

where $\mathbf{c}_{S,t}^* = \mathbf{c}_S$, $\mathbf{r}_t^* = \mathbf{r}$ for $t = 1, \ldots, k$, and

$$\mathbf{c}_{H,t}^* = \begin{cases} \mathbf{c}_H & 1 \leq t \leq k-1 \\ \mathbf{c}_H - \beta \mathbf{c}_S & t = k \end{cases} \qquad \mathbf{c}_{P,t}^* = \begin{cases} \mathbf{c}_P & 1 \leq t \leq k-1 \\ \mathbf{c}_P - \beta \mathbf{R}_B^\top \mathbf{r} + \beta \mathbf{R}_B^\top \mathbf{R}_S^\top \mathbf{c}_S & t = k. \end{cases}$$

As a result, an optimal PA policy is obtained by solving a single optimization problem

$$\begin{aligned}
\max_{\mathbf{S}, \boldsymbol{\pi}} \quad & \min_{\mathbf{d}_{[1:k]} \in \mathcal{U}^k} \ \max_{\mathbf{s}_{[1:k]}, \mathbf{x}_{[1:k]}} \ \mathbf{P}_\infty^{\mathrm{PA}}\Big(\boldsymbol{\pi}, \mathbf{d}_{[1:k]}, \mathbf{x}_{[1:k]}\Big) \\
\text{s.t.} \quad & (\mathbf{s}_{[1:k]}, \mathbf{x}_{[1:k]}) \in \mathcal{X}(\boldsymbol{\pi}, \mathbf{d}_{[1:k]}, \mathbf{0}, \mathbf{0}) \\
& \mathbf{w}_1 \ \geq \ \mathbf{S}
\end{aligned} \tag{20}$$

where the last constraint ensures that the solution is replicable over time periods. Based on the solution of (20), an infinite PA policy determines order quantity as

$$\mathbf{s}_t = \begin{cases} \mathbf{w}_1 - \mathbf{u}^{\mathbf{s}}_{t-1} + \mathbf{R}_S \mathbf{R}_B \mathbf{u}^{\mathbf{d}}_{t-1} & t = nk+1,\ k = 0,1,2,\ldots \\ \mathbf{w}_+ \sum_{\tau=1}^{l-1} \mathbf{W}_{\tau,l} \mathbf{d}_{nk+\tau} & t = nk+l, l > 1,\ k = 0,1,2,\ldots. \end{cases} \tag{21}$$

We next present our main result for infinite-horizon cases.

Theorem 4. *For a infinite-horizon multi-station network and the uncertainty set $\mathcal{U}^\infty$, if Assumption 1 holds, then the infinite periodic-affine policy (21) is optimal to the DP in* (11).

In summary, we generalized the periodic-affine policies into multi-station networks and infinite horizon problems. In both these cases, we presented periodic-affine algorithms and showed that the theoretical properties hold. We next discuss a numerical case study to demonstrate the practical applicability of these findings and the performance of the proposed policies.

6. Discussion: Insights and Implications

In this section, we present various implications of our modeling and solution approach and demonstrate the following advantages of our modeling approach and solution algorithm:

- *Practical Relevance*: The relevance of an approach and corresponding algorithms hinges on the ability to model features in a real-world setting and provide implementable solutions. We demonstrate that our approach is able to achieve this. In particular, we show that PA policies perform well in large-scale and data-driven environments by studying the case of a major pharmacy retailer in India. We also demonstrate that we are able to model the service level guarantees by using the robust optimization approach. This ensures that the demands are satisfied for all scenarios in an uncertainty set. We also demonstrate the robustness to mis-specification and study the performance for a spectrum of various cost parameters.
- *Generalizability and Extendability*: It is also important for the approach to be generalizable and extendable in order to accommodate higher-dimensional versions of the problem and newer types of demand information. We achieve this by modifying the uncertainty set based on the available demand information, and by showing that our approach naturally extends to multi-dimensional settings. In particular, we incorporate correlation information in computing the optimal PA policy, and demonstrate our algorithm on the high-dimensional real-world case study.
- *Computational Tractability*: An algorithm suited for real applications needs to be tractable and implementable. We demonstrate tractability of the PA policies by presenting empirical evidence on the computational times on simulated data as well as on the data from the pharmacy retailer.

Next we elaborate on each of these advantages.

6.1. Practical Relevance: Case study of a Pharmacy retailer

We analyze the sales data of a leading pharmacy retailer in India to probe the performance of the policies in a real-world setting. A common problem in forecasting demands is that sales records do not necessarily imply customers' demands, because product shortage is not reflected in sales data. However, since pharmacy retailers in India face a prohibitively large penalty for unmet demand, we can interpret the sales records as demands for this numerical study.

The data consists of more than 1.5 million transactions over 40 days (end of September to early November of 2016) for 228 different products. To reduce the problem size, we analyzed the 20 most-popular products, comprising nearly 80% of all transactions. Hierarchical clustering (Maechler et al. 2016) is used to bundle the products into groups, within which demands are highly correlated. Moving averages and residuals are extracted from the sales records, and used as nominal means and variances to define data-driven uncertainty sets for each group. Penalty and holding costs are not revealed in the sales records. Therefore, we fix penalty and holding cost rates, and compute penalty and holding costs as a product of the corresponding rates and net profit per unit.

Performance of PA policies. We compute ordering policies for each product group. We begin by assuming that the sales of different product groups are independent of each other, and later consider the more realistic case of correlated sales. We compare the performance of three policies: PA, affine policies with affine approximation (Aff-approx), and base-stock policies for these product groups. For the base-stock policies, the order-up-to level at each period is determined myopically by using the nominal means and variances, assuming normally distributed demands. On the other hand, Aff-approx assumes that the processing activities $\mathbf{x}_t$ at each time t, which are inner decision variables in the multi-period problem, are affine functions of demand $\mathbf{d}_{[1:t]}$ up to time t. That is, Aff-approx converts the two-stage problem into a static linear optimization problem, making it computationally tractable (see (Ben-Tal et al. 2005, Lorca et al. 2016) for more details).

Table 1 displays performance of these three policies for different values of the penalty cost rate. We observe that PA performs better than Aff-approx in terms of worst-case performance. In Section 6.3, we will show that this is also consistent with synthetic data. On the other hand, if the penalty cost is low, both PA and Aff-approx are not effective and the base-stock policy outperforms them. However, PA yields better lower percentile performance than the other policies for increased penalty cost. This is because PA maximizes the worst-case profit. We also observe that under significant penalty costs, PA not only protects the worst-case performance and lower percentiles (improves by 19% over base-stock at 5th percentile for cost of 10) but also leads to better average profit and historical backtesting than the other policies.

Penalty cost rate	Policy	Worst	5% quantile	25% quantile	Median	Historical
	Aff-approx	2.68	3.50	3.64	3.74	3.81
0.2	PA	2.72	3.64	3.77	3.85	3.97
	Base-stock	N/A	3.80	4.16	4.37	4.67
	Aff-approx	1.96	2.87	3.27	3.51	3.86
1.0	PA	2.49	3.30	3.41	3.48	3.56
	Base-stock	N/A	3.27	3.66	3.96	4.25
	Aff-approx	1.49	3.31	3.60	3.79	4.04
5.0	PA	1.78	3.37	3.64	3.82	3.83
	Base-stock	N/A	2.85	3.27	3.58	3.68
	Aff-approx	1.26	3.10	3.45	3.68	4.15
10.0	PA	1.39	3.28	3.62	3.84	3.77
	Base-stock	N/A	2.76	3.19	3.49	3.40

Note: all values are multiplied by 10^6.

Table 1 **Performance of policies for different penalty cost rates. Percentiles and medians are calculated from 1000 samples. The last column is from the historical sales data.**

Robustness to model misspecification. Given that all these policies are implemented using the nominal mean and variance inferred from past records, it is important to measure their robustness to errors in model calibration. For this, we consider demand realizations to have mean greater than (Figure 2a), same as (Figure 2b), or less than (Figure 2c) their nominal values for varying holding and penalty costs. We observe that when the realized demand distribution differs from

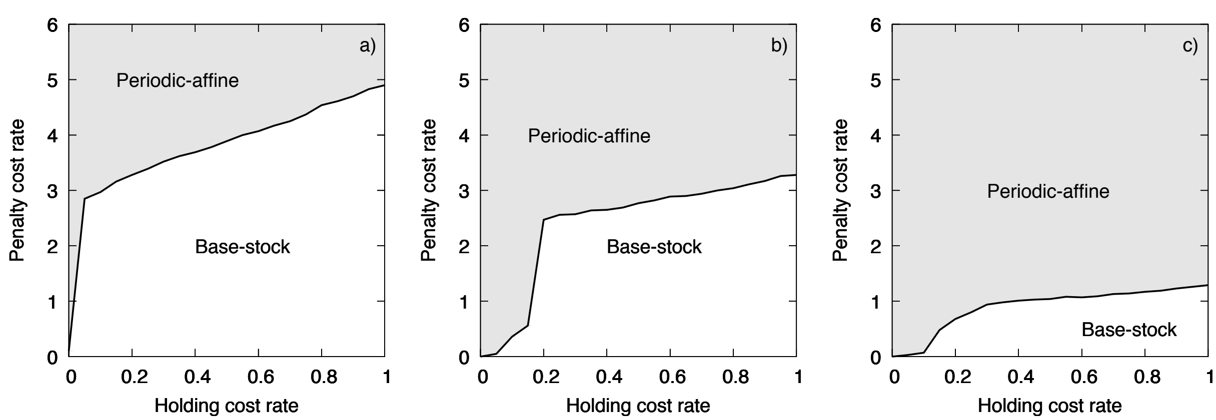

Figure 2 **Phase diagram of periodic-affine and the base-stock policies. The realized demand means are (a) increased by 5%, (b) not changed, and (c) decreased by 5% from the nominal values.**

the assumed one, the region of parameters (*phase*), for which PA outperforms base-stock policy changes. For example for a holding cost of 0.1 and penalty of 1, the PA policy outperforms base-stock, if the mean of the assumed demand coincides with the realized one (see Figure 2b). However, only 5% increase of the means is sufficient for the base-stock to prevail (see Figure 2a).

Performance dependence on holding and penalty cost. As the costs vary, we observe a *phase transition* between a phase where PA outperforms the base-stock policies and a phase in which the base-stock policy outperforms. The phase diagrams in Figure 2 allow the decision maker to select the policy based on the given cost and demand structure. In fact, for pharmacy retailers, who face a substantial penalty with unsatisfied demand, it shows that our proposed PA policy is preferable. On the other hand, if the decision maker is committed to a certain ordering policy (e.g. contractually), the phase diagrams in Figure 2 can suggest suitable changes to the cost structure in order to make the policy superior.

Impact of high penalty cost. When analyzing backlogged demands and inputs for different values of the penalty cost, Figure 3 shows that the three policies react differently for high penalty costs. First, the base-stock policy does not effectively control the backlogged demands. Although the penalty cost is accounted for in the newsvendor quantile to avoid high backlogs, increasing it slightly decreases the amount of backlogged demands. On the other hand, Aff-approx determines order quantities more conservatively than PA. Under high penalty cost, Aff-approx satisfies nearly all customer demands by ordering an excessive amount of input. This causes a significantly larger holding cost, and leads to less profit than PA. Finally, PA controls both leftover input and backlogged demands. As the penalty cost increases, PA not only reduces backlogged demands (same as Aff-approx), but also maintains much lower input levels than Aff-approx. This leads to a higher profit than the other policies.

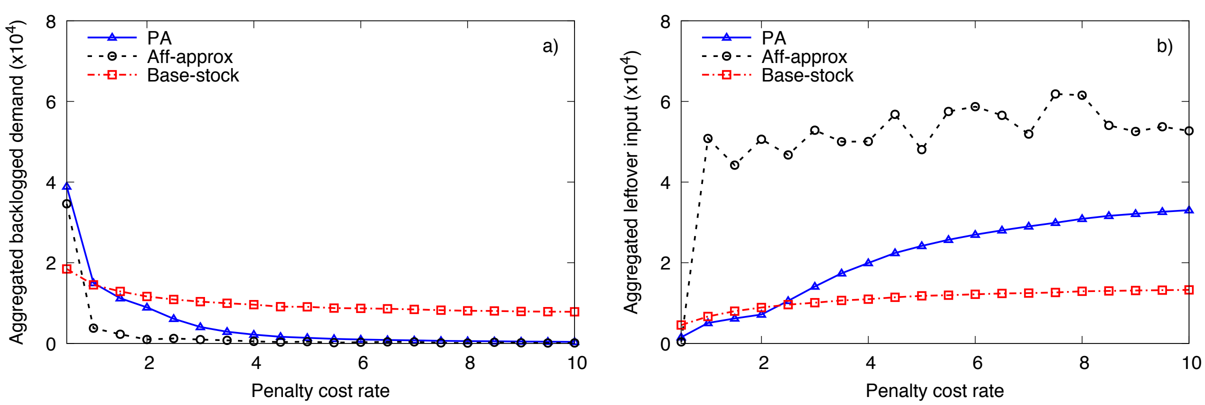

Figure 3 Impact of penalty cost on (a) backlogged demands, and (b) leftover resources on random samples.

Summary. In summary, the case-study with pharmacy retailer's data demonstrates significant increase in performance (about 17%) for the periodic-affine policies. It also allows decision makers to identify the optimal policy based on their respective cost and demand structures.

6.2. Generalizability and Extentability: Modeling Correlation and Solving multi-station newsvendor problems

To adequately discuss the performance of PA under a multi-station setting, we take demand correlation information into account. We model correlations in demand using the *correlated uncertainty sets* presented in (3) by using a factor model approach. For the numerical analysis, we consider two products over 15 time periods with subperiods of length 5.

The benefits of modeling correlation become apparent, when comparing the following two policies: multi-station PA (PA-single) for each product using marginal mean and variance, and multi-station PA (PA-multi) using the correlated uncertainty set. For comparison, we compute the *relative performance* (RP) of PA-multi over PA-single as

$$\text{RP} \; = \; \frac{\text{profit of PA-multi} - \text{profit of PA-single}}{\text{profit of PA-single}}.$$

After the two policies are implemented, we generate random demand with nominal mean and covariance and evaluate the relative performance for each sample. Figure 4 displays this relative performance for different correlation coefficient ρ. We observe that the median RP for every ρ is positive. However, the behavior differs for positively or negatively correlated products. While for

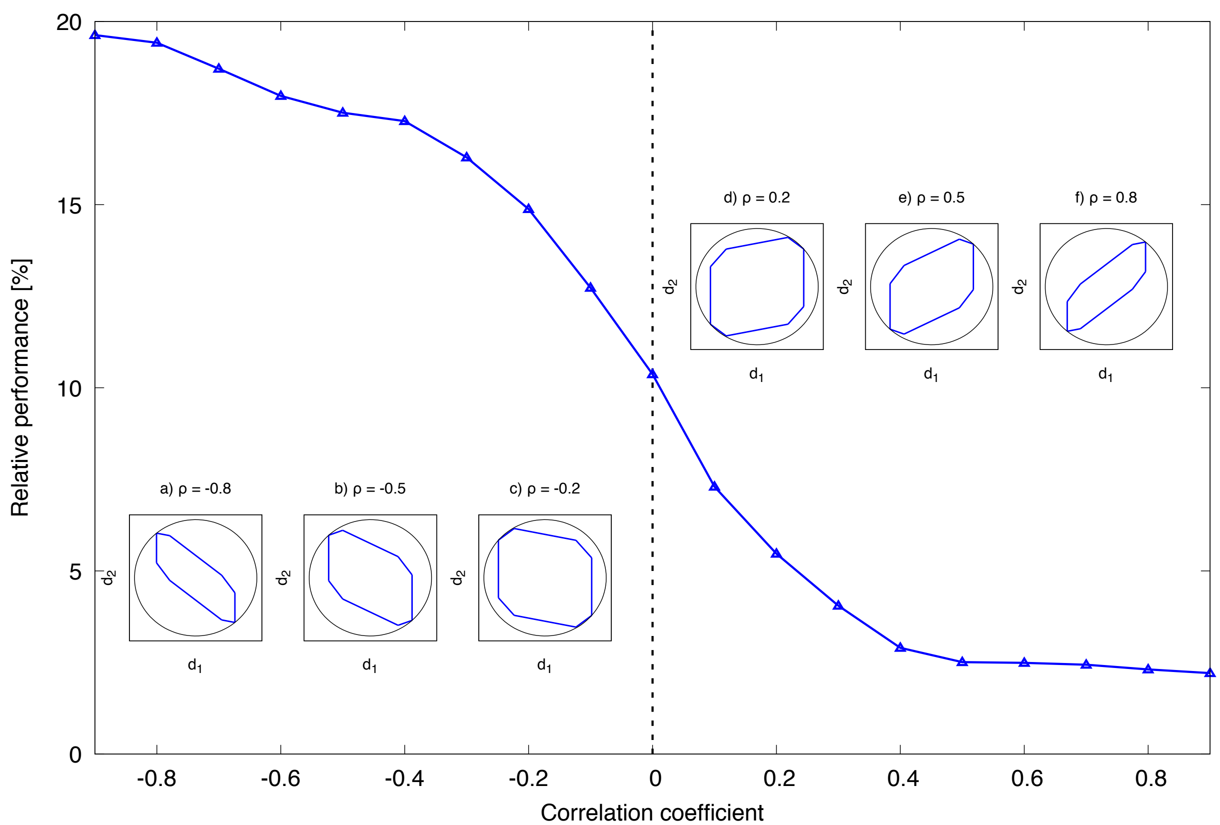

Figure 4 **Relative performance for different correlations ρ. Inserts are the corresponding uncertainty sets.**

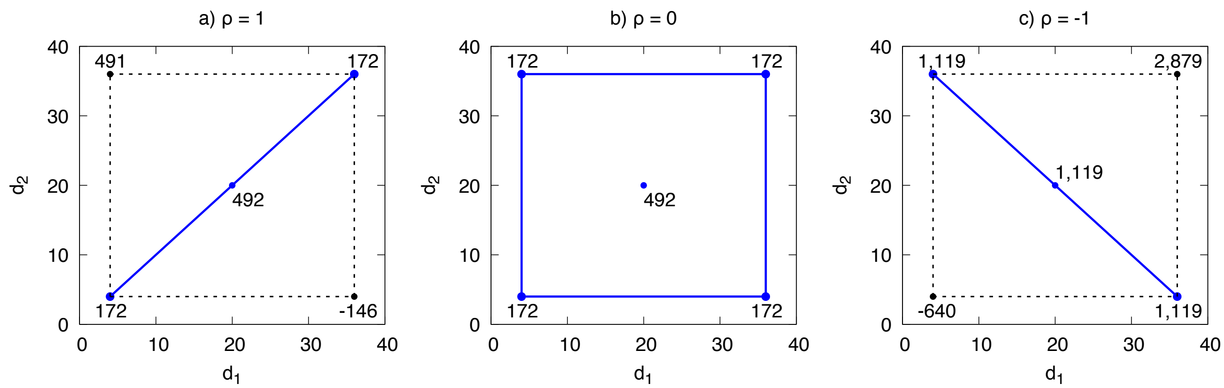

Figure 5 An illustrative example for correlation and multi-station modeling: Two products are (a) perfectly correlated, (b) uncorrelated, and (c) perfectly anti-correlated. Uncertainty sets are defined as within the blue contours along with the profit of optimal policies for extreme and nominal points.

highly correlated products the RP slightly decreases with growing ρ, significant improvements are made for negatively correlated products. In fact, more RP increases by more than 17%, when the products have a strong negative correlation ($\rho = -0.9$).

This observation can be interpreted by the structure of the uncertainty sets, as shown in the inserts of Figure 4. Positively correlated products lead to sets that allow both uncertain demands d_1 and d_2 to be concurrently at their maximum or minimum value. As ρ decreases, the area of the polyhedron shrinks, however the extreme points are unaffected. However, when ρ becomes negative, i.e. the products are negatively correlated, if one of the uncertain demands can take its maximum value, the other is forced to its lowest, and vice versa. This effect forces an increase in RP as $\rho \to -1$. The extreme cases are illustrated in the example of Figure 5. For perfectly correlated products, Figure 5a shows that even though the uncertainty set is dramatically shrunk, the worst-case profit cannot not improved over the uncorrelated demand setting (Figure 5b), because the worst-case is often captured when both demands are high or low. However, for negatively correlated demands, the uncertainty set does not contain this region (high/high or low/low), allowing for substantial improvement in worst-case profit, as shown in Figure 5c.

Case study of a pharmacy retailer – revisited. Here, we account for correlation amongst the product groups and compare the following five policies in Table 2: PA-multi, PA-single, Aff-approx with correlated uncertainty set (Aff-approx-multi), Aff-approx for each group (Aff-approx-single), and the myopic base-stock policy. For lower quantiles, we observe that the base-stock policy performs poorly compared to the single-station models. Both PA-single and PA-multi yield significantly greater profit in lower quantiles than the base-stock policy. The multi-station framework and correlated demand uncertainty sets offer better performance than single-station framework. In particular, PA-multi achieves at least 7% more profit than PA-single for moderate choice of the

Penalty cost rate	Policy	Worst	5% quantile	25% quantile	Median	Historical
0.2	Aff-approx-single	6.20	9.09	9.65	10.03	10.60
	Aff-approx-multi	7.03	8.65	9.72	10.31	9.51
	PA-single	6.32	8.59	8.97	9.21	9.66
	PA-multi	7.17	9.24	10.24	10.71	11.12
	Base-stock	N/A	8.50	9.97	10.74	10.81
1.0	Aff-approx-single	4.07	6.44	7.59	8.15	8.39
	Aff-approx-multi	5.29	8.22	8.83	9.34	10.02
	PA-single	5.69	7.76	8.07	8.31	8.97
	PA-multi	6.64	8.65	9.30	9.70	10.54
	Base-stock	N/A	6.70	8.10	9.21	8.98
5.0	Aff-approx-single	2.63	5.19	6.82	7.90	7.21
	Aff-approx-multi	4.20	7.44	8.44	9.01	9.37
	PA-single	3.44	7.57	8.33	8.77	9.75
	PA-multi	4.95	7.82	8.73	9.43	9.25
	Base-stock	N/A	5.04	6.61	7.76	7.03
10.0	Aff-approx-single	1.69	5.06	6.67	7.64	6.88
	Aff-approx-multi	3.46	7.07	8.17	8.85	9.88
	PA-single	2.11	7.26	8.24	8.81	9.01
	PA-multi	3.98	7.04	8.25	8.98	8.87
	Base-stock	N/A	4.61	6.27	7.40	6.42

Note: all values are multiplied by 10^6.

Table 2 Performance of policies for two correlated product groups for different penalty cost rates. Percentiles and medians are calculated from 1000 samples. The last column is from the historical sales data.

penalty cost. However, for extremely high penalty cost rates (e.g. 10), PA-single performs slightly better than PA-multi for lower quantiles, even though the worst case objective value is lower. This is due to samples that are generated outside of the correlated uncertainty sets.

Summary. In summary, Table 2 demonstrates that capturing correlation using the periodic-affine policies in multi-station setting outperforms the base-stock policies for moderate penalties. Since the demand of the studied pharmaceutical product groups is positively correlated, variability of the base-stock policy increases, causing a sizable degradation of the profit when compared to periodic-affine policies.

6.3. Computational Tractability

In order to focus on computational performance in a sterile environment, we consider the following simulation environment. We simulate three cases with duration $T \in \{10, 15, 20\}$ with a subperiod consisting of 5 time periods. We randomly generate 100 instances of single-station newsvendor problems for each T. Nominal means are generated by autoregression processes AR(1) and nominal coefficient of variations are chosen uniformly in (0.3, 0.5). Unless modified, cost parameters are $c_S = 20$, $r = 120$, $c_H = c_P = 20$, and all variability parameters are set to 2. We then compare the PA policies with affine policies (Aff-exact) and affine policies with affine approximation (Aff-approx). Aff-exact finds the optimal parameters by solving the multi-period model (6–8) with cut generation (see Section 3).

For the three policies, Table 3 displays the computation times and worst objective values on the same uncertainty set. We observe that the computation of PA is significantly faster than Aff-exact, because PA is tractable. However, the worst objective values of PA are very close (within 0.1%) to Aff-exact, while Aff-approx consistently deviates by $\geq 10\%$ from the others. Indeed, only 13 out of the 300 artificial instances have greater worst objective values in Aff-exact. Moreover, there is only one instance in which PA loses more than 1% of optimality. This implies that Assumption 1, which is a sufficient optimality condition of PA, holds for fairly general settings. This means that PA is as competitive as Aff-exact in worst-case values.

Policy	Aff-exact			PA			Aff-approx		
Time periods T	10	15	20	10	15	20	10	15	20
Computation time [sec]	3.8	88.5	1388.3	0.18	0.24	0.29	0.02	0.09	0.22
Worst objective value	11,035	16,833	22,440	11,033	16,831	22,438	9,523	14,601	19,479
Difference to Aff-exact [%]				-0.015	-0.011	-0.006	-13.39	-13.03	-12.96

Table 3 **Average performance of three policies on randomly generated instances.**

Comparison of PA and Affine policies on synthetic data. We next compare the performance of PA and affine policies for a spectrum of parameters. Figure 6 shows that the gap between the two policies are different for holding and for penalty cost. Figure 6a shows that the gap between PA and Aff-approx decreases as holding cost increases. However, PA protects the worst-case profit significantly better than Aff-approx as penalty cost increases, shown in Figure 6b. These results

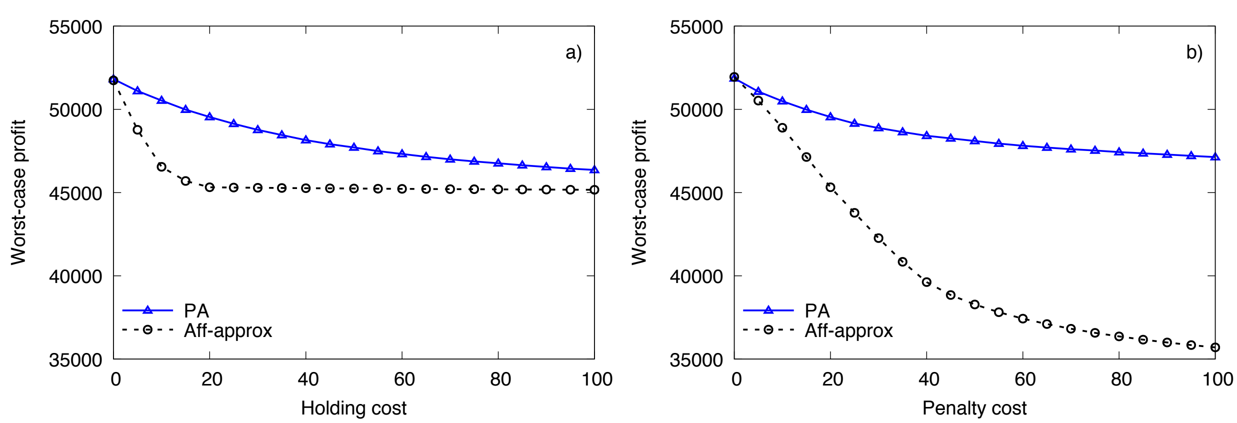

Figure 6 **Impact of (a) holding cost, and (b) penalty cost on worst-case performance.**

suggest that depending of the holding and penalty costs, PA orders the proper amount of input to meet demands, i.e., both on-hand input and unsatisfied demands are well-controlled. Therefore, its worst-case performance is far less affected by larger cost parameters than Aff-approx. On the other hand, Figure 6b indicates that Aff-approx may not manage the leftover resources and backlogged demands as well as PA, which yields poor performance for higher penalty cost.

The significant relative decrease of worst objective values between the two policies of up to 35%, as shown in Figure 6 and also observed in Table 3, implies that the degree of suboptimality of Aff-approx may render it inferior to PA. Although theoretical bounds of affine approximation have been proposed (Bertsimas and Goyal 2012, Bertsimas and Bidkhori 2015), these results indicates that the affine approximation cannot be successfully applied to general settings, despite its computational advantage of tractability.

Summary. In summary, the experiment on artificial data reveals the optimality and tractability of periodic-affine policies, while the other methods did not achieve both properties. In addition, we observe that periodic-affine policies perform substantially better for larger T than Aff-approx in terms of protecting its worst-case performance at high penalty cost. This implies that PA is particularly useful when a decision maker faces a massive penalty for unsatisfied demands, such as in the case of the pharmacy retailer in India.

7. Conclusions

In this paper, we consider the problem of optimal control in multi-period and multi-stage newsvendor networks. To this end, we introduce a new class of adaptive policies called *periodic-affine policies*. These policies are data-driven and incorporate the correlation amongst products, which is an instrumental feature of real-world settings. These policies also remain robust to parameter mis-specification. For this, we model the uncertain demand via sets, which can incorporate correlations, and can be generalized to multi-product settings and extended to multi-period problems. This approach offers a natural framework to study current competing policies of base-stock, affine, and approximative approaches with respect to their profit, sensitivity to parameters and assumptions, and computational scalability. We showed that the periodic-affine policies are sustainable, i.e. time consistent, because they warrant optimality both within subperiods and over the entire planning horizon.

We presented empirical evidence of tractability and robustness which makes our approach well-suited for real-world applications. We demonstrate the advantages of our approach by considering the problem of one of India's largest pharmacy retailers using their sales data. We show that the periodic-affine policies are capable to increase the profits by up to 17% over base-stock policies. This study reveals capturing the demand correlation can sizably affect the performance. Furthermore, we offer a phase diagram for managers to select the optimal policy based on their cost and demand structures.

In future, we intend to incorporate time-dependent uncertainty sets to more accurately model seasonal demand. This step forward will lend itself well to incorporates returns, i.e. feedback from satisfied demand that can guide the next period's decisions.

Acknowledgments

We are grateful to Jan A. Van Mieghem for insightful comments.

References

Arrow K, Harris T, Marschak J (1951) Optimal inventory policy. *Econometrica* XIX:250–272.

Baboli A, Fondrevelle J, Tavakkoli-Moghaddam R, Mehrabi A (2011) A replenishment policy based on joint optimization in a downstream pharmaceutical supply chain: Centralized vs. decentralized replenishment. *Internat. J. Adv. Manufacturing Tech* 57.

Bandi C, Bertsimas D (2012) Tractable stochastic analysis in high dimensions via robust optimization. *Math. Programming* 134(1):23–70.

Ben-Tal A, Golany B, Nemirovski A, Vial J (2005) Retailer-supplier flexible commitments contracts: a robust optimization approach. *Manufacturing & Service Oper. Management* 7(3):248–271.

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Math. Programming* 99(2):351–376.

Bertsimas D, Bidkhori H (2015) On the performance of affine policies for two-stage adaptive optimization: a geometric perspective. *Math. Programming* 153(2):577–594.

Bertsimas D, Goyal V (2012) On the power and limitations of affine policies in two-stage adaptive optimization. *Math. Programming* 134(2):491–531.

Bertsimas D, Iancu D, Parillo P (2010) Optimality of affine policies in multistage robust optimization. *Math. Oper. Res.* 35(2):363–394.

Bertsimas D, Sim M (2004) The price of robustness. *Oper. Res.* 52(1):35–53.

Bertsimas D, Thiele A (2006) Robust and data-driven optimization: Modern decision making under uncertainty. *Tutorials in Oper. Res.* 95–122.

Bienstock D, Özbay N (2008) Computing robust base-stock levels. *Discrete Optim.* 5(2):389–414.

Clark A, Scarf H (1960) Optimal policies for a multi-echelon inventory problem. *Management Sci.* 6:475–490.

Crawford GS, Shum M (2005) Uncertainty and learning in pharmaceutical demand. *Econometrica* 73(4):1137–1173.

Federgruen A, Zipkin P (1986) An inventory model with limited production capacity and uncertain demands, i: The average cost criterion. *Math. Oper. Res.* 11:193–207.

Fu M (1994) Sample path derivatives for (s, s) inventory systems. *Oper. Res.* 42:351–364.

Fu M, Glover F, April J (2005) Simulation optimization: a review, new developments, and applications. *Proc. of the 2005 Winter Simulation Conf.*, 83–95.

Fu M, Healy K (1997) Techniques for simulation optimization: An experimental study on an (s, S) inventory system. *IIE Trans.* 29(3):191–199.

Gallego G, Moon I (1993) The distribution free newsboy problem: Review and extensions. *J. Oper. Res. Soc.* 44:825–834.

Glasserman P, Tayur S (1995) Sensitivity analysis for base-stock levels in multiechelon production-inventory systems. *Management Sci.* 41(2):263–281.

Graves S, Willems S (2000) Optimizing strategic safety-stock placement in supply chains. *Manufacturing & Service Oper. Management* 2(1):68–83.

Guerrero WJ, Yeung T, Guéret C (2013) Joint-optimization of inventory policies on a multi-product multi-echelon pharmaceutical system with batching and ordering constraints. *European Journal of Operational Research* 231(1):98–108.

Huh W, Janakiraman G (2008) A sample-path approach to the optimality of echelon order-up-to policies in serial inventory systems. *Oper. Res. Lett.* 36(5):547–550.

Iyengar GN (2005) Robust dynamic programming. *Math. Oper. Res.* 30(2):257–280.

Kapuscinski R, Tayur S (1999) Optimal policies and simulation-based optimization for capacitated production inventory systems. *Quantitative Models for Supply Chain Management*, 7–40 (Springer).

Karlin S (1960) Dynamic inventory policy with varying stochastic demands. *Management Sci.* 6:231–258.

Kasugai H, Kasegai T (1961) Note on minimax regret ordering policy – static and dynamic solutions and a comparison to maximin policy. *J. Oper. Res. Soc. of Japan* 3:155–169.

Kerrigan E, Maciejowski J (2004) Properties of a new parametrization for the control of constrained systems with disturbances. *Proc. of the 2004 Amer. Control Conf.* 5:4669–4674.

Kuhn D, Wiesmann W, Georghiou A (2011) Primal and dual linear decision rules in stochastic and robust optimization. *Math. Programming* 130(1):177–209.

Langenhoff L, Zijm W (1990) An analytical theory of multi-echelon production/distribution systems. *Statistica Neerlandica* 44(3):149–174.

Löfberg J (2003) Approximations of closed-loop minimax mpc. *Proc. of the 42nd IEEE Conf. on Decision Control* 2:1438–1442.

Lorca A, Sun XA, Litvinov E, Zheng T (2016) Multistage adaptive robust optimization for the unit commitment problem. *Oper. Res.* 64(1):32–51.

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K (2016) *cluster: Cluster Analysis Basics and Extensions*. R package version 2.0.5.

Morton T (1978) The non-stationary infinite horizon inventory problem. *Management Sci.* 24:1474–1482.

Muharremoglu A, Tsitsiklis J (2008) A singe-unit decomposition approach to multi-echelon inventory systems. *Oper. Res.* 56:1089–1103.

Nohadani O, Roy A (2017) Robust optimization with time-dependent uncertainty in radiation therapy. *IISE Trans. Healthcare Systems Engineering* 7(2):81–92.

Rikun A (2011) *Applications of robust optimization to queueing and inventory systems.* Ph.D. thesis, Massachusetts Institute of Technology.

Rosling K (1989) Optimal inventory policies for assembly systems under random demands. *Oper. Res.* 37:565–579.

Scarf H (1958) *Studies in the Mathematical Theory of Inventory and Production (eds. K.J. Arrow and S. Karlin and H. Scarf)*, chapter A Min-Max Solution of An Inventory Problems (Stanford University Press, Stanford, CA).

Scarf H (1960) *Mathematical Methods in the Social Sciences (eds. K.J. Arrow and S. Karlin and P. Suppes)*, chapter The Optimality of (s,S) policies in the dynamic inventory problem (Stanford University Press, Stanford, CA).

Sethi S, Cheng F (1997) Optimality of (s,S) policies in inventory models with markovian demand. *Oper. Res.* 45(6):931–939.

Uthayakumar R, Priyan S (2013) Pharmaceutical supply chain and inventory management strategies: optimization for a pharmaceutical company and a hospital. *Operations Research for Health Care* 2(3):52–64.

Van Mieghem JA, Rudi N (2002) Newsvendor networks: Inventory management and capacity investment with discretionary activities. *Manufacturing & Service Oper. Management* 4(4):313–335.

Appendix 1. Proofs of Auxilliary Results

Proof of Proposition 1. First suppose that the nominal mean $\boldsymbol{\mu}$ increases by $\boldsymbol{\Delta\mu} \geq \mathbf{0}$. Then there exists $\boldsymbol{\Delta}\mathbf{s}^* \geq \mathbf{0}$ and $\boldsymbol{\Delta}\mathbf{x}^* \geq \mathbf{0}$ that solves the nominal problem with deterministic demand $\boldsymbol{\Delta\mu}$, with non-negative profit $p \geq 0$. Thus the worst-case profit with $\boldsymbol{\mu} + \boldsymbol{\Delta\mu}$ increases at least by p. On the other hand, one can show that the set $\mathcal{U}$ increases in any of $\lambda_1, \ldots, \lambda_l$, followed by that the objective value decreases. $\square$

Proof of Proposition 2. For fixed $\{\mathbf{w}_t, \mathbf{W}_{\tau,t}\}$ and $\mathbf{d}_{[1:T]}$, the inner maximization in problem (6) with respect to $\mathbf{X}_{[1:T]}$ is a linear program in which $\{\mathbf{w}_t, \mathbf{W}_{\tau,t}\}$ are on the right-hand side. It follows that for any $\mathbf{d}_{[1:T]}$, the inner maximization problem (6) is concave in $\{\mathbf{w}_t, \mathbf{W}_{\tau,t}\}$. Hence the objective function in (6) is concave as well, because it is a pointwise infimum of concave functions. Moreover, the problem is always feasible with assigning zero vectors and matrices to $\{\mathbf{w}_t, \mathbf{W}_{\tau,t}\}$. Finally, applying strong duality to constraint (7) shows that a feasible set of $\{\mathbf{w}_t, \mathbf{W}_{\tau,t}\}$ is a polyhedron and hence, convex. $\square$

Proof of Proposition 3. $\Phi(s_0,d_0) := \max_{\pi} V\Big(\pi(w_t,W_{\tau,t});s_0,d_0\Big)$, where $V\Big(\pi(w_t,W_{\tau,t});s_0,d_0\Big)$ is defined as

$$\begin{aligned}
V\Big(\pi(w_t,W_{\tau,t});s_0,d_0\Big) &:= \min_{d_{[1:T]}\in\mathcal{U}} \max_{x_{[1:T]},s_{[1:T]}} P\Big(\pi(w_t,W_{\tau,t}),d_{[1:T]},x_{[1:T]};s_0,d_0\Big)\\
&= \min_{d_{[1:T]}\in\mathcal{U}} \max_{x_{[1:T]},s_{[1:T]}} \Bigg[-c_S\Bigg(\sum_{t=1}^{T}s_t\Bigg) - c_H\sum_{t=1}^{T}\Bigg(s_0+\sum_{\tau=1}^{t}(s_\tau-x_\tau)\Bigg)\\
&\qquad\qquad -c_P\sum_{t=1}^{T}\Bigg(d_0+\sum_{\tau=1}^{t}(d_\tau-x_\tau)\Bigg)+r\Bigg(\sum_{t=1}^{T}x_t\Bigg)\Bigg],
\end{aligned}$$

where the inner maximization problem has a feasible set $\mathcal{X}(\pi,d_{[1:T]},s_0,d_0)$.

We first claim that the following equation holds for any fully-affine policy $\pi(w_t,W_{\tau,t})$ such that $s_0\le w_1$,

$$V\Big(\overline{\pi}(\overline{w}_t,\overline{W}_{\tau,t});s_0,d_0\Big) = V\Big(\pi(w_t,W_{\tau,t});0,0\Big)+c_Ss_0+(r-c_S)d_0, \tag{22}$$

where $\overline{\pi}(\overline{w}_t,\overline{W}_{\tau,t})$ is defined as (10). For any demand realization $d_{[1:T]}\in\mathcal{U}$, let $S_t=S_t(\pi,d_{[1:T]})$ and $X_t=X_t(\pi,d_{[1:T]})$ be an aggregated order quantity and the corresponding optimal processing activity at time t with zero initial input and demand, i.e., it solves the inner maximization problem of $V(\pi;0,0)$ with $d_{[1:T]}$. Let $\overline{S}_t=\overline{S}_t(\overline{\pi},d_{[1:T]})$ be an order quantity by the policy $\overline{\pi}$ for $V(\overline{\pi};s_0,d_0)$, and define $\overline{X}_t=\overline{X}_t(d_{[1:T]})=X_t(d_{[1:T]})+d_0$. Then by (10), $\overline{S}_t=S_t-s_0+d_0$ for every $t=1,\dots,T$, and thus

$$\begin{aligned}
&P\Big(\pi(w_t,W_{\tau,t}),d_{[1:T]},x_{[1:T]};0,0\Big)+c_Ss_0+(r-c_S)d_0\\
&\quad=\Big(-c_SS_T-c_H\sum_{t=1}^{T}(S_t-X_t)-c_P\sum_{t=1}^{T}(D_t-X_t)+rX_T\Big)+(r-c_S)d_0+c_Ss_0\\
&\quad=-c_S(S_T-s_0+d_0)-c_H\sum_{t=1}^{T}(S_t+d_0)-c_P\sum_{t=1}^{T}(d_0+D_t)+(c_H+c_P)\sum_{t=1}^{T}(X_t+d_0)+r(X_t+d_0)\\
&\quad=-c_S\overline{S}_T-c_H\sum_{t=1}^{T}(s_0+\overline{S}_t-\overline{X}_t)-c_P\sum_{t=1}^{T}(d_0+D_t-\overline{X}_t)+r\overline{X}_T\\
&\quad=P\Big(\overline{\pi}(\overline{w}_t,\overline{W}_{\tau,t}),d_{[1:T]},\overline{x}_{[1:T]};s_0,d_0\Big).
\end{aligned}$$

Since $\overline{X}_t(d_{[1:T]})$ is a feasible solution of the inner maximization problem in $V(\overline{\pi};s_0,d_0)$, the LHS of (22) is greater than or equal to the RHS. Similar argument can be made to show the contrary and this concludes the proof of (22).

Now suppose $\overline{\pi}^*(\overline{w}_t^*,\overline{W}_{\tau,t}^*)$ is not optimal, and let $\overline{\varphi}^*(\overline{v}_t^*,\overline{V}_{\tau,t}^*)$ be an optimal solution of $\Phi(s_0,d_0)$. Now one can easily check that by (22),

$$\begin{aligned}
\Phi(s_0,d_0)=V(\overline{\varphi}^*;s_0,d_0)&=V(\varphi^*;0,0)+c_Ss_0+(r-c_S)d_0\\
&\le V(\pi^*;0,0)+c_Ss_0+(r-c_S)d_0=V(\overline{\pi}^*;s_0,d_0)<\Phi(s_0,d_0),
\end{aligned}$$

which makes a contradiction. Finally, the proof of Eq. (22) directly shows that both $\Phi(0,0)$ and $\Phi(s_0,d_0)$ shares a common worst-case sceneario among $\mathcal{U}$. $\square$

Proof of Proposition 4. Let $\overline{\pi}_{\text{PA}} = (\overline{\pi}_1,\ldots,\overline{\pi}_N)$, where each $\overline{\pi}_j$ solves the j^{th} subproblem. Then the worst-case scenario $(d^*_{I_1},\ldots,d^*_{I_N})$ solves an optimization problem

$$\min_{d_{I_1},\ldots,d_{I_N}} \max_{x_{I_1},\ldots,x_{I_N}} \left[\sum_{j=1}^N P_j(\overline{\pi}_j, d_{I_j}, x_{I_j}; u^s_{t_{j-1}}, u^d_{t_{j-1}})\right]. \tag{23}$$

Since $\overline{\pi}_{\text{PA}}$ is a well-defined periodic-affine policy, (23) is rewritten as

$$\begin{aligned}
&\min_{d_{I_1},\ldots,d_{I_N}} \max_{x_{I_1},\ldots,x_{I_N}} \left[\sum_{j=1}^N P_j(\overline{\pi}_j, d_{I_j}, x_{I_j}; u^s_{t_{j-1}}, u^d_{t_{j-1}})\right] \\
&= \min_{d_{I_1},\ldots,d_{I_N}} \max_{x_{I_1},\ldots,x_{I_N}} \left[\sum_{j=1}^N \widetilde{P}_j(\overline{\pi}_j, d_{I_j}, x_{I_j}; u^s_{t_{j-1}}, u^d_{t_{j-1}})\right] \\
&= \min_{d_{I_1},\ldots,d_{I_N}} \max_{x_{I_1},\ldots,x_{I_N}} \left[\sum_{j=1}^N P^{\text{PA}}_j(\overline{\pi}_j, d_{I_j}, x_{I_j})\right] \\
&= \sum_{j=1}^N \left[\min_{d_{I_j}\in\mathcal{U}^j} \max_{x_{I_j}} P^{\text{PA}}_j(\overline{\pi}_j, d_{I_j}, x_{I_j})\right] \\
&= \sum_{j=1}^N P^{\text{PA}}_j(\overline{\pi}_j, d^*_{I_j}, x_{I_j}),
\end{aligned}$$

where $\widetilde{P}_j$ and P^{PA}_j are defined in (16) and (17). As a result, the overall objective function is separable for each subproblem, and hence, the worst-case scenario consists of those of the subperiods. $\square$

Proof of Theorem 1. Since every fully-affine policy is feasible to the DP problem (11), proving $V^*_{\text{FA}} \le V^*_{\text{DP}}$ directly follows. Thus it suffices to show $V^*_{\text{PA}} = V^*_{\text{DP}}$, and let $\overline{\pi}_{\text{PA}} = (\overline{\pi}_1,\ldots,\overline{\pi}_N)$ be an output of the periodic-affine algorithm. That is, $\overline{\pi}_j$ is an affine-IBS policy associated with $\pi_j = \pi_j(w^{(j)}_t, W^{(j)}_{\tau,t})$, where π_j solves (18) for each $j=1,\ldots,N$. Define $V_j(u^s,u^d)$ as a worst-case optimal profit from the j^{th} subperiod to the last period, where u^s and u^d are current on-hand input and backlogged demand. Our framework justifies using the (robust) optimality equation and a (worst-case) value function approach in robust dynamic programming scheme; we refer Iyengar (2005) to readers for technical details. We will show for every $j=1,\ldots,N$,

(a) $V_j(u^s,u^d)$ is concave in (u^s,u^d), and
(b) $V_j(u^s,u^d) = V_j(0,0) + c_S u^s + (r-c_S)u^d$ for every $0 \le u^s \le w^{(j)}_1$, $u^d \ge 0$.

by mathematical induction.

Now consider $j=N$ and suppose $u^s \le w^{(N)}_1$. Then for fixed $d_{I_N} \in \mathcal{U}^N$, $V_N(u^s,u^d)$ can be written as

$$\begin{aligned}
V_N(u^s,u^d) := \max_{\pi\in\Pi_{\text{aff}}(\mathcal{U}^N)} \min_{d_{I_N}} \max_{x_{I_N},s_{I_N}} \quad & P_N(\pi, d_{I_N}, x_{I_N}; u^s, u^d) \\
\text{s.t.} \quad & (x_{I_N}, s_{I_N}) \in \mathcal{X}(\pi, d_{I_N}, u^s, u^d),
\end{aligned}$$

where $\pi = \pi(w_t, W_{\tau,t})$ and the constraints in $\mathcal{X}(\pi, d_{I_N}, u^s, u^d)$ can be rearranged so that the right hand sides are linear in $(u^s, u^d, w_t, W_{\tau,t})$. Since P_N is concave in $(x_{I_N}, w_t, W_{\tau,t}, u^s, u^d)$ and $\mathcal{X}(\pi, d_{I_N}, u^s, u^d)$ defines a polyhedron for any π, u^s, and u^d, the objective function within the min operator is concave in $(u^s, u^d, w_t, W_{\tau,t})$ by concavity preservation under maximization. Since a pointwise infimum of concave functions are concave and applying concavity preservation under maximization again to the outermost max operator, we finally have that $V_N(u^s, u^d)$ is concave in (u^s, u^d). On the other hand, (b) follows directly from Proposition 3 for $j = N$.

Now suppose that both (a) and (b) hold for any $1 < j \le N$, and let $u^s \le w_1^{(j-1)}$ and $u^d \ge 0$. Then from the optimality equation, we have

$$V_{j-1}(u^s, u^d) = \max_{\pi \in \Pi_{\text{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}} \in \mathcal{P}(\mathcal{U}^{j-1})} \left[P_{j-1}\Big(\pi, d_{I_{j-1}}, x_{I_{j-1}}; u^s, u^d\Big) + V_j(u_j^s, u_j^d) \right]. \tag{24}$$

Since $V_j(u_j^s, u_j^d)$ is concave in (u_j^s, u_j^d) and (u_j^s, u_j^d) can be expressed as an affine function of $(w_t, W_{\tau,t}, u^s, u^d)$, applying the above argument shows that $V_{j-1}(u^s, u^d)$ is also concave in (u^s, u^d). In addition, we have

$$\begin{aligned}
V_{j-1}(u^s, u^d) &= \max_{\pi \in \Pi_{\text{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[P_{j-1}\Big(\pi, d_{I_{j-1}}, x_{I_{j-1}}; u^s, u^d\Big) + V_j(u_j^s, u_j^d) \right] \\
&\le \max_{\pi \in \Pi_{\text{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[P_{j-1}\Big(\pi, d_{I_{j-1}}, x_{I_{j-1}}; u^s, u^d\Big) + V_j(0,0) + c_S u_j^s + (r - c_S) u_j^d \right] \\
&= \max_{\pi \in \Pi_{\text{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[P_{j-1}\Big(\pi, d_{I_{j-1}}, x_{I_{j-1}}; u^s, u^d\Big) + c_S u_j^s + (r - c_S) u_j^d \right] + V_j(0,0) \\
&= \max_{\pi \in \Pi_{\text{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[\widetilde{P}_{j-1}\Big(\pi, d_{I_{j-1}}, x_{I_{j-1}}; u^s, u^d\Big) + c_S u^s + (r - c_S) u^d \right] + V_j(0,0) \\
&= c_S u^s + (r - c_S) u^d + \max_{\pi \in \Pi_{\text{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[\widetilde{P}_{j-1}\Big(\pi, d_{I_{j-1}}, x_{I_{j-1}}; 0, 0\Big) \right] + V_j(0,0) \\
&\le V_{j-1}(u^s, u^d).
\end{aligned}$$

The first inequality comes from that both (a) and (b) hold for V_j, and the third equality is from the definition of $\widetilde{P}_{j-1}$. Finally, the last inequality is a worst-case profit from j^{th} subperiod with a policy $\overline{\pi}_{\text{PA}}$, by Assumption 1. Since $\overline{\pi}_{\text{PA}}$ is a feasible policy to the DP, the last inequality follows. This shows that whenever $u^s \le w_1^{(j-1)}$, then the value function $V_{j-1}(u^s, u^d)$ is achieved with a policy $(\overline{\pi}_{j-1}, \ldots, \overline{\pi}_N)$. Finally we have

$$\begin{aligned}
V_{j-1}(u^s, u^d) &= c_S u^s + (r - c_S) u^d + \max_{\pi \in \Pi_{\text{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[\widetilde{P}_{j-1}\Big(\pi, d_{I_{j-1}}, x_{I_{j-1}}; 0, 0\Big) \right] + V_j(0,0) \\
&= V_{j-1}(0,0) + c_S u^s + (r - c_S) u^d,
\end{aligned}$$

and thus (b) holds for $j - 1$. By Assumption 1, $\overline{\pi}_{\text{PA}} = (\overline{\pi}_1, \ldots, \overline{\pi}_N)$ satisfies $u_{t_j}^s \le w_1^{(j+1)}$ for every $j = 1, \ldots, N-1$ and every realization of demands, and hence, $\overline{\pi}_{\text{PA}}$ is Bellman-optimal to the DP (11) and this concludes with $V_{\text{PA}}^* = V_{\text{DP}}^*$. $\square$

Proof of Proposition 5. It suffices to show that an optimal periodic-affine policy is indeed a fully-affine policy. Using the same notations in Theorem 1 and without loss of generality, we may assume that $N=2$ and let $\overline{\pi}^*_{\mathrm{PA}} = (\overline{\pi}_1, \overline{\pi}_2)$, where $\pi_j = (w_t^{(j)}, W_{\tau,t}^{(j)})$ for $j=1,2$. By definition of periodic-affine policies, we only need to check that if an order quantity at time t_1+1 is affine in $\mathcal{U} = \mathcal{U}^1 \times \mathcal{U}^2$. Recall that $\overline{\pi}^*_{\mathrm{PA}}$ determines order quantity at t_1+1 as

$$\begin{aligned} w_1^{(2)} - u_{t_1}^s + u_{t_1}^d &= w_1^{(2)} - \max\left(\sum_{t=1}^{t_1}(s_t - d_t),\ 0\right) + \max\left(\sum_{t=1}^{t_1}(d_t - s_t),\ 0\right) \\ &= w_1^{(2)} + \left(\sum_{t=1}^{t_1}(d_t - s_t)\right). \end{aligned}$$

It is affine in $\mathcal{U}$, since s_t is affine in $d_{[1:t_1]}$, and this concludes the proof. $\square$

Proof of Theorem 2. We use the value function $V_j(u^s, u^d)$ defined in Theorem 1. From concavity $V_j(u^s, u^d)$ and using (b), we have

$$V_j(u^s, u^d) \le V_j(0,0) + c_S u^s + (r - c_S) u^d$$

for every $u^s \ge 0$ and $u^d \ge 0$. We will show that

$$V_j(u^s, u^d) \le c_S u^s + (r - c_S) u^d + \sum_{k=j}^{N} \widetilde{f}_k^* \qquad \forall j = 1, \ldots, N \tag{25}$$

by induction, and plugging $j=1$ and $u^s = u^d = 0$ into (25) concludes the proof.

From $V_N(0,0) = \widetilde{f}_N^*$, we have that (25) holds for $j=N$. Now suppose $1 < j \le N$. Then from the optimality equation we have

$$\begin{aligned} V_{j-1}(u^s, u^d) = &\max_{\pi \in \Pi_{\mathrm{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[P_{j-1}\left(\pi, d_{I_{j-1}}, x_{I_{j-1}}; u^s, u^d\right) + V_j(u_j^s, u_j^d)\right] \\ \le &\max_{\pi \in \Pi_{\mathrm{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[P_{j-1}\left(\pi, d_{I_{j-1}}, x_{I_{j-1}}; u^s, u^d\right) + c_S u_j^s + (r - c_S) u_j^d + \sum_{k=j}^{N} \widetilde{f}_k^*\right] \\ \le \ & c_S u^s + (r - c_S) u^d + \max_{\pi \in \Pi_{\mathrm{aff}}(\mathcal{U}^{j-1})} \min_{d_{I_{j-1}}} \max_{x_{I_{j-1}}} \left[\widetilde{P}_{j-1}\left(\pi, d_{I_{j-1}}, x_{I_{j-1}}; 0, 0\right)\right] + \sum_{k=j}^{N} \widetilde{f}_k^* \\ = \ & c_S u^s + (r - c_S) u^d + \sum_{k=j-1}^{N} \widetilde{f}_k^*, \end{aligned}$$

where the first inequality holds from the induction hypothesis and the third equality is from definition of $\widetilde{P}_{j-1}$. One can show that the maximization problem in the third equality is concave in u^s and u^d, as similar in the proof of Theorem 1 and this concludes the proof. $\square$

Proof of Theorem 3. All the proofs of Theorem 1 and 2 can be extended into multi-station networks, by using a basis matrix $\mathbf{R}_B$ to replace $c_S u^s$ and $(r - c_S) u^d$ terms in the proof with $\mathbf{c}_S^\top \mathbf{u}^s$ and $(\mathbf{R}_B^\top \mathbf{r} - \mathbf{R}_B^\top \mathbf{R}_S^\top \mathbf{c}_S)^\top \mathbf{u}^d$, respectively. This expressions are still linear in $\mathbf{u}^s$ and $\mathbf{u}^d$, hence all the arguments in the proof are valid. $\square$

Proof of Theorem 4. It suffices to show for single-station cases, since it is straightforward to extend the result to general multi-station networks, as in Theorem 3. Note that the optimality equation for the infinite horizon problem is written as

$$V_\infty(u^s,u^d) = \max_{\pi\in\Pi_{\text{aff}}(\mathcal{U}^k)} \min_{d_{[1:k]}} \max_{x_{[1:k]}} \left[P\Big(\pi, d_{[1:k]}, x_{[1:k]}; u^s, u^d\Big) + \beta^k V_\infty(\overline{u}^s, \overline{u}^d_j) \right], \tag{26}$$

where $\overline{u}^s$ and $\overline{u}^d$ denotes on-hand input and backorders after k periods (one stage).

We impose mild conditions so that the optimality equation (26) defines a contraction mapping and there exists V_∞ which is the unique fixed point. (See Iyengar (2005) for details.) Hence the value iteration algorithm is well-defined, and let $V_n(u^s,u^d)$ be a value function after n iterations. Recalling that $\overline{u}^s$ and $\overline{u}^d$ are expressed as linear functions of u^s and u^d, one can show by applying concavity preservation under maximization as similar in Theorem 1 that if $V_n(u^s,u^d)$ is concave, then so $V_{n+1}(u^s,u^d)$ is. Since we can start with any bounded continuous function for the value iteration algorithm, we conclude that $V_\infty(u^s,u^d)$ is concave in (u^s,u^d).

In this setting, there exists a stationary optimal policy $\pi_\infty = (\pi,\pi,\dots)$ where $\pi = \pi(u^s,u^d)$ is defined for each subperiod of length k. By Proposition 3, we can see that $V_\infty(u^s,u^d) = V_\infty(0,0) + c_S u^s + (r-c_S)u^d$ for $u^s \le w_1$ and with concavity of V_∞, we have

$$V_\infty(u^s,u^d) \le V_\infty(0,0) + c_S u^s + (r-c_S)u^d$$

for every $u^s \ge 0$ and $u^d \ge 0$.

Let $V_\infty(\pi_\infty)$ be a worst-case objective value under policy π_∞, and V^*_∞ be an optimal value of the DP problem. Since π_∞ is feasible to the DP by Assumption 1, we have $V_\infty(\pi_\infty) \le V^*_\infty$. On the other hand,

$$\begin{aligned}
V^*_\infty &= \max_{\pi\in\Pi_{\text{aff}}(\mathcal{U}^k)} \min_{d_{[1:k]}} \max_{x_{[1:k]}} \left[P\Big(\pi, d_{[1:k]}, x_{[1:k]}; 0,0\Big) + \beta^k V_\infty(\overline{u}^s, \overline{u}^d_j) \right] \\
&\le \max_{\pi\in\Pi_{\text{aff}}(\mathcal{U}^k)} \min_{d_{[1:k]}} \max_{x_{[1:k]}} \left[P\Big(\pi, d_{[1:k]}, x_{[1:k]}; u^s, u^d\Big) + \beta^k \left(c_S\overline{u}^s + (r-c_S)\overline{u}^d + V_\infty(0,0) \right) \right] \\
&= \max_{\pi\in\Pi_{\text{aff}}(\mathcal{U}^k)} \min_{d_{[1:k]}} \max_{x_{[1:k]}} P^{\text{PA}}_\infty\Big(\pi, d_{[1:k]}, x_{[1:k]}\Big) + \beta^k V_\infty(0,0) \\
&= V_\infty(\pi_\infty),
\end{aligned}$$

by Assumption 1 (this step is similar to Theorem 1), and this implies that an optimal value to the DP is achieved by π_∞, where the stationary policy π is obtained by solving the optimization problem (20). □