

A Hierarchy of Suboptimal Policies for the Multi-period, Multi-echelon, Robust Inventory Problem

Dimitris J. Bertsimas • Dan A. Iancu • Pablo A. Parrilo

Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E40-147, Cambridge, MA 02139, USA

Operations Research Center, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E40-130, Cambridge, MA 02139, USA

Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 77 Massachusetts Avenue, 32D-726, Cambridge, MA 02139, USA

dbertsim@mit.edu • daniancu@mit.edu • parrilo@mit.edu

1 Introduction

In this research, we will be concerned with the operational problem of managing a multi-echelon, multi-product supply chain, in which inventories are renewed periodically over a finite planning horizon, while trying to meet unknown demands from customers and minimize a combination of ordering, holding and backlogging costs. This problem is certainly not new in the field. Multi-period models have been traditionally studied under distributional assumptions for the demand, in formulations for both single and multi-echelon, with finite or infinite horizon, and more general cost structures (see Zipkin (2000) and references therein). Several researchers have also provided distribution-free (robust) formulations for the problem (see Gallego et al. (2001), Ben-Tal et al. (2005), Bertsimas and Thiele (2006) and references therein), and have discussed their performance, tradeoffs, assumptions and general modeling philosophy. Therefore, the goal of our current work is not to *justify* the use of robust (and, more generally, distribution-free) techniques. Rather, we take the modeling approach as a given, and investigate the performance and tractability of a new class of ordering policies, parameterized directly in the observed demands, in the context of a general class of supply chain problems. The results presented here are based on our two journal papers, Bertsimas et al. (2009a,b), to which we refer the interested reader for complete details.

Our contributions can be summarized as follows. For the case of a single echelon, with any convex inventory cost and linear ordering cost, we prove that ordering policies that are *affine in past demands* are optimal. Furthermore, we prove that a certain affine relaxation of the inventory costs is also possible, without any loss of optimality, which gives rise to efficient algorithms for computing the optimal affine policies when the inventory costs are piecewise affine. In particular, the user needs to solve a single Linear Programming Problem (LP), with $O(M \cdot T^2)$ variables and $O(M \cdot T^2)$ constraints, where T is the planning horizon, and M is the maximal number of pieces in the inventory cost (*e.g.*, for Newsvendor costs, $M = 2$). To the best of our knowledge, this is the first result of its kind, and it provides intuition and motivation for the widespread advocacy of such policies in both theory and applications. Our theoretical results are tight (if the conditions in the problem are slightly perturbed, then simple counterexamples can be found), and the proof of the theorem itself is atypical, consisting of a forward

induction and making use of polyhedral geometry to construct the optimal affine policies. For complete details, we refer the interested reader to our journal paper, Bertsimas et al. (2009b).

The proof also delivers insight into the structure of the optimal ordering policies. More precisely, with the order quantities $q_k, \forall k \in \{1, \dots, T\}$, depending affinely on the past observed demands, w_1, \dots, w_{k-1} , the affine coefficients corresponding to a particular demand, w_t , can be interpreted as decomposing the satisfaction of that demand by means of future orders, $q_{t+1}, q_{t+2}, \dots, q_T$. This result resembles that of Muharremoglu and Tsitsiklis (2007), but is obtained under very different techniques and assumptions.

In the case of a general supply chain with a single demand node, we prove that the optimal ordering policies depend multi-linearly on the past-observed demands, and computing them requires solving an exponential Linear Program (LP). Motivated by this result, we consider ordering policies that depend polynomially on the past-observed demands. Since the resulting optimization problems become non-convex, we provide an efficient convex relaxation, allowing the user to compute policies of fixed degree by solving a single Semidefinite Programming Problem (SDP). Our approach is general purpose, *i.e.* it can be applied “out-of-the-box” to any supply chain configuration, allowing for any linear combination of constraints on order quantities and inventories. It is simple, requiring few model inputs from the user, and it allows for an explicit trade-off between computational efficiency and quality of the objective, by means of the degree of the polynomials. Complete details and proofs can be found in our journal paper, Bertsimas et al. (2009a).

We provide numerical results for such polynomial policies of degree up to 3 for the case of a serial supply chain, and compare their performance with the optimal value of the problem (obtained by solving the exponential LP). Our numerical results are very promising, indicating that quadratic policies significantly reduce the optimality gap, while cubic policies are always within 1% of optimality (with a median for the relative optimality gap of less than 0.01% in all our simulations).

2 Optimality of Affine Policies in the Demands for the Single Echelon Problem

The model here is the following: over a planning horizon T , unknown demands w_t arrive at the (unique) echelon, henceforth referred to as the *retailer*, and are satisfied from the on-hand inventory, denoted by x_t at the beginning of period t . Unmet demands can be backlogged, and the retailer can replenish inventory by placing orders u_t , at the beginning of each period t , for a cost of c_t per unit of product. The orders are immediately available (zero lead-time), but there are capacities on how much can be ordered: $L_t \leq u_t \leq U_t$. After the demand w_t is realized, the retailer incurs a cost $h_t(x_{t+1})$ that depends on the net inventory at the end of the period, $x_{t+1} = x_t + u_t - w_t$. The only restriction imposed on h_t is that it is convex (one such example is the classical Newsvendor cost, $h_t(x_{t+1}) = H_t \cdot \max\{0, x_{t+1}\} + B_t \cdot \max\{-x_{t+1}, 0\}$).

The goal is to find ordering policies $\{u_k(\cdot)\}_{1 \leq k \leq T}$ minimizing the worst-case cost:

$$(P) \quad \min_{u_1} \left[c_1 \cdot u_1 + \max_{w_1} \left[h_1(x_2) + \cdots + \min_{u_T} (c_T \cdot u_T + \max_{w_T} h_T(x_{T+1})) \cdots \right] \right]$$

$$\text{s.t.} \quad x_{k+1} = x_k + u_k - w_k,$$

$$L_k \leq u_k \leq U_k, \quad \forall k \in \{1, 2, \dots, T\},$$

where the initial inventory x_1 is assumed fixed, and the demands $w_k \in \mathcal{W}_k \stackrel{\text{def}}{=} [\underline{w}_k, \bar{w}_k] \subset \mathbf{R}$.

Letting J_{mM} denote the optimal cost in (P), and $\mathbf{w}^k = [w_1, \dots, w_{k-1}]$ the history of demands known at the beginning of time k , our main result can be summarized in the following theorem:

Theorem 1. *For every time step $k = 1, \dots, T$, the following quantities exist:*

- an affine ordering policy, $q_k(\mathbf{w}^k) \stackrel{\text{def}}{=} q_{k,0} + \sum_{t=1}^{k-1} q_{k,t} \cdot w_t$,
- an affine stage cost, $z_k(\mathbf{w}^{k+1}) \stackrel{\text{def}}{=} z_{k,0} + \sum_{t=1}^k z_{k,t} \cdot w_t$,

such that the following properties are obeyed:

$$L_k \leq q_k(\mathbf{w}^k) \leq U_k, \quad \forall \mathbf{w}^k \in \mathcal{W}_1 \times \cdots \times \mathcal{W}_{k-1} \quad (1)$$

$$z_k(\mathbf{w}^{k+1}) \geq h_k \left(x_1 + \sum_{t=1}^k (q_t(\mathbf{w}^t) - w_t) \right), \quad \forall \mathbf{w}^{k+1} \in \mathcal{W}_1 \times \cdots \times \mathcal{W}_k \quad (2)$$

$$J_{mM} = \max_{w_1, \dots, w_k} \left[\sum_{t=1}^k (c_t \cdot q_t(\mathbf{w}^t) + z_t(\mathbf{w}^{t+1})) + J_{k+1}^* \left(x_1 + \sum_{t=1}^k (q_t(\mathbf{w}^t) - w_t) \right) \right]. \quad (3)$$

The interpretation of the above statements is as follows. (1) confirms the existence of an affine policy $q_k(\mathbf{w}^k)$ that is robustly feasible, i.e. that obeys the order capacity constraints, no matter what the realization of the demands may be. (2) states the existence of an affine cost $z_k(\mathbf{w}^{k+1})$ that is always larger than the convex cost $h_k(x_{k+1})$ incurred when the affine policies $\{q_t(\cdot)\}_{1 \leq t \leq k}$ are used. (3) guarantees that, despite using the (suboptimal) affine control law $q_k(\mathbf{w}^k)$, and incurring a (potentially larger) affine stage cost $z_k(\mathbf{w}^{k+1})$, the overall objective function value is not increased.

The proof of the theorem proceeds by forward induction on the time step, and examines the adversarial problem (*i.e.* the maximization solved by the demands). It is shown that, due to the geometrical structure of the feasible set, only k extreme demand sequences can actually cause the worst-case cost to occur at time k . Since the affine order quantity at time k has exactly k degrees of freedom (*i.e.* coefficients), these can be chosen so as to mimic the performance of the (truly) optimal policy at the k extreme demand sequences, thereby preserving the worst-case cost. A similar construction is done for the affine costs, $z_k(\mathbf{w}^{k+1})$.

Apart from the theoretical result that affine policies $q_k(\mathbf{w}^k)$ are optimal for Problem (P), the theorem also has *computational implications*. When the convex costs $h_k(x_{k+1})$ are piecewise affine, with at most M pieces, a set of optimal affine policies $\{q_k(\mathbf{w}^k)\}_{1 \leq k \leq T}$ can be computed by solving a Linear Programming problem with $O(MT^2)$ variables and $O(MT^2)$ constraints (refer to Ben-Tal et al. (2005) for details).

Furthermore, the following characterization of the affine ordering policy is a direct result:

Theorem 2. *If the affine orders $q_t(\mathbf{w}^t)$ computed by our procedure are implemented at every time step t , and we let $x_k(\mathbf{w}^k) = x_1 + \sum_{t=1}^{k-1} (q_t(\mathbf{w}^t) - w_t) \stackrel{\text{def}}{=} x_{k,0} + \sum_{t=1}^{k-1} x_{k,t} \cdot w_t$ denote the affine dependency of the inventory x_k on the history of demands, \mathbf{w}^k , then:*

1. *If a certain demand w_t is fully satisfied by time $k \geq t + 1$, i.e. $x_{k,t} = 0$, then all the (affine) orders q_r placed after time k will not depend on w_t .*
2. *Every demand w_t is at most satisfied by the future orders q_k , $k \geq t + 1$, and the coefficient $q_{k,t}$ represents what fraction of the demand w_t is satisfied by the order q_k .*

If we think of $\{q_k\}_{k \geq t+1}$ as future orders that are partially satisfying the demand w_t , then every future order quantity $q_k(\mathbf{w}^k)$ will satisfy exactly a fraction of the demand w_t , and every demand will be at most satisfied by the sequence of orders following after it appears. This interpretation bears some similarity with the unit decomposition approach of Muharremoglu and Tsitsiklis (2007), where every unit of supply can be interpreted as satisfying a particular unit of the demand. Here, we are accounting for fractions of the total demand, as being satisfied by future order quantities.

3 General Supply Chain

As an extension of our earlier model, we next examine the case of a general supply chain. More precisely, consider a network consisting of J echelons $j = 1, \dots, J$, managed over a planning horizon of T periods. To simplify the exposition here, assume there is a single product transferred in the network, and zero lead time in shipping the product between the different echelons (the model can be extended to handle multiple items and deterministic, positive lead times, by suitably increasing the state and the control). The evolution of the supply chain can be modeled as the following constrained linear dynamical system:

$$\mathbf{x}(k+1) = A \cdot \mathbf{x}(k) + B \cdot \mathbf{u}(k) + C \cdot \mathbf{w}(k) \quad (4)$$

$$\mathbf{F} \geq E_x \cdot \mathbf{x}(k) + E_u \cdot \mathbf{u}(k), \quad (5)$$

where the state variable, $\mathbf{x}(k) \in \mathbf{R}^J$, is the vector of all local inventories at the beginning of time k , the control $\mathbf{u}(k) \in \mathbf{R}^{n_u}$ denotes all orders placed in period k , at all echelons, and $\mathbf{w}(k) \in \mathbf{R}^{n_w}$ are the unknown demands in period k (we assume that unmet demands can be backlogged, and, in the spirit of robust optimization, we will take the vector $\mathbf{w}(k)$ to lie in a suitable uncertainty set, $\mathcal{W}(k) \subseteq \mathbf{R}^{n_w}$). A, B, C are matrices of appropriate dimensions, describing the connections between the echelons, the way in which orders affect inventories, and the points where the external demands arrive in the system. The constraints (5) capture all the conditions obeyed by the inventories and orders (e.g., total orders placed in period t , at echelon j , by all downstream echelons, are at most equal to on-hand inventory at j , $x_j(t)$; other examples are nonnegativity and capacity constraints).

To complete the description of the system, for every time-step $k \in \{1, \dots, T\}$, we consider costs that are piecewise affine in the inventories $\mathbf{x}(k)$ and orders $\mathbf{u}(k)$, that is, $g_k(\mathbf{x}(k), \mathbf{u}(k)) = \max_{i=1, \dots, r_k} [\mathbf{c}_x(k, i)^T \cdot \mathbf{x}(k) + \mathbf{c}_u(k, i)^T \cdot \mathbf{u}(k) + c(k, i)_0]$, where r_k is a fixed number. The goal is to find feasible ordering policies $\mathbf{u}(k)$ minimizing the cost in the worst-case scenario:

$$\min_{\mathbf{u}(1)} \left[g_1(\mathbf{x}(1), \mathbf{u}(1)) + \max_{\mathbf{w}(1)} \left[\min_{\mathbf{u}(2)} \left[g_2(\mathbf{x}(2), \mathbf{u}(2)) + \dots + \max_{\mathbf{w}(T+1)} g_{T+1}(\mathbf{x}(T+1), \mathbf{u}(T+1)) \right] \dots \right] \right]. \quad (6)$$

For simplicity, we consider the case of a single external demand arriving in every time-period, *i.e.* $w(k) \in \mathcal{W}_k = [\underline{w}_k, \bar{w}_k] \subset \mathbf{R}$ (e.g. the serial or assembly network). While results similar to Theorem 1 do not hold, one can show that multi-affine policies of degree at most $T - 1$ in the disturbances are optimal for the problem. Furthermore, the maximum in (6) is always reached at the extreme points of the uncertainty set, *i.e.* $\{\underline{w}_1, \bar{w}_1\} \times \{\underline{w}_2, \bar{w}_2\} \times \dots \times \{\underline{w}_T, \bar{w}_T\}$. Hence, the optimal value can be found by solving a large Linear Program, exponential in the problem horizon, T (see Ben-Tal et al. (2005), Bemporad et al. (2003)).

Motivated by these results, consider to ordering policies that depend polynomially on the observed disturbances. More precisely, for a specified degree d , the j -th component of the control vector $\mathbf{u}(k)$ is a polynomial of the form:

$$\mathbf{u}_j(k)(w_1, w_2, \dots, w_{k-1}) = \sum_{i=1}^{\binom{d+k}{d}} \ell_i^{(k,j)} \cdot w_1^{d_1} w_2^{d_2} \dots w_{k-1}^{d_{k-1}}, \text{ with } d \geq \sum_{i=1}^{k-1} d_i. \quad (7)$$

The new decision variables become the coefficients $\ell_i^{(k,j)}$, $k \in \{1, \dots, T\}$, $j \in \{1, \dots, n_u\}$, $i \in \{1, \dots, \binom{d+k}{d}\}$. A typical constraint (5) can now be rewritten equivalently as $0 \leq f(\mathbf{w}), \forall \mathbf{w} \in [\underline{w}_1, \bar{w}_1] \times \dots \times [\underline{w}_T, \bar{w}_T]$, where $f(w_1, w_2, \dots, w_T)$ is some polynomial in variables w_1, \dots, w_T , with coefficients given by our decision variables, $\ell_i^{(k,j)}$. Thus, letting d_1, d_2, \dots, d_T be the degrees of $f(\mathbf{w})$ in the variables w_1, \dots, w_T , respectively, and $d_M \stackrel{\text{def}}{=} \max_{i=1, \dots, T} d_i$, we distinguish the following sufficient conditions for such a constraint to hold:

$$\begin{aligned} d_M \text{ odd} &\Rightarrow \exists S_i(\mathbf{w}), T_i(\mathbf{w}), i = 1, \dots, T, \deg(S_i) = \deg(T_i) = d_M - 1, S_i, T_i \text{ SOS,} \\ &\text{such that } f(\mathbf{w}) = \sum_{i=1}^T (w_i - \underline{w}_i) \cdot S_i(\mathbf{w}) + (\bar{w}_i - w_i) \cdot T_i(\mathbf{w}) \end{aligned} \quad (8)$$

$$\begin{aligned} d_M \text{ even} &\Rightarrow \exists S(\mathbf{w}), T_i(\mathbf{w}), i = 0, \dots, T, \deg(S) = d_M, \deg(T_i) = d_M - 2, S, T_i \text{ SOS :} \\ &\text{such that } f(\mathbf{w}) = S(\mathbf{w}) + \sum_{i=0}^T (w_i - \underline{w}_i)(\bar{w}_i - w_i) \cdot T_i(\mathbf{w}). \end{aligned} \quad (9)$$

In the above, the conditions S_i, T_i SOS signify that the polynomials S_i, T_i are sums of polynomials which are perfect squares. Since testing whether a polynomial can be written as a sum of squares (SOS) is equivalent to a linear matrix inequality (see Parrilo (2000)), the problem of finding polynomial policies of the form (7) that obey all constraints (5) can be written as a Semi-definite Programming Problem (SDP) (see Bertsimas et al. (2009a) for details).

To test our framework, we consider the model corresponding to a serial supply chain, and compute polynomial policies of the form (7) by solving the resulting SDPs. For benchmarking purposes, we also solve the exponential

LP yielding the optimal value of the problem. This is done over 100 different problem instances, and we record different statistics for the relative optimality gaps. Our results, presented in Table 1, indicate that, for horizons of $T = 7$ and number of echelons up to $J = 5$, policies of higher degree result in strict improvements, with cubic policies always resulting in relative optimality gaps smaller than 1% (see Figure 1(a) for a typical case). Also, increasing the problem size (i.e., number of echelons, J) does not affect the reductions in gaps, and the computational requirements do not increase drastically (see Figure 1(b), which corresponds to quadratic policies).

Table 1: Relative gaps (in %) for polynomial policies

J	Degree $d = 1$					Degree $d = 2$					Degree $d = 3$				
	avg	std	mdn	min	max	avg	std	mdn	min	max	avg	std	mdn	min	max
2	1.87	1.48	1.47	0.00	8.27	1.38	1.16	1.11	0.00	6.48	0.06	0.14	0.01	0.00	0.96
3	1.47	0.89	1.27	0.16	4.46	1.08	0.68	0.93	0.14	3.33	0.04	0.06	0.00	0.00	0.32
4	1.14	2.46	0.70	0.05	24.63	0.67	0.53	0.53	0.01	2.10	0.04	0.07	0.00	0.00	0.38
5	0.35	0.37	0.21	0.03	1.85	0.27	0.32	0.15	0.00	1.59	0.02	0.03	0.00	0.00	0.15

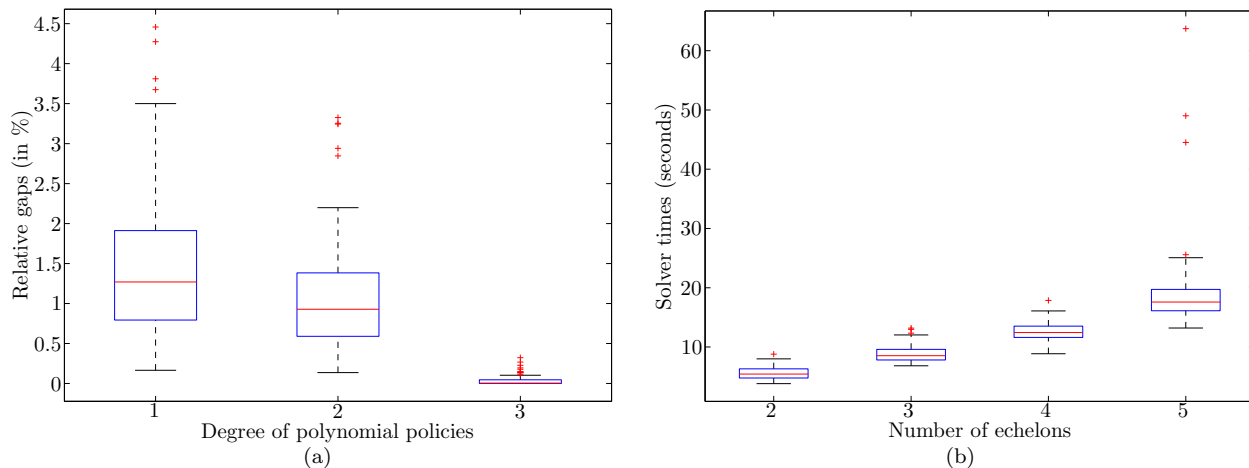


Figure 1: Performance of polynomial policies. (a) compares the three policies for problems with $J = 3$ echelons, and (b) shows the solver times needed to compute quadratic policies for different problem sizes.

References

- Bemporad, A., F. Borrelli, M. Morari. 2003. Min-max control of constrained uncertain discrete-time linear systems. *IEEE Transactions on Automatic Control* **48**(9) 1600–1606.
- Ben-Tal, A., B. Golany, A. Nemirovski, J.-P. Vial. 2005. Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Operations Management* **7**(3) 248–271. doi: <http://dx.doi.org/10.1287/msom.1050.0081>.
- Bertsimas, D., D. Iancu, P. Parrilo. 2009a. A hierarchy of near-optimal policies for multi-stage adaptive optimization. Submitted for publication.
- Bertsimas, D., D. Iancu, P. Parrilo. 2009b. Optimality of affine policies in robust optimization. URL <http://arxiv.org/abs/0904.3986>. Submitted for publication.
- Bertsimas, D., A. Thiele. 2006. A robust optimization approach to inventory theory. *Oper. Res.* **54**(1) 150–168. doi: <http://dx.doi.org/10.1287/opre.1050.0238>.
- Gallego, G., J. Ryan, D. Simchi-Levi. 2001. Minimax analysis for finite-horizon inventory models. *IIE Transactions* **33**(10) 861–874.
- Muharremoglu, A., J. N. Tsitsiklis. 2007. A single-unit decomposition approach to multi-echelon inventory systems. *Operations Research* in press URL <http://citeseer.ist.psu.edu/591628.html>.
- Parrilo, P.A. 2000. Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. Ph.D. thesis, California Institute of Technology.
- Zipkin, P. 2000. *Foundations of Inventory Management*. McGraw Hill.