

# Enforcing Service-Level Constraints in Supply Chains with Assembly Operations

Carmen Del Vecchio and Ioannis Ch. Paschalidis *Senior Member, IEEE,*

*IEEE Transactions on Automatic Control*, in print.

**Abstract**—We study a feedforward supply network that involves assembly operations. We compute optimal stock levels which minimize inventory costs and maintain stockout probabilities below given desirable levels (service-level constraints). To that end, we develop large deviations approximations for inventory costs and service level constraints and formulate the stock level selection problem as a nonlinear programming problem which can be solved using standard techniques. This results in significant computational savings when compared to exhaustive search using simulation. Our distributional assumptions are general enough to include temporal dependencies in the demand and production processes. We leverage the solution of the inventory control problem in the design of supply contracts under explicit service-level constraints.

**Index Terms**—Supply chain management, Assembly systems, Inventory control, Service levels, Supply Contracts, Large Deviations.

## I. INTRODUCTION

We study supply networks that involve assembly operations. For brevity, we will refer to these systems as *assembly networks*. Our primary objective is to develop *efficient supplier inventory control policies* that provide *explicit buyer satisfaction guarantees*.

We adopt a modified *echelon base-stock* policy and seek to optimize echelon stock levels in order to minimize supplier expected inventory costs while providing to the buyer *Quality of Service (QoS)* guarantees. QoS is captured by introducing a constraint in the inventory cost minimization problem requiring stockout probabilities to be below a given desirable level. We refer to such constraints as *QoS* or *service-level* constraints. Our analysis allows demand and production to be modeled by *autocorrelated* stochastic processes; thus, capturing strong *temporal dependencies* that are often observed in these processes. We rely on *large deviations techniques* which yield asymptotically tight approximations of the stockout probabilities as they approach zero. Large deviations for such general stochastic processes have already been employed by Paschalidis and Liu [1] and Bertsimas and Paschalidis [2] in related inventory control settings (refer to these papers for an extensive literature review).

The outcome of our approach is a deterministic (nonlinear) optimization problem whose objective function (inventory cost) and constraints (service-level constraints) are given by closed-form analytical expressions obtained by large deviations asymptotics and limited simulation. Solving this problem using standard optimization techniques yields the echelon stock levels we seek. In short, our approach reduces a complex stochastic optimization problem into a deterministic optimization problem which can be solved by standard techniques. Building on the solution of the supplier inventory control problem, we approach the problem of *designing optimal supply contracts* where QoS guarantees are part of the contract agreements (see Corbett and Tang [3]). From the analytical expression of the supplier inventory problem we are able to estimate the derivative of the inventory cost with respect to demand model parameters and

Research partially supported by the NSF under grants DMI-0300359, DMI-0330171, ECS-0426453, and CNS-0435312, and by the ARO under the ODDR&E MURI2001 Program Grant DAAD19-01-1-0465.

C. Del Vecchio is with the Dipartimento di Ingegneria, Università degli Studi del Sannio, 82100 Benevento, Italy, e-mail: c.delvecchio@unisannio.it

I. Ch. Paschalidis is with the Center for Information and Systems Engineering, and the Department of Manufacturing Engineering, Boston University, Boston, MA 02215, e-mail: yannis@bu.edu, url: http://ionia.bu.edu/.

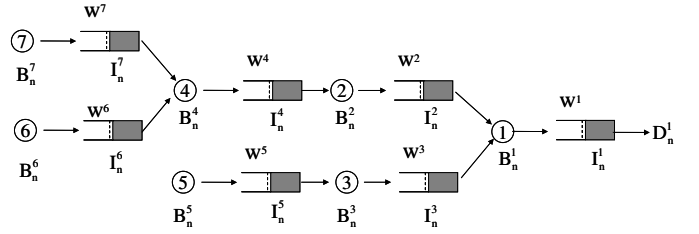


Fig. 1. A supply chain with assembly operations.

use it to evaluate contract incentives in order to achieve a demand configuration that can lead to lower overall supply chain costs.

Among the main contributions of our work we consider: (i) extending the results of Paschalidis and Liu [1] to assembly networks where correlation among stages makes the stockout analysis more complex (see also Song et al. [4]); (ii) enabling the optimization of supply contract parameters to benefit both the supplier and the buyer.

The remainder of this paper is organized as follows: in Section II, we provide the detailed model we consider. In Section III, we present our main large deviations result and formulate the inventory control problem as a deterministic optimization problem. The supply contract model is described in Sec. IV and conclusions are in Sec. V.

## II. THE MODEL OF THE ASSEMBLY NETWORK

The supplier's production process consists of an assembly network that produces a single product and is formed by  $N$  production facilities or stages (Fig. 1). Buyer demand is met from the finished goods inventory maintained in front of stage 1 and is backordered if inventory is not available. Each stage has exactly one successor stage so that the assembly system forms a tree rooted at stage 1. Except from the most upstream stages (the leaves of the tree) which are fed with an infinite supply of raw material, every stage requires one unit of intermediate product (i.e., a component) from each immediate predecessor stage in order to assemble one product unit.

We will use the following notation.  $S(i)$  denotes the unique immediate successor of stage  $i$ ,  $i = 1, \dots, N$ , where  $S(1) = 0$ .  $A(i)$  denotes the set of all successors (immediate and more downstream) of stage  $i$ , where  $A(1) = \{0\}$ .  $P(i)$  denotes the set of immediate predecessors of stage  $i$ , where  $P(i) = \emptyset$  for all stages  $i$  that are leaves of the tree.  $B(i)$  denotes the set of all predecessors (immediate and more upstream) of stage  $i$ , where again  $B(i) = \emptyset$  for all leaf stages  $i$ .  $\mathcal{F}$  denotes the set of all leaf stages and we assume that it contains at least stage  $N$ . Finally,  $\mathcal{N} = \{1, \dots, N\}$  denotes the set of all stages.

We assume a periodic review policy where time is divided into times slots of equal duration.  $B_n^i$  denotes the amount that the facility of stage  $i$  can produce during time slot  $n$ .  $D_n^1$  denotes the amount of external orders arriving at stage 1 during time slot  $n$ . Each stage has an inventory buffer for its final product and from which the successor stage draws material for its production. We let  $I_n^i$  denote the inventory in front of stage  $i$  at the beginning of time slot  $n$ . In intermediate stages,  $i = 2, \dots, N$ , the inventory is constrained to be nonnegative. Inventory at stage 1 can take negative values to denote backordered demand. We denote by  $X_n^i$  the sum of the inventory of stage  $i$  and of all its successors at time  $n$ , i.e.,  $X_n^i = I_n^i + \sum_{k \in A(i)} I_n^k$ . We will refer to this quantity as the *stage- $i$  echelon inventory*. We assume that all stages adopt a modified echelon base-stock policy: stage  $i$  produces when  $X_n^i$  falls below a certain threshold  $w^i$  and idles otherwise. It is implied that  $w^i \geq w^k$  for all  $i \in \mathcal{N}$  and  $k \in P(i)$ .

The demand process  $\{D_n^1\}$  and the production capacity processes  $\{B_n^i\}$  are mutually independent, possibly autocorrelated, arbitrary

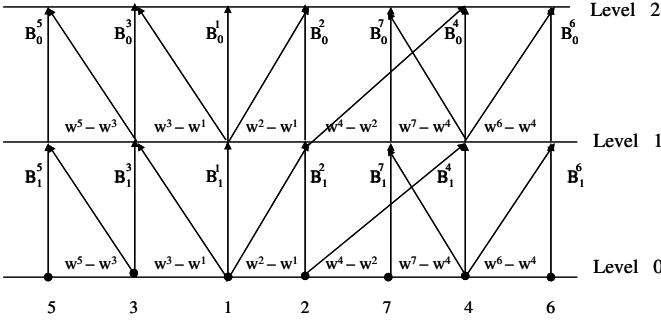


Fig. 2. A two-level graph for the system of Fig. 1.

stationary stochastic processes satisfying certain mild technical conditions (a large deviations principle), that hold for renewal processes, Markov-modulated processes, and, in general, stationary processes with mild mixing conditions. For stability purposes we assume that

$$\mathbf{E}[D_n^1] < \min_{i=1, \dots, N} \mathbf{E}[B_n^i]. \quad (1)$$

We seek safety stocks  $w^i$  for each stage in order to minimize supplier expected inventory costs and guarantee that for all stages  $i$  the steady-state stockout probabilities  $\mathbf{P}[X_n^i \leq 0]$ , at some arbitrary time slot  $n$ , do not exceed desirable small values  $\varepsilon_i$ , respectively.

We define the shortfall of the echelon inventory as:  $Y_n^i = w^i - X_n^i$ , for all  $i \in \mathcal{N}$ . The dynamics of the system can be written as:

$$Y_{n+1}^i = \max \{0, Y_n^i + D_n^1 - B_n^i, \max_{k \in P(i)} \{Y_n^k + D_n^1 - (w^k - w^i)\}\}, \quad i \in \mathcal{N} \setminus \mathcal{F}, \quad (2)$$

$$Y_{n+1}^i = \max \{0, Y_n^i + D_n^1 - B_n^i\}, \quad i \in \mathcal{F}. \quad (3)$$

### III. LARGE DEVIATION ANALYSIS

Next, we establish a large deviations result for the steady state probability  $\mathbf{P}[Y^1 \geq w^1]$ , which is equal to the steady-state stockout probability  $\mathbf{P}[X^1 \leq 0]$ . The line of development follows [1], it is however more involved due to the assembly operations. On a notational remark,  $\Lambda_X(\cdot)$  and  $\Lambda_X^*(\cdot)$  denote the limiting log-moment generating function and the large deviations rate function, respectively, of the process  $X$ .  $\Lambda^+(a)$  (resp.,  $\Lambda^-(a)$ ) denotes the function which equals  $\Lambda^*(a)$  for  $a$  above (resp., below) the mean of  $X$  and zero otherwise.

#### A. Sample path characterization of stockout events

In what follows we will be dropping the subscript  $n$  when referring to steady-state quantities and we will be using

$$\mathcal{O}^i = \left\{ (\xi_i, \xi_{A(i)}) \mid \xi_i \in [0, 1], \xi_{A(i)} \in [0, 1]^{|A(i)|}, \xi_i + \xi_{A(i)}^T \mathbf{e} = 1 \right\}$$

for all  $i \in \mathcal{N}$ , where  $|A(i)|$  denotes the cardinality of the set  $A(i)$ ,  $\mathbf{e}$  is the vector of all ones, and prime denotes transpose.

Suppose that at time 0, the finished goods inventory and the components inventory are all equal to the safety stock level, i.e.,  $Y_0^i = 0$  for all  $i \in \mathcal{N}$ . After some algebra one can show that the shortfall for each stage  $i$  at time 2 satisfies  $Y_2^i = \max\{0, D_1^1 - r_{2,1}^i, D_0^1 + D_1^1 - r_{2,2}^i\}$ . Fig. 2 depicts a two-level graph for the system of Fig. 1. Points 1, ..., 7 at the bottom of the graph correspond to stages of the assembly system. The quantities  $r_{2,m}^i$  can be interpreted as the length of the shortest path from point  $i$  at level 0 to level  $m$ . Generalizing, the shortfall of stage 1 at time slot  $n$  is given by:

$$Y_n^1 = \max \left\{ 0, \max_{1 \leq m \leq n} \left[ \sum_{j=1}^m D_{n-j}^1 - r_{n,m}^1 \right] \right\},$$

where  $r_{n,m}^1$  is the length of the shortest path from point 1 at level 0 to level  $m$  in a corresponding  $n$  level graph constructed as Fig. 2.

Let us denote by  $\{\hat{D}_n^1\}$  the time reversed stochastic process obtained from the demand process  $\{D_n^1\}$ . In particular for any  $k \in \mathbb{Z}$ ,  $(\hat{D}_1^1, \hat{D}_2^1, \dots, \hat{D}_k^1)$  has the same distribution as  $(D_k^1, D_{k-1}^1, \dots, D_1^1)$ . Similarly, let  $\{\hat{B}_n^i\}$ ,  $i = 1, 2, \dots, N$ , denote the time reversed production processes. Then, it can be shown that  $Y_n^1$  has the same distribution as the right hand side of the following equation:

$$Y_n^1 \stackrel{D}{=} \max \left\{ 0, \max_{1 \leq m \leq n} \left[ \sum_{j=1}^m \hat{D}_j^1 - \min_{\substack{0 \leq m_i \leq m, l_h \in \{0,1\} \\ l_h=0 \Rightarrow m_h=0, m_k, l_k=0 \forall k \in B(h) \\ \forall i l_h=1, h \in P(i) \Rightarrow l_k=0 \forall k \in P(i) \setminus \{h\}}} \left( \sum_{i=1}^{m_1} \hat{B}_i^1 + \sum_{h \in P(1)} l_h (w^h - w^1) \right) + \sum_{i=k_2+1}^{k_2+m_2} \hat{B}_i^2 + \sum_{h \in P(2)} l_h (w^h - w^2) + \dots + \sum_{i=k_N+1}^{k_N+m_N} \hat{B}_i^N \right] \right\}, \quad (4)$$

where " $\stackrel{D}{=}$ " denotes equality in distribution, and  $k_i = |A(i)| + \sum_{j \in A(i)} m_j$  (with  $k_1 \triangleq 0$ ). The last constraint in the minimization above has the following interpretation: for any stage  $i$  and for any stage  $h$  which is a predecessor of  $i$ ,  $l_h = 1$  implies  $l_k = 0$  for all other stages  $k$  that are also predecessors of stage  $i$  ( $k \in P(i) \setminus \{h\}$ ). Due to (1) a steady-state distribution exists for  $Y_n^1$ ; let  $Y^1$  be distributed according to this distribution. In particular,  $Y_n^1$  converges to  $Y^1$  as  $n \rightarrow \infty$ . In the sequel, we will denote by  $G_m$  the argument of the maximum, i.e.,  $Y^1 = \max_{m \geq 0} G_m$ .

#### B. Lower Bound

For any stage  $i$  let  $A(i) = \{1, \dots, |A(i)|\}$ ; we will be also writing  $\xi_{A(i)}$  for the vector  $(\xi_1, \dots, \xi_{|A(i)|})$ .

**Proposition III.1** Assume the hedging points  $w^1, w^2, \dots, w^N$  satisfy  $w^i = \beta_i w^1$ ,  $i = 2, \dots, N$ , where  $\beta_i$  are constants and  $1 \leq \beta_h \leq \beta_i$  for all  $h \in A(i)$  ( $\beta_1 \triangleq 1$ ). The steady state shortfall  $Y^1$  satisfies

$$\liminf_{w^1 \rightarrow \infty} \frac{1}{w^1} \log \mathbf{P}[Y^1 \geq w^1] \geq -\theta_{G,1}^*, \quad (5)$$

$$\theta_{G,1}^* = \min \left[ \inf_{a>0} \frac{1}{a} \inf_{x_0, x_1=a} (\Lambda_{D^1}^{*+}(x_0) + \Lambda_{B^1}^{*-}(x_1)), \min_{2 \leq i \leq N} \left\{ \inf_{a>0} \frac{1}{a} \inf_{\substack{x_0 - \sum_{h \in A(i)} \xi_h x_h - \xi_i x_i = a \beta_i \\ (\xi_i, \xi_{A(i)}) \in \mathcal{O}^i}} (\Lambda_{D^1}^{*+}(x_0) + \sum_{h \in A(i)} \xi_h \Lambda_{B^h}^{*-}(x_h) + \xi_i \Lambda_{B^i}^{*-}(x_i)) \right\} \right]. \quad (6)$$

*Proof:* (Outline) For any  $m \geq 0$  we have

$$\frac{1}{w^1} \log \mathbf{P}[Y^1 \geq w^1] = \frac{1}{w^1} \log \mathbf{P}[\max_{m \geq 0} G_m \geq w^1] \geq \frac{1}{w^1} \log \mathbf{P}[G_m \geq w^1].$$

Choose  $a > 0$  and write  $w^1 = ma$ . Then for any node  $i$  we have  $w^h - w^i = m(\beta_h - \beta_i)a$  for all  $h \in P(i)$ . We obtain

$$\frac{1}{w^1} \log \mathbf{P}[Y^1 \geq w^1] \geq \frac{1}{ma} \log \mathbf{P}[G_m \geq ma], \quad (7)$$

and since we are interested in the regime where  $w^1 \rightarrow \infty$ , it suffices to analyze the behavior of the right hand side of (7) for large values of  $m$ . To that end, select  $x_i \geq 0$ ,  $l_h \in \{0, 1\}$ , and  $0 \leq m_i \leq m$ , for all  $i \in \mathcal{N}$  and  $h \in P(i)$  such that  $\sum_{i=1}^N (m_i + \sum_{h \in P(i)} l_h) = m$ ,

$$mx_0 - m_1x_1 - \sum_{h \in P(1)} l_h m(\beta_h - 1)a - m_2x_2 + \dots - \sum_{h \in P(2)} l_h m(\beta_h - \beta_2)a - \dots - m_N x_N = ma, \quad (8)$$

$l_h = 0$  implies  $m_h = 0$  and  $m_k = 0$ ,  $l_k = 0$  for all  $k \in B(h)$ , and for any stage  $i$  and  $h \in P(i)$  for which  $l_h = 1$  it is implied that  $l_k = 0$  for all  $k \in P(i) \setminus \{h\}$ . To obtain a lower bound on  $\mathbf{P}[G_m \geq ma]$  we construct particular sample path scenarios characterized by  $x_i$ ,  $l_i$  and  $m_i$  that lead to  $G_m \geq ma$ . For each sample path we determine a lower bound for the stockout probability (see [1]); after some algebra, the tightest lower bound reduces to (5). ■

### C. Upper Bound

**Proposition III.2** Assume the hedging points  $w^1, w^2, \dots, w^N$  satisfy  $w^i = \beta_i w^1$ ,  $i = 2, \dots, N$ , where  $\beta_i$  are constants and  $1 \leq \beta_h \leq \beta_i$  for all  $h \in A(i)$  ( $\beta_1 \triangleq 1$ ). The steady state shortfall  $Y^1$  satisfies

$$\limsup_{w^1 \rightarrow \infty} \frac{1}{w^1} \log \mathbf{P}[Y^1 \geq w^1] \leq -\bar{\theta}_{G,1}^*, \quad (9)$$

$$\bar{\theta}_{G,1}^* \triangleq \min(\theta_1^*, \beta_2 \theta_2^*, \dots, \beta_N \theta_N^*) \quad (10)$$

$$\theta_i^* \triangleq \begin{cases} \theta \geq 0 \\ \sup_{(\xi_i, \xi_{A(i)}) \in \mathcal{O}^i} (\Lambda_{D^i}(\theta) + \sum_{h \in A(i)} \xi_h \Lambda_{B^h}(-\theta) + \xi_i \Lambda_{B^i}(-\theta)) < 0 \end{cases} \theta. \quad (11)$$

*Proof:* We consider all sample paths that can lead to a value larger than  $w^1$ . In particular, the first such sample path corresponds to  $l_i = 0$  for all  $i \in P(1)$  and the  $i$ th sample path corresponds to  $l_h = 1$  for all  $h \in A(i)$  and  $l_k = 0$  for all  $k \in \mathcal{N} \setminus A(i)$ . The detailed proof follows [1] in establishing the corresponding upper bound. ■

The following result is due to convex duality; see [1, Prop. 5.4].

**Proposition III.3** It holds  $\theta_{G,1}^* = \bar{\theta}_{G,1}^*$ .

### D. Main Theorem

Summarizing Propositions III.1, III.2, and III.3 we obtain the following main theorem.

**Theorem III.4** Assume the hedging points  $w^1, w^2, \dots, w^N$  satisfy  $w^i = \beta_i w^1$ ,  $i = 2, \dots, N$ , where  $\beta_i$  are constants and  $1 \leq \beta_h \leq \beta_i$  for all  $h \in A(i)$  ( $\beta_1 \triangleq 1$ ). The steady state shortfall  $Y^1$  satisfies

$$\lim_{w^1 \rightarrow \infty} \frac{1}{w^1} \log \mathbf{P}[Y^1 \geq w^1] = -\theta_{G,1}^*, \quad (12)$$

where  $\theta_{G,1}^*$  is given by either (6) or (10).

Consider next the form of  $\theta_{G,1}^*$  provided by (10). Consider the case  $\theta_{G,1}^* = \beta_i \theta_i^*$  for some  $i = 1, \dots, N$ , where  $\beta_1 \triangleq 1$ . To avoid degenerate cases assume that all production processes  $B^i$  have distinct log-moment generating functions and that  $\beta_i \neq \beta_j$  for all  $i, j \in \mathcal{N}$ . Let  $\xi_j^*$ ,  $j \in A(i) \cup \{i\}$ , be the optimal solution of the optimization problem  $\sup_{(\xi_i, \xi_{A(i)}) \in \mathcal{O}^i} (\Lambda_{D^i}(\theta) + \sum_{h \in A(i)} \xi_h \Lambda_{B^h}(-\theta) + \xi_i \Lambda_{B^i}(-\theta)) = 0$  at  $\theta = \theta_i^*$ . It can be seen that one of the  $\xi_j^*$ 's,  $j \in A(i) \cup \{i\}$  is equal to 1. In particular,  $\xi_i^*$  is equal to 1; otherwise, i.e., if  $\xi_j^* = 1$  for some  $j \in A(i)$ , then  $\theta_j^* = \theta_i^*$  and  $\beta_j \theta_j^*$  will be the minimizer in the definition of  $\theta_{G,1}^*$  since  $\beta_j \theta_j^* < \beta_i \theta_i^*$  as  $\beta_j < \beta_i$  for  $j \in A(i)$ . Consequently,  $\theta_i^*$  is the largest root of the equation  $\Lambda_{D^i}(\theta) + \Lambda_{B^i}(-\theta) = 0$ . This observation has an important implication. Specifically, the stockout probability at stage 1 is characterized by the distributions of the demand and the production process at stage  $i$ . We can say that stage  $i$  is the ‘‘bottleneck’’ since it is the one whose production process determines

the tail of the stockout probability. This is though a ‘‘stochastic bottleneck’’ since it is specified by distributions rather than means.

Thm. III.4 can be easily generalized to yield the steady-state stockout probability of the echelon inventory at stages  $i = 2, \dots, N$ .

### E. Refining the Large Deviations Asymptotics

Thm. III.4 suggests an approximation of the stockout probability at stage 1. To make this approximation more accurate, especially for large stockout probabilities, we introduce a prefactor in front of the exponential. We will use the following refined approximation:

$$\mathbf{P}[Y^1 \geq w^1] \approx f_1(w^1, \beta) e^{-\theta_{G,1}^* w^1}, \quad (13)$$

where the prefactor  $f_1(w^1, \beta)$  is a function of  $w^1$  and  $\beta = (\beta_2, \dots, \beta_N) = (\frac{w^2}{w^1}, \dots, \frac{w^N}{w^1})$ . To compute  $f_1(w^1, \beta)$ , we evaluate by simulation the stockout probability  $\mathbf{P}[Y^1 \geq w^1]$  at several sample safety stock vectors  $\mathbf{w}_1, \dots, \mathbf{w}_K$  and set  $f_1(w_k^1, \beta_k) = \mathbf{P}[Y^1 \geq w_k^1] / e^{-\theta_{G,1}^* w_k^1}$ , for each sample point, where  $\beta_k = (\frac{w_k^2}{w_k^1}, \dots, \frac{w_k^N}{w_k^1})$ . Thus, we obtain a data set consisting of  $K$  pairs  $((w_k^1, \beta_k); f_1(w_k^1, \beta_k))$ . Then, we derive a piecewise linear function  $f_1(w^1, \beta)$  so that  $f_1(w^1, \beta) e^{-\theta_{G,1}^* w^1}$  ‘‘fits’’ the true value of  $\mathbf{P}[Y^1 \geq w^1]$  at the points in the data set. Some practical details on the appropriate choice of the sample points  $\mathbf{w}_1, \dots, \mathbf{w}_K$  and on the selection of the function  $f_1(w^1, \beta)$  can be found in [1]; the observations reported there can be straightforwardly applied to the assembly system.

### F. Approximating the expected inventory cost

We now turn our attention to approximating expected inventory costs. We assume a linear cost structure. Let  $h_i$  be the holding cost for echelon- $i$  inventory for all  $i = 1, \dots, N$ . Noting that expected echelon- $i$  inventory is given by  $h_i(E[I^i] + \sum_{h \in A(i) \setminus \{1\}} E[I^h] + E[(I^1)^+])$  where  $(I^1)^+ = \max(I^1, 0)$ , the total expected inventory cost is given by:

$$\left( \sum_{i=1}^N h_i \right) \mathbf{E}[(I^1)^+] + \sum_{i=2}^N h_i \left( \mathbf{E}[I^i] + \sum_{j \in A(i) \setminus \{1\}} \mathbf{E}[I^j] \right).$$

We have

$$\mathbf{E}[(I^1)^+] = \mathbf{E}[(w^1 - Y^1)^+] = w^1 - \mathbf{E}[Y^1] + \mathbf{E}[\max(0, Y^1 - w^1)].$$

Using the tail distribution of  $Y^1$  given in Eq. (13) we obtain

$$\mathbf{E}[\max(0, Y^1 - w^1)] = \int_0^\infty \mathbf{P}[Y^1 - w^1 > y] dy \approx f_1(w^1, \beta) \frac{e^{-\theta_{G,1}^* w^1}}{\theta_{G,1}^*}.$$

For all  $i \geq 2$  we have  $I^i = (w^i - Y^i) - (w^{S(i)} - Y^{S(i)})$ , which implies:

$$\mathbf{E}[I^i] = (w^i - \mathbf{E}[Y^i]) - (w^{S(i)} - \mathbf{E}[Y^{S(i)}])$$

Thus, expected inventory costs can be approximated by

$$\sum_{i=1}^N h_i (w^i - \mathbf{E}[Y^i]) + \left( \sum_{i=1}^N h_i \right) f_1(w^1, \beta) \frac{e^{-\theta_{G,1}^* w^1}}{\theta_{G,1}^*}. \quad (14)$$

To obtain an analytical approximation for the inventory cost we are left with computing  $\mathbf{E}[Y^i]$ . In the following proposition we establish some structural properties for  $\mathbf{E}[Y^i]$ ; to this end for each stage  $i$  ( $i \in \mathcal{N}$ ) we define  $\Delta_i \triangleq (w^i - w^{S(i)})$ .

**Proposition III.5** Consider the assembly system and let  $w^1, \dots, w^N$  be the corresponding hedging points satisfying  $0 \leq w^k \leq w^i$  with  $k \in A(i)$ . For all stages  $i$  which do not have any predecessor (i.e.,  $i \in \mathcal{F}$ ),  $\mathbf{E}[Y^i]$  is a constant function of  $(\Delta_2, \dots, \Delta_N)$ . Furthermore, for all stages which have at least one predecessor (i.e.,  $i \in \mathcal{N} \setminus \mathcal{F}$ ),  $\mathbf{E}[Y^i]$  is a function of  $\Delta_k$  for all  $k \in B(i)$ , and it is convex and monotonically nonincreasing in every coordinate.

Motivated by these properties, we approximate  $\mathbf{E}[Y^i]$  by using a piecewise linear convex function  $g_i$  which is a function of  $\Delta_k$  for all  $k \in B(i)$ . Specifically, we use the following approximation  $\mathbf{E}[Y^i] = g_i(\Delta_{B(i)(1)}, \dots, \Delta_{B(i)(|B(i)|)})$ , where  $B(i)^{(j)}$  denotes the  $j$ th element of  $B(i)$ . As we did with the prefactors of the stockout probabilities we can select a number of sample points  $\mathbf{w}_1, \dots, \mathbf{w}_K$  and construct a piecewise linear convex function that “fits”  $\mathbf{E}[Y^i]$  at those points.

### G. The inventory optimization problem

We can now pose the problem of optimizing expected inventory costs subject to maintaining service level constraints. Using the approximating expressions for expected inventory cost and stockout probabilities we have

$$\begin{aligned} \min \quad & \sum_{i=1}^N h_i(w^i - g_i(\Delta_{B(i)(1)}, \dots, \Delta_{B(i)(|B(i)|)})) + \\ & \left( \sum_{i=1}^N h_i \right) f_1\left(w^1, \frac{w^2}{w^1}, \dots, \frac{w^N}{w^1}\right) e^{-\frac{\theta_{G,1}^* w^1}{\theta_{G,1}^* w^1}} \\ \text{s.t.} \quad & f_i\left(w^i, \frac{w^{B(i)(1)}}{w^i}, \dots, \frac{w^{B(i)(|B(i)|)}}{w^i}\right) e^{-\theta_{G,i}^* w^i} \leq \varepsilon_i, \quad i = 1, \dots, N, \\ & w^i \geq w^k, \quad \forall i \text{ and } \forall k \in P(i). \end{aligned} \quad (15)$$

This problem can be solved analytically using standard nonlinear programming techniques.

We have compared the solution obtained from the above formulation with the one obtained using an exhaustive search over all possible stock levels in a number of assembly systems. The simulation results, which we omit due to space limitations, confirm the validity of our approximations and reveal that our analytical approach is very accurate. In particular, our policy is on the order of 1.5% away from the optimal even for  $\varepsilon_i$  on the order of 1-5%. (Due to the nature of our approximations the accuracy improves as  $\varepsilon_i$  become smaller.) Note that to obtain the base stock policies through simulation it is necessary to explore all the possible combinations of  $w^i$  and select the one that yields the lowest cost within the desired service levels constraints. Since stockouts are controlled by  $\varepsilon_i$  to be relatively rare events, estimating their probability requires long simulation runs; from several hours to several days exploring an interval “centered” at the analytical solution. Instead, the solution of the nonlinear programming problem can be achieved in no longer than a minute for a 5-stage assembly system, while the “pre-processing” operations (the prefactor evaluation and the approximation of  $\mathbf{E}[Y^i]$ ) took on the order of a couple of hours. Clearly the proposed analytical approach leads to significant computational savings.

## IV. THE SUPPLY CONTRACT MODEL

We consider a buyer and a supplier that interact as follows. The buyer determines (a model of) the demand  $D_n^1$  according to market conditions and associates a cost (not necessarily monetary) to any change in demand parameters. The supplier production process is an assembly network following a modified base stock production policy. Safety stocks are set to minimize inventory costs – the cost function in problem (15) – using as input the buyer’s demand model, the service level agreement – the constraint in problem (15) – and (a model of) the production capacities  $B_n^i$ . We will explore how demand parameters can be adjusted according to both supplier and buyer cost structures so that any generated savings can be shared among them.

We assume that the buyer’s demand is an  $m$ -state *Markov-modulated process (MMP)*, with transition probability matrix  $\mathbf{P}_D$  and demand levels at each state given by the vector  $\mathbf{r}_D$ . The supplier’s production capacity at each node  $i$  of the assembly network is modeled by an  $m^i$ -state MMP with transition probability matrix  $\mathbf{P}_{B^i}$  and capacities at each state given by the vector  $\mathbf{r}_{B^i}$ ,  $i = 1, \dots, N$ . We

assume that any mutually agreed adjustments in  $\mathbf{r}_D$  keep the mean demand  $\mathbf{E}[D]$  constant; otherwise the buyer would be unable to satisfy demand in the long term. We denote by  $\overline{\mathbf{E}[D]}$  this constant value.

Let  $C_R(\mathbf{r}_D)$ , the buyer cost associated with any change of  $\mathbf{r}_D$ , be

$$C_R(\mathbf{r}_D) = \sum_{i=1}^m \zeta_i (r_{D,i} - \bar{r}_{D,i})^2, \quad (16)$$

where  $\zeta_i$  expresses the buyer’s cost for changing the demand level in the  $i^{\text{th}}$  state of the demand process.  $\bar{\mathbf{r}}_D$  is the vector of demand levels initially determined and to which the buyer associates zero cost. The supplier’s inventory cost function,  $C_M(\mathbf{r}_D)$ , is the optimal value function of problem (15). Our objective is to find the vector  $\mathbf{r}_D$  that solves the following problem:

$$\begin{aligned} \min f(\mathbf{r}_D) &= C_R(\mathbf{r}_D) + C_M(\mathbf{r}_D) \\ \text{over } \Xi &= \{\mathbf{r}_D \mid \mathbf{E}[D] = \overline{\mathbf{E}[D]}, r_{D,i} \geq 0 \forall i\}. \end{aligned} \quad (17)$$

Problem (17) is a nonlinear optimization problem over a convex set and can be solved using the conditional gradient method. To that end, we need the cost function gradient. The gradient of the buyer’s cost can be easily derived from (16). We can evaluate the gradient of the supplier inventory costs using finite differences.

It is interesting to observe that the above algorithm can be used in a *distributed* fashion. In particular, the buyer can take charge of finding new demand vectors on feasible descent directions. At each iteration, the buyer presents the supplier with the new demand levels  $\mathbf{r}_D^k$  and the supplier responds with the gradient  $\nabla C_M(\mathbf{r}_D^k)$  of its cost function. This gradient information can be provided in the form of appropriate incentives to the buyer. The buyer uses this gradient information to find a feasible direction and compute  $\mathbf{r}_D^{k+1}$ . Note that neither the supplier nor the buyer need to know each other’s cost structures.

## V. CONCLUSIONS

We studied the inventory control problem for a single class assembly network which operates under a modified echelon base-stock policy. We developed an approach to find close-to-optimal echelon stock levels that minimize inventory costs while guaranteeing stockout probabilities stay below some predefined levels. Relying upon large deviations techniques, we reduced the safety stock selection to a deterministic nonlinear optimization problem.

We also used our inventory control approach to analyze how a supplier can interact with a buyer to reach a mutually beneficial mode of operations. This interaction takes the form of a supply contract that enforces explicit QoS guarantees. The joint optimization algorithm we proposed can be used in a distributed fashion with neither the supplier nor the buyer revealing their corresponding cost structures.

## REFERENCES

- [1] I. C. Paschalidis and Y. Liu, “Large deviations-based asymptotics for inventory control in supply chains,” *Operations Research*, vol. 51, no. 3, pp. 437–460, 2003.
- [2] D. Bertsimas and I. C. Paschalidis, “Probabilistic service level guarantees in make-to-stock manufacturing systems,” *Operations Research*, vol. 49, no. 1, pp. 119–133, 2001.
- [3] C. Corbett and C. Tang, “Designing supply contracts: contract type and information asymmetry,” in *Quantitative Models for Supply Chain Management* (S. Tayur, R. Ganeshan, and M. Magazine, eds.), pp. 269–297, Kluwer, 1999.
- [4] J. S. Song and S. H. Xu and B. Liu “Order-fulfillment performance measures in an assemble-to-order system with stochastic leadtimes” *Operations Research*, vol. 47, pp. 131–149, 1999.