Analysis and Evaluation of An Assemble-to-Order System with Batch Ordering Policy and Compound Poisson Demand *

Yao Zhao

Department of Management Science and Information Systems

Rutgers University. Newark, NJ

September 1, 2008

Abstract

We consider a multi-product and multi-component Assemble-to-Order (ATO) system where the external demand follows compound Poisson processes and component inventories are controlled by continuous-time batch ordering policies. The replenishment lead-times of components are stochastic, sequential and exogenous. Each element of the bill of material (BOM) matrix can be any non-negative integer. Components are committed to demand on a first-come-first-serve basis. We derive exact expressions for key performance metrics under either the assumption that each demand must be satisfied in full (non-split orders), or the assumption that each unit of demand can be satisfied separately (split orders). We also develop an efficient sampling method to estimate these metrics, e.g., the expected delivery lead-times and the order-based fill rates. Based on the analysis and a numerical study of an example motivated by a real-world application, we characterize the impact of the component interaction on system performance, demonstrate the efficiency of the numerical method and quantify the impact of order splitting.

Key words: Assemble-to-Order system, batch ordering policy, compound Poisson demand, order splitting.

*Research supported in part by a Faculty Research Grant from Rutgers Business School–Newark and New Brunswick.

1 Introduction

Assemble-To-Order (ATO) systems are becoming increasingly important for today's manufacturing firms as many companies strive to increase product variety and responsiveness to demand without carrying the expensive finish goods inventories. A key challenge for ATO systems is the efficient management of the component inventories, which has been the main focus of recent studies. We refer the reader to Song and Zipkin (2003) and Xu (2001) for excellent reviews of the motivation, examples and related literature of ATO systems.

Batch demand and batch ordering policies are common in many real world ATO systems. In such a system, a customer may demand multiple units of a product, each of which is assembled from multiple components with different quantities. The component inventories are replenished in batches (e.g., truck load), to achieve economies of scale in production and transportation. While in some cases, customers prefer to receive each unit of product as soon as it becomes available (split orders); in others, customers prefer to receive all units simultaneously (non-split orders).

In this paper, we consider ATO systems where external demand follows independent compound Poisson processes and the component inventories are controlled by continuous-time batch ordering policies. Components are committed to demand on a first-come-first-serve (FCFS) basis. The replenishment lead-times of components are stochastic, sequential and exogenous (Svoronos and Zipkin 1991). Our objective is two-fold: (1) deriving exact expressions for the key system performance metrics, i.e., the delivery lead-times and order-based fill rates, under either the split orders or the non-split orders assumption; (2) developing algorithms that can evaluate these performance metrics for ATO systems of large sizes.

Research on ATO systems with compound Poisson demand is quite extensive, see, e.g., Song and Zipkin (2003), Xu (2001) and Hausman, Lee and Zhang (1998) for literature reviews. For constant lead times, we refer to Song (1998, 2002) for systems with continuous-time base-stock policies, and to Hausman, et al. (1998), Zhang (1997), Agrawal and Cohen (2001) and de Kok (2003) for systems with periodic-review base-stock policies. For i.i.d. lead times, we refer to Lu, Song and Yao (2003), Lu and Song (2005) and Lu (2007, 2008) for systems with continuous-time base-stock policies. Cheng, et al. (2002) studies the performance of an Configure-to-Order (CTO) system with stochastic sequential lead times. To link the base-stock levels of components to the order-based fill rates, the authors assume that at most one component can be out of stock at any time. Zhao (2008) provides an exact analysis for supply chains where there is at most one directed path between every two stages. External demand follows compound Poisson processes, and the lead-times are stochastic and sequential. However, the focus of the paper is still on base-stock policies.

Due to the analytical and numerical challenges, research on batch ordering policies for ATO systems in particular, and for assembly systems in general, is limited. Ernst and Pyke (1992) studies an assembly system with multiple components and one final product. The component inventories are replenished in batches. The authors propose approximations for the expected cost functions and develop algorithms to compute the reorder points subject to a service level constraint. Chen (2000) studies a multi-stage assembly system with constant lead-times and one product at each stage. Given the batch sizes at all stages, it is shown that the batch ordering policy is optimal provided that the batch sizes satisfy certain regularity conditions. To evaluate and optimize the performance for the assembly system, an equivalent serial supply chain is constructed based on Rosling (1989). Plambeck (2005) considers a batch ordering ATO system with capacitated suppliers under the expediting assumption, that is, the component orders can be expedited instantaneously at higher costs. Further assuming high volume of demand, it is shown that the control problem of a multi-component ATO system separates into independent control problem of each component. Benjaafar and Elhafsi (2006) studies an ATO system with a single product but multiple demand classes. Under Markovian assumptions on demand and supply, the paper proves structural results for the optimal ordering and allocation policies. It further compares the performance between the optimal policy and some simple heuristic policies. Elhafsi (2008) generalizes the results to compound Poisson demand.

Song (2000) studies a multi-component and multi-product ATO system with constant lead-times where demand follows a multivariate compound Poisson process and component inventories are managed by continuous-time batch ordering policies. It is shown that under certain general conditions, the inventory position vector of the components has a uniform equilibrium distribution. Therefore, the key performance measures, e.g., the expected order-based backorders and fill-rate, of a batch-ordering ATO system can be expressed as the average of the counter-parts of multiple base-stock systems. Two challenges remain: (1) the number of base-stock systems corresponding to a batch ordering ATO system is exponential in the number of components; (2) how to incorporate stochastic sequential lead times?

Zhao and Simchi-Levi (2006) addresses these challenges in an ATO system with Poisson demand and unit BOM matrix (i.e., each element of the BOM matrix can be either zero or one). They presented an exact analysis for the ATO systems with batch ordering policy and stochastic sequential lead times. They also developed an efficient numerical method based on Monte Carlo simulation to evaluate system performance. Compound Poisson demand introduces additional complexities to the exact analysis and computation because of the components' interaction caused by common demand size processes. In addition, the delivery lead-times and fill-rates are now demand size (or demand unit) dependent (Zipkin 1991, Simchi-Levi and Zhao 2005); so Proposition 5.2 of Zhao and Simchi-Levi (2006), which lays the foundation for the numerical method, does not hold for compound Poisson demand.

This paper generalizes the existing literature to a class of ATO systems with *both* batch ordering policies and compound Poisson demand. We considered the case of the non-split orders as well as the case of split orders. Section 2 presents the model and notations. In Section 3, we characterize key system performance metrics (e.g., delivery lead times and fill rates) as explicit functions of lead times, demand sizes and interarrival times. For the special case of a single product and two components, we provide an analysis for the delivery lead times that allow their probability distributions to be determined exactly. Based on the analysis, we show that ignoring the dependence among components results in over-estimating the delivery lead times. In Section 4, we show that although the numerical method of Zhao and Simchi-Levi (2006) does not directly apply here, a modification does which leads to computationally efficient algorithms for systems of large sizes. The algorithms are essentially constrained by their requirements of memory due to the large BOM matrix. A numerical example motivated by a real-world problem is presented in Section 5. The numerical example also allows us to quantify the impact of order splitting. This is valuable because as Diks, et al. (1996) points out, the impact of order-splitting has never been explicitly studied.

2 The Model

Most notations and assumptions here follow Zhao and Simchi-Levi (2006) except for those on the demand process and the BOM matrix. For clarity, we summarize the notations and assumptions as follows: let $\mathcal{I}$ be the product set and $\mathcal{J}$ be the component set. For any component $j \in \mathcal{J}$, the continuous-time batch ordering policy has the reorder point (r_j and batch size Q_j). A continuous-time batch-ordering policy works as follows: whenever the inventory position (net inventory plus inventory on order) drops to or below a reorder point, an order of an integer number of the batch size is placed to raise the inventory position up to the smallest integer above the reorder point. We refer the reader to Zipkin (2000) for more discussions on this policy. We assume that the replenishment lead time of component $j \in \mathcal{J}$ is stochastic and sequential, and denoted by L_j. We further assume that the assembly cycle time is negligible with respect to the replenishment lead-times.

Demand follows independent compound Poisson processes with arrival rates λ^i and random sizes D^i, $i \in \mathcal{I}$, where $P\{D^i \geq 1\} = 1$. Demands are satisfied on a FCFS basis. For any demand that cannot be satisfied immediately, we assume it is fully backlogged. Consider one unit of a demand, if some of its required components are in stock but others are not, we put the in-stock components aside as "committed stock" (we refer the reader to Song and Zipkin 2003 for more explanation of this assumption).

Define $A = [a_j^i]$ to be the BOM matrix, i.e., assembling product i requires a_j^i units of component j. a_j^i is a non-negative integer. For convenience, we define $\mathcal{I}_j$ to be the set of products that require component j, $\mathcal{I}_j = \{i \in \mathcal{I} | a_j^i \geq 1\}$; and $\mathcal{J}^i$ to be the set of components required by product i, $\mathcal{J}^i = \{j \in \mathcal{J} | a_j^i \geq 1\}$. We finally define vector $\overline{X} = \{X^i, i \in \mathcal{I}\}$ to be the delivery lead-times of the products.

For the ease of exposition, we define the following notations. Let $\mathcal{S}_j = \{r_j+1, r_j+2, \ldots, r_j+Q_j\}$, $\mathcal{S}^i = \bigotimes_{j \in \mathcal{J}^i} \mathcal{S}_j$ and $\mathcal{S} = \bigotimes_{j \in \mathcal{J}} \mathcal{S}_j$. In a similar vein, let $\mathcal{Q}_j = \{1, 2, \ldots, Q_j\}$, $\mathcal{Q}^i = \bigotimes_{j \in \mathcal{J}^i} \mathcal{Q}_j$ and $\mathcal{Q} = \bigotimes_{j \in \mathcal{J}} \mathcal{Q}_j$.

3 Performance Analysis

For constant lead times and compound Poisson demand, Song (2000) shows that the inventory position vector of all the components are uniformly distributed in $\mathcal{S}$ under the assumption that the Markov chain of the inventory position vector of the components is aperiodic and irreducible. We refer to Song (2000) for sufficient conditions under which this assumption holds. Even if the assumption does not hold for some ATO systems, e.g., single-product assembly systems, we can still study these systems by assuming randomized initial inventory positions (Song 2000), which leads to uniformly distributed inventory position vector in $\mathcal{S}$.

All these results hold for stochastic sequential lead times and compound Poisson demand considered in this paper. This is true because the inventory position vector does not depend on the replenishment process (e.g., the lead times) but only depends on demand processes and the ordering policy (which remain unchanged from this paper to Song (2000)).

We first study the systems under the assumption of non-split orders. Consider a component $j \in \mathcal{J}$. First note that the demand process faced by component j is the superposition of the demand processes of all products $i \in \mathcal{I}_j$. Therefore, it is a compound Poisson process with size D_j such that $P\{D_j = D^i a^i_j\} = \frac{\lambda^i}{\sum_{l \in \mathcal{I}_j} \lambda^l}$. It is convenient to assign a priority list to different units of each demand so that inventories will be committed to these demand units according to this list. Clearly, under the non-split orders assumption, a demand is not satisfied until the last unit of this demand is satisfied.

Consider a product $i \in \mathcal{I}$ and one of its component $j \in \mathcal{J}^i$. We ask the following key question (Zhao and Simchi-Levi 2006): suppose a demand of size y' for product i arrives at time t, when is the corresponding order of the component j placed that completely satisfies this demand (i.e., satisfy the last unit of $y' a^i_j$)? Clearly, $y' \geq 1$ as $P\{D^i \geq 1\} = 1$ for all $i \in \mathcal{I}$. For simplicity, we denote $y = y' a^i_j$.

To answer this question, we note that because the demand process faced by component j is compound Poisson, the demand process counting backward starting from the arrival time of any demand is a compound Poisson process and is independent of the starting time. Hence, we suppress the product index i in the following analysis unless otherwise mentioned. We define the following notations. For the component j, we count *backwards* starting at t and let $D_{j,k}$, $k = 1, 2, \ldots$ be the size of the kth most recent demand arrival prior to t. Note that $D_{j,k}, k = 1, 2, \ldots$ can represent demands for different products that require component j. Let $IP_{j,k}$ be the inventory position right after the arrival of the kth most recent demand and the corresponding ordering decision. Finally, let $V_{j,k}$ be the kth most recent interarrival time, $k = 1, 2, \ldots$, e.g., $V_{j,1}$ is the time between the most recent demand arrival prior to t and t. For simplicity, we index the kth most recent demand arrival prior to t by the kth demand arrival.

It is more convenient to consider the complement of $IP_{j,k}$, $IP^c_{j,k}$, where $IP^c_{j,k} = r_j + Q_j - IP_{j,k}$.

By the continuous-time batch ordering policy, we must have,

$$IP^c_{j,k} \; = \; [IP^c_{j,k+1} + D_{j,k}] \bmod Q_j. \tag{1}$$

Here, $x \bmod y$ is the remainder on dividing x by y.

Lemma 3.1 *$IP^c_{j,k+1}$ is uniquely determined by $IP^c_{j,k}$ and $D_{j,k}$, where $k = 1, 2, \ldots$,*

Proof. Given $0 \le IP^c_{j,k} < Q_j$ and $D_{j,k} \ge 1$, Eq. (1) implies that for some $m = 0, 1, \ldots$, the following equations hold,

$$IP^c_{j,k+1} = IP^c_{j,k} - D_{j,k} + mQ_j \quad \text{and} \quad IP_{j,k} = IP_{j,k+1} - D_{j,k} + mQ_j. \tag{2}$$

Note that $0 \le IP^c_{j,k+1} < Q_j$, then m must satisfy,

$$0 \le IP^c_{j,k} - D_{j,k} + mQ_j < Q_j.$$

Clearly, there must exist a unique $m \ge 0$ so that the above inequalities are satisfied. The proof is now completed. ∎

Due to the batch ordering policies, the corresponding order of the component j that satisfies the last unit (i.e., the yth unit) of the demand realized at t, must be placed at one of the demand arrival times either at or prior to t. Let $IP_{j,1} = r_j + q_j$ where $q_j \in \mathcal{Q}_j$. We define $K_j(r_j + q_j, y)$ to be the index of the demand arrival at which the corresponding order is placed, where $K_j(r_j + q_j, y) = 0$ indicates that the corresponding order is placed at time t. To characterize the probability distribution of $K_j(r_j + q_j, y)$, we make the following observation.

Observation 3.2 *For $k = 1, 2, \ldots$, if $IP_{j,k} \ge \sum_{l=1}^{k-1} D_{j,l} + y$, then the corresponding order must be placed either at or prior to the kth demand arrival, that is, $K_j(r_j + q_j, y) \ge k$; otherwise, the corresponding order must be placed after the kth demand arrival, that is $K_j(r_j + q_j, y) < k$.*

This observation follows immediately from the assumption of the FCFS rule and the non-crossing property of the stochastic sequential lead-times. Indeed, this observation holds for any inventory policy that places orders only upon demand arrivals as long as the FCFS rule and non-crossing property hold. Based on this observation, we design the following procedure to identify $K_j(r_j + q_j, y)$ by looking backward from time t and checking on the inventory positions $IP_{j,k}$ for $k = 1, 2, \ldots$.

1. Right before t, let the inventory position $IP_{j,1} = r_j + q_j$. If $IP_{j,1} < y$, then it follows from Observation 3.2 that $K_j(r_j + q_j, y) < 1$. Since $K_j(r_j + q_j, y) \geq 0$, then $K_j(r_j + q_j, y) = 0$. Stop. Otherwise, if $IP_{j,1} \geq y$, then $K_j(r_j + q_j, y) \geq 1$. Calculate $IP_{j,2}$ by $IP_{j,1}$ and $D_{j,1}$ and continue.

2. If $IP_{j,2} < D_{j,1} + y$, then $K_j(r_j + q_j, y) < 2$ (Observation 3.2). It follows from the fact of $K_j(r_j + q_j, y) \geq 1$ that $K_j(r_j + q_j, y) = 1$. Stop. Otherwise, if $IP_{j,2} \geq D_{j,1} + y$, then $K_j(r_j + q_j, y) \geq 2$. Calculate $IP_{j,3}$ by $IP_{j,2}$ and $D_{j,2}$ and continue.

3. In general, given that $IP_{j,k} \geq \sum_{l=1}^{k-1} D_{j,l} + y$ for a $k \leq r_j + Q_j$, if $IP_{j,k+1} < \sum_{l=1}^{k} D_{j,l} + y$, then $K_j(r_j + q_j, y) = k$. Otherwise, if $IP_{j,k+1} \geq \sum_{l=1}^{k} D_{j,l} + y$, then $K_j(r_j + q_j, y) \geq k$. Calculate $IP_{j,k+1}$ by $IP_{j,k}$ and $D_{j,k}$ and continue.

Remarks:

- In each step of $k > 1$, $IP_{j,k}$ can be uniquely determined by $IP_{j,k-1}$ and $D_{j,k-1}$. This is true because the complement of $IP_{j,k}$, $IP^c_{j,k}$, is uniquely determined by $IP^c_{j,k-1}$ and $D_{j,k-1}$ (Lemma 3.1).

- $K_j(r_j + q_j, y) \leq r_j + Q_j$, due to Observation 3.2 and the fact that $P\{D^i \geq 1\} = 1, \forall i \in \mathcal{I}$.

- At step k, if $IP_{j,k} \geq \sum_{l=1}^{k-1} D_{j,l} + y$ and $IP_{j,k+1} < \sum_{l=1}^{k} D_{j,l} + y$, then an order must be placed right after the kth demand arrival. This is true because the above two inequalities imply that $IP_{j,k} > IP_{j,k+1} - D_{j,k}$, and therefore by Eq. (2), $IP_{j,k} = IP_{j,k+1} - D_{j,k} + mQ_j$ for some $m > 0$.

- For a k, if $IP_{j,k} \geq \sum_{l=1}^{k-1} D_{j,l} + y$, then $IP_{j,k-1} \geq \sum_{l=1}^{k-2} D_{j,l} + y$ must hold either in case of $IP_{j,k} - D_{j,k-1} = IP_{j,k-1}$ or in case of $IP_{j,k} - D_{j,k-1} < IP_{j,k-1}$. Furthermore, if $IP_{j,k} < \sum_{l=1}^{k-1} D_{j,l} + y$, then $IP_{j,k+1} < \sum_{l=1}^{k} D_{j,l} + y$ must hold either in case of $IP_{j,k+1} - D_{j,k} = IP_{j,k}$ or in case of $IP_{j,k+1} - D_{j,k} < IP_{j,k}$. Therefore, there exists a unique k such that $K_j(r_j + q_{j,y}) = k$.

The following Proposition summarizes the above results.

Proposition 3.3 *Under the assumption of non-split orders, consider a demand of size y arrives at time t and sees the inventory position of component $j \in \mathcal{J}$ at $IP_{j,1} = r_j + q_j$. Then the corresponding* *order of this component that satisfies the last unit of* y *is placed at time* $t - T_j(K_j(r_j + q_j, y))$*, where* $T_j(K_j(r_j + q_j, y)) = \sum_{k=1}^{K_j(r_j+q_j,y)} V_{j,k}$*, and* $K_j(r_j + q_j, y) \leq r_j + Q_j$ *is uniquely determined by,*

$$K_j(r_j + q_j, y) = \begin{cases} 0 & \text{if } r_j + q_j < y \\ k & \text{if } IP_{j,k} \geq \sum_{l=1}^{k-1} D_{j,l} + y \text{ and } IP_{j,k+1} < \sum_{l=1}^{k} D_{j,l} + y. \end{cases} \tag{3}$$

Due to the replenishment lead-time, the corresponding order of component j *is replenished at* $t - T_j(K_j(r_j + q_j, y)) + L_j$*, and the delay of this component is* $[L_j - T_j(K_j(r_j + q_j, y))]^+$*.*

Clearly, $K_j(r_j + q_j, y)$ is statistically different for different y or $r_j + q_j$.

In the special case of unit demand, i.e., Poisson demand processes, the joint distribution $P\{K_j(r_j + q_j, 1) = k_j, K_{\tilde{j}}(r_{\tilde{j}} + q_{\tilde{j}}, 1) = k_{\tilde{j}}\} = \frac{1}{Q_j}\frac{1}{Q_{\tilde{j}}}$ (see Zhao and Simchi-Levi 2006). But for compound Poisson demand, the joint distribution is much more complex. To see this, let's consider a component $j \in \mathcal{J}$ and a product $i \in \mathcal{I}_j$. We first derive the marginal probability distribution for $K_j(IP_{j,1}, D^i a_j^i)$. We need the following Lemma.

Lemma 3.4 *In steady state,* $IP_{j,k}$ *is independent of the demand sizes* $D_{j,l}$ *for* $l > k$*.*

Proof. We consider $P\{D_{j,k+1} = y_1, \ldots, D_{j,k+l} = y_l, IP_{j,k} = r_j + q_j\}$ in steady state, where $q_j \in \mathcal{Q}_j$. Because $IP_{j,k+l+1}$ is uniquely determined by $D_{j,k+1}, \ldots, D_{j,k+l}$ and $IP_{j,k}$ (Lemma 3.1), we can define $IP_{j,k+l+1} = \phi(D_{j,k+1}, \ldots, D_{j,k+l}, IP_{j,k})$. Then we must have

$$\begin{aligned} & P\{D_{j,k+1} = y_1, \ldots, D_{j,k+l} = y_l, IP_{j,k} = r_j + q_j\} \\ = {} & P\{D_{j,k+1} = y_1, \ldots, D_{j,k+l} = y_l, IP_{j,k+l+1} = \phi(y_1, \ldots, y_l, r_j + q_j)\} \\ = {} & P\{D_{j,k+1} = y_1, \ldots, D_{j,k+l} = y_l\} P\{IP_{j,k+l+1} = \phi(y_1, \ldots, y_l, r_j + q_j)\} \\ = {} & \frac{1}{Q_j} P\{D_{j,k+1} = y_1, \ldots, D_{j,k+l} = y_l\} \\ = {} & P\{D_{j,k+1} = y_1, \ldots, D_{j,k+l} = y_l\} P\{IP_{j,k} = r_j + q_j\}. \end{aligned}$$

The last two equalities are due to the steady state assumption. The proof is now completed. ∎

For the ease of exposition, we denote $K_j^i = K_j(IP_{j,1}, D^i a_j^i)$ and $D_{j,0} = D^i a_j^i$. Note that in steady-state, $IP_{j,1}$ is uniformly distributed in $\mathcal{S}_j$ and is independent of the demand sizes $D_{j,k}, k \geq 0$ (Lemma 3.4). By Proposition 3.3, we condition on $IP_{j,1} = r_j + q_j$ and arrive at,

$$P\{K_j^i = 0\} = \frac{1}{Q_j} \sum_{q_j=1}^{Q_j} P\{D_{j,0} > r_j + q_j\} = \frac{1}{Q_j} \sum_{q_j=1}^{Q_j} P\{D^i a_j^i > r_j + q_j\}. \tag{4}$$

For $0 < k \le r_j + Q_j$, conditioning on $IP_{j,k+1} = r_j + q_j$ and $D_{j,k} = z$ yields,

$$P\{K_j^i = k\} \;=\; \frac{1}{Q_j}\sum_{q_j=1}^{Q_j}\sum_{z=1}^{\infty} P\{D_{j,k} = z\}P\{r_j + q_j - z < \sum_{l=0}^{k-1} D_{j,l} \le IP_{j,k}\}, \tag{5}$$

where $IP_{j,k} = r_j + Q_j - IP^c_{j,k}$ and $IP^c_{j,k} = [Q_j - q_j + z] \bmod Q_j$. If $z \le q_j - 1$, then $P\{r_j + q_j - z < \sum_{l=0}^{k-1} D_{j,l} \le IP_{j,k}\} = 0$ since $IP_{j,k} = r_j + q_j - z$. In addition, if we define $z = mQ_j + d$ where $d = q_j, q_j+1, \ldots, Q_j + q_j - 1$, then $IP_{j,k}$ varies from $r_j + Q_j$, $r_j + Q_j - 1$ to $r_j + 1$ as d varies from q_j, $q_j + 1$ to $Q_j + q_j - 1$. Combining these facts with Eq. (5) yields,

$$\begin{aligned} P\{K_j^i = k\} \quad &= \quad \frac{1}{Q_j}\sum_{q_j=1}^{Q_j}\sum_{m=0}^{\infty}\sum_{d=q_j}^{Q_j+q_j-1} P\{D_{j,k} = mQ_j + d\} \\ &\quad \times P\{r_j + q_j - mQ_j - d < \textstyle\sum_{l=0}^{k-1} D_{j,l} \le r_j + Q_j + q_j - d\}. \end{aligned} \tag{6}$$

To characterize the joint distribution, we note that for the demand of product i that arrives at time t, K_j^i for different components $j \in \mathcal{J}^i$ are clearly dependent because these components face the common demand process of product i. To demonstrate the dependence, we consider two components $j, \tilde{\jmath} \in \mathcal{J}^i$ and a special case in which both of these components face the identical demand processes, that is, $\mathcal{I}_j = \mathcal{I}_{\tilde{\jmath}}$. Therefore, the demand size processes $\{D_{j,k}, k = 0, 1, \ldots\}$ and $\{D_{\tilde{\jmath},k}, k = 0, 1, \ldots\}$ are identical, and the interarrival processes $\{V_{j,k}, k = 1, 2, \ldots\}$ and $\{V_{\tilde{\jmath},k}, k = 1, 2, \ldots\}$ are identical. For simplicity, we use the demand process, i.e., the demand size process and the interarrival process, associated with the component j for both components.

Because the inventory position vector of all components is uniformly distributed in $\mathcal{S}$, $IP_{j,1}$ is independent of $IP_{\tilde{\jmath},1}$. Hence,

$$\begin{aligned} P\{K_j^i = 0, K_{\tilde{\jmath}}^i = 0\} &= P\{D_{j,0} > IP_{j,1}, D_{j,0} > IP_{\tilde{\jmath},1}\} \\ &= \frac{1}{Q_j Q_{\tilde{\jmath}}}\sum_{q_j=1}^{Q_j}\sum_{q_{\tilde{\jmath}}=1}^{Q_{\tilde{\jmath}}} P\{D^i a_j^i > r_j + q_j, D^i a_j^i > r_{\tilde{\jmath}} + q_{\tilde{\jmath}}\}. \end{aligned} \tag{7}$$

For $k > 0$, since $IP_{j,k+1}$ is independent of $IP_{\tilde{\jmath},k+1}$, and both of them are independent of future demands,

$$P\{K_j^i = k, K_{\tilde{\jmath}}^i = k\} = P\{IP_{j,k+1} - D_{j,k} < \sum_{l=0}^{k-1} D_{j,l} \le IP_{j,k},\ IP_{\tilde{\jmath},k+1} - D_{j,k} < \sum_{l=0}^{k-1} D_{j,l} \le IP_{\tilde{\jmath},k}\}$$

$$= \frac{1}{Q_j Q_{\tilde{j}}} \sum_{q_j=1}^{Q_j} \sum_{q_{\tilde{j}}=1}^{Q_{\tilde{j}}} \sum_{y=\max\{q_j, q_{\tilde{j}}\}}^{\infty} P\{D_{j,k} = y\} \times$$
$$\times P\{r_j + q_j - y < \sum_{l=0}^{k-1} D_{j,l} \leq IP_{j,k},\ r_{\tilde{j}} + q_{\tilde{j}} - y < \sum_{l=0}^{k-1} D_{j,l} \leq IP_{\tilde{j},k}\}, \quad (8)$$

where $IP_{j,k} = r_j + Q_j - [Q_j - q_j + y] \bmod Q_j$ and $IP_{\tilde{j},k} = r_{\tilde{j}} + Q_{\tilde{j}} - [Q_{\tilde{j}} - q_{\tilde{j}} + y] \bmod Q_{\tilde{j}}$.

We now consider $P\{K_j = k_j, K_{\tilde{j}} = k_{\tilde{j}}\}$ for $0 < k_j < k_{\tilde{j}}$. Since $IP_{\tilde{j},k_{\tilde{j}}+1}$ is independent of $IP_{j,k_{\tilde{j}}+1}$, and $IP_{\tilde{j},k_{\tilde{j}}+1}$ is independent of the future demand sizes $D_{j,l}$ for $l = k_{\tilde{j}}, \ldots, 0$, thus $IP_{\tilde{j},k_{\tilde{j}}+1}$ is independent of IP_{j,k_j+1}. Combining this fact with Lemma 3.4, we can condition on $IP_{\tilde{j},k_{\tilde{j}}+1}$ and IP_{j,k_j+1}, and arrive at

$$P\{K_j^i = k_j, K_{\tilde{j}}^i = k_{\tilde{j}}\} =$$
$$= P\{IP_{j,k_j+1} - D_{j,k_j} < \sum_{l=0}^{k_j-1} D_{j,l} \leq IP_{j,k_j},\ IP_{\tilde{j},k_{\tilde{j}}+1} - D_{j,k_{\tilde{j}}} < \sum_{l=0}^{k_{\tilde{j}}-1} D_{j,l} \leq IP_{\tilde{j},k_{\tilde{j}}}\}$$
$$= \frac{1}{Q_j Q_{\tilde{j}}} \sum_{q_j=1}^{Q_j} \sum_{q_{\tilde{j}}=1}^{Q_{\tilde{j}}} P\{r_j + q_j - D_{j,k_j} < \sum_{l=0}^{k_j-1} D_{j,l} \leq IP_{j,k_j}, r_{\tilde{j}} + q_{\tilde{j}} - D_{j,k_{\tilde{j}}} < \sum_{l=0}^{k_{\tilde{j}}-1} D_{j,l} \leq IP_{\tilde{j},k_{\tilde{j}}}\}. \quad (9)$$

In view of Eq. (6), further conditioning on D_{j,k_j} and $D_{j,k_{\tilde{j}}}$ yields,

$$P\{K_j^i = k_j, K_{\tilde{j}}^i = k_{\tilde{j}}\} =$$
$$= \frac{1}{Q_j Q_{\tilde{j}}} \sum_{q_j=1}^{Q_j} \sum_{q_{\tilde{j}}=1}^{Q_{\tilde{j}}} \sum_{m_j=0}^{\infty} \sum_{d_j=q_j}^{Q_j+q_j-1} P\{D_{k_j} = m_j Q_j + d_j\} \sum_{m_{\tilde{j}}=0}^{\infty} \sum_{d_{\tilde{j}}=q_{\tilde{j}}}^{Q_{\tilde{j}}+q_{\tilde{j}}-1} P\{D_{k_{\tilde{j}}} = m_{\tilde{j}} Q_{\tilde{j}} + d_{\tilde{j}}\} \times$$
$$\times P\{r_j + q_j - m_j Q_j - d_j < \sum_{l=0}^{k_j-1} D_{j,l} \leq r_j + Q_j + q_j - d_j,$$
$$r_{\tilde{j}} + q_{\tilde{j}} - m_{\tilde{j}} Q_{\tilde{j}} - d_{\tilde{j}} < \sum_{l=k_j+1}^{k_{\tilde{j}}-1} D_{j,l} + m_j Q_j + d_j + \sum_{l=0}^{k_j-1} D_{j,l} \leq r_{\tilde{j}} + Q_{\tilde{j}} + q_{\tilde{j}} - d_{\tilde{j}}\}. \quad (10)$$

Once we have the joint distribution of K_j^i and $K_{\tilde{j}}^i$, we can easily write out the joint probability density function of $T_j(K_j^i)$ and $T_{\tilde{j}}(K_{\tilde{j}}^i)$. Let's assume $k_j \leq k_{\tilde{j}}$ without loss of generality. Since the demand size process of the component j is independent of its inter-arrival time process, it follows from Zhao and Simchi-Levi (2006) that conditioning on $K_j^i = k_j$ and $K_{\tilde{j}}^i = k_{\tilde{j}}$ yields,

$$P\{T_j(k_j) = t_j, T_{\tilde{j}}(k_{\tilde{j}}) = t_{\tilde{j}}\} \ = \ P\{T_j(k_j) = t_j\} P\{T_j(k_{\tilde{j}} - k_j) = t_{\tilde{j}} - t_j\}. \quad (11)$$

Eqs. (7)-(11) imply that the K^i_j and $K^i_{\tilde{j}}$, and the $T_j(K^i_j)$ and $T^i_{\tilde{j}}(K_{\tilde{j}})$ are highly dependent due to the common demand size process and the common interarrival time process faced by components j and $\tilde{j}$. For the more general case where components j and $\tilde{j}$ satisfy $\mathcal{I}_j \bigcap \mathcal{I}_{\tilde{j}} \neq \emptyset$ but $\mathcal{I}_j \neq \mathcal{I}_{\tilde{j}}$, the exact form of the joint distribution of K_j and $K_{\tilde{j}}$, and of $T_j(K_j)$ and $T_{\tilde{j}}(K_{\tilde{j}})$ are much more complex. In what follows, we shall not characterize the joint probability density functions analytically for the general case, but rather, we develop efficient numerical methods to estimate the key system performance metrics (see Section 4). For this purpose, we need the exact sample-path expressions for the system performance metrics.

Suppose that the demand size of product $i \in \mathcal{I}$ that arrives at time t is D^i, we denote $\overline{IP}^i = (IP_{j,1}, j \in \mathcal{J}^i)$ to be the inventory position vector of the components $j \in \mathcal{J}^i$ seen by this demand. Clearly, $\overline{IP}^i$ is uniformly distributed in $\mathcal{S}^i$ (Song 2000). Let $X^i(\overline{IP}^i, D^i)$ be the delivery lead-time of this demand given $\overline{IP}^i$ and D^i. By Proposition 3.3,

$$X^i(\overline{IP}^i, D^i) \;=\; \max_{j \in \mathcal{J}^i}\{[L_j - T_j(K_j(IP_{j,1}, D^i a^i_j))]^+\}. \tag{12}$$

Let $X^i(z)$ be the delivery lead-time conditioning on $D^i = z$. Then, the expected delivery lead-time $E(X^i(z))$ for $i \in \mathcal{I}$ can be characterized by

$$E(X^i(z)) \;=\; \frac{1}{\prod\limits_{j \in \mathcal{J}^i} Q_j} \sum_{\overline{q}^i \in \mathcal{Q}^i} E(X^i(\overline{r}^i + \overline{q}^i, z)), \tag{13}$$

where $\overline{r}^i = (r_j, j \in \mathcal{J}^i)$ and $\overline{q}^i = (q_j, j \in \mathcal{J}^i)$. The fill rate for a target service time $\tau \geq 0$ is given by

$$\begin{aligned} P\{X^i(z) \leq \tau\} \;&=\; \frac{1}{\prod\limits_{j \in \mathcal{J}^i} Q_j} \textstyle\sum_{\overline{q}^i \in \mathcal{Q}^i} P\{X^i(\overline{r}^i + \overline{q}^i, z) \leq \tau\} \\ &=\; \frac{1}{\prod\limits_{j \in \mathcal{J}^i} Q_j} \textstyle\sum_{\overline{q}^i \in \mathcal{Q}^i} P\{L_j - T_j(K_j(r_j + q_j, z a^i_j)) \leq \tau, \forall j \in \mathcal{J}^i\}. \end{aligned} \tag{14}$$

In the special case of Poisson demand and unit BOM matrix (i.e., $a^i_j = 0$ or 1, $\forall i, j$), the random vector $(K_j(IP_{j,1}, D^i a^i_j),\ j \in \mathcal{J}_i)$ is uniformly distributed in $\mathcal{S}^i$ (see Proposition 5.2 in Zhao and Simchi-Levi 2006). However, this result does not hold for compound Poisson demand and general BOM matrix because $K_j(r_j + q_j, D^i a^i_j)$ now depends on demand size $D^i a^i_j$ and the batch ordering policy (Proposition 3.3). Indeed, as we have shown before, the marginal and joint probability distribution of $K_j(IP_{j,1}, D^i a^i_j), j \in \mathcal{J}_i$ for compound Poisson demand become much more complex than their counter-parts for Poisson demand.

We next consider the system under the assumption of split orders. As opposed to the assumption of non-split orders, each unit in a demand can now be satisfied separately. We now ask the following key question (Zhao and Simchi-Levi 2006): suppose a demand of size y arrives at time t for a component $j \in \mathcal{J}$, and the demand sees the inventory position of this component at $r_j + q_j$, then when is the corresponding order of this component placed that satisfies the nth unit of this demand? where $1 \le n \le y$.

Notice that inventories are committed to each demand unit in the same way under either split orders assumption or non-split orders assumption. Thus, we can answer the above question by considering a demand of size n in the non-split order case. That is,

Observation 3.5 *Under the assumption of split orders, the corresponding order of the component j that satisfies the nth unit of the demand that arrives at time t, is placed at time $t - T_j(K_j(r_j+q_j, n))$, where $T_j(\cdot)$ and $K_j(r_j + q_j, n)$ are defined in Proposition 3.3. Due to the replenishment lead-time, the corresponding order of the component j is replenished at $t - T_j(K_j(r_j + q_j, n)) + L_j$, and the delay of this component is $[L_j - T_j(K_j(r_j + q_j, n))]^+$.*

Clearly, $K_j(r_j + q_j, n)$ is statistically different if n or $r_j + q_j$ is different.

The performance measures of the ATO systems under the assumption of split orders can be characterized in a similar way as those under the assumption of non-split orders. By Eq. (12), the delivery lead-time for the nth unit of a demand for product $i \in \mathcal{I}$ is

$$X^i(\overline{IP}^i, n) \; = \; \max_{j \in \mathcal{J}^i} \{[L_j - T_j(K_j(IP_{j,1}, na^i_j))]^+\}, \tag{15}$$

where $\overline{IP}^i = (IP_{j,1}, j \in \mathcal{J}^i)$ is uniformly distributed in $\mathcal{S}^i$. By Eqs. (13)-(14), the expected delivery lead-time and the fill rate for the nth unit in a demand for product i are given by

$$E(X^i(n)) = \frac{1}{\prod_{j \in \mathcal{J}^i} Q_j} \sum_{\overline{q}^i \in \mathcal{Q}^i} E(X^i(\overline{r}^i + \overline{q}^i, n)), \tag{16}$$

$$\begin{aligned} P\{X^i(n) \le \tau\} &= \frac{1}{\prod_{j \in \mathcal{J}^i} Q_j} \sum_{\overline{q}^i \in \mathcal{Q}^i} P\{X^i(\overline{r}^i + \overline{q}^i, n) \le \tau\} \\ &= \frac{1}{\prod_{j \in \mathcal{J}^i} Q_j} \sum_{\overline{q}^i \in \mathcal{Q}^i} P\{L_j - T_j(K_j(r_j + q_j, na^i_j)) \le \tau, \forall j \in \mathcal{J}^i\}. \end{aligned} \tag{17}$$

We now compare the system performances between the case of split orders and the case of non-split orders. To gain insight into the impact of order splitting, we first consider a single-product single-component system. We name the component by j for convenience, and assume that a demand of size y arrives at time t. In the case of split orders, the delivery lead-time for the nth unit is $[L_j - T_j(K_j(r_j + q_j, n))]^+$ where $n = 1, 2, \ldots, y$; while in the case of non-split orders, the delivery lead-times are identical for all units, that is $[L_j - T_j(K_j(r_j + q_j, y))]^+$. Because $T_j(K_j(r_j+q_j, n)) \geq_{st} T_j(K_j(r_j+q_j, y))$ (by Proposition 3.3) where $\geq_{st}$ denotes the stochastic order, the delivery lead-time in the former case is stochastically smaller than or equal to those in the later case for any unit of demand.

To analyze the waiting time of the corresponding order, we introduce the following notation. Let $L_{j,k}$ be the lead-time of the order triggered by the kth most recent demand arrival prior to t, where $k = 1, 2, \ldots$. To simplify the notation, we uncondition on $IP_{j,1} = r_j + q_j$ and replace $K_j(r_j + q_j, y)$ by $K_j(y)$. In the case of split orders, the corresponding order that satisfies the nth unit is replenished at time $t - T_j(K_j(n)) + L_{j,K_j(n)}$. Therefore, the waiting time of this order is

$$[t - (t - T_j(K_j(n)) + L_{j,K_j(n)})]^+ \; = \; [T_j(K_j(n)) - L_{j,K_j(n)}]^+. \tag{18}$$

In the case of non-split orders, the corresponding order that satisfies the nth unit is replenished at the same time as in the case of split orders, but the nth unit will not be filled until the last unit (the yth unit) of the same demand is satisfied. Thus, the total waiting time of the corresponding order that satisfies the nth unit is

$$\begin{aligned} & [\max\{t, t - T_j(K_j(y)) + L_{j,K_j(y)}\} - (t - T_j(K_j(n)) + L_{j,K_j(n)})]^+ \\ = \; & [[L_{j,K_j(y)} - T_j(K_j(y))]^+ + T_j(K_j(n)) - L_{j,K_j(n)}]^+. \end{aligned} \tag{19}$$

Comparing Eqs. (18)-(19), we can see that the waiting time of the corresponding order is stochastically smaller under the assumption of split orders than that under the assumption of non-split orders. In addition, the waiting time distribution of the corresponding order is easier to characterize under the assumption of split orders than under the assumption of non-split orders. This is true because the waiting time distribution of the former only depends on the marginal distribution of L_j, while the waiting time distribution of the latter depends on the joint probability distribution of $L_{j,k}, k = 1, 2, \ldots$. Since characterizing the joint distribution of $L_{j,k}$ requires an extension of the "stochastic sequential lead time" model, we shall leave the waiting time distribution for a future study.

Applying the same logic to each component, the insights of the single-product single-component system transfer to ATO systems. To quantify the impact of order splitting, we conduct a numerical study in Section 5.

Finally, we study the impact of component dependence on system performance.

Proposition 3.6 *For all $j \in \mathcal{J}^i$, define $K'_j(\cdot,\cdot)$ to be independent copy of $K_j(\cdot,\cdot)$, that is, $K'_j(\cdot,\cdot)$ has the same marginal distribution as $K_j(\cdot,\cdot)$ but $K'_j(\cdot,\cdot)$ are mutually independent. Then for any $\overline{l}$, $\overline{q} \in \mathcal{Q}$ and $\tau_j \geq 0, j \in \mathcal{J}$, as well as for all $i \in \mathcal{I}$, we must have,*

$$P\{l_j - T_j(K_j(r_j+q_j,n)) \leq \tau_j, j \in \mathcal{J}^i\} \geq P\{l_j - T_j(K'_j(r_j+q_j,n)) \leq \tau_j, j \in \mathcal{J}^i\} \tag{20}$$

$$\geq \Pi_{j \in \mathcal{J}^i} P\{l_j - T_j(K'_j(r_j+q_j,n)) \leq \tau_j\}. \tag{21}$$

Proof. To prove inequality (20), we condition on the interarrival times of all products. First note that the demand sizes of all products are associated because they are independent random variables (Part (d) of Theorem 5.2.2 of Tong 1980). Furthermore, $K_j(r_j+q_j,n)$ is a non-increasing function of the demand sizes of all products, which implies that $l_j - T_j(K_j(r_j+q_j,n))$ is non-decreasing function of the demand sizes. It follows from Theorem 5.2.3 that conditioning on the interarrival times of all products, $l_j - T_j(K_j(r_j+q_j,n)), \forall j \in \mathcal{J}^i$ are associated random variables for any $i \in \mathcal{I}$, and therefore, inequality (20) follows from Theorem 5.2.4 of Tong (1980). Unconditioning on the interarrival times of all products yields the desired result.

To prove inequality (21), we further observe that the interarrival times of all products are independent, thus they are associated. Conditioning on $K'_j(r_j+q_j,n) = k_j$ for all $j \in \mathcal{J}^i$, $T_j(k_j)$s are associated because they are non-decreasing functions of the interarrival times. Finally, by Theorem 5.2.4 of Tong (1980), unconditioning on $K'_j(r_j+q_j,n)$ yields the desired result. ∎

Proposition 3.6 implies that ignoring the dependence among demand sizes or the dependence among interarrival times or both results in stochastically larger delivery lead-times. Proposition 3.6 applies to ATO systems under either the split orders assumption or the non-split orders assumption. In the former, the delivery lead-times are associated with demand of different sizes, while in the latter, the delivery lead-times are associated with different units in one demand.

4 Numerical Method

In this section, we develop a numerical method based-on Monte Carlo simulation to estimate the expected delivery lead-times and the order-based fill rates for the ATO systems under either split orders or non-split orders assumption. As we point out in Section 1, because Proposition 5.2 of Zhao and Simchi-Levi (2006) does not hold here, the numerical method by that paper does not directly apply. Interestingly, the numerical method can be modified to treat compound Poisson processes and general BOM matrix, as shown in this section.

The key difference between Poisson demand and compound Poisson demand is the demand size processes, which significantly complicates the analysis of the joint probability distribution of the delivery lead-times but can be randomly sampled without significantly more effort. Furthermore, the logic of the numerical method by Zhao and Simchi-Levi (2006) applies here with some modification on the way that K_j is calculated.

Due to the similarity between Eqs. (13)-(14) and Eqs. (16)-(17), we shall focus on the systems under the assumption of non-split orders in the rest of this section. The numerical method can be easily modified to handle split-orders. To develop efficient numerical method to estimate $E(X^i(z))$ and $P\{X^i(z) \leq \tau\}$ for a particular z, we first re-write Eqs. (13)-(14) as follows,

$$E(X^i(z)) \;=\; E(\frac{1}{\prod\limits_{j\in\mathcal{J}^i} Q_j}\sum_{\overline{q}^i\in\mathcal{Q}^i} X^i(\overline{r}^i+\overline{q}^i,z)) \;=\; E(\tilde{X}^i(z)), \tag{22}$$

$$P\{X^i(z)\leq\tau\} \;=\; E(\frac{1}{\prod\limits_{j\in\mathcal{J}^i} Q_j}\sum_{\overline{q}^i\in\mathcal{Q}^i} 1_{\{L_j-T_j(K_j(r_j+q_j,za^i_j))\leq\tau,\forall j\in\mathcal{J}^i\}}) \;=\; E(\Xi^i(\tau,z)), \tag{23}$$

where,

$$\tilde{X}^i(z) = \frac{1}{\prod\limits_{j\in\mathcal{J}^i} Q_j}\sum_{\overline{q}^i\in\mathcal{Q}^i} X^i(\overline{r}^i+\overline{q}^i,z) \tag{24}$$

$$\Xi^i(\tau,z) = \frac{1}{\prod\limits_{j\in\mathcal{J}^i} Q_j}\sum_{\overline{q}^i\in\mathcal{Q}^i} 1_{\{L_j-T_j(K_j(r_j+q_j,za^i_j))\leq\tau,\forall j\in\mathcal{J}^i\}}. \tag{25}$$

and $1_{\{L_j-T_j(K_j(r_j+q_j,za^i_j))\leq\tau,\forall j\in\mathcal{J}^i\}}$ is the indicator function of the event $\{L_j-T_j(K_j(r_j+q_j,za^i_j))\leq\tau,\forall j\in\mathcal{J}^i\}$. Note that $\tilde{X}^i(z)$ is an expansion of $X^i(z)$ in the form of the total probability conditioning on the inventory position vector. Because $\tilde{X}^i(z)$ and $X^i(z)$ have the same mean, we shall estimate the former.

To estimate $\tilde{X}^i(z)$, we utilize Eq. (24) and thus Eq. (12). Note that the stochastic sequential lead time assumption is embedded in Eq. (12), which allows us to use a simple sampling method as follows: To get a sample of $\tilde{X}^i(z)$, we keep track of one demand arrival of product i as well as the corresponding replenishment of each component that satisfies this demand. By Eq. (12), we only need to generate one sample for the lead time of each component, and at most $\max_{j\in\mathcal{J}}\{r_j + Q_j\}$ many samples for each product demand arrivals.

For random variables $X^i(z)$, $X^i(\overline{r}^i + \overline{q}^i, z)$, $\tilde{X}^i(z)$, $\Xi^i(\tau, z)$, L_j, $T_j(\cdot)$ and $K_j(\cdot,\cdot)$, we use the following notation to denote their samples, $x^i(z)$, $x^i(\overline{r}^i + \overline{q}^i, z)$, $\tilde{x}^i(z)$, $\xi^i(\tau, z)$, l_j, $t_j(\cdot)$ and $k_j(\cdot,\cdot)$. Eqs. (24)-(25) imply that we can generate a common sample of the lead-times and demand processes for $X^i(\overline{r}^i + \overline{q}^i, z)$ with different $\overline{q}^i \in \mathcal{Q}^i$, as follows,

1. *Sampling lead times.* Generate a sample of L_j, l_j, independently for all $j \in \mathcal{J}$.

2. *Sampling product demand arrivals.* For each product $i \in \mathcal{I}$, sample the demand inter-arrival times and demand sizes for $\max_{j\in\mathcal{J}}\{r_j + Q_j\}$ many demand arrivals.

3. *Calculating demand arrival process for each component.* For each component $j \in J$, superimpose the demand processes of all products $i \in \mathcal{I}_j$ to determine the demand process for component j. Then determine $k_j(r_j + q_j, za^i_j)$ according to Proposition 3.3, and consequently $t_j(k_j(r_j + q_j, za^i_j))$ for each $q_j \in \mathcal{Q}_j$.

4. *Calculating delays for each component.* Finally, for each component j, compute $l_j - t_j(k_j(r_j + q_j, za^i_j))$ for each $q_j \in \mathcal{Q}_j$.

Based on the common sample, we design an efficient numerical method, namely Method A-CP (where CP stands for Compound Poisson demand), to compute

$$\tilde{x}^i(z) \;=\; \frac{1}{\prod_{j\in\mathcal{J}^i} Q_j} \sum_{\overline{q}^i\in\mathcal{Q}^i} x^i(\overline{r}^i + \overline{q}^i, z), \tag{26}$$

where $x^i(\overline{r}^i + \overline{q}^i, z) = \max_{j\in\mathcal{J}^i}\{[l_j - t_j(k_j(r_j + q_j, za^i_j))]^+\}$.

Method A-CP is based on the following observation. Given a particular sample $l_j - t_j(k_j(r_j + q_j, za^i_j)) > 0$, there may exist multiple $\overline{q}^i \in \mathcal{Q}^i$ such that $x^i(\overline{r}^i + \overline{q}^i, z) = l_j - t_j(k_j(r_j + q_j, za^i_j))$.

For instance, suppose in one sample, $l_1 - t_1(k_1(r_1 + q^0, za^i_j)) = \max_{\overline{q}^i \in \mathcal{Q}^i}\{l_j - t_j(k_j(r_j + q_j, za^i_j))\}$ for a q^0, then for all $\overline{q}^i$ such that $q_1 = q^0$, $x^i(\overline{r}^i + \overline{q}^i, z) = l_1 - t_1(k_1(r_1 + q^0, za^i_j))$. We refer the reader to Zhao and Simchi-Levi (2006) for more discussions.

Method A-CP:

1. Sort the real numbers $[l_j - t_j(k_j(r_j + q_j, za^i_j))]^+, \forall q_j \in \mathcal{Q}_j$ and $\forall j \in \mathcal{J}^i$ into a non-increasing sequence, and denote the sequence by $\delta^i_n, n = 1, 2, \cdots, \sum_{j \in \mathcal{J}^i} Q_j$.

2. Set $n = 1$, $\tilde{x}^i(z) = 0$ and $R_j = Q_j, \forall j \in \mathcal{J}^i$.

3. If $\delta^i_n = 0$, output $\tilde{x}^i(z)$ and stop.

   Otherwise, identify the corresponding component associated with δ^i_n, namely, j_n.

   Then add all $x^i(\overline{r}^i + \overline{q}^i, z)$ which equals to δ^i_n by

$$\tilde{x}^i(z) \;=\; \tilde{x}^i(z) + \delta^i_n \times \prod_{j \neq j_n, j \in \mathcal{J}^i} R_j / \prod_{j \in \mathcal{J}^i} Q_j, \tag{27}$$

   and delete δ^i_n by

$$R_{j_n} \;=\; R_{j_n} - 1. \tag{28}$$

   If $R_{j_n} = 0$, output $\tilde{x}^i(z)$ and stop; otherwise, $n = n + 1$, and repeat this step. ∎

Clearly, Method A-CP always stops before n reaches $\sum_{j \in \mathcal{J}^i} Q_j$ because R_j for some $j \in \mathcal{J}^i$ will reach zero before n reaches $\sum_{j \in \mathcal{J}^i} Q_j$. In view of Eq. (24), the variance of a random sample of $\tilde{X}^i(z)$ is bounded from above by the maximum variance of $X^i(\overline{r}^i + \overline{q}^i, z)$ over all $\overline{q}^i \in \mathcal{Q}^i$.

We now analyze the computational complexity of Method A-CP. Given the sample size, the computing time for generating the lead-times, inter-arrival times, and superimposing demand processes is at most proportional to $|\mathcal{J}| + |\mathcal{I}| \times \max_{j \in \mathcal{J}}\{r_j + Q_j\} + |\mathcal{I}| \times \sum_{j \in \mathcal{J}}(r_j + Q_j)$. Calculating $k_j(r_j + q_j, za^i_j)$ for all q_j and j requires a computing time at most proportional to $|\mathcal{J}| \max_{j \in \mathcal{J}}\{r_j + Q_j\} \max_{j \in \mathcal{J}} Q_j$. Computing $[l_j - t_j(k_j(r_j + q_j, za^i_j))]^+, \forall q_j, \forall j$ and sorting these numbers require a computing time proportional to $\sum_{j \in \mathcal{J}^i}(r_j + Q_j)$ and $(\sum_{j \in \mathcal{J}^i} Q_j) log(\sum_{j \in \mathcal{J}^i} Q_j))$ respectively. Finally, Step 3 takes a computing time at most proportional to $\sum_{j \in \mathcal{J}} Q_j$. Thus, the overall computational effort for generating a sample of $\tilde{X}^i(z)$ for all $i \in \mathcal{I}$ is at most proportional to $|\mathcal{I}| \times \sum_{j \in \mathcal{J}}(r_j + Q_j) + |\mathcal{J}| \max_{j \in \mathcal{J}}\{r_j + Q_j\} \max_{j \in \mathcal{J}} Q_j + |\mathcal{I}| \times (\sum_{j \in \mathcal{J}} Q_j) log(\sum_{j \in \mathcal{J}} Q_j)$.

We next design an efficient numerical method, namely Method B-CP, to compute

$$\xi^i(\tau, z) \;=\; \frac{1}{\prod\limits_{j\in\mathcal{J}^i} Q_j} \sum_{\overline{q}^i\in\mathcal{Q}^i} 1_{\{l_j - t_j(k_j(r_j+q_j, za^i_j))\le\tau, \forall j\in\mathcal{J}^i\}}. \tag{29}$$

Method B-CP:

1. For each component $j \in \mathcal{J}_i$, count the number of $q_j \in \mathcal{Q}_j$ so that $l_j - t_j(k_j(r_j + q_j, za^i_j)) \le \tau$, and denote it by Q'_j,

2. Calculate $\xi^i(\tau, z) = \prod_{j\in\mathcal{J}^i} Q'_j / \prod_{j\in\mathcal{J}^i} Q_j$. ∎

Method B-CP is based on the fact that $1_{\{l_j - t_j(k_j(r_j+q_j, za^i_j))\le\tau, \forall j\in\mathcal{J}^i\}} = 1$ if and only if $l_j - t_j(k_j(r_j + q_j, za^i_j)) \le \tau$ for all $j \in \mathcal{J}^i$. Given the sample size, the computational complexity of Method B-CP is proportional to $|\mathcal{I}| \times \sum_{j\in\mathcal{J}} Q_j + |\mathcal{I}| \times |\mathcal{J}|$. The variance of a random sample of $\Xi^i(\tau, z)$ generated by Method B-CP is less than the maximum variance of the indicator functions $1(X^i(\overline{r}^i + \overline{q}^i, z) \le \tau)$ over $\overline{q}^i \in \mathcal{Q}^i$.

5 A Numerical Study

In order to study the efficiency of the proposed numerical method and the impact of order splitting, we consider a numerical example motivated by a real world problem: the Dell *Dimension 3000* desktop computer for small business (see http://zhao.rutgers.edu/ for more information). The computer has 27 categories of components or softwares which can be customized; among which, 15 categories are non software related. We choose to focus on 12 out of these 15 categories which seem to be necessary for most computers. Each category has multiple options for customization. The categories (number of options) are, processor (4), memory (3), keyboard (3), mouse (3), hard disk (2), CD or DVD (10), Monitor (2), sound card (2), floppy & memory keys (7), speakers (4), modem (3) and wireless routers (4). The total number of options is 47.

Due to the lack of demand and supply data, we construct the example as follows. We first create demand types (i.e., products) in the following way: We assume that there exists a demand type that chooses all the base-line options. We also assume that for every category, there exists a demand type that chooses an option other than the baseline only in that category; We further assume that for every two categories, there exits a demand that chooses options other than the baselines only in those two categories. Thus, the total number of demand types (or products) is 567.

For each component, we assume that the replenishment lead-time follows Erlang distribution with $E(L_j)$ being randomly generated from $uniform(0.1, 1)$ and n_j from $uniform\{4, 5, \ldots, 16\}$ (see Zipkin 2000 pg. 457 for the definition of n_j). The re-order points $r_j = \tilde{r}_j \times \beta, j \in \mathcal{J}$ where $\tilde{r}_j, j \in \mathcal{J}$ are randomly generated from $uniform\{1, 2, 3, 4\}$ and $\beta \in \{1, 2, 4, 6, 8\}$. The ordering quantities $Q_j, j \in \mathcal{J}$ are randomly generated by $uniform\{4, 5, 6, 7, 8\}$. Therefore, the maximum possible $r_j + Q_j$ is 40, which is the largest value that we can choose due to the memory limit of my laptop computer.

For each demand type (or product), the demand size follows uniform distribution in $\{1, 2, \ldots, u^i\}$ where u^i is randomly generated by $uniform\{2, 3, \ldots, 8\}$. The demand arrival rate λ^i is generated randomly by $uniform(0.1, 1) \times \tilde{\lambda}$, where $\tilde{\lambda} = 0.02$ is chosen so that the weighted fill-rates across all products (see Eqs. (35) and (37)) fall in a range of practical interests. Due to the large number of products, the cumulative arrival rates for a component may be significant unless the arrival rate of each product is sufficiently small. Finally, it is reasonable to assume, in this example, that each element of the BOM matrix, a^i_j, equals either 1 or 0.

The computation is conducted on a laptop computer with Pentium 4 processor, 1.68 GHZ, and 256 MB memory. Generating 10000 samples of the delivery lead-time $X^i(z)$ for all $i \in \mathcal{I}$ and a particular z takes no more than 10 minutes. In the special case of Poisson demand and base-stock policy, Zhao and Simchi-Levi (2006) reports that 3-4 minutes computing time is needed for a slightly smaller example on the same computer. Comparing these examples, it is clear that the increment in the computing time from base-stock policy and Poisson demand to batch-ordering policy and compound Poisson demand is moderate, and the numerical method proposed here can handle problems of real world sizes.

To quantify the impact of order splitting in ATO systems, we compute the expected delivery lead-time and the order-based fill-rate for a randomly picked demand unit in both the case of split orders and the case of non-split orders. More specifically, in the case of split orders, we compute, for each $i \in \mathcal{I}$,

$$E(X^i) = \sum_{z \geq 1} P\{D^i \geq z\}/E(D^i) \times E(X^i(z)), \tag{30}$$

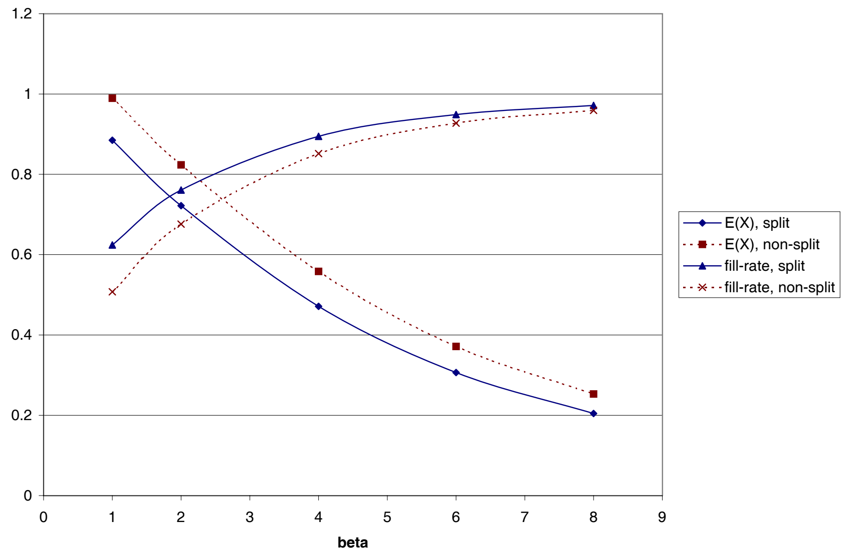

Figure 1: The impact of order splitting.

$$P\{X^i \leq \tau\} = \sum_{z \geq 1} P\{D^i \geq z\}/E(D^i) \times P\{X^i(z) \leq \tau\}, \tag{31}$$

where $\tau \geq 0$ is the committed service time, and $P\{D^i \geq z\}/E(D^i)$ is the probability that a randomly picked unit is the z^{th} unit of a demand (Sigman 2001). In the case of non-split orders, we compute, for each $i \in \mathcal{I}$,

$$E(\hat{X}^i) = \sum_{z \geq 1} zP\{D^i = z\}/E(D^i) \times E(X^i(z)), \tag{32}$$

$$P\{\hat{X}^i \leq \tau\} = \sum_{z \geq 1} zP\{D^i = z\}/E(D^i) \times P\{X^i(z) \leq \tau\}, \tag{33}$$

where $zP\{D^i = z\}/E(D^i)$ is the probability that a randomly picked unit is in a demand of size z (Sigman 2001). Finally, we compute the *weighted* expected delivery lead-times and fill-rates for a randomly picked demand unit as follows:

$$E(X) = \sum_{i\in\mathcal{I}}[E(X^i) \times \lambda^i E(D^i)]/\sum_{i\in\mathcal{I}}[\lambda^i E(D^i)] \tag{34}$$

$$P\{X \le \tau\} = \sum_{i\in\mathcal{I}}[P\{X^i \le \tau\} \times \lambda^i E(D^i)]/\sum_{i\in\mathcal{I}}[\lambda^i E(D^i)] \tag{35}$$

$$E(\hat{X}) = \sum_{i\in\mathcal{I}}[E(\hat{X}^i) \times \lambda^i E(D^i)]/\sum_{i\in\mathcal{I}}[\lambda^i E(D^i)] \tag{36}$$

$$P\{\hat{X} \le \tau\} = \sum_{i\in\mathcal{I}}[P\{\hat{X}^i \le \tau\} \times \lambda^i E(D^i)]/\sum_{i\in\mathcal{I}}[\lambda^i E(D^i)]. \tag{37}$$

For $\tau = 1$, the numerical results are demonstrated in Table 1 and Figure 1. The 95% confidence intervals of the weighted expected delivery lead-times (or the fill-rates) have lengths less than 0.02 (2%, respectively). First, we notice that the weighted expected delivery lead-times (the fill-rates) in the case of split orders are always smaller (greater, respectively) than the counter-parts in the case of non-split orders. In the range of fill-rates of interests, we observe that as the weighted fill-rate increases, the gap between the fill-rates of the split and non-split cases decreases, but the percentage difference in the weighted expected delivery lead-times increases. Finally, the impact of order splitting can be quite substantial, e.g., the second row/last column of Table 1 shows a 11.67% difference in fill-rates, and the last row/second last column shows a nearly 20% difference in the expected delivery lead-times.

Table 1: Impact of order splitting

β	$E(X)$ (split)	$E(\hat{X})$ (non-split)	$P\{X \le \tau\}$ (split)	$P\{\hat{X} \le \tau\}$ (non-split)	$[E(X) - E(\hat{X})]$ $/E(\hat{X})$	$P\{X \le \tau\} - P\{\hat{X} \le \tau\}$
1	0.885	0.99	62.41%	50.74%	−10.58%	11.67%
2	0.722	0.824	76.07%	67.60%	−12.35%	8.47%
4	0.471	0.558	89.48%	85.15%	−15.56%	4.33%
6	0.306	0.372	94.89%	92.76%	−17.52%	2.13%
8	0.204	0.253	97.15%	95.93%	−19.16%	1.22%

6 Conclusion

ATO systems with both batch ordering policy and batch demand can often be found in practice. These systems are also known to be challenging both analytically and computationally. In this paper, we demonstrate that while the combination of batch ordering policy and compound Poisson demand increases significantly the analytical complexity, such a system of real-world sizes is still numerically tractable given that the memory requirement for storing the BOM matrix can be met. We should point out that the same numerical method can also be used to estimate inventory costs for the case of split orders. However, for the case of non-split orders, one requires an additional piece of information – the joint probability distribution of consecutive lead-times for each component (Section 3).

In practice, of course, economies of scale in production or transportation costs may drive batch ordering policies across a supply chain beyond the assembly system. Therefore, one important direction for future research is to extend the analysis to general multi-level supply chains that include both assembly and distribution operations. As Axsater (2003) points out, exact evaluation of multi-level distribution systems alone with batch ordering policies and compound Poisson demand is computationally prohibitive. Therefore, we expect that accurate approximations should be developed to efficiently evaluate and optimize these supply chains.

Acknowledgments: The author would like to thank the Editor and two referees for their constructive suggestions which have greatly improved the paper.

References

Agrawal, N. & M. A. Cohen (2001). Optimal material control in an assembly system with component commonality. *Naval Research Logistics*, **48**, 409-429.

Axsater, S. (2003). Supply chain operations: serial and distribution inventory systems. A. G. de Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science*, Vol. 11: Supply Chain Management. Elsevier (North-Holland), Amsterdam. The Netherlands.

Benjaafar, S. & M. Elhafsi (2006). Production and Inventory Control of a Single Product Assemble-to-Order Systems with Multiple Customer Classes. *Management Science*, **52**, 1896-1912.

Chen, F. (2000). Optimal policies for multi-echelon inventory problems with batch ordering. *Operations Research*, **48**, 376-389.

Cheng, F., M. Ettl, G. Lin & D. D. Yao (2002). Inventory-service optimization in Configure-to-Order systems. *Manufacturing & Service Operations Management*. **4**, 114-132.

De Kok, Ton G. (2003). Evaluation and optimization of strongly ideal Assemble-To-Order systems. Working Paper, Technische Universiteit Eindhoven. The Netherlands.

Diks, E. B., A. G. de Kok & A. G. Lagodimos (1996). Multi-Echelon systems: a service measure perspective. *European Journal of Operational Research*, **95**, 241-263.

Elhafsi, M. (2008). Optimal Integrated Production and Inventory Control of an Assemble-to-Order System with Multiple Non-Unitary Demand Classes. To appear in *European Journal of Operational Research*.

Ernst, R. & D. F. Pyke (1992). Component part stocking policies. *Naval Research Logistics*, **39**, 509-529.

Esary, J. D., F. Proschan & D. W. Walkup (1967). Association of random variables, with applications. *Annals of Mathematical Statistics*, **38**, 1466-1474.

Hausman, W. H., H. L. Lee, & A. X. Zhang (1998). Joint demand fulfillment probability in a multi-item inventory system with independent order-up-to policies. *European Journal of Operational Research*, **109**, 646-659.

Law, A. M. & W. D. Kelton (1991). *Simulation Modeling and Analysis*, 2nd edition. McGraw-Hill, Inc. New York.

Lu, Y. (2007). Estimation of average backorder for an assemble-to-order system with random batch demands through extreme statistics. *Naval Research Logistics*, **54**, 33-45.

Lu, Y. (2008). Performance analysis for assemble-to-order systems with general renewal arrivals and random batch demands. *European Journal of Operational Research*, **185**, 635-647.

Lu, Y. & J. S. Song (2005). Order-based cost optimization in Assemble-to-order systems. *Operations Research*, **53**, 151-169.

Lu, Y., Song, J. S. & D. D. Yao (2003). Order fill rate, lead-time variability and advance demand information in an Assemble-to-order system. *Operations Research*, **51**, 292-308.

Plambeck, E. L. (2005). Near optimal control of an Assemble-to-Order system with leadtime constraints and fixed shipping costs for components. Working Paper. Graduate School of Business, Stanford University. Stanford.

Rosling, K. (1989). Optimal inventory policies for assembly systems under random demand. *Operations Research*, **37**, 565-579.

Shaked, M. & J. G. Shanthikumar (1994). *Stochastic Orders and Their Applications*. Academic

Press, New York.

Sigman, K. (2001). Lecture notes on $GI/M/1$ queue. Columbia University. New York, NY.

Simchi-Levi, D. & Y. Zhao (2005). Safety stock positioning in supply chains with stochastic lead-times. *Manufacturing and Service Operations Management*, **7**, 295-318.

Song, J. S. (1998). On the order fill rate in a multi-item, base-stock inventory system. *Operations Research*, **46**, 831-845.

Song, J. S. (2000). A note on Assemble-to-Order systems with batch ordering. *Management Sciences*, **46**, 739-943.

Song, J. S. (2002). Order-Based backorders and their implications in multi-item inventory systems. *Management Science*, **48**, 499-516.

Song, J. S. & P. Zipkin (2003). Supply chain operations: Assemble-to-order systems. A. G. de Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science*, Vol. 11: Supply Chain Management. Elsevier (North-Holland), Amsterdam. The Netherlands.

Svoronos, A. & P. Zipkin (1991). Evaluation of one-for-one replenishment policies for multiechelon inventory systems. *Management Science*, **37**, 68-83.

Tong, Y. L. (1980). *Probability inequalities in multivariate distributions*. Academic Press, New York.

Xu, S. H. (2001). Dependence analysis of assemble-to-order systems. Chapter 11 in J. Song and D. Yao, eds. *Supply Chain Structures: Coordination, Information and Optimization*. Kluwer Academic Publishes, Norwell, MA.

Zhang, A. X. (1997). Demand fulfillment rates in an assemble-to-order system with multiple products and dependent demands. *Production and Operations Management*, **6**, 309-323.

Zhao, Y. & D. Simchi-Levi (2006). Performance analysis and evaluation of Assemble-to-Order systems with stochastic sequential lead-times. *Operations Research*, **54**, 706-724.

Zhao, Y. (2008). Evaluation and optimization of installation base-stock policies in supply chains with compound Poisson processes. *Operations Research*, **56**, 437-452.

Zipkin, P. (1991). Evaluation of base-stock policies in multiechelon inventory systems with compound-Poisson demands. *Naval Research Logistics*, **38**, 397-412.

Zipkin, P (2000). *Foundations of Inventory Management*. McGraw Hill, Boston.