U N I V E R S I T Ä T   H A M B U R G

Monotonicity of Base Stock Policies

Nha-Nghi de la Cruz, Hans Daduna

Preprint–No. 2015-02   Oktober 2015

DEPARTMENT MATHEMATIK

SCHWERPUNKT MATHEMATISCHE STATISTIK
UND STOCHASTISCHE PROZESSE

BUNDESSTR. 55, D – 20146 HAMBURG

Monotonicity of Base Stock Policies

Nha-Nghi de la Cruz
Faculty of Business Administration, University of Hamburg, Von-Melle-Park 5, 20146 Hamburg, Germany
Email: nha-nghi.huynh@uni-hamburg.de

Hans Daduna
Department of Mathematics, University of Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany
Email: daduna@math.uni-hamburg.de

Abstract

We analyze monotonicity of base stock levels in multi - item inventory - production systems where arriving demand triggers production of a new unit. Rubio and Wein (Rubio, R., Wein, L.M.: Setting base stock levels using product-form queueing networks. Management Science **42**(2), 259–268 (1996)) used open Jackson networks to describe such integrated models. They conjectured that the base stock level should increase with the utilization in the production system and confirmed this for a single - product system via numerical investigations. We present a basic analytical proof of the general presumption for single- and multi - item systems utilizing stochastic orderings and discuss the implications for resetting base stock levels when load and/or capacity in the manufacturing network change. We further develop a new algorithm to evaluate the optimal base stock levels.

Keywords: Queueing, Stochastic processes, Inventory, Stochastic Ordering

1 Introduction

For a manufacturing system which produces items of different types on a make-to-stock basis, we consider a queueing network model to determine an optimal base stock policy. A base stock policy prescribes a target total inventory position for each product to balance backorder and inventory holding costs. In the system, produced items are stored in a finished good (FG) inventory to serve exogenous demand. If there are finished goods at stock when demand arrives, it is satisfied immediately, otherwise it is backordered (negative FG). In both cases, an order for producing a new item is placed instantaneously, which is counted as work-in-process (WIP). Consequently, the sum of FG and WIP inventory is maintained at a fixed level for each product in this CONWIP-like system. Related models are base stock systems with multiple production stages each holding its individual buffer stock (see Buzacott et al. (1992) and Lee and Zipkin (1992)).

Using networks of queues to investigate complex inventory systems found some interest at the end of the 20th century; see the references in Toktay et al. (2000), Rubio and Wein (1996), and the work on similar models in Zazanis (1994) and Spearman and Zazanis (1992). A survey of related models for integrated inventory-production systems is in Krishnamoorthy et al. (2011).

Our starting point is the paper of Rubio and Wein Rubio and Wein (1996) whose base stock model consists of a multi-item inventory and a replenishment network of the Jacksonian type. They show that it is sufficient to study the steady state WIP to determine the optimal base stock level.

We focus on monotonicity behavior of optimal base stock policies within the parameters of the network and discuss the following important monotonicity property for optimally setting the target base stock level: Whenever the demand intensity for a product increases and/or somewhere in the network the production capacity decreases, the optimal base stock level increases (we use "increasing" for "non decreasing", and similarly "decreasing").

Although this property seems to be intuitive and natural, it is not easy to prove. In fact, Rubio and Wein provided a proof only for the case of a single-item inventory-production system where all stations have the same utilization (in their terminology: a balanced network). For the single-product case in an unbalanced network, they conjectured that such monotonicity should hold as well, relying on numerical experiments (see Rubio and Wein (1996), p. 263). The main objective of the present paper is to prove the monotonicity property even for a more general system, i.e., the manufacturing system does not have to be balanced nor reduced to the singe-product case. This advises the inventory controller on how to increase (decrease) the target base stock level when demand increases (decreases) and/or capacity is reduced (expanded).

Our paper is structured as follows: In section 2 we present the inventory-production model. In section 3 we describe Jackson networks with different classes of customers representing the different types of WIP in the replenishment network. We prove the monotonicity behavior of base stock levels related to the system's utilization in section 4. In section 5, we provide new numerical insights for the balanced and unbalanced single-item case and prove a general monotonicity property suggested by the numerical results.

For queueing networks we only cite facts and refer for details to the literature. For the numerical evaluations, we restrict our presentation mainly to the case of exponential single server queues. This is only for simplicity of presentation. As it will be seen, the algorithm and the experiments can be easily extended to more general systems.

2 Manufacturing queueing system with backordering

We consider a multi-item manufacturing system with products u, $u = 1, ..., U$. In the system, we distinguish between the FG and WIP inventory which contain the finished and unfinished goods, respectively. Unsatisfied demand is captured by the backorder level for each type of good.

If an exogenous demand for a specific good arrives, the corresponding WIP inventory increases in the same amount as the FG inventory for this good decreases. This ensures that the demanded item is eventually reproduced. On the other hand, finished produced items are transferred from the manufacturing network to the FG inventory, i.e., are converted from WIP to FG items. Let $A_u(t)$ be the cumulative demand of product u in $[0, t]$ and $D_u(t)$ the amount of finished items moved to the FG inventory until t. For the WIP inventory $N_u(t)$ we therefore have: $N_u(t) = A_u(t) - D_i(t)$. For $t = 0$, we assume $N_u(t) = 0$.

The WIP processes $(N_u(t) : t \geq 0)$ for $u = 1, \ldots, U$, are determined by the manufacturing system, which will be described in terms of stochastic networks of generalized Jacksonian type in section 3 below.

Let $Z_u(t)$ denote the FG inventory level of product u, where negative levels count backordered items. Without loss of generality, the initial level is set as the base stock level $z_u, u = 1, ..., U$.

Demand reduce the FG inventory whereas finished produced items increment it: $Z_u(t) = z_u + D_i(t) - A_u(t)$. Therefore, the base stock level is equal to the total inventory position: $N_u(t) + Z_u(t) = z_u$.

The WIP level can grow arbitrarily high because it also includes the backordered products whereas $Z_u(t)$ is bounded from above by z_u. $Z_u(t)$ can be divided into two parts: the on-hand inventory $Z_u^O(t)$ and backorders $Z_u^B(t)$. Thus, $Z_u(t) = Z_u^O(t) - Z_u^B(t)$, and we have $Z_u^O(t) \cdot Z_u^B(t) = 0$.

Shortfalls occur if demand is larger than the base stock level or, differently said, if the WIP in the replenishment network exceeds the base stock. The expected backorder level is: $E[Z_u^B(t)] = \sum_{n=z_u}^{\infty}(n - z_u)P(N_u(t) = n)$.

Using also the fact that the FG inventory is the difference of the on-hand inventory and backorders, we receive: $z_u = E[N_u(t)] + E[Z_u^O(t)] - E[Z_u^B(t)]$. Consequently, we obtain for the expected on-hand-inventory: $E[Z_u^O(t)] = z_u - E[N_u(t)] + \sum_{n=z_u}^{\infty}(n - z_u)P(N_u(t) = n)$.

The cost function encompasses item specific costs per time unit: for items of type $u = 1, ..., U$, the holding cost c_u for WIP inventory, the holding cost h_u for FG inventory, the backorder cost b_u. We denote by $C_u(z_u)$ the total cost per time unit which originates from type u items when the base stock level for type u items is set to z_u (in the steady state system).

We are interested in the long run overall costs per time unit, which by exploiting the ergodic theorem for our Markovian system process can be computed via the stationary expected costs. Let N_u denote a random variable distributed according to the total number of WIP items of type u under the stationary WIP distribution, which will be given explicitly below.

It turns out that the main decision variables are N_u, the overall WIP of product $u = 1, ..., U$, (see Rubio and Wein (1996), p. 261)

$$\begin{aligned}\sum_{u}^{U} E[C_u(z_u)] &= \sum_{u}^{U} \Big((h_u + b_u) \Big[z_u P(N_u \leq z_u - 1) + \sum_{n=z_u}^{\infty} nP(N_u = n) \Big] \\ &\qquad -b_u z_u + (c_u - h_u)E[N_u] \Big).\end{aligned}$$

Control variables are stock sizes $z_u \geq 0$, for product type $u = 1, ..., U$. Rubio and Wein showed that the cost minimizing base stock levels $z_u^*, u = 1, ..., U$, can be expressed in terms of the WIP of product u and the respective inventory and backorder cost per item and time unit only. These are the smallest integers $z_u^* \geq 0$, such that for the stationary WIP in the replenishment network holds

$$P(N_u \leq z_u^*) \geq \frac{b_u}{b_u + h_u}, \qquad u = 1, ..., U. \tag{1}$$

3 Multi-class Jackson network with exponential service times

To describe the behavior of the WIP in the replenishment system, we use a standard multi-class Jackson network in which N_u is the total population size of item u. We adapt the notation of Chao et al. (see Chao et al. (1999), chap. 5) and consider a multi-class Jackson network with J stations, numbered by $j = 1, ..., J$. We distinguish in the network products u, $u = 1, ..., U$ (customer classes). Set n_{ju} the number of class u customers at node j, the total population size at node j as $n_j = n_{j1} + ... + n_{jU}$, and define $\overline{n_j} = (n_{j1}, ..., n_{jU})$ and $\overline{n} = (\overline{n_1}, ..., \overline{n_J})$, the joint class occupation vector. Arrivals of class u at node j from outside of the network follow a Poisson process with rate λ_{ju}. $r_{ju,ku}$ is the probability that class u customers departing from node j join node k, and $r_{ju,0}$ is the probability that a class u customer leaves the system after finishing his service at node j with $\sum_{k=1}^{J} r_{ju,ku} + r_{ju,0} = 1, u = 1, ..., U$. Since demand of product u triggers a class u arrival at the replenishment network leaving as an output of item u, customers in the network do not change their class (which would be possible in the general setting of Chao et al. (see Chao et al. (1999), chap. 5).

α_{ju}, the total arrival rate of class u at node j, is obtained as the solution (assumed to be unique) of the so called traffic equations

$$\alpha_{ju} = \lambda_{ju} + \sum_{k=1}^{J} \alpha_{ku} r_{ku,ju}, \quad j = 1, ..., J; u = 1, ..., U. \tag{2}$$

At node j all customers require an amount of service which is exponentially distributed with rate μ_j. The server at node j provides service at rate $\Phi_j(n_j)$, which is non-decreasing in n_j and $\Phi_j(n_j) > 0$ if $n_j \geq 1$. This results in a versatile construction of service regimes. For example, if the network consists of first-come-first-served (FCFS) multi-sever queues with $s_j \geq 1$ servers, we can write $\Phi_j(n_j) = \min\{n_j, s_j\}$. We define the class u utilization of node j by $\rho_{ju} = \alpha_{ju}/\mu_j$.

The stationary distribution π of the multi-class Jackson network for the joint class occupation process with values $\overline{n} = (\overline{n_1}, ..., \overline{n_J})$ is with normalization constants (see Chao et al. (1999), p. 127) $B_j = \sum_{n=0}^{\infty} \left((\sum_{u=1}^{U} \rho_{ju})^n / \prod_{\ell=1}^{n} \Phi_j(\ell) \right) < \infty$

$$\pi(\overline{n}) = \prod_{j=1}^{J} \pi_j(\overline{n_j}) \text{ with } \pi_j(\overline{n_j}) = B_j^{-1} \frac{n_j!}{n_{j1}! ... n_{jU}!} \prod_{\ell=1}^{n_j} \Phi_j(\ell)^{-1} \prod_{u=1}^{U} \rho_{ju}{}^{n_{ju}}. \tag{3}$$

The joint class occupation process (which does not count for positions in the queue) with values $\overline{n} = (\overline{n_1}, ..., \overline{n_J})$ is in general not Markovian. However, it suffices as state description for the relevant quantities for our investigations.

Example 1. Deterministic replenishment schedule. *Our model encompasses the important case where each class (product) has its dedicated fixed sequence of service stations to visit in the replenishment procedure: Each customer (product) class follows its own route inside the network. Such networks with class dependent fixed routing are called Kelly networks, which exhibit stationary distributions with product form structure as introduced before. In such* *a network, class u customers arrive from outside in a Poisson process with rate λ_u at entrance node $u[1]$. The node a class u customer visits in his m^{th} stage of his path is denoted by $u[m]$, $m = 1, ..., M(u)$, and the probability that a class u customer leaves the network after receiving service at his last stage $M(u)$ is 1.*

Remark 1. *So far, we have only considered networks in which customers have exponential service requirements. For arbitrary service requirements, we refer to BCMP networks (see Chao et al. (1999), chap. 6), where the service requirement of a customer u at station j may have an arbitrary distribution, depending on class and node. Denote by S_{ju} a random variable distributed as class u customers' service request at node j with mean $E(S_{ju}) = \mu_{ju}^{-1}$. In this situation, we need to abandon the FCFS-property and employ the so called symmetric service disciplines (see Chao et al. (1999), chap. 6, p. 150). We will not need details for our development here and therefore sketch only an example for the reader's convenience.*

In the context of our replenishment manufacturing system, we could think of a server network with $\Phi_j(n_j)$ increasing in (n_j), where all arriving customers (products) have the same priority and the server's effort is equally shared among all the customers present at the same node. The service completion rate is $\Phi_j(n_j) \cdot \frac{n_{ju}}{n_j}\mu_{ju}$ for customers of class u at node j, $u = 1, ..., U$. The utilizations in the network are $\rho_{ju} = \frac{\alpha_{ju}}{\mu_{ju}}, u = 1, ..., U, j = 1, \ldots, J$. If the service requirements are exponentially distributed, the stationary joint queue length distribution is (3). Otherwise, the stationary distribution of this BCMP network with single servers under processor-sharing with $\sum_{u=1}^{U} \rho_{ju} < 1$ and unique solutions of the traffic equations α_{ju}, $j = 1, ..., J, u = 1, ..., U$, is (see Chao et al. (1999), Theorem 6.1 and Theorem 6.2)

$$\pi(\overline{n}) = \prod_{j=1}^{J} \pi_j(\overline{n_j}) \text{ with } \pi_j(\overline{n_j}) = B_j^{-1} \frac{n_j!}{n_{j1}! \ldots n_{jU}!} \prod_{\ell=1}^{n_j} \Phi_j(\ell)^{-1} \prod_{u=1}^{U} {\rho_{ju}}^{n_{ju}}, \tag{4}$$

where $B_j = \sum_{n=0}^{\infty} \left((\sum_{u=1}^{U} \rho_{ju})^n / \prod_{\ell=1}^{n} \Phi_j(\ell) \right)$.

The formulas displayed in (3) *and* (4) *are identical but the cautious reader should recall the difference of $\rho_{ju} = \alpha_{ju}/\mu_j$ in* (3) *and $\rho_{ju} = \alpha_{ju}/\mu_{ju}$ in* (4).

4 Monotonicities

We apply stochastic order to show that the optimal base stock level increases with the utilization in the system. We first provide definitions of the used orderings, see Mueller and Stoyan (2002). Let N and N' be two $\mathbb{N}$-valued random variables, F_N and $F_{N'}$ their cumulative distribution functions, and f_N and $f_{N'}$ their probability density functions, respectively.

Definition 1. Usual Stochastic Order *N' is larger than N in the usual stochastic order, written $N \leq_{st} N'$, if $F_N(n) \geq F_{N'}(n)$ $\forall n \in \mathbb{N}$.*

Definition 2. Likelihood Ratio Order *N' is larger than N in likelihood ratio order, written $N \leq_{lr} N'$, if $f_N(n)f_{N'}(k) \leq f_N(k)f_{N'}(n), \forall k \leq n \in \mathbb{N}$.*

4.1 Single-product network with single-server nodes

We firstly analyze the special case of a single-product network with exponential FCFS single-server nodes with $U = 1$ and $s_j = 1, j = 1, ..., J$. Hence, the capacity functions reduce to $\Phi_j(n_j) = 1$ if $n_j \geq 1$. Denote by $\alpha = (\alpha_j, j = 1, ..., J)$ the solution (assumed to be unique) of the traffic equations $\alpha_j = \lambda_j + \sum_{i=1}^{J} \alpha_i r_{i,j}$, $j = 1, ..., J$, where we suppressed index u.

With $(X_j, j = 1, ..., J)$ representing the stationary state of the joint queue length process, Jackson's theorem states that for $\rho_j := \alpha_j/\mu_j < 1$, $j = 1, ..., J$, we have for $(n_1, ..., n_J) \in \mathbb{N}^J$

$$\pi(n_1, ..., n_J) := P(X_j = n_j, j = 1, ..., J) = \prod_{j=1}^{J} (1 - \rho_j)\rho_j^{n_j}. \tag{5}$$

Product (5) implies that the stationary distribution of the total WIP is a convolution of geometric distributions with parameters $1-\rho_j, j = 1, ..., J$. So for single-product systems, (1) is a statement on $N := X_1 + X_2 + \cdots + X_J$.

Proposition 1. *If the utilization ρ_j at some station j is increased, then z^* increases.*

Proof. Consider two J-station networks where $X := (X_j, j = 1, ..., J)$ and $X' := (X'_j, j = 1, ..., J)$ represent the stationary states of the joint queue length processes. Moreover, assume for the utilizations $\rho_j \leq \rho'_j, j = 1, ..., J$.

For geometrically distributed random variables $X_j \sim geo^{\circ}(1 - \rho_j)$ and $X'_j \sim geo^{\circ}(1 - \rho'_j)$ holds: if $\rho_j < \rho'_j$, then X'_j is larger than X_j in likelihood ratio order, $X_j \leq_{lr} X'_j$ (see Mueller and Stoyan (2002), p. 63). Since the likelihood ratio order is stronger than the usual stochastic order, we have: $X_j \leq_{lr} X'_j \Rightarrow X_j \leq_{st} X'_j$ (see Mueller and Stoyan (2002), p. 12). From the convolution property of $\leq_{st}$ (see Mueller and Stoyan (2002), p.7), it results

$$N := X_1 + X_2 + \cdots + X_J \leq_{st} N' := X'_1 + X'_2 + \cdots + X'_J. \tag{6}$$

From the definition of the usual stochastic order, we conclude

$$z^* = \inf_z \left(P(N \leq z) \geq \frac{b}{b+h} \right) \leq z'^* = \inf_z \left(P(N' \leq z) \geq \frac{b}{b+h} \right).$$

Thus, the optimal base stock level z^* increases in the $\rho_j, j = 1, ..., J$. □

Recall: For $\rho_j = \rho$ we obtain the main result of Rubio and Wein (1996).

As an illuminating example, consider a replenishment network with fixed routing $(r_{j,i})$ and demand intensities λ_j, where service intensity can be adjusted. The proven monotonicity implies that increasing service intensity somewhere results in less stock needed to maintain the cost level. One can also easily prove that the optimal base stock levels grows with the backorder-to-holding cost ratio.

4.2 Multi-product network with multi-server nodes

Rubio & Wein's conjecture proposes a similar monotonicity of base stock policies for replenishment networks with exponential multi-server nodes, which we extend to an even much more general system described in Remark 1.

Proposition 2. *Consider a generalized multi-class Jackson or BCMP network with queue length dependent capacities $\Phi_j(n_j)$ increasing in the queue lengths (n_j). If the service discipline at node j is FCFS, all customers have an exponential service time request with node dependent mean μ_j^{-1}; if the service discipline at node j is processor sharing, a customer of class u may have a general class and node dependent service time request with mean μ_{ju}^{-1}.*

The utilization of class u customers at j is $\rho_{ju} = \alpha_{ju}/\mu_j$ for FCFS node j, and $\rho_{ju} = \alpha_{ju}/\mu_{ju}$ for processor sharing node j.

If class u utilization ρ_{ju} is increased at some station j, then all z_u^* increase.

Proof. The basic structure of the proof resembles that of Proposition 1 but is more involved. To determine z_u^* in (1), we need the total number N_u of class u customers in the network. We start with the total number N_{ju} of class u customers at node j. The marginal distribution for the number of type u customer at node j is from (3), resp. (4),

$$P\left(N_{ju} = n_{ju}\right) = \sum_{\substack{v=1,\dots,U,\\ v\neq u}} \sum_{n_{jv}=0}^{\infty} \pi_j\left(n_{j1}, \dots, n_{jU}\right)$$

$$= \sum_{\substack{v=1,\dots,U,\\ v\neq u}} \sum_{n_{jv}=0}^{\infty} B_j^{-1} \frac{(n_{j1}+n_{j2}+\dots+n_{jU})!}{n_{j1}!n_{j2}!\dots n_{jU}!} \prod_{v=1}^{U} {\rho_{jv}}^{n_{jv}} \left(\prod_{l=1}^{n_{j1}+\dots+n_{ju}} \frac{1}{\Phi_j(l)}\right)$$

$$= B_j^{-1} \rho_{ju}^{n_{ju}} \frac{1}{n_{ju}!} \sum_{m=0}^{\infty} \left(\prod_{l=1}^{m+n_{ju}} \frac{1}{\Phi_j\left(l\right)}\right) \frac{(m+n_{ju})!}{m!} \sum_{\substack{(n_{jv}: v\in\{1,\dots,U\}\setminus\{u\})\\ \left(\sum_{v=1,\dots,U,\, v\neq u} n_{jv}\right)=m}} \frac{m!}{\prod_{v=1,\dots,U,\, v\neq u} n_{jv}!} \prod_{\substack{v=1,\dots,U,\\ v\neq u}} \rho_{jv}^{n_{jv}}$$

$$= B_j^{-1} \rho_{ju}^{n_{ju}} \frac{1}{n_{ju}!} \sum_{m=0}^{\infty} \left(\prod_{l=1}^{m+n_{ju}} \frac{1}{\Phi_j\left(l\right)}\right) \frac{(m+n_{ju})!}{m!} \left(\sum_{\substack{v=1\\ v\neq u}}^{U} \rho_{jv}\right)^m .$$

In order to show that the likelihood ratio order holds, we assume that at some node j the class u utilization changes from ρ_{ju} to ρ'_{ju} with $\rho_{ju} < \rho'_{ju}$. We rewrite the probability that there are n_{ju} type u customers at node j with $A_j\left(n_{ju}\right) := \sum_{m=0}^{\infty} \left(\prod_{l=1}^{m+n_{ju}} \frac{1}{\Phi_j(l)}\right) \frac{(m+n_{ju})!}{m!} \left(\sum_{v=1,\ v\neq u}^{U} \rho_{jv}\right)^m$

as

$$\pi_{ju}\left(n_{ju}\right) = P\left(N_{ju} = n_{ju}\right) = B_j^{-1} {\rho_{ju}}^{n_{ju}} \frac{1}{n_{ju}!} A_j(n_{ju}) \,,$$

and similarly with ρ'_{ju}

$$\pi'_{ju}\left(n_{ju}\right) = P\left(N'_{ju} = n_{ju}\right) = B_j'^{-1} {\rho'_{ju}}^{n_{ju}} \frac{1}{n_{ju}!} A_j(n_{ju}) \,.$$

Hence, we obtain $\rho_{ju} \le \rho'_{ju} \Leftrightarrow {\rho_{ju}}^{n_{ju}-k_{ju}} \le {\rho'_{ju}}^{n_{ju}-k_{ju}}$ for $0 \le k_{ju} \le n_{ju}$ from

$$\frac{\pi_{ju}(n_{ju})}{\pi_{ju}(k_{ju})} \le \frac{\pi'_{ju}(n_{ju})}{\pi'_{ju}(k_{ju})} \Leftrightarrow \frac{{\rho_{ju}}^{n_{ju}} \frac{1}{n_{ju}!} A_j\left(n_{ju}\right)}{{\rho_{ju}}^{k_{ju}} \frac{1}{k_{ju}!} A_j\left(k_{ju}\right)} \le \frac{{\rho'_{ju}}^{n_{ju}} \frac{1}{n_{ju}!} A_j\left(n_{ju}\right)}{{\rho'_{ju}}^{k_{ju}} \frac{1}{k_{ju}!} A_j\left(k_{ju}\right)}.$$

Thereafter, N'_{ju} is larger than N_{ju} in likelihood ratio order and also in usual stochastic order, and we conclude as above from the product form structure of π that $N_u \le_{st} N'_u$. Thus, a higher base stock $z_u'^*$ is needed to fulfill the optimality condition (1) with larger $\rho'_{ju} > \rho_{ju}$ at node j. We conclude our arguments for the multi-product system with the remark that (1) is a statement on $N_u =: N_{1u} + \cdots + N_{Ju}$ and $N'_u =: N'_{1u} + \cdots + N'_{Ju}$. □

Remark 2. *For simplicity of presentation, we focused on the queueing regimes FCFS for multi-server exponential nodes and processor sharing for nodes with general class and node dependent service time requests. It is easy to see that Proposition 2 is valid for any symmetric service discipline at nodes with general class and node dependent service time requests and for any non-symmetric service discipline at nodes with exponential node dependent service time requests (for a precise definition see Chao et al. (1999), Theorem 6.1 and Chao et al. (1999), p. 119-120 and Theorem 5.4, respectively).*

Example 2. *Consider a system where for each demand class the production schedule is a fixed sequence of stations which results in fixed arrival rates α_{ju} (of class u customers at station j). Proposition 2 says: Increasing service intensity somewhere results in less stock needed to maintain the cost level.*

Remark 3. *If we consider a multi-product network with single-server nodes, the marginal distribution of type u customers at node j can be simplified to*

$$P\left(N_{ju} = n_{ju}\right) = B_j^{-1} \rho_{ju}^{n_{ju}} \frac{1}{n_{ju}!} \sum_{m=0}^{\infty} \frac{(m+n_{ju})!}{m!} \left(\sum_{\substack{v=1 \\ v\neq u}}^{U} \rho_{jv} \right)^m$$

$$= B_j^{-1} {\rho_{ju}}^{n_{ju}} \left(\frac{1}{1 - \sum_{\substack{v=1, \\ v\neq u}}^{U} \rho_{jv}} \right)^{n_{ju}+1} = \frac{1 - \sum_{v=1}^{U} \rho_{jv}}{1 - \sum_{\substack{v=1 \\ v\neq u}}^{U} \rho_{jv}} \left(\frac{\rho_{ju}}{1 - \sum_{\substack{v=1, \\ v\neq u}}^{U} \rho_{jv}} \right)^{n_{ju}} ,$$

i.e., the number N_{ju} of type u customers at node j is geometrically distributed with success probability $p_{ju} = \frac{1-\sum_{v=1}^{U}\rho_{jv}}{1-\sum_{v=1,\ v\neq u}^{U}\rho_{jv}}$. Due to the product form equilibrium, the type u queue lengths behave independently over nodes and the probability that there are in total n_u type u customers in the system is $P(N_u = n_u) =$

$$\sum_{n_{1u}+...+n_{Ju}=n_u} P(N_{ju} = n_{ju}, u = 1, ..., U) = \sum_{n_{1u}+...+n_{Ju}=n_u} \prod_{j=1}^{J} p_{ju}\,(1-p_{ju})^{n_{ju}}\,.$$

5 Algorithmic evaluation of optimal base stock levels

The numerical explorations in Rubio and Wein (1996), in which inequation (1) was numerically solved in the direct way, are limited to the balanced case of single server FCFS nodes (see Rubio and Wein (1996), equation (4)). Our approach also applies to networks with unbalanced parameter settings. For clarity of exposition, we restrict our presentation to single server FCFS nodes.

The main idea of our new algorithm for determining optimal base stock levels is to use the so-called Buzen's algorithm for computing normalization constants in closed networks of queues (Gordon-Newell networks). Moreover, we exploit the fact that for a given Jackson network (our manufacturing network) with total population size N the probability $P(N = n)$ is proportional to the normalization constant in a suitably constructed Gordon-Newell network.

To be more precise: Buzen's algorithm enables a straightforward and numerically stable calculation of normalization constants in a closed Gordon-Newell network consisting of J single-server stations and a fixed number of customers N Buzen (1973). Utilizing the data of the prescribed Jackson network, we then apply the property that a Gordon-Newell network can be constructed such that its normalization constant $G(N, J)$ behaves proportionally to $P(N = n)$ of the given Jackson network (see Chen and Yao (2001), p. 20, Theorem 2.3). Therefore, we can use the following inequalities to calculate the optimal base stock level of a balanced and unbalanced single-item system, respectively (see A):

$$(1-\rho)^J \sum_{n=0}^{z^*} G(n, J) \geq \frac{b}{b+h}, \qquad \prod_{j=1}^{J}(1-\rho_j) \sum_{n=0}^{z^*} G(n, J) \geq \frac{b}{b+h} \tag{7}$$

5.1 Numerical evaluation and results

Rubio and Wein Rubio and Wein (1996) discovered that for balanced networks the optimal base stock level increases almost linearly with the number of stations as long as $\rho \leq 0.9$ holds for the utilization. On the other hand, their Figures 1 and 2 (see Rubio and Wein (1996), p. 264) suggest a nearly exponential increase of the optimal base stock level in the utilization for $\rho \leq 0.9$.

Exploiting the numerical stability of Buzen's algorithm, we found that the optimal base stock level also grows linearly with the number of stations in a balanced network under nearly heavy traffic conditions (see Figure 1). Moreover, we studied the interrelation of base stock level and

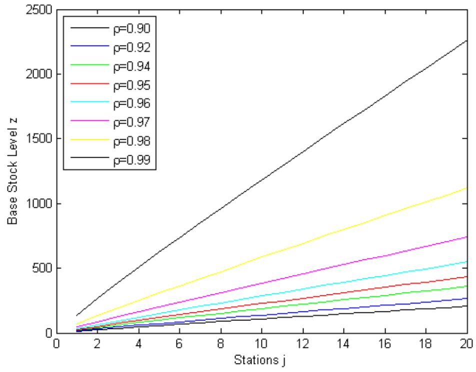

Figure 1: Optimal Base Stock Level in Balanced Networks under Heavy Traffic

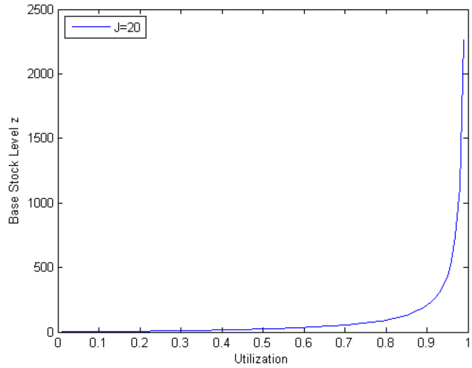

Figure 2: Optimal Base Stock Level in Balanced Networks

utilization for $0.01 \leq \rho \leq 0.99$ and can confirm that the base stock level grows exponentially with the utilization in a network with balanced stations. See Figure 2 for a typical scenario with 20 stations.

This shows that our algorithm works well even in critical parameter domains. We remark that utilizations greater than 0.9 should be avoided in practice. However, this parameter domain is of special interest when applying heavy traffic (diffusion) approximations for performance prediction.

Figure 3 demonstrates a typical behavior of the optimal base stock level when the number of stations in an unbalanced network increases from 1 to 20. In this investigation, we considered utilizations within the range of $0.4 \leq \rho \leq 0.99$. We start in the four reported experiments with a single station with $\rho_1 = 0.4$ (dark blue), $\rho_1 = 0.6$ (green), $\rho_1 = 0.8$ (red, light blue) and add successively stations with higher utilizations. For $j = 1, \ldots, 19$ a newly added station $j+1$ has utilization $\rho_{j+1} = \rho_j + 0.01$ (dark blue, green, light blue), or $\rho_{j+1} = \rho_j + 0.005$ (red). A detailed list of the utilizations for the dark blue setting is given in Table 1 (Case $1a$).

The somewhat surprising conclusion of the experiments is that even for the case of unbalanced networks (with linear increase of utilizations for added nodes), the optimal base stock level increases almost linearly as long as $\rho_j \leq 0.9$ is maintained (dark blue, green, red). On the other hand, the light blue curve indicates the dramatic change of the safety requirements in the range of a nearly heavy traffic regime. This property also holds for mixed networks in which not all utilizations are pairwise different, which can be easily evaluated with the help of Buzen's algorithm. Our numerical investigations reported in Figure 4 show, as expected, that the base stock level increases (decreases) as soon as the utilization of some single station increases (decreases). We demonstrate two different settings.

Firstly, we reconsider the dark blue line from Figure 3 with linearly increasing utilizations and maintain the utilizations for all stations apart from station 11 where we reset $\rho_{11} = 0.8$ (see

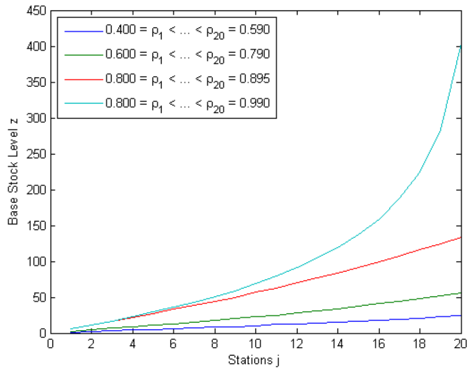

Figure 3: Optimal Base Stock Level in Unbalanced Networks

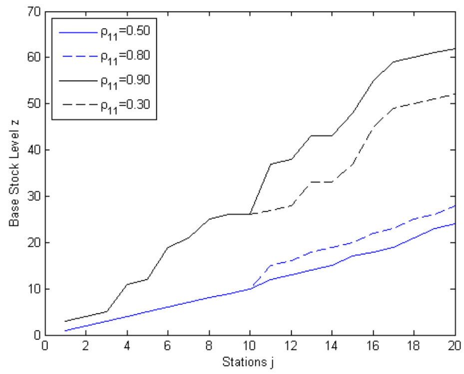

Figure 4: Monotonicity Behavior in Unbalanced and Irregular Networks

Table 1, Case 1b). This increases the dark blue line to the dark blue dashed line for $j = 11, \ldots, 20$ in Figure 4. Secondly, we consider an irregular network (see Table 1, Case 2a) which results in the upper black curve. We then decrease the utilization of station 11 where we reset $\rho_{11} = 0.3$ and maintain the utilizations for all other stations (see Table 1, Case 2b). This decreases the black line to the dashed black line for $j = 11, \ldots, 20$. In this numerical research, we found

Station	1	2	3	4	5	6	7	8	9	10
Case 1a	0.40	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.49
Case 1b	0.40	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.49
Case 2a	0.70	0.19	0.40	0.83	0.34	0.85	0.66	0.77	0.48	0.30
Case 2b	0.70	0.19	0.40	0.83	0.34	0.85	0.66	0.77	0.48	0.30

Station	**11**	12	13	14	15	16	17	18	19	20
Case 1a	**0.50**	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59
Case 1b	**0.80**	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59
Case 2a	**0.90**	0.48	0.80	0.15	0.80	0.86	0.78	0.48	0.60	0.40
Case 2b	**0.30**	0.48	0.80	0.15	0.80	0.86	0.78	0.48	0.60	0.40

Table 1: Utilizations of the numerical experiments

that the advantage of Buzen's algorithm is twofold: First, it facilitates the direct comparison of balanced and unbalanced networks, also under nearly heavy traffic conditions, and secondly, it permits the calculation of base stock levels in systems with both balanced and unbalanced nodes.

We emphasize that our algorithm can be directly transformed to the case of multi-server networks and the more general BCMP and Kelly networks by substituting the original Buzen's algorithm by the adapted algorithms available in the performance evaluation literature, e.g. Bruell and Balbo (1980). Another variant is to use the extended mean value analysis for computing the normalization constants, (see Reiser and Lavenberg (1980) and Akyildiz and Bolch (1983)).

5.2 A general monotonicity property

Recalling the observed monotonicity of the optimal base stock level, which is nearly linear for $\rho_j < 0.9$ in the homogeneous balanced (Rubio and Wein) and unbalanced (our Figure 3) case, and the increasing property of the optimal base stock level in the inhomogeneous network (our Figure 4), leads to the conjecture that the observed interrelation between the base stock level and number of stations holds in general. We summarize this as our

Proposition 3. *If a new station with any utilization is added to the single-item inventory-production network with single server nodes while the already existing stations maintain their utilizations, the optimal base stock level increases.*

Proof. We consider the general case, thus, the second formula from (7) and compute utilizing $G(0, J) = 1$ and $\sum_{\emptyset} a_n = 0$

$$\prod_{j=1}^{J}(1-\rho_j)\sum_{n=0}^{z} G(n,J) - \prod_{j=1}^{J+1}(1-\rho_j)\sum_{n=0}^{z} G(n,J+1)$$

$$\begin{aligned}
&= \prod_{j=1}^{J}(1-\rho_j)\sum_{n=0}^{z} G(n,J) - \prod_{j=1}^{J}(1-\rho_j)\sum_{n=0}^{z} G(n,J+1) \\
&\qquad +\rho_{J+1}\prod_{j=1}^{J}(1-\rho_j)\sum_{n=0}^{z} \overbrace{G(n,J+1)}^{(\star)} \\
&= \prod_{j=1}^{J}(1-\rho_j)\sum_{n=0}^{z} G(n,J) - \prod_{j=1}^{J}(1-\rho_j)\left[\sum_{n=0}^{z} G(n,J) + \rho_{J+1}\sum_{n=0}^{z-1} G(n,J+1)\right] \\
&\qquad +\rho_{J+1}\prod_{j=1}^{J}(1-\rho_j)\sum_{n=0}^{z} G(n,J+1) \\
&= \rho_{J+1}\prod_{j=1}^{J}(1-\rho_j)G(z,J+1) > 0,
\end{aligned}$$

where we used Buzen's formula to transform $(\star)$.
This implies that with $J+1$ stations a higher base stock level is needed to satisfy the required minimal cost criterion. □

Remark 4. *For single item production-inventory systems, we have $\mu_{ju} = \mu_j$. Therefore, Proposition 3 is also valid for the case of single product systems with general service time distributions at the nodes whenever $\Phi_j(n_j) = 1$ for $n_j \geq 1$ holds. This follows directly from formulas* (3) *and* (4)*.*

6 Conclusion

We have proved the conjectured monotonicity of optimal base stock levels for an inventory system, where replenishment is provided by a complex production system, modeled as a generalized Jackson network. As Rubio and Wein pointed out, this is a versatile class of models, and we extended the applicability of their model even to general systems where their required conditions, *either homogeneous (equal utilizations) or all utilizations are pairwise different*, do not hold. Our proof relies on the interplay of Jackson network theory and stochastic order theory.

The conjecture of the monotonicity implies: If the (general unbalanced) system produces items of different types, then, whenever the demand for any product type increases and/or anywhere in the network production capacity is reduced, an optimal decision of the inventory controller would be to hold more items on stock.

It is easy to see that the monotonicity proved here holds as well in product form networks with more general structure (see Chao et al. (1999), chap. 6).

Our ongoing work in this field comprises monotonicity problems in general inventory - production networks and provides analytical explanations of the networks' behavior.

A Base stock algorithm

Initialization:

Set number of stations $J \geq 1$;
set utilizations $\rho_j, j = 1, ..., J$;
set backorder-to-holding cost ratio $c = \frac{b}{b+h}$;
set threshold $t = \frac{c}{\prod_{j=1}^{J}(1-\rho_j)}$;
set base stock level $z = 0$;
set $g(0, j) = 1, j = 1, ..., J$;
set $s = 1$;

Iteration:

1. If $s \geq t$, stop and set $z_{opt} = 0$; else go to 2;
2. Update $z := z + 1$;
3. Calculate $g(z, j), j = 1, ..., J$, with $g(z, 1) = \rho_1^z$ and $g(z, j) = g(z, j - 1) + \rho_j \cdot g(z - 1, j)$ for $j = 2, ..., J$;
4. Update $s := s + g(z, J)$;
5. If $s \geq t$, stop and set $z_{opt} = z$; else go to 2;

References

Akyildiz, I. F., Bolch, G., 1983. Erweiterung der Mittelwertanalyse zur Berechnung der Zustandswahrscheinlichkeiten für geschlossene und gemischte Netze (Extension of the mean value analysis to compute stationary distributions of closed and mixed networks). In: Messung, Modellierung und Bewertung von Rechensystemen. Springer, pp. 267–276.

Bruell, S. C., Balbo, G., 1980. Computational algorithms for closed queueing networks. North Holland.

Buzacott, J. A., Price, S. M., Shanthikumar, J. G., 1992. Service level in multistage MRP and base stock controlled production systems. In: Fandel, G., Gulledge, T., Jones, A. (Eds.), New Directions for Operations Research in Manufacturing. Springer, pp. 445–463.

Buzen, J. P., 1973. Computational algorithms for closed queueing networks with exponential servers. Communications of the ACM 16 (9), 527–531.

Chao, X., Miyazawa, M., Pinedo, M., 1999. Queueing networks: customers, signals and product form solutions. Wiley, Chichester.

Chen, H., Yao, D. D., 2001. Fundamentals of queueing networks: Performance, asymptotics, and optimization. Vol. 46. Springer.

Krishnamoorthy, A., Lakshmy, B., Manikandan, R., 2011. A survey on inventory models with positive service time. Opsearch 48 (2), 153–169.

Lee, Y.-J., Zipkin, P., 1992. Tandem queues with planned inventories. Operations Research 40 (5), 936–947.

Mueller, A., Stoyan, D., 2002. Comparison methods for stochastic models and risks. Wiley, Chichester.

Reiser, M., Lavenberg, S. S., 1980. Mean-value analysis of closed multichain queuing networks. Journal of the ACM 27 (2), 313–322.

Rubio, R., Wein, L. M., 1996. Setting base stock levels using product-form queueing networks. Management Science 42 (2), 259–268.

Spearman, M. L., Zazanis, M. A., 1992. Push and pull production systems: issues and comparisons. Operations Research 40 (3), 521–532.

Toktay, L. B., Wein, L. M., Zenios, S. A., 2000. Inventory management of remanufacturable products. Management Science 46 (11), 1412–1426.

Zazanis, M. A., 1994. Push and pull systems with external demands. In: Proc. 32nd Allerton Conf. on Communication, Control and Computing. Allerton, IL, pp. 691–698.