Understanding the Performance of the Long Chain and Sparse Designs in Process Flexibility

David Simchi-Levi
Engineering Systems Division, Department of Civil and Environmental Engineering, and the Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, dslevi@mit.edu

Yehua Wei
Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, y4wei@mit.edu

The long chain has been an important concept in the design of flexible processes. This design concept, as well as other sparse designs, have been applied by the automotive and other industries as a way to increase flexibility in order to better match available capacities with variable demands. Numerous empirical studies have validated the effectiveness of these designs. However, there is little theory that explains the effectiveness of the long chain, except when the system size is large, i.e., by applying an asymptotic analysis.

Our attempt in this paper is to develop a theory that explains the effectiveness of long chain designs for finite size systems. First, we uncover a fundamental property of long chains, supermodularity, that serves as an important building block in our analysis. This property is used to show that the marginal benefit, i.e., the increase in expected sales, increases as the long chain is constructed, and the largest benefit is always achieved when the chain is closed by adding the last arc to the system. Then, supermodularity is used to show that the performance of the long chain is characterized by the difference between the performances of two open chains. This characterization immediately leads to the optimality of the long chain among 2-flexibility designs. Finally, under independent and identically distributed (i.i.d.) demand, this characterization gives rise to three developments: (i) an effective algorithm to compute the performances of long chains using only matrix multiplications; (ii) a result that the gap between the fill rate of full flexibility and that of the long chain increases with system size, thus implying that the effectiveness of the long chain relative to full flexibility increases as the number of products decreases; (iii) a risk-pooling result implying that the fill rate of a long chain increases with the number of products, but this increase converges to zero exponentially fast.

Subject classifications: flexible manufacturing; capacity pooling; facility planning: design; network flow; supermodular function; random walk.

Area of review: Operations and Supply Chains.

History: Received May 2011; revisions received September 2011, February 2012; accepted March 2012. Published online in Articles in Advance September 26, 2012.

1. Introduction

For many manufacturing firms, the ability to match demand and supply is key to their success. Failure to do so could lead to loss of revenue, reduced service levels, impact on reputation, and decline in the company’s market share. Unfortunately, recent developments such as intense market competition, product proliferation, and the increase in the number of products with short life cycle have created an environment where customer demand is volatile and unpredictable. In such an environment, traditional operations strategies such as building inventory, investing in capacity buffers, or increasing committed response time to consumers do not offer manufacturers a competitive advantage. Therefore, many manufacturers have started to adopt an operations strategy known as process flexibility to better respond to market changes without significantly increasing cost, inventory, or response time (see Simchi-Levi 2010).

Process flexibility is defined as the ability to “build different types of products in the same manufacturing plant or on the same production line at the same time” (Jordan and Graves 1995, p. 577). For example, in “full” (process) flexibility, each plant is capable of producing all products. In this case, when the demand for one product is higher than expected while the demand for a different product is lower than expected, a flexible manufacturing system can quickly make adjustments by shifting production capacities appropriately. By contrast, in a “dedicated” strategy (sometimes called “no flexibility”), each plant is responsible for a single product and hence does not have the same ability to match supply with demand.

Because of its effectiveness in responding to uncertainties, process flexibility has gained significant attention in several industries, in particular, in the automotive industry. Indeed, the plants for most of the automobile giants are much more flexible (in terms of process flexibility) today compared to twenty years ago (Boudette 2006). Evidently, it is often too expensive to achieve a high degree of flexibility, for example full flexibility, and as a result, sparse or partial flexibility, is implemented instead. One set of sparse flexibility designs is the 2-flexibility designs.
A flexibility design is a 2-flexibility design if each plant produces exactly two products and demand for each product can be satisfied from exactly two plants.

Of course, there are many ways to implement sparse designs and the challenge is to identify an effective one. An important concept analyzed in the literature and applied in practice by various companies is the concept of the long chain. The first to observe the power of the long chain were Jordan and Graves (1995) who, through empirical analysis, showed that the long chain design can provide almost as much benefit as full flexibility. In particular, Jordan and Graves (1995) found that for randomly generated demand, the expected amount of demand that can be satisfied by a long chain design is very close to that of a full flexibility design. Unfortunately, with a few exceptions (see Chou et al. 2010b), there is very little theory to explain why long chain works so well. One objective of our research is to provide an answer to this question.

Throughout the paper, we consider balanced manufacturing systems, i.e., manufacturing systems with an equal number of plants and products, and all plants have the same capacity. Although this is rarely the case in the real world, as Jordan and Graves (1995) argues, understanding balanced systems can provide insights into very realistic scenarios.

Given a balanced manufacturing system, a flexibility design \(\mathcal{A}\) is represented by the arc set of a directed bipartite graph, where an arc from plant node \(i\) to product node \(j\) implies that plant \(i\) is capable of producing product \(j\). For example, if \(\mathcal{A}\) is a dedicated design, then \(\mathcal{A}\) has exactly \(n\) arcs such that each plant node is incident to one arc and each product node is incident to one arc. By contrast, if \(\mathcal{A}\) is a full flexibility design, then \(\mathcal{A}\) has arcs connecting every plant node to all product nodes.

Because \(\mathcal{A}\) is represented by a bipartite graph, applying standard graph theory notation, we define an undirected cycle in \(\mathcal{A}\) to be a set of arcs that forms a cycle when the arc directions are ignored. A flexibility design \(\mathcal{A}\) is a long chain if its arcs form exactly one undirected cycle containing all plant and product nodes (see Figure 1 for an example). A closed chain is defined as an induced subgraph in \(\mathcal{A}\) that forms an undirected cycle, whereas an open chain is an induced subgraph in \(\mathcal{A}\) that forms an undirected line (one arc less than an undirected cycle). In Figure 1, an example of an open and a closed chain is presented. It can be seen that any 2-flexibility design, where each product/plant node is incident to two arcs, is the union of a number of closed chains.

This paper studies the performance of flexibility designs on balanced systems with stochastic demand. In this model, given a random demand instance \(d\), the level of sales associated with a flexibility design \(\mathcal{A}\) is characterized by the maximum demand satisfied without exceeding plant capacities and without violating the constraints associated with the flexibility design. This quantity can be determined by solving a simple max-flow problem. We use \(D\) to denote the random distribution of demand, and the performance of a flexibility design is measured by the expected sales over all random realizations. In what follows, we use the terms performance and expected sales interchangeably. Over the years, expected sales has been a popular metric for studying and evaluating the effectiveness of flexibility designs; see the survey by Chou et al. (2008). Finally, the fill rate of a flexibility design is defined as the ratio between the performance of a given flexibility design and the total expected demand.

The first part of our paper, §4, is motivated by an observation that has been made in the literature (Graves 2008, Hopp et al. 2004) regarding the performance of the long chain for a balanced system when product demands are independent and identically distributed (i.i.d.). The observation states that if one starts with a dedicated design and adds arcs to create the long chain, the incremental benefits, or the change in performance, associated with each added arc is increasing.

To illustrate this observation, consider an example with six plants and six products, where the demand for each product is equal to either 0.8, 1, or 1.2 with equal probabilities, and the capacity of each plant is one. Then, we start with a dedicated flexibility design (the dashed arcs in Figure 2(a)), and add arcs \((1, 2), (2, 3), \ldots, (5, 6),\) and \((6, 1)\) one at a time, until we complete the long chain. Each time we add such an arc, we determine the expected sales associated with the resulting design at that time. Figure 2(b) displays the performance of the flexibility designs at different stages, as well as the incremental benefits when a new arc is added. As you can see, the incremental benefits increase as we add more arcs. The biggest impact, surprisingly, occurs when we add the last arc and close the long chain. To formally prove this observation, we identify in §4 an important property of the long chain, supermodularity, and apply it as the long chain is constructed.

Our second main result, §5, characterizes the performance of a long chain as the difference between the performances of two open chains. Interestingly, this new result also depends on the supermodularity property identified in §4. A direct consequence of this characterization of the

Figure 1. Configurations for flexibility designs.
performance of the long chain is that the long chain design is optimal among all 2-flexibility designs.

Finally, in §6, we further apply this characterization in three different ways. First, we develop an effective method to compute the performance of long chains of finite size. Second, we show that the difference between the fill rate of full flexibility and that of the long chain increases with system size, thus implying that the effectiveness of the long chain is closer to that of full flexibility as the number of products decreases. Lastly, we prove a risk-pooling result showing that the fill rate of the long chain increases with the number of products, but this increase converges to zero exponentially fast.

Next, we provide a literature review. This is followed by §3 where we introduce notations and definitions.

2. Literature Review

The study of process flexibility, also known as “mix flexibility” or “short-term flexibility,” began in the 1980s. We refer readers to the survey of Sethi and Sethi (1990) for a detailed introduction and literature review on process flexibility prior to the 1990s. This research typically focused on the benefits, challenges, and trade-offs between fully flexible and dedicated systems. Unfortunately, most companies are not interested in full flexibility because of its enormous implementation cost.

The seminal paper of Jordan and Graves (1995) is the first to consider the design and effectiveness of limited degree of (process) flexibility. Applying numerical analysis (simulation) to a stochastic demand model, Jordan and Graves demonstrate two important insights in process flexibility. First, they show that a sparse flexibility design, which they called the long chain, can provide almost as much benefit as full flexibility. Second, they introduce the concept of chaining, which generalizes the idea of the long chain, and show that this concept leads to sparse designs that perform extremely well in numerical studies.

The paper of Jordan and Graves (1995) paved ways for two important lines of research in the studies of process flexibility. The first is to apply and extend the ideas of Jordan and Graves to other settings. Indeed, since the publication of Jordan and Graves (1995), there has been extensive research that has applied the concept of the long chain and limited flexibility in a variety of applications. For instance, Graves and Tomlin (2003) study process flexibility in multistage supply chains; Sheikhzadeh et al. (1998) and Gurumurthi and Benjaafar (2004) in queueing systems; Iravani et al. (2005) in queueing networks and Wallace and Whitt (2005) in call centers. We also note the work of Hopp et al. (2004), Bish et al. (2005), Muriel et al. (2006), Mak and Shen (2009), Chou et al. (2010a), and Bassamboo et al. (2010) for research on flexibility in other settings.

The other line of research tries to establish a theoretical foundation that explains the effectiveness of limited degree of flexibility, and in particular, the insights provided by the work of Jordan and Graves (1995). Surprisingly, very little theory has been developed thus far. Two of the rare exceptions are Chou et al. (2010b, 2011). In Chou et al. (2010b), the authors develop a method to compute the ratio between the performance of the long chain to that of full flexibility in asymptotic regime. They also present a constraint sampling method to find a sparse flexibility design that performs almost as well as full flexibility. In a companion paper, Chou et al. (2011) prove that when demands are bounded by a constant, there exists a sparse flexibility with graph expansion property that achieves sales close to that of full flexibility in the worst-case demand scenario. Moreover, Chou et al. (2011) describe a heuristic to find such structures. We note that both the constraint sampling method and results in Chou et al. (2011) hold in the general system where the number of plants and products need not be equal, with arbitrary plant capacities and random distribution of demands. Finally, Aksin and Karaesmen (2007) show that there is a decrease in marginal benefit associated with the increase in either the degree of flexibility or the capacities of the manufacturing plants.
For more information about the literature on flexibility, we refer readers to surveys by Buzzacott and Mandelbaum (2008) and Chou et al. (2008). The first survey introduces different applications of flexibility, and the second focuses on the recent theoretical advancements in process flexibility.

3. Definitions and Notation

Consider a balanced manufacturing system of size \( n \) facing random demand. In such a system, there are \( n \) plants, each with unit capacity and \( n \) different products. Throughout the paper, the vector \( \mathbf{D} \) is used to denote the random demand distribution, and \( \mathbf{d} \) is a particular random instance. Because in practice demand is never negative, this paper assumes that \( \mathbf{D} \) is nonnegative.

For a balanced system of size \( n \), we say that its demand, \( \mathbf{D} \), is exchangeable if \( [D_1, \ldots, D_n] \) equals to \( [D_{\sigma(1)}, \ldots, D_{\sigma(n)}] \) in distribution for any \( \sigma \) that is a permutation of \( \{1, 2, \ldots, n\} \). We note that any i.i.d. demand is exchangeable but not all exchangeable demand are i.i.d. For example, consider a random vector \( \mathbf{D} = [D_1, \ldots, D_n] \) that is uniformly distributed on the linear polyhedron

\[
\left\{ (x_1, \ldots, x_n) \left| \sum_{i=1}^n x_i = n, x_i \geq 0, \forall i = 1, \ldots, n \right. \right\}.
\]

Clearly \( \mathbf{D} \) is exchangeable, but the random variables in \( \mathbf{D} \) are not independent, because they always sum up to \( n \).

Next, we define several classes of flexibility designs for balanced manufacturing systems. For any integer \( n \geq 2 \), the dedicated design for a balanced system of size \( n \), \( \mathcal{D}_n \), is defined as \( \mathcal{D}_n = \{(i, i) \mid i = 1, 2, \ldots, n\} \); the long chain flexibility design for a balanced system of size \( n \), \( \mathcal{C}_n \), is defined as \( \mathcal{C}_n = \mathcal{D}_n \cup \{(i, i + 1) \mid i = 1, 2, \ldots, n - 1\} \cup \{(n, 1)\} \); and full flexibility design for a balanced system of size \( n \), \( \mathcal{F}_n \), is defined as \( \mathcal{F}_n = \{(i, j) \mid i, j = 1, 2, \ldots, n\} \). In flexibility designs, we refer to an arc \( (i, i) \) as a dedicated arc and arc \( (i, j), i \neq j \) as a flexible arc.

Given a random instance of the demand, \( \mathbf{d} \), the maximum sales that can be achieved by a flexibility design \( \mathcal{A} \), denoted by \( P(\mathbf{d}, \mathcal{A}) \), is defined as

\[
P(\mathbf{d}, \mathcal{A}) = \max \sum_{1 \leq i, j \leq n} f_{ij} \text{ s.t. } \sum_{i=1}^n f_{ij} \leq d_j, \quad \forall 1 \leq j \leq n, \quad \sum_{j=1}^n f_{ij} \leq 1, \quad \forall 1 \leq i \leq n, \quad f_{ij} \geq 0, \quad \forall (i, j) \in \mathcal{A}, \quad f_{ij} = 0, \quad \forall (i, j) \notin \mathcal{A}, \quad \mathbf{f} \in \mathbb{R}^n.
\]

Throughout the paper, this optimization problem is referred to as the optimization problem associated with \( P(\mathbf{d}, \mathcal{A}) \), or simply, \( P(\mathbf{d}, \mathcal{A}) \), when there is no ambiguity. It is not difficult to see that this optimization problem is a max-flow problem, and as a result, we refer to \( f_{ij} \) as the flow on arc \((i, j)\).

Under random demands \( \mathbf{D} \), we define the performance, also referred to as expected sales, of \( \mathcal{A} \) to be \( E[P(\mathbf{D}, \mathcal{A})] \), where \( E[\cdot] \) is the expectation of a random variable. For succinctness, we also use \([\mathcal{A}]\) to denote this quantity, when the random vector \( \mathbf{D} \) is given. Finally, for any integer \( k \geq 0 \), we define open chain \( \mathcal{L}_k \) as \( \mathcal{L}_k = \mathcal{D}_k \cup \{ (i, i + 1) \mid i = 1, \ldots, k - 1 \} \). One can think of \( \mathcal{L}_k \) as the open chain that connects plant \( 1 \) to product \( k \). Note that \( \mathcal{L}_k \) is simply \( \mathcal{C}_k \setminus \{(k, 1)\} \).

4. Supermodularity in Long Chains

In this section, we establish an important building block in our analysis by identifying the supermodularity of flexible arcs in the long chain design. Throughout the section, we consider a balanced system of size \( n \) facing exchangeable random demand \( \mathbf{D} \). In this section we first define and prove the supermodularity property (§4.1), and then apply the supermodularity property (§4.2) to show that the incremental benefits are nondecreasing when the long chain is constructed.

4.1. Proof of Supermodularity

We start by formally defining the notion of supermodularity.

**Definition 1.** A function \( f(x, y) \) is said to be supermodular in \( x \) and \( y \) if for any real numbers \( x', x'', y', y'' \),

\[
f(\max\{x', x''\}, \max\{y', y''\}) + f(\min\{x', x''\}, \min\{y', y''\}) \geq f(x', y') + f(x'', y').
\]

Next, consider a design \( \mathcal{A} \), a demand instance \( \mathbf{d} \), and two specific arcs \( \alpha \) and \( \beta \) with given nonnegative capacities \( u_\alpha \) and \( u_\beta \). Define

\[
P_{\alpha, \beta}(u_\alpha, u_\beta, \mathbf{d}, \mathcal{A}) = \max \sum_{i, j} f_{ij} \text{ s.t. } \sum_i f_{ij} \leq d_j, \quad \sum_j f_{ij} \leq 1, \quad f_{ij} \geq 0, \quad \forall (i, j) \in \mathcal{A}, \quad f_{ij} = 0, \quad \forall (i, j) \notin \mathcal{A}, \quad \mathbf{f} \in \mathbb{R}^n.
\]

We prove that if \( \mathcal{A} \subset \mathcal{C}_n \), then for any two flexible arcs \( \alpha \) and \( \beta \) in \( \mathcal{A} \), \( P_{\alpha, \beta}(u_\alpha, u_\beta, \mathbf{d}, \mathcal{A}) \) is supermodular in \( u_\alpha \) and \( u_\beta \). Note that if \( \alpha \) and \( \beta \) are not in \( \mathcal{A} \), then
Plants $X$ to the upper bound for the flow on an arc connecting product $C$ (that is, the arcs in the transportation problem, we set the weight of each plant to product arc to zero. The upper bound (capacity) on the flow on an arc in a balanced system of size five. Additional node $s$ (1981). Define $G_4$ equivalent to a max-weight circulation problem, which contains a max-weight circulation problem, which contains $A$ in $A_4$.

Finally, we set the lower bound for the flow on every arc in $A$ bound for the flow on every arc in $A$. Bound for the flow on each product arc (i.e., $X_4$ contains 2k–1 arcs).

**Lemma 1.** Let $\alpha$ and $\beta$ be two flexible arcs in $C_n$. Then $\alpha$ and $\beta$ are in series in $G(C_n)$, where $G(C_n)$ is the underlying graph of the max-weight circulation problem for $P_{\alpha, \beta}(u_a, u_b, d, C_n)$.

**Proof of Lemma 1.** Let $C$ be an arbitrary undirected cycle in $G(C_n)$. If $C$ does not contain node $s$, then $C$ must be the undirected cycle that contains every plant to product arcs in $C_n$. If $X_4$ contains 2k–1 arcs. This assumption implies that all arcs in $X_4$ are plant to product arcs (i.e., $X_4 \subset C_n$). The structure of $C_n$ implies that $X_4$ must contain an odd number of arcs. Moreover, the path in $X_5 \cup \{\alpha\} \cup \{\beta\}$ has alternating directions for every two consecutive arcs, and therefore $\alpha$ and $\beta$ have the same direction in $C$. This is illustrated by Figure 4. Because this is true for any arbitrary undirected cycle $C$, $\alpha$ and $\beta$ are in series in $G(C_n)$.

Lemma 1 allows us to apply the following important result of Gale and Politof (1981). They show that if two arcs, $\alpha$ and $\beta$, in the underlying graph are in series, then the optimal flow of the max-weight circulation is supermodular with respect to the capacities of both arcs.

**Theorem 1.** Let $A$ be a flexibility design for a balanced system of size $n$, and $A \subset C_n$. For any flexible arcs $\alpha, \beta$ in $A$, $P_{\alpha, \beta}(u_a, u_b, d, A)$ is supermodular in $u_a$ and $u_b$. Hence,

$$P(d, A) + P(d, A \setminus \{\alpha, \beta\}) \geq P(d, A \setminus \{\alpha\}) + P(d, A \setminus \{\beta\}).$$

**Proof of Theorem 1.** By construction, $P_{\alpha, \beta}(u_a, u_b, d, A)$ can be computed by solving the max-weight circulation problem. Because $A \subset C_n$, the set arcs in $G(A)$ is a subset of the set of arcs in $G(C_n)$. By Lemma 1, $\alpha$ and $\beta$ are in series in $G(C_n)$. Thus, $\alpha$ and $\beta$ are in series in $G(A)$.

Applying the main theorem in Gale and Politof (1981), we have that $P_{\alpha, \beta}(u_a, u_b, d, A)$ is supermodular in $u_a$ and $u_b$.

Hence

$$P(1, 1, d, A) + P(0, 0, d, A) \geq P(0, 1, d, A) + P(1, 0, d, A)$$

$$\Rightarrow P(d, A) + P(d, A \setminus \{\alpha, \beta\}) \geq P(d, A \setminus \{\alpha\}) + P(d, A \setminus \{\beta\}).$$

$\square$
Because the theorem holds for any realization of demand, it must be also true in expectation. Thus,

**Corollary 1.** For any flexible arcs \( \alpha, \beta \) in \( \mathcal{A} \subset \mathcal{E}_a \),
\[
E[P_{\alpha,\beta}(u_{\alpha}, u_{\beta}, \mathbf{D}, \mathcal{S})] \text{ is supermodular in } u_{\alpha} \text{ and } u_{\beta} \text{ for any random distributions } \mathbf{D}.
\]

The corollary thus suggests that any two flexible arcs in the long chain complement each other. That is, the existence of one flexible arc increases the marginal benefit that can be gained when the other flexible arc is added.

Interestingly, the supermodular result of Gale and Politof (1981) can be extended for two sets of arcs \( X \) and \( Y \), where any pair of arcs in \( X \cup Y \) are in series with each other. Although that was not stated in the paper of Gale and Politof (1981), it was proven by Granot and Veinott Jr. (1985) in a more general setting. Here, we state Corollary 2, which is a special case from (Granot and Veinott Jr. 1985, Theorem 17).

**Corollary 2.** Let \( \mathcal{A} \) be a flexibility design for a balanced system of size \( n \), and \( \mathcal{A} \subset \mathcal{E}_a \). For any \( X, Y \subseteq S \), where \( S \) is the set of all flexible arcs in \( \mathcal{A} \), and demand instance \( \mathbf{d} \),
\[
P(\mathbf{d}, \mathcal{A} \setminus (X \cap Y)) + P(\mathbf{d}, \mathcal{A} \setminus (X \cup Y)) \geq P(\mathbf{d}, \mathcal{A} \setminus X) + P(\mathbf{d}, \mathcal{A} \setminus Y).
\]

### 4.2. Incremental Benefits in Long Chains

In this section, we apply Corollary 1 to formally prove the observation made in the introduction that the incremental benefits associated with adding arcs to the long chain is increasing. Consider the following sequence of flexibility designs: \( \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_n \), where we define \( \mathcal{L}_1 = \mathcal{D}_n \) and \( \mathcal{L}_k = \mathcal{L}_{k-1} \cup \{(i, i) \mid i = k + 1, \ldots, n\} \). In words, \( \mathcal{L}_1 \) is simply the open chain from plant 1 to product 1 plus the dedicated arcs connecting plants \( i \) to products \( i \) for all \( k < i \leq n \). Finally, recall that \( \mathcal{E}_a \) is the long chain of size \( n \).

In the example of Graves (2008) and the table in Figure 2 of this paper, one starts at \( \mathcal{L}_1 \) and adds arcs sequentially to create \( \mathcal{L}_2, \ldots, \mathcal{L}_n, \mathcal{E}_a \). Now, we apply the supermodularity result to show that the incremental benefit, \( [\mathcal{L}_a^k] - [\mathcal{L}_a^{k-1}] \), is nondecreasing with \( k \).

**Theorem 2.** For any balanced system of size \( n \) with exchangeable demand, we have
\[
[\mathcal{L}_1^a] - [\mathcal{L}_2^a] \leq [\mathcal{L}_3^a] - [\mathcal{L}_2^a] \leq \cdots \leq [\mathcal{L}_n^a] - [\mathcal{L}_{n-1}^a] \leq [\mathcal{E}_a] - [\mathcal{L}_1^a].
\]

**Proof of Theorem 2.** Fix any \( 1 \leq k \leq n - 1 \). Let \( \alpha = (1, 2), \beta = (k, k + 1) \). By Corollary 1, we have
\[
E[P_{\alpha,\beta}(1, 1, \mathbf{D}, \mathcal{L}_{k+1}^a)] + E[P_{\alpha,\beta}(0, 0, \mathbf{D}, \mathcal{L}_{k+1}^a)] \geq E[P_{\alpha,\beta}(1, 0, \mathbf{D}, \mathcal{L}_{k+1}^a)] + E[P_{\alpha,\beta}(0, 1, \mathbf{D}, \mathcal{L}_{k+1}^a)].
\]

Setting \( u_{\alpha} = 0 \) is equivalent to deleting arc \( \alpha \) in the optimization problem associated with \( P_{\alpha,\beta}(u_{\alpha}, u_{\beta}, \mathbf{d}, \mathcal{S}) \), whereas setting \( u_{\beta} = 1 \) implies that this arc exists in the same model and its capacity is redundant. As a result, we have that,
\[
E[P_{\alpha,\beta}(1, 1, \mathbf{D}, \mathcal{L}_{k+1}^a)] = E[P(\mathbf{D}, \mathcal{L}_{k+1}^a)] = [\mathcal{L}_{k+1}^a],
\]
and
\[
E[P_{\alpha,\beta}(0, 0, \mathbf{D}, \mathcal{L}_{k+1}^a)] = E[P(\mathbf{D}, \mathcal{L}_k^a)] = [\mathcal{L}_k^a].
\]

Let \( \mathbf{D}_\alpha = [D_2, D_3, \ldots, D_n, D_1] \), then
\[
E[P_{\alpha,\beta}(1, 0, \mathbf{D}, \mathcal{L}_{k+1}^a)] = E[P(\mathbf{D}_\alpha, \mathcal{L}_{k+1}^a)] = E[P(\mathbf{D}, \mathcal{L}_k^a)] = [\mathcal{L}_k^a],
\]
and
\[
E[P_{\alpha,\beta}(0, 1, \mathbf{D}, \mathcal{L}_{k+1}^a)] = E[P(\mathbf{D}_\alpha, \mathcal{L}_{k+1}^a)] = E[P(\mathbf{D}, \mathcal{L}_k^a)].
\]

where the second to last equality in (4) and (5) holds because the random vector \( \mathbf{D} \) is exchangeable. Substituting Equations (2)–(5) into Inequality (1), we obtain that
\[
[\mathcal{L}_k^a] - [\mathcal{L}_k^a] \geq [\mathcal{L}_k^a] - [\mathcal{L}_{k-1}^a],
\]
for \( k = 2, \ldots, n - 1 \).

Finally, to show \([\mathcal{E}_a] - [\mathcal{L}_{n-1}^a] \leq [\mathcal{E}_a] - [\mathcal{L}_1^a] \), let \( \alpha = (1, 2), \beta = (n, 1) \) and let \( \mathbf{D}_\alpha = [D_2, D_3, \ldots, D_n, D_1] \). Then
\[
E[P_{\alpha,\beta}(1, 1, \mathbf{D}, \mathcal{E}_a)] = [\mathcal{E}_a],
\]
\[
E[P_{\alpha,\beta}(0, 0, \mathbf{D}, \mathcal{E}_a)] = [\mathcal{L}_a^a],
\]
\[
E[P_{\alpha,\beta}(0, 1, \mathbf{D}, \mathcal{E}_a)] = E[P(\mathbf{D}_\alpha, \mathcal{L}_{n-1}^a)] = [\mathcal{L}_{n-1}^a],
\]
and
\[
E[P_{\alpha,\beta}(0, 1, \mathbf{D}, \mathcal{E}_a)] = E[P(\mathbf{D}_\alpha, \mathcal{L}_{n-1}^a)] = [\mathcal{L}_a^a].
\]

By Corollary 1,
\[
E[P_{\alpha,\beta}(1, 1, \mathbf{D}, \mathcal{E}_a)] + E[P_{\alpha,\beta}(0, 0, \mathbf{D}, \mathcal{E}_a)] \geq E[P_{\alpha,\beta}(1, 0, \mathbf{D}, \mathcal{E}_a)] + E[P_{\alpha,\beta}(0, 1, \mathbf{D}, \mathcal{E}_a)],
\]
we have that \([\mathcal{E}_a] - [\mathcal{L}_a^a] \geq [\mathcal{E}_a] - [\mathcal{L}_{n-1}^a] \). This completes the proof.

Observe that the proof of Theorem 2 requires the application of the supermodularity result (Theorem 1), which holds deterministically for any fixed demand instance. By contrast, Theorem 2 holds only stochastically under exchangeable demand but does not hold for any fixed demand instance.

### 5. Characterizing the Performance of the Long Chain

In this section, we show that in a balanced system of size \( n \) with exchangeable demand, the performance of the long chain can be characterized by the difference between the performances of two open chains. Like the previous section, we start the section by developing several properties of the long chain when the demand is deterministic.
5.1. Long Chain with Deterministic Demand

In this subsection, we fix an arbitrary demand instance $d$. Throughout the subsection, when some integer $k$ appears in a statement, we are in fact referring to some $i \in \{1, \ldots, n\}$ congruent to $k$ modulo $n$. For example, if $f_{n+1,n+2}$ appears in a statement, then we are referring to plant 3; and if $f_{n+1,n+2}$, the flow from plant $n+1$ to product $n+2$ appears in a statement, then we are referring to $f_{1,2}$, the flow from plant 1 to product 2.

First, we start with the following lemma.

**Lemma 2.** Suppose $P(d, e_{n}\setminus\{\alpha\}) = P(d, e_{n})$, where $\alpha$ is a flexible arc in $e_{n}$. Then, for any set $X \subseteq S$, where $S$ is the set of all flexible arcs in $e_{n}$, we have that

$$P(d, e_{n}\setminus(X \cup \{\alpha\})) = P(d, e_{n}X).$$

**Proof of Lemma 2.** If $\alpha \in X$, the result is trivial as $X \cup \{\alpha\} = X$. Otherwise, by Corollary 2,

$$P(d, e_{n}\setminus(X \cup \{\alpha\})) + P(d, e_{n}) \geq P(d, e_{n}\setminus X) + P(d, e_{n}\setminus\{\alpha\})$$

$$\Rightarrow P(d, e_{n}\setminus(X \cup \{\alpha\})) \geq P(d, e_{n}\setminus X),$$

since $P(d, e_{n}) = P(d, e_{n}\setminus\{\alpha\})$.

But by definition of $P(\cdot)$, $P(d, e_{n}\setminus(X \cup \{\alpha\})) \leq P(d, e_{n}\setminus X)$, hence

$$P(d, e_{n}\setminus(X \cup \{\alpha\})) = P(d, e_{n}\setminus X).$$

Next, we show that the sales associated with $e_{n}$ can be expressed as a sum of $n$ quantities, where each quantity is the difference of the sales associated with two open chains in $e_{n}$.

**Theorem 3.** For any fixed demand instance $d$ on balanced system of size $n$, we have

$$P(d, e_{n}) = \sum_{i=1}^{n} (P(d, e_{n}\setminus\{(i, i + 1)\}) - P(d, e_{n}\setminus\{(i - 1, i), (i, i), (i, i + 1)\})).$$

**Proof of Theorem 3.** For each $1 \leq k_{1} \leq k_{2} \leq n$, define $X_{k_{1} \rightarrow k_{2}} = \{(i, i) \mid i = k_{1}, k_{1} + 1, \ldots, k_{2}\} \cup \{(i, i + 1) \mid i = k_{1}, k_{1} + 1, \ldots, k_{2} - 1\}$, and for each $1 \leq k_{2} < k_{1} \leq n$, define $X_{k_{1} \rightarrow k_{2}} = \{(i, i) \mid i = k_{1}, k_{1} + 1, \ldots, n, 1, 2, \ldots, k_{2}\} \cup \{(i, i + 1) \mid i = k_{1}, k_{1} + 1, \ldots, k_{2} - 1\}$. One can think of $X_{k_{1} \rightarrow k_{2}}$ as the open chain connecting plant $k_{1}$ to product $k_{2}$ in the balanced system of size $n$. Also, because demand instance $d$ is fixed, for the sake of succinctness, we use $P(\cdot)$ to denote $P(d, \cdot)$. Finally, we define $\alpha_{i} = (i, i + 1)$ and $\beta_{i} = (i, i)$ for $i = 1, 2, \ldots, n$ (note that $\alpha_{n} = (n, 1)$ as $n + 1$ is congruent with one modular $n$).

By definition of $\alpha$ and $\beta$, we can rewrite $e_{n}\setminus\{(i, i + 1)\}$ and $e_{n}\setminus\{(i - 1, i), (i, i), (i, i + 1)\}$ as $e_{n}\setminus\{\alpha_{i}\}$ and $e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\}$. For any $1 \leq i \leq n$, because $e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}\} = \{\beta_{i}\} \cup e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\}$, where $\cup$ represents the symbol for disjoint union,

$$P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\})$$

$$= P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}\}) + \min\{1, d_{i}\}. \quad (6)$$

Lemma 9 in Appendix B shows that there is some $i^{*}$ such that $P(e_{n}) = P(e_{n}\setminus\{\alpha_{i^{*}}\})$. Without loss of generality, we assume that $i^{*} = n$, because we can always relabel each plant (and product) $i$ by $i - i^{*}$. Now, we have that for $i = 2, \ldots, n - 1$,

$$P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\})$$

$$= P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}\}) + \min\{1, d_{i}\}$$

(by Equation (6))

$$= P(e_{n}\setminus\{\alpha_{i}, \alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \alpha_{i}\}) + \min\{1, d_{i}\}$$

(by Lemma 2).

Because $e_{n}\setminus\{\alpha_{i}, \alpha_{i}\} = X_{1 \rightarrow i^{*}} \cup X_{(i+i-1) \rightarrow n}$, and $e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \alpha_{i}\} = X_{1 \rightarrow i^{*}} \cup X_{(i+i-1) \rightarrow n} \cup \{\beta_{i}\}$, we have for $i = 2, \ldots, n - 1$,

$$P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\})$$

$$= P(e_{n}\setminus\{\alpha_{i}, \alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \alpha_{i}\}) + \min\{1, d_{i}\}$$

(by Equation (6))

$$= P(e_{n}\setminus\{\alpha_{i}, \alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \alpha_{i}\}) + \min\{1, d_{i}\}$$

(by Lemma 2).

$$= P(X_{1 \rightarrow i^{*}}) - P(X_{1 \rightarrow (i^{*}-1)}). \quad (7)$$

Also,

$$P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\})$$

$$= P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}\}) + \min\{1, d_{i}\}$$

(by Equation (6))

$$= P(e_{n}\setminus\{\alpha_{i}, \alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}\}) + \min\{1, d_{i}\}$$

(by Lemma 2)

$$= \min\{1, d_{i}\}, \quad (8)$$

and

$$P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\})$$

$$= P(X_{1 \rightarrow n}) - P(X_{1 \rightarrow (n-1)}). \quad (9)$$

Now, applying Equations (7)–(9), we obtain that

$$\sum_{i=1}^{n} (P(e_{n}\setminus\{\alpha_{i}\}) - P(e_{n}\setminus\{\alpha_{i-1}, \alpha_{i}, \beta_{i}\}))$$

$$= \min\{1, d_{i}\} + \sum_{i=2}^{n} (P(X_{1 \rightarrow i^{*}}) - P(X_{1 \rightarrow (i^{*}-1)}))$$

$$= \min\{1, d_{i}\} + P(X_{1 \rightarrow n}) - P(X_{1 \rightarrow 1})$$

$$= P(X_{1 \rightarrow n})$$

$$= P(e_{n}). \quad \square$$
We note that $\mathcal{C}_n \setminus \{(i, i + 1)\}$ is an open chain connecting plant $i + 1$ to product $i$, and $\mathcal{C}_n \setminus \{(i - 1, i), (i, i), (i, i + 1)\}$ is an open chain connecting plant $i + 1$ to product $i - 1$. Also, Theorem 3 can be extended to a more general setting. This extension of Theorem 3 is presented in Appendix A.

5.2. Characterization and Optimality of the Long Chain

With Theorem 3, we can now characterize the performance of the long chain using the performances of open chains.

**Theorem 4.** For any balanced system of size $n$ with exchangeable demand $D$, we have

$$\mathbb{E}_n = n(\mathbb{L}_n - \mathbb{L}_{n-1}).$$

**Proof of Theorem 4.** Theorem 3 states that for any $d$, that is an instance of $D$,

$$P(d, \mathcal{C}_n) = \sum_{i=1}^{n} \left(P(d, \mathcal{C}_n \setminus \{(i, i + 1)\}) - P(d, \mathcal{C}_n \setminus \{(i - 1, i), (i, i), (i, i + 1)\}) \right).$$

Because $D$ is exchangeable, for any $1 \leq i \leq n$,

$$\mathbb{E}[P(D, \mathcal{C}_n \setminus \{(i, i + 1)\})] = \mathbb{L}_n,$$

$$\mathbb{E}[P(D, \mathcal{C}_n \setminus \{(i - 1, i), (i, i), (i, i + 1)\})] = \mathbb{L}_{n-1}.$$

Thus, integrating over all random instances of $D$ on Equation (10), we have

$$\mathbb{E}_n = n(\mathbb{L}_n - \mathbb{L}_{n-1}). \quad \square$$

Theorem 4 provides insights on the performance of long chains. Indeed, it relates the expected performance of a long chain, $\mathbb{E}_n$, with the difference in the expected performances of two open chains, $\mathbb{L}_n$ and $\mathbb{L}_{n-1}$, which are much easier to compute and analyze.

An immediate corollary of Theorem 4 is that the long chain is optimal among all 2-flexibility designs.

**Corollary 3.** Consider a balanced system of size $n$ with exchangeable demand. Let $\mathbb{F}_2$ be the set of all 2-flexibility designs of the system. That is, $\mathbb{F}_2$ is the set of all flexibility designs where each plant node and each product node are incident to exactly two arcs. Then, we have

$$\mathbb{E}_n = \arg \max_{\mathcal{A} \in \mathbb{F}_2} \mathbb{A}.$$

In words, the long chain maximizes expected sales among all 2-flexibility designs in the system.

**Proof of Corollary 3.** Consider a 2-flexibility design $\mathcal{A} \in \mathbb{F}_2$. Recall that a closed chain in $\mathcal{A}$ is an induced subgraph in $\mathcal{A}$ that forms an undirected cycle. Because every node is incident to two arcs in $\mathcal{A}$, $\mathcal{A}$ must be a collection of several closed chains, which we denote as $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_k$. Let $n_i$ be the number of products and plants in the closed chain $\mathcal{C}_i$. Because the system size is $n$, $\sum_{i=1}^{k} n_i = n$. Now, by Theorem 4, we have

$$\mathbb{A} = \sum_{i=1}^{k} n_i (\mathbb{L}_n - \mathbb{L}_{n-1})$$

$$= \sum_{i=1}^{k} n_i (\mathbb{L}_n - \mathbb{L}_{n-1}) + \mathbb{E}[\min\{1, D_1\}]$$

(by definition of $\mathbb{L}_n$)

$$\leq \sum_{i=1}^{k} n_i (\mathbb{L}_n - \mathbb{L}_{n-1}) + \mathbb{E}[\min\{1, D_1\}]$$

(by Theorem 2)

$$= \sum_{i=1}^{k} n_i (\mathbb{L}_n - \mathbb{L}_{n-1})$$

$$= n(\mathbb{L}_n - \mathbb{L}_{n-1}) = \mathbb{E}_n. \quad \square$$

6. Applications to Systems with i.i.d. Demand

In this section, we present three interesting applications of Theorem 4 under i.i.d. demand. The organization of this section is as follows. The first subsection describes an effective method of computing the performance of the long chain; the second subsection studies the effectiveness of the long chains relative to that of full flexibility; and the last subsection presents a risk-pooling result associated with long chains.

Throughout this section, we consider balanced systems with i.i.d. demand. Because we consider systems of arbitrary sizes, we let $D$ be an infinite random vector with i.i.d. entries, where $D_i$ is the random demand for product $i$ generated by a given distribution $D$ for all $i \geq 1$. Note that for any $n \geq 2$, $\mathbb{E}[D, \mathcal{C}_n]$ (also denoted by $\mathbb{E}_n$), only depends on the first $n$ entries of the random vector $D$.

6.1. Computing the Performance of the Long Chain

This section presents a method to compute $\mathbb{E}_n$, the performance of the long chain, using matrix multiplication. First, we introduce Algorithm 1, a greedy algorithm that finds the optimal solution of the linear program associated with $P(d, \mathcal{L}_n)$, where $d$ is an instance of $D$. Then, we apply Algorithm 1 to develop an efficient procedure to compute the quantity $\mathbb{L}_n - \mathbb{L}_{n-1}$, which is equal to $\mathbb{E}_n/n$ by Theorem 4.

**Algorithm 1** (Finding optimal solution $f^*$ for $P(d, \mathcal{L}_n)$)

1. procedure SOLVE ($P(d, \mathcal{L}_n)$)
2. $f^*_{1,1} \leftarrow \min\{1, d_1\}$
3. for $k = 2, \ldots, n$ do
4. $f^*_{k-1,k} \leftarrow \min\{1 - f^*_{k-1,k-1}, D_k\}$
5. $f^*_{k,k} \leftarrow \min\{1, D_k - f^*_{k-1,k}\}$
6. end for
7. return $f^*$
8. end procedure.
We note that a similar greedy style algorithm for computing the maximum sales in an open chain was also used by Chou et al. (2010b). We omit the proof for the correctness of Algorithm 1, because $P_i(d, \mathcal{F}_i)$ is simply a max-flow problem on a path and it is well known that this problem can be solved using a greedy algorithm.

Given a random demand vector $D$, let $F_j$ be the random flow on arc $(i,j)$ returned by Algorithm 1, for $1 \leq i,j \leq n$. For each integer $1 \leq k \leq n - 1$, define $W_k = 1 - F_k$, and $W_0 = 0$. One can think of $W_k$ as the remaining capacity in plant $k$ after the production of product $k$ at plant $k$ is determined.

To develop a method to compute the performance of the long chain, assume that the support of $D$ lies in $\{i/N \mid i = 0, 1, 2, \ldots\}$ for some $N \geq 1$. Under this assumption, we let $p_i = \mathbb{P}[D = i/N]$, for any $i = 0, 1, \ldots, 2N - 1$, and $p_{2N} = \mathbb{P}[D \geq 2]$, where $\mathbb{P}[\cdot]$ denotes the probability mass function.

Because the support of $D$ lies in $\{i/N \mid i = 0, 1, 2, \ldots\}$ and $0 \leq F_{kk} \leq \sum F_i \leq 1$, we have that the support of $F_{kk}$ lies in $\{i/N \mid i = 0, 1, 2, \ldots, N\}$. But $W_k = 1 - F_{kk}$, so $W_k$ also have a support of $\{i/N \mid i = 0, 1, 2, \ldots, N\}$. As a result, the distribution of $W_k$ can be described by a row vector $q^k$ with $N + 1$ elements, where $q^k_i = \mathbb{P}[W_k = i/N]$, for $i = 0, 1, \ldots, N$. Then, we have

**Lemma 3.** $q^{k+1} = q^k A = q^k A^{k+1}$ for $0 \leq k \leq n - 1$, where

\[
A = \begin{bmatrix}
\sum_{i=N}^{2N} p_i & p_{N-1} & p_{N-2} & \cdots & p_1 & p_0 \\
\sum_{i=N+1}^{2N} p_i & p_N & p_{N-1} & \cdots & p_2 & p_0 + p_1 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
p_{2N-1} + p_{2N} & p_{2N-2} & \cdots & p_{N+1} & p_{N+1} & p_{N+1} \\
p_{2N} & p_{2N-1} & p_{2N-2} & \cdots & p_{N} & p_{N+1} \\
\end{bmatrix}
\]

and $q^0 = [1 \ 0 \ 0 \ \cdots \ 0]$.

**Proof for Lemma 3.** Because $W_0$ is zero with probability one, $q^0_i = [1 \ 0 \ 0 \ \cdots \ 0]$. Because the demand is independent and $W_k$ only depends on $D_1, \ldots, D_k$, $W_k$ is independent of $D_{k+1}$. Hence, we have,

$q^{k+1}_i = \mathbb{P}[W_{k+1} = i] = \sum_{j=0}^{N} \mathbb{P}[W_k = j] \mathbb{P}[D_{k+1} = N - i + j]$

$= \sum_{j=0}^{N} q^k_j p_{N-i+j}, \quad \text{for } 1 \leq i \leq N - 1,$

$q^{k+1}_0 = \mathbb{P}[W_{k+1} = 0] = \sum_{j=0}^{N} \mathbb{P}[W_k = j] \mathbb{P}[D_{k+1} \geq N + j]$

$= \sum_{j=0}^{N} q^k_j \sum_{i=N+j}^{2N} p_i.$

This implies that $q^{k+1} = q^k A$.

A direct consequence of Lemma 3 is that the following matrix multiplications can be used to determine the performance of the long chain, when demands are i.i.d. and the support of a product demand is a subset of $\{i/N \mid i = 0, 1, 2, \ldots\}$.

**Theorem 5.** $[\mathcal{L}_n]/n = \lfloor \mathcal{L}_n \rfloor = q^{n-1} \pi = q^0 A^{n-1} \pi$, where $\pi$ is a vector of size $N + 1$ and

$\pi_i = \sum_{j=1}^{N+i} j p_j + (N + i) \sum_{j=N+i+1}^{2N} p_j, \quad \forall 1 \leq i \leq N.$

**Proof for Theorem 5.** By Algorithm 1, $[\mathcal{L}_n] - [\mathcal{L}_{n-1}]$ can be written as the expectation of $F_{n-1} + F_{n}$, which is equal to $\mathbb{E} \left[ \min \{1 + W_{n-1}, D_n \} \right]$, thus

$\mathbb{E} \left[ \min \{1 + W_{n-1}, D_n \} \right] = \sum_{i=0}^{N} \mathbb{P}[W_{n-1} = i] \mathbb{E} \left[ \min \left\{ D_n, \frac{N+i}{N} \right\} \right] = \sum_{i=0}^{N} q^i \left( \sum_{j=1}^{N+i} j p_j + (N + i) \sum_{j=N+i+1}^{2N} p_j \right).$

Hence, we have that $[\mathcal{L}_n] - [\mathcal{L}_{n-1}] = q^{n-1} \pi$. Apply Theorem 4, and we are done.

The matrix multiplication method developed here to compute the performance of the long chain is polynomial in $N$ and $n$. Indeed, computing $q^k A^{n-1} \pi$ requires $O(nN^2)$ operations if one sequentially evaluates $q^k A^i$ for $i = 1, \ldots, k$, or $O(N^{2.807} \log n)$ operations if one starts by determining $A^{n-1}$ using the classical algorithm from Strassen (1969). To the best of our knowledge, this is the first polynomial time algorithm that computes the performance of a finite size long chain exactly when demand is discrete and i.i.d. The other known algorithm to compute the performance of the long chain exactly is to solve the max-flow problem for all demand instances and sum them to determine the expected performance. Unfortunately, this method is exponential in $n$.

The matrix multiplication method can be applied for general i.i.d. demands as an approximation algorithm to compute the performance of long chains. In this case, one can approximate the performance of the long chain by discretizing the demand distribution on the set of $\{i/N \mid i = 0, 1, 2, \ldots\}$ for some integer $N$. Clearly, as $N$ increases, the error of the approximation decreases while the running time grows. Specifically, it is straightforward to show that the error of the approximation is bounded by $n/2N$. However, our computational experience suggests that the error is much smaller than this bound.
Moreover, the matrix multiplication method is fairly fast even for large \( N \). For example, when \( N = 1,000 \) and \( n = 100 \), \( q_0^0 A^{n-1} \pi \) can be computed within two seconds using Matlab on a standard 2.1 GHz laptop. Hence, even for general i.i.d. demands, the matrix multiplication method can quickly approximate the performance of a large size long chain very accurately.

Figure 5 presents computational results obtained using the matrix multiplication method for three different i.i.d. demand distributions:

1. **Normal.** Demand for a product is a discretized normal random variable with mean 1 and standard deviation of 0.33 on the support set of \( \{i/14 \mid i = 0, 1, \ldots, 28\} \); this distribution was originally applied in Chou et al. (2010b) for their analysis of asymptotic behavior of long chains;

2. **Uniform.** Demand for a product is uniformly distributed on the set \( \{i/10 \mid i = 0, 1, 2, \ldots, 9, 11, 12, \ldots, 20\} \);

3. **Asymmetric.** Demand for a product is equal to 4 with probability 0.4, one with probability 0.5 and two with probability 0.1.

For each distribution, Figure 5 depicts \( [\mathcal{F}_n]/n \) (the per product performance of full flexibility), \( [\mathcal{C}_n]/n \) (the per product performance of the long chain), and \( [\mathcal{C}_n]/[\mathcal{F}_n] \) (the ratio between the performance of the long chain and the performance of full flexibility design) for \( n = 1, \ldots, 30 \).

Figure 5 reveals several interesting observations. First, \( [\mathcal{F}_n]/n - [\mathcal{C}_n]/n \), i.e., the gap between the fill rates of full flexibility and the long chain, is increasing, whereas the ratio, \( [\mathcal{C}_n]/[\mathcal{F}_n] \), is decreasing. A similar observation on the ratio, using simulation results, is reported in Chou et al. (2008). In addition, Figure 5 suggests that the quantity \( [\mathcal{C}_n]/n \), the fill rate of the long chain, is increasing but converges to a constant very quickly. These observations are discussed in detail in the next two subsections.

Finally we discuss two important extensions of the matrix multiplication method: (i) computing the per product performance of the long chain for infinite size system, and (ii) dealing with nonidentical but independent demand distributions.

Observe that the matrix \( A \) is the transition matrix of a Markov chain with states \( i/N \) for each \( i = 0, 1, \ldots, N \). It can be shown that in the matrix \( A \), the communication class that contains state 0 is irreducible and aperiodic. Then, by the Perron-Frobenius theorem (see Grimmett and Stirzaker 1992) we have that \( \lim_{n \to \infty} q_0^0 A^{n-1} = q_*^* \), where \( q_*^* A = q_*^* \) and \( q_*^* > 0 \). Thus, to compute \( \lim_{n \to \infty} [\mathcal{C}_n]/n \), one can solve for \( q_*^* \) by finding the eigenvectors of \( A \), and then compute \( q_*^* \pi \), which equals to \( \lim_{n \to \infty} [\mathcal{C}_n]/n \). This provides another method for computing \( \lim_{n \to \infty} [\mathcal{C}_n]/n \) in addition to the result of Chou et al. (2010b). Interestingly, their procedure also involves discretizing demand and solving a system of linear equations.

**Figure 5.** The performance of long chains vs. the performance of full flexibility.
We also note that this matrix multiplication method can be applied for long chain in balanced systems with independent but nonidentical product demands (and even non-identical plant capacities). In that case, the performance of the long chain, \([\mathcal{E}_n]\), can be no longer characterized using \((\mathcal{L}_n - \mathcal{L}_{n-1})\). Instead, we have that \([\mathcal{E}_n]\) is equal to

\[
\sum_{i=1}^{n} ([\mathcal{E}_n \setminus (i, i+1)] - [\mathcal{E}_n \setminus ((i-1, i), (i, i), (i, i+1)])].
\]

Similar to the multiplication procedures we described, for each \(i\), \([\mathcal{E}_n \setminus (i, i+1)] - [\mathcal{E}_n \setminus ((i-1, i), (i, i), (i, i+1)])\) can be evaluated by computing \(q^0 \prod_{k=1}^{i-1} A(k)\) for \(n-1\) \(N \times N\) matrices \(A(1), A(2), \ldots, A(n-1)\). Because computing \(q^0 \prod_{k=1}^{i-1} A(k)\) requires \(O(n^2)\) operations, computing the sum of \((\mathcal{L}_n - \mathcal{L}_{n-1})\) for \(1 \leq i \leq n\) requires \(O(n^3)\) operations.

### 6.2. Long Chain vs. Full Flexibility

In Figure 5, it was observed that the gap between the fill rates of full flexibility and that of the long chain, \(\mathcal{L}_n/n - [\mathcal{E}_n]/n\), is increasing, whereas the ratio, \([\mathcal{E}_n]/\mathcal{L}_n\), is decreasing. In this section, we will formally prove the first part of the observation, and discuss some partial results related to the second part. We start by defining two random walks in §6.2.1. These random walks are applied to analyze the difference between the fill rates of the long chain and full flexibility in §6.2.2, as well as the ratio of the fill rate of long chain to that of full flexibility in §6.2.3.

#### 6.2.1. Random Walks.

We define two random walks, \(W_i\) and \(\tilde{W}_i\), as follows:

**Definition 3.** Let \(W_0 = \tilde{W}_0 = 0\). For \(i \geq 1\), define

\[
W_i = \begin{cases} 
0 & \text{if } W_{i-1} + 1 - D_{i-1} < 0 \\
1 & \text{if } W_{i-1} + 1 - D_{i-1} > 1 \\
W_{i-1} + 1 - D_{i-1} & \text{otherwise,}
\end{cases}
\]

\[
\tilde{W}_i = \begin{cases} 
0 & \text{if } \tilde{W}_{i-1} + 1 - D_{i-1} < 0 \\
\tilde{W}_{i-1} + 1 - D_{i-1} & \text{otherwise.}
\end{cases}
\]

\(W_i\) and \(\tilde{W}_i\) are generalized random walks with random steps \(1 - D_1, \ldots, 1 - D_i\) and different sets of reflecting boundaries. The random walk \(W_i\) has reflecting boundaries of zero and one, whereas \(\tilde{W}_i\) has a reflecting boundary only at zero. For any fixed vector \(d\) that is an instance of \(D\), define \(W_i(d)\) (and \(\tilde{W}_i(d)\)) to be the instance of \(W_i\) (and \(\tilde{W}_i\)) corresponding to \(d\). Figure 6 illustrates an example of \(W_i(d)\) and \(\tilde{W}_i(d)\), with \(i = 6\) and \(d = [0.6, 0.2, 1.2, 1.9, 1.4, 0.5]\).

The next lemma states several simple observations regarding \(W_i(d)\) and \(\tilde{W}_i(d)\) for any fixed vector \(d\).

**Lemma 4.** For any fixed vector \(d\) and \(i \geq j \geq 1\),

\[
W_i(d) \leq W_j(d), \quad W_{i-j}(d) \leq W_j(d), \quad \text{and} \quad \tilde{W}_{i-j}(d) \leq \tilde{W}_j(d),
\]

where \(d' = [d_j, d_{j+1}, \ldots]\).

**Proof of Lemma 4.** Because \(\tilde{W}_j\) has no reflecting boundary at one, \(W_j(d) \leq W_j(d)\). For Equation (12), observe that \(W_{i-j}(d')\) has the same step lengths as the last \(i - j\) steps of \(W_i(d)\). Because \(W_j(d)\), the position of the random walk \(W_i(d)\) after \(j\) steps, is greater or equal to zero, we have that \(W_{i-j}(d') \leq \tilde{W}_j(d')\). Similarly, we can also show that \(\tilde{W}_{i-j}(d') \leq \tilde{W}_j(d')\).

The rest of this subsection establishes the connections between these random walks and sales (and performances) of the long chain and full flexibility. These connections would be used for the comparisons between the long chain and full flexibility in §6.2.2.
Similarly to §5.1, in the rest of this subsection, when some integer \( k \) appears in a statement, we will in fact be referring to some \( i \in \{1, \ldots, n\} \) congruent to \( k \) modulo \( n \). First, we show the relationship between \( P(d, \mathcal{E}_n) \) and a random walk on \( d \).

**Lemma 5.** Let \( d^{i|n} = [d_i, d_{i+1}, \ldots, d_n, d_1, \ldots, d_{i-1}] \), then

\[
P(d, \mathcal{E}_n) = \sum_{i=1}^{n} \min \{1 + W_{n-1}(d^{i|n}), d_{i-1}\}.
\]

**Proof for Lemma 5.** For each \( 1 \leq i \leq n \), it is not difficult to check that \( \min \{1 + W_{n-1}(d^{i|n}), d_{i-1}\} \) is equal to the quantity \( f^*_{n+1} + f^*_{n-1} \), returned by Algorithm 1 on \( P(d^{i|n}, \mathcal{L}_n) \). By Algorithm 1’s greedy property, \( f^*_{n+1} + f^*_{n-1} = P(d^{i|n}, \mathcal{L}_n) - P(d^{i|n}, \mathcal{L}_{n-1}) \), which implies \( \min \{1 + W_{n-1}(d^{i|n}), d_{i-1}\} = P(d^{i|n}, \mathcal{L}_n) - P(d^{i|n}, \mathcal{L}_{n-1}) \). Finally, by definition, \( P(d^{i|n}, \mathcal{L}_n) = P(d, \mathcal{E}_n \setminus \{(i-2, i-1), (i-1, i-1), (i-1, i)\}) \) and

\[
\min \{1 + W_{n-1}(d^{i|n}), d_{i-1}\} = P(d, \mathcal{E}_n \setminus \{(i-2, i-1), (i-1, i-1), (i-1, i)\})
\]

Substitute Equation (14) to Theorem 3, and we have

\[
P(d, \mathcal{E}_n) = \sum_{i=1}^{n} \min \{1 + W_{n-1}(d^{i|n}), d_{i-1}\}.
\]

We note that a similar observation to Lemma 5 was stated in Chou et al. (2010b). In particular, Chou et al. observed that \( \min \{1 + W_{n-1}(d), d_{i-1}\} = P(d, \mathcal{L}_n) - P(d, \mathcal{L}_{n-1}) \). Our Lemma 5 goes a step further by applying Theorem 3.

Establishing the relationships between \( P(d, \mathcal{F}_n) \) and a random walk on \( d \) is more difficult. We do this by proving a lemma that shows that the sales associated with \( \mathcal{F}_n \) is equal to the sales of \( \mathcal{E}_n \) under a new demand \( \tau(d) \), which is a linear transformation of \( d \). Specifically, we define \( \tau(d_i) = (d_i + (n-1))/n \), for \( i = 1, 2, \ldots, n \).

**Lemma 6.** For any demand instance \( d \),

\[
P(\tau(d), \mathcal{E}_n) = P(\tau(d), \mathcal{F}_n),
\]

where \( \tau(d_i) = (d_i + (n-1))/n \), for \( i = 1, 2, \ldots, n \).

**Proof for Lemma 6.** By duality of linear programs,

\[
P(\tau(d), \mathcal{E}_n) = \min \sum_{1 \leq i \leq n} p_i + \sum_{1 \leq j \leq n} q_j \tau(d_j) \quad \text{(VC)}
\]

s.t. \( p_i + q_j \geq 1, \quad \forall (i, j) \in \mathcal{E}_n \)

\( p_i \geq 0, q_j \geq 0, \quad \forall 1 \leq i \leq n, 1 \leq j \leq n \)

\( p, q \in \mathbb{R}^n \).

The linear program denoted by (VC) is an LP-relaxation of a min-weight bipartite vertex cover problem. Because the LP-relaxation of min-weight bipartite vertex cover is tight, it has an optimal solution \((p^*, q^*)\), where entries in \( p^* \) and \( q^* \) are either zero or one. Let \( S = \{i \mid p^*_i = 0\} \) and \( S^* = \{j \mid q^*_j = 1\} \). Note that \( N(S) \subseteq S^* \), where \( N(S) \) is the set of neighbors of \( S \) in \( \mathcal{E}_n \). First, suppose \( S \neq \emptyset \), and \( S^* \neq \{1, 2, \ldots, n\} \). Then, we must have \( |S^*| - 1 \geq |N(S)| - 1 \geq |S| \). Let \( p_i^0 = 1, q_j^0 = 0 \) for all \( 1 \leq i, j \leq n \). Clearly \((p^0, q^0)\) is a feasible solution of (VC). Also as \( |S^*| - 1 \geq |S| \),

\[
n \leq (n - |S|) + (|S^*| - 1)
\]

\[
< (n - |S|) + \sum_{j \in S^*} \frac{(n-1)}{n} \quad \text{(because } |S| < n)\]

\[
\leq \sum_{1 \leq i \leq n} p_i^0 + \sum_{j \in S^*} \frac{d_j + (n-1)}{n} \quad \text{(because } d_j \geq 0)\]

\[
= \sum_{1 \leq i \leq n} p_i^0 + \sum_{1 \leq j \leq n} \tau(d_j) q_j^0.
\]

But \( \sum_{1 \leq i \leq n} p_i^0 + \sum_{1 \leq j \leq n} \tau(d_j) q_j^0 = n \), and this contradicts the optimality of \((p^*, q^*)\). Thus, one must have that either \( S = \emptyset \) or \( S^* = \{1, 2, \ldots, n\} \). Therefore, \( P(\tau(d), \mathcal{E}_n) = \min \{n, \sum_{1 \leq j \leq n} \tau(d_j)\} = P(\tau(d), \mathcal{F}_n) \).

Now, we can prove the lemma that establishes the relationship between \( P(d, \mathcal{F}_n) \) and \( \tilde{W} \).

**Lemma 7.** Let \( d^{i|n} = [d_i, d_{i+1}, \ldots, d_n, d_1, \ldots, d_{i-1}] \), then

\[
P(\tau(d), \mathcal{F}_n) = \sum_{i=1}^{n} \min \{1 + \tilde{W}_{n-1}(d^{i|n}), d_{i-1}\}.
\]

**Proof of Lemma 7.** By Lemma 5 and 6, we have

\[
P(\tau(d), \mathcal{F}_n) = \sum_{i=1}^{n} \min \{1 + W_{n-1}(\tau(d^{i|n})), \tau(d_{i-1})\}.
\]

Note that for any \( 1 \leq j \leq n-1 \), and \( 1 \leq i \leq n \),

\[
W_j(\tau(d^{i|n})) \leq \sum_{i \leq k \leq i+j} \max \left\{0, 1 - \frac{d_k + (n-1)}{n}\right\}
\]

\[
\leq (n-1) \cdot \frac{1}{n} < 1.
\]

This implies that \( W_j(\tau(d^{i|n})) \) never touches the reflecting boundary at one. Hence,

\[
W_{n-1}(\tau(d^{i|n})) = \tilde{W}_{n-1}(\tau(d^{i|n})) \quad \forall 1 \leq i \leq n.
\]

Because \( 1 - \tau(d_k) = (1/n)(1 - d_k) \), for any \( 1 \leq k \leq n \), we have

\[
\tilde{W}_j(\tau(d^{i|n})) = \frac{1}{n} \tilde{W}_j(d^{i|n}) \quad \forall 1 \leq j \leq n-1.
\]

Thus,

\[
P(\tau(d), \mathcal{F}_n) = \sum_{i=1}^{n} \min \{1 + \tilde{W}_{n-1}(\tau(d^{i|n})), \tau(d_{i-1})\}
\]

(by Equation (15))
On the other hand,

\[
\sum_{i=1}^{n} \min \left\{ 1 + \frac{\bar{W}_{n-1}(d^{i/n})}{n}, \frac{d_{i-1} + n - 1}{n} \right\} = \sum_{i=1}^{n} \left( \frac{n-1}{n} + \min \left\{ 1 + \frac{\bar{W}_{n-1}(d^{i/n})}{n}, \frac{d_{i-1}}{n} \right\} \right)
\]

(by Equation (16))

\[
= n - 1 + \frac{1}{n} \sum_{i=1}^{n} \min \{1 + \bar{W}_{n-1}(d^{i/n}), d_{i-1}\}.
\]

On the other hand, we have that

\[
P(\tau(d), \bar{F}_n) = \min \left\{ n, \sum_{1 \leq i \leq n} d_i + n - 1 \right\}
\]

\[
= n - 1 + \frac{1}{n} \sum_{1 \leq i \leq n} d_i
\]

\[
= n - 1 + \frac{1}{n} P(d, \bar{F}_n).
\]

Therefore, we have that

\[
n - 1 + \frac{1}{n} \sum_{i=1}^{n} \min \{1 + \bar{W}_{n-1}(d^{i/n}), d_{i-1}\} = n - 1 + \frac{1}{n} P(d, \bar{F}_n),
\]

which implies

\[
P(d, \bar{F}_n) = \sum_{i=1}^{n} \min \{1 + \bar{W}_{n-1}(d^{i/n}), d_{i-1}\}. \quad \square
\]

Integrating the equations in Lemma 5 and Lemma 7 over all instances in D, we obtain the Lemma 8, which relates the expectation of the two random walks with the performance of the long chain and that of full flexibility.

**Lemma 8.** Under i.i.d. demand, we have

\[
\frac{[\bar{F}_n]}{n} = \mathbb{E} \left[ \min \{1 + W_{n-1}(D), D \} \right]
\]

\[
\frac{[\bar{F}_n]}{n} = \mathbb{E} \left[ \min \{1 + \bar{W}_{n-1}(D), D \} \right].
\]

**Proof of Lemma 8.** Because D is i.i.d., we have that for any \(1 \leq i \leq n\),

\[
\mathbb{E} \left[ \min \{1 + W_{n-1}(D^{i/n}), D_{i-1} \} \right] = \mathbb{E} \left[ \min \{1 + W_{n-1}(D), D \} \right]
\]

\[
\mathbb{E} \left[ \min \{1 + \bar{W}_{n-1}(D^{i/n}), D_{i-1} \} \right] = \mathbb{E} \left[ \min \{1 + \bar{W}_{n-1}(D), D \} \right].
\]

After integrating the equations in Lemma 5 and Lemma 7 over D, we get

\[
\frac{\bar{F}_n}{n} = n \mathbb{E} \left[ \min \{1 + W_{n-1}(D), D \} \right]
\]

\[
\frac{\bar{F}_n}{n} = n \mathbb{E} \left[ \min \{1 + \bar{W}_{n-1}(D), D \} \right]. \quad \square
\]

### 6.2.2. Difference in Fill Rates

With Lemma 8 at hand, we now prove that the quantity \(\frac{\bar{F}_n}{n} - \frac{[\bar{F}_n]}{n}\) is non-decreasing with n.

**Theorem 6.** For any integer \(n \geq 2\) and i.i.d. demand,

\[
\frac{[\bar{F}_n]}{n} - \frac{\bar{F}_n}{n} \leq \frac{[\bar{F}_{n+1}]}{n+1} - \frac{\bar{F}_{n+1}}{n+1} \leq \min \{1, \mathbb{E}[D] \} - \gamma,
\]

where \(\gamma = \lim_{k \to \infty} \mathbb{E} \left[ \frac{\bar{F}_k}{k} \right] / k\).

**Proof of Theorem 6.** First, we show that for any demand instance \(d\) of D, we have,

\[
\begin{align*}
&\min \{1 + \bar{W}_{n-1}(d^{2}), d_{n+1} \} - \min \{1 + W_{n-1}(d^{2}), d_{n+1} \} \\
&\leq \min \{1 + \bar{W}_{n}(d), d_{n+1} \} - \min \{1 + W_{n}(d), d_{n+1} \}, \quad (17)
\end{align*}
\]

where \(d^{2} = \{d_2, d_1, \ldots \}\). One can think of \(W_{n-1}(d^{2}) \) (and \(\bar{W}_{n-1}(d^{2})\)) as the walk that started one time unit later than \(W_{n}(d) \) (and \(\bar{W}_{n}(d)\)). To prove Inequality (17), consider the following two cases.

**Case 1.** \(W_{i}(d^{2}) = 1\) for some \(1 \leq i \leq n - 1\). Then \(W_{i}(d^{2}) = W_{i+1}(d^{2}) = 1\), and by the definition of W, we must have that \(W_{n-1}(d^{2}) = \bar{W}_{n-1}(d^{2})\). By Lemma 4, \(\bar{W}_{n-1}(d^{2}) \leq \bar{W}_{n}(d)\), and therefore we have

\[
\begin{align*}
&\min \{1 + W_{n-1}(d^{2}), d_{n+1} \} - \min \{1 + W_{n}(d), d_{n+1} \} \\
&\leq \min \{1 + \bar{W}_{n}(d), d_{n+1} \} - \min \{1 + W_{n}(d), d_{n+1} \},
\end{align*}
\]

which implies that Inequality (17) holds.

**Case 2.** \(W_{i}(d^{2}) < 1\) for all \(1 \leq i \leq n - 1\). By definition of W and \(\bar{W}\), we have that \(W_{n-1}(d^{2}) = \bar{W}_{n-1}(d^{2})\). By Lemma 4, \(W_{n}(d) \leq \bar{W}_{n}(d)\), and therefore we have

\[
\begin{align*}
&\min \{1 + W_{n-1}(d^{2}), d_{n+1} \} - \min \{1 + W_{n}(d), d_{n+1} \} \\
&\leq \min \{1 + \bar{W}_{n}(d), d_{n+1} \} - \min \{1 + W_{n}(d), d_{n+1} \},
\end{align*}
\]

which again implies that Inequality (17) holds.

Because demand is i.i.d., after integrating over D for both sides of Inequality (17), we get

\[
\begin{align*}
&\mathbb{E} \left[ \min \{1 + W_{n-1}(D), D_{n+1} \} \right] - \mathbb{E} \left[ \min \{1 + W_{n}(D), D_{n+1} \} \right] \\
&\leq \mathbb{E} \left[ \min \{1 + \bar{W}_{n}(D), D_{n+1} \} \right] - \mathbb{E} \left[ \min \{1 + W_{n}(D), D_{n+1} \} \right].
\end{align*}
\]

Applying Lemma 8, we have that

\[
\begin{align*}
&\frac{[\bar{F}_n]}{n} - \frac{\bar{F}_n}{n} \leq \frac{[\bar{F}_{n+1}]}{n+1} - \frac{\bar{F}_{n+1}}{n+1}, \quad \text{for all } n \geq 2. \quad (18)
\end{align*}
\]

Finally, Equation (18) implies that

\[
\begin{align*}
&\frac{[\bar{F}_n]}{n} - \frac{\bar{F}_n}{n} = \lim_{k \to \infty} \left( \frac{[\bar{F}_k]}{k} - \frac{\bar{F}_k}{k} \right) \\
&= \lim_{k \to \infty} \left( \mathbb{E} \left[ \min \left\{ \frac{W_{n-1}(D)}{k}, 1 \right\} \right] \right) - \gamma \\
&= \min \{\mathbb{E}[D], 1\} - \gamma.
\end{align*}
\]
where the last equality holds because of the weak law of large numbers.

Note that the fill rate of $\mathcal{E}_n$ (and $\mathcal{F}_n$) is equal to $[\mathcal{E}_n]/nE[D]$ (and $[\mathcal{F}_n]/nE[D]$). Thus, Theorem 6 implies that the smaller the system size, the smaller the gap between the fill rate of full flexibility and that of the long chain. This suggests that the long chain is more effective relative to full flexibility for smaller size systems. Moreover, Theorem 6 can be used to bound the gap between the fill rate of full flexibility and that of the long chain for systems of any size. For this purpose, we point out that Chou et al. (2010b) shows that for many i.i.d. demand with $E[D] = 1$, $\gamma$ is close to one, implying that for any size system, the performance of the long chain is close to that of full flexibility.

For example, when $D$ is normal with mean 1 and standard deviation of 0.33, Chou et al. (2010b) shows that in this case $\gamma = 0.96$. Therefore, for this demand distribution, we have that the gap between the fill rate of full flexibility and that of the long chain for systems of any size is at most 4%.

### 6.2.3. Performance Ratio

The difference between the fill rate of the long chain and that of full flexibility is only one metric to evaluate the effectiveness of the long chain. A different metric, discussed in Chou et al. (2008, 2010b), is to consider the ratio of the performance of long chain to that of full flexibility. Partial results related to the ratio are discussed next.

Applying the same argument as in the proof of Theorem 6, one can show that

$$\min\{1 + W_{n-1}(d^2), d_{n+1}\} \leq \min\{1 + W_n(d), d_{n+1}\}.$$  \hspace{1cm} (19)

Unfortunately, one cannot integrate this inequality over $D$ to obtain $[\mathcal{E}_n]/[\mathcal{F}_n] \geq [\mathcal{E}_{n+1}]/[\mathcal{F}_{n+1}]$, as expectation does not preserve over multiplication and division. Indeed, it is not known whether $[\mathcal{E}_n]/[\mathcal{F}_n] \geq [\mathcal{E}_{n+1}]/[\mathcal{F}_{n+1}]$ holds, though this inequality has been observed empirically in Chou et al. (2008) and this paper (see Figure 5). Note that $[\mathcal{E}_n]/[\mathcal{F}_n] \geq [\mathcal{E}_{n+1}]/[\mathcal{F}_{n+1}]$ for all $n \geq 2$ implies Theorem 6 and hence is a stronger statement.

Of course, if $[\mathcal{E}_n]/[\mathcal{F}_n] \geq [\mathcal{E}_{n+1}]/[\mathcal{F}_{n+1}]$ holds, then it follows that

$$\frac{[\mathcal{E}_n]}{[\mathcal{F}_n]} \geq \lim_{k \to \infty} \frac{[\mathcal{E}_n]}{[\mathcal{F}_n]} = \lim_{k \to \infty} \frac{\gamma}{\min\{E[D], 1\}}, \quad \forall n \geq 2. \hspace{1cm} (20)$$

Where as before $\gamma = \lim_{k \to \infty} \frac{[\mathcal{E}_1]}{[\mathcal{F}_1]}$, and $\lim_{k \to \infty} \frac{[\mathcal{F}_1]}{[\mathcal{F}_n]} = \min\{E[D], 1\}$ by the weak law of large numbers. This would provide a lower bound on the ratio of the performance of the long chain to that of full flexibility for any system size.

Again, using the example from Chou et al. (2010b), $\gamma = 0.96$ when $D$ is normal with mean 1 and standard deviation of 0.33. Thus, if Inequality (19) holds, it would indicate that the performance of the long chain is at least 96% of that of full flexibility for any size system.

Although we do not have a proof for $[\mathcal{E}_n]/[\mathcal{F}_n] \geq \gamma$ when $E[D] = 1$, we provide a lower bound for $[\mathcal{E}_n]/[\mathcal{F}_n]$ that is almost equal to $\gamma$.

**Corollary 4.** Suppose demand is i.i.d. and $E[D] = 1$, then

$$\frac{[\mathcal{E}_n]}{[\mathcal{F}_n]} \geq 1 - \frac{(1 - \gamma)n}{[\mathcal{F}_n]},$$

where $\gamma = \lim_{k \to \infty} \frac{[\mathcal{E}_1]}{k}$.

To power the lower bound in Corollary 4, let $\delta_n = n/[\mathcal{F}_n] - 1$, which implies that $1 - (1 - \gamma)n/[\mathcal{F}_n] = \gamma - \delta_n(1 - \gamma)$. It can be shown that $n/[\mathcal{F}_n]$ is non-decreasing with $n$ (by applying for example Lemma 8), and hence, $\delta_n$ is non-increasing. Thus, if $\delta_n \approx 0$ for some small integer $k$, then Corollary 4 provides a lower bound for $[\mathcal{E}_n]/[\mathcal{F}_n]$ that is close to $\gamma$ for all $n \geq k$. Indeed, for many distributions with $E[D] = 1$, $\delta_n \approx 0$ for small $k$. For example, suppose the distribution of $D$ is normal with mean 1 and standard deviation 0.33, the long chain of any size greater than 3 is normal with mean 1 and standard deviation 0.33, the long chain of any size greater than 3 is close to one, implying that for any size system, the performance of the long chain is at least 96% of that of full flexibility for any size system.

That is, when demand is normal with mean 1 and standard deviation 0.33, the long chain of any size greater than 3 achieves at least 95.68% of the performance of full flexibility. Next, we provide a proof for Corollary 4.

**Proof for Corollary 4.** By Theorem 6,

$$\left[\frac{\mathcal{F}_i}{\mathcal{F}_{i+1}}\right]_i \leq \frac{[\mathcal{E}_i]}{[\mathcal{E}_{i+1}]} - \frac{[\mathcal{E}_{i+1}]}{[\mathcal{F}_{i+1}]} \quad \forall i \geq 1 \hspace{1cm} (21)$$

Now, add Inequality (21) for all $i \geq n$, we have that

$$\frac{[\mathcal{F}_n]}{n} - \lim_{k \to \infty} \frac{[\mathcal{F}_i]}{k} \leq \frac{[\mathcal{E}_i]}{n} - \frac{[\mathcal{E}_i]}{k}.$$  \hspace{1cm} (22)

But

$$\lim_{k \to \infty} \frac{[\mathcal{E}_1]}{k} = 1, \quad \lim_{k \to \infty} \frac{[\mathcal{E}_1]}{k} = \gamma,$$

and substituting those into Inequality (22), we have

$$\frac{[\mathcal{F}_n]}{n} - 1 \leq \frac{[\mathcal{E}_n]}{n} - \gamma \Rightarrow \frac{[\mathcal{E}_n]}{[\mathcal{F}_n]} \geq 1 - \frac{(1 - \gamma)n}{[\mathcal{F}_n]}. \quad \square$$
6.3. Risk Pooling of the Long Chain

In this subsection we focus on the per product performance (and fill rate, which is linearly proportional to the per product performance) of the long chain as a function of system size. We start by showing that \( \left[ E_n \right] / n \) is nondecreasing with \( n \) under i.i.d. demand.

**Theorem 7.** Under i.i.d. demand \( D \), we have \( \left[ E_n \right] / n \leq \left[ E_{n+1} \right] / n + 1 \), for any integer \( n \geq 2 \).

**Proof of Theorem 7.** Because \( D \) is i.i.d., the first \( n \) and \( n + 1 \) entries in \( D \) is exchangeable for balanced system of size \( n \) and \( n + 1 \). Thus, by Theorem 2, we have that \( \left[ E_n \right] - \left[ E_{n+1} \right] \leq \left[ E_{n+1} \right] - \left[ E_{n+2} \right] \), which is equivalent to \( \left[ E_n \right] \geq \left[ E_{n+1} \right] - \left[ E_{n+2} \right] \). Hence implies \( \left[ E_n \right] \geq \left[ E_{n+1} \right] - \left[ E_{n+2} \right] \). Apply Theorem 4 completes the proof.

The theorem thus states that \( \left[ E_n \right] / n \) as well as the fill rate associated with a long chain, increases with the number of products, \( n \). This phenomenon is analogous to the classical risk pooling effect associated with demand aggregation, except that here we aggregate across capacities.

Interestingly, Figure 5 suggests that the fill rate of the long chain quickly converges to a constant. This is shown in the next theorem, where we prove that the convergence rate is exponential for arbitrary i.i.d., nondegenerate demands.

**Theorem 8.** When demands are i.i.d. and nondegenerate, there exist constants \( c < 0 \) and \( K > 0 \) such that

\[
\frac{\left[ E_{n+1} \right] - \left[ E_n \right]}{n + 1} \leq Ke^n,
\]

for any \( n \geq 2 \).

**Proof of Theorem 8.** From Lemma 8 we have, \( \left[ E_n \right] / n = E[\min(1 + W_n(D), D_n)] \). Recall that \( D^2 = [D_2, D_3, \ldots] \).

We have,

\[
\frac{\left[ E_{n+1} \right] - \left[ E_n \right]}{n + 1} = E[\min(1 + W_{n+1}(D), D_{n+1})] - E[\min(1 + W_n(D^2), D_{n+1})]
\]

\[
= E[\min(1 + W_n(D), D_{n+1})] - E[\min(1 + W_{n-1}(D^2), D_{n+1})]
\]

\[
\leq P[W_n(D) \neq W_{n-1}(D^2)]
\]

where the last inequality is true because \( \min(1 + W_n(D), D_{n+1}) - \min(1 + W_{n-1}(D^2), D_{n+1}) \) never exceeds one. Note that for any particular random instance \( d \), \( W_n(d) = W_{n-1}(d^2) \) if \( W_n(d) = 0 \) for some \( 1 \leq i \leq n \) or \( W_i(d^2) = 1 \) for some \( 1 \leq i \leq n - 1 \). Thus,

\[
P[W_n(D) \neq W_{n-1}(D^2)]
\]

\[
\leq P[W_i(D) > 0, W_i(D^2) < 1, \forall 1 \leq i \leq n].
\]

Therefore, we have

\[
\frac{\left[ E_{n+1} \right] - \left[ E_n \right]}{n + 1} \leq P[W_i(D) > 0, W_i(D^2) < 1, \forall 1 \leq i \leq n].
\]

Now, because \( D \) is nondegenerate and i.i.d., there exists some \( t \) such that \( p = P[\sum_{j=k+1}^{n} (D_j - 1) \geq 1] > 0 \). If some instance \( d \) satisfies the condition \( W_i(d) > 0, W_i(d^2) < 1, \forall 1 \leq i \leq n \), then we must have that \( \sum_{j=k+1}^{n} (D_j - 1) < 1 \) for any \( 1 \leq k \leq (n - 1)/t \). Hence,

\[
\frac{\left[ E_{n+1} \right] - \left[ E_n \right]}{n + 1} \leq P[W_i(D) > 0, W_i(D^2) < 1, \forall 1 \leq i \leq n]
\]

\[
\leq \left[ \frac{k+1}{t} \sum_{j=2+(k-1)t}^{n} (D_j - 1) < 1, \forall 1 \leq k \leq \left[ \frac{n-1}{t} \right] \right]
\]

\[
= (1 - p)^{(n-1)/t}
\]

\[
\leq Ke^n,
\]

for some constants \( K > 0 \) and \( c < 0 \).

Figure 5 and Theorem 8 show that \( \left[ E_n \right] / n \approx \left[ E_{n+1} \right] / (n + t) \) for any \( t \) provided that \( n \) is large. Hence, it implies that in a system with a large number of plants and products, it is not necessary to have a long chain design that connects all the plants and products. A collection of several chains, each of which with a large number of plants and products can be as effective.

Finally, we note that Theorem 8 can be applied to show that \( \left[ E_n \right] / \left[ F_n \right] \approx \left[ E_{n+1} \right] / \left[ F_{n+1} \right] \) when \( n \) is large and \( D \) has mean 1 and finite variance. To see this, one needs to apply the central limit theorem to show that \( \left[ F_{n+1} \right]/(n + 1) - \left[ F_n \right]/n \geq K_1(1/\sqrt{n} - 1/\sqrt{n+1}) \) for some constant \( K_1 > 0 \). Then, we have that

\[
\frac{\left[ E_{n+1} \right]}{\left[ F_{n+1} \right]} - \frac{\left[ E_n \right]}{\left[ F_n \right]}
\]

\[
\leq \left( \frac{\left[ E_n \right]}{n} + Ke^n \right) \left( \frac{\left[ F_n \right]}{n} + K_1 \left( \frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} \right) \right)
\]

because \( Ke^n \ll K_1(1/\sqrt{n} - 1/\sqrt{n+1}) \) for large \( n \),

\[
\leq \left( \frac{\left[ E_n \right]}{\left[ F_n \right]} \frac{\left[ F_n \right]}{n} \right) \quad \text{for large } n,
\]

\[
= \left( \frac{\left[ E_n \right]}{\left[ F_n \right]} \right) \quad \text{for large } n.
\]

7. Conclusion

This paper presents two theoretical results that complement existing literature on long chain flexibility designs. The first result, the supermodularity of flexible arcs, provides theoretical justification to the idea of “closing the chain,” a concept that was empirically observed by authors such as Hopp et al. (2004) and Graves (2008). The supermodularity property reveals that flexible arcs in the long chain complement each other, and this provides some explanation for the power of the long chain. In addition, the supermodularity property motivates the second key result of the paper: the characterization of the performance of the long chain as the difference between the performances of two open chains.
This characterization of the performance of the long chain leads to four important developments: First, we show that the long chain is optimal among all 2-flexibility systems, a property of the long chain that has been widely held but, to the best of our knowledge, never proven before. Second, it allows the development of a simple matrix multiplication algorithm to determine the performance of the long chain. Third, it leads to a proof that the difference between the fill rate of full flexibility design and the fill rate of the long chain is increasing with the number of products. This suggests that, relative to full flexibility design, the long chain is more effective for smaller size systems. Finally, it motivates a risk-pooling result where the fill rate of a long chain increases with system size. This increase in fill rate, however, converges to zero exponentially fast. Thus, our analysis suggests that although the long chain is the optimal 2-flexibility design, a design with several closed clusters method from Jordan and Graves (1995) to create an approximately balanced system and then follow the insights and guidelines developed here. Thus, the design principles emphasized in this paper are also important for unbalanced systems as well. This includes the importance of closing the chain, and that a design with several closed chains, where each chain connects a substantial number of plants and products, performs well for large size systems.

Appendix A. Extension for the Characterization

In this section, we present an extension to Theorem 3 in a more general model where we add different capacities across the plant, different profits, and capacity constraint for each plant to product link. We start off the section by defining \( P(p, c, d, u, \mathcal{A}) \), for any flexible design \( \mathcal{A} \), and nonnegative vectors \( p, c, d, \) and \( u \):

\[
P(p, c, d, u, \mathcal{A}) = \max \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} f_{ij}
\]

s.t. \( \sum_{i=1}^{n} f_{ij} \leq d_j, \quad \forall 1 \leq j \leq n \)

\( \sum_{j=1}^{n} f_{ij} \leq c_i, \quad \forall 1 \leq i \leq n \)

\( 0 \leq f_{ij} \leq u_{ij}, \quad \forall (i, j) \in \mathcal{A} \)

\( f_{ij} = 0, \quad \forall (i, j) \notin \mathcal{A} \)

\( f \in \mathbb{R}^n \).

One can interpret \( P(p, c, d, u, \mathcal{A}) \) as the maximum profit achieved by flexibility design \( \mathcal{A} \), given demand \( d \), plants capacity vector \( c \), linear profit vector \( p \), and a flexibility capacity vector \( u \) (that is, the production for each arc \( (i, j) \) in \( \mathcal{A} \) is bounded by \( u_{ij} \)). Also, we assume that for each \( 1 \leq i \leq n \), \( p_{ij} \geq p_{ji} \) for all \( j \), and \( u_{ij} = c_i \). This assumption is intuitive in practice, as plant \( i \) can be thought of as primary production plant for product \( i \).

Now, we state Corollary 5, an extension of Theorem 3.

**Corollary 5.** Let \( \alpha_k = (k, k+1) \), \( \beta_k = (k, k) \) for any integer \( k = 1, 2, \ldots, n \). For nonnegative vectors \( p \in \mathbb{R}^n \), \( c \in \mathbb{R}^n \), \( d \in \mathbb{R}^n \), and \( u \in \mathbb{R}^n \), if \( p_{ij} \geq p_{ji} \) and \( u_{ij} = c_i \forall 1 \leq i \leq n \), we have

\[
P(p, c, d, u, \mathcal{A}) = \sum_{i=1}^{n} (P(p, c, d, u, \mathcal{A} \setminus \{\alpha_i\}) - P(p, c, d, u, \mathcal{A} \setminus \{\alpha_i, \beta_i, \alpha_i\})).
\]

**Proof for Corollary 5.** Similar to §4.1, \( P(p, c, d, u, \mathcal{A}) \) is equivalent to a max-weight circulation problem. Thus, we again apply Gale and Politof (1981) and have

\[
P(p, c, d, \mathcal{A}) + P(p, c, d, \mathcal{A} \setminus \{\alpha_i\}) \geq P(p, c, d, \mathcal{A} \setminus \{\alpha_i\}) + P(p, c, d, \mathcal{A} \setminus \{\alpha_i\}),
\]

(A1)

for any \( \mathcal{A} \subset \mathcal{A} \) and any \( 1 \leq i < n \). With Equation (A1), we can prove an extension of Lemma 2 in §5.1, which states that

\[
P(p, c, d, \mathcal{A} \setminus \{\alpha_i\}) = P(p, c, d, \mathcal{A} \setminus X)
\]

(A2)

for any \( X \in \{\alpha_1, \alpha_2, \ldots, \alpha_n\} \) and any \( 1 \leq i < n \).

Because \( p_{ij} \geq p_{ji} \) and \( u_{ij} = c_i \forall 1 \leq i \leq n \), we can apply Lemma 10 in Appendix B and have that \( P(p, c, d, u, \mathcal{A} \setminus \{\alpha_i\}) = P(p, c, d, u, \mathcal{A} \setminus \{\alpha_i\}) \), for some \( 1 \leq i < n \). With these, we have established all the results used for proving Theorem 3 in this extended model. The rest of the proof can be completed following the proof of Theorem 3. \( \square \)

Given any random demand distribution \( D \), we can integrate the equation presented in Corollary 5 over all instances of \( D \). This allows us to obtain a characterization for the expected profit of a long chain, which we present next.

**Corollary 6.** Let \( \alpha_k = (k, k+1) \), \( \beta_k = (k, k) \) for any integer \( k = 1, 2, \ldots, n \). For nonnegative vectors \( p \in \mathbb{R}^n \), \( c \in \mathbb{R}^n \), \( d \in \mathbb{R}^n \), and \( u \in \mathbb{R}^n \), if \( p_{ij} \geq p_{ji} \) and \( u_{ij} = c_i \forall 1 \leq i \leq n \), we have

\[
\mathbb{E}[P(p, c, d, u, \mathcal{A})] = \sum_{i=1}^{n} (\mathbb{E}[P(p, c, d, u, \mathcal{A} \setminus \{\alpha_i\})] - \mathbb{E}[P(p, c, d, u, \mathcal{A} \setminus \{\alpha_i, \beta_i, \alpha_i\})]).
\]
Appendix B. Characteristics of the Optimal Solutions to $P(d, \mathcal{C}_n)$

We prove that with a fixed demand instance $d$, there is always an optimal solution for $P(d, \mathcal{C}_n)$ that does not use all the flexible arcs in $\mathcal{C}_n$.

**Lemma 9.** Let $\alpha_i = (i, i+1)$ for $i = 1, \ldots, n-1$, and $\alpha_n = (n, 1)$. Then, there exists some $1 \leq i^* \leq n$ such that $P(d, \mathcal{C}_n \setminus \alpha_{i^*}) = P(d, \mathcal{C}_n)$, for any demand instance $d$.

**Proof of Lemma 9.** Recall that

$$ P(d, \mathcal{C}_n) = \max \sum_{i,j \in \mathcal{C}_n} f_{ij} $$

s.t.

$$ \sum_{i,j \in \mathcal{C}_n} f_{ij} \leq d_j, \quad \forall 1 \leq j \leq n $$

$$ \sum_{j=1}^n f_{ij} \leq 1, \quad \forall 1 \leq i \leq n $$

$$ f_{ij} \geq 0, \quad \forall (i, j) \in \mathcal{C}_n $$

$$ f_{ij} = 0, \quad \forall (i, j) \notin \mathcal{C}_n $$

$$ f \in \mathbb{R}^2 $$

Thus, the optimization problem associated with $P(d, \mathcal{C}_n)$ is linear, bounded, and feasible, which implies that it must have an optimal solution. Let $f^*$ be an optimal solution of $P(d, \mathcal{C}_n)$. If $f_{ij}^* = 0$ for some $(i, j) \in \{\alpha_1, \ldots, \alpha_n\}$, then there is some $i^*$ such that $f_{i^*j}^* = 0$ and this implies that $P(d, \mathcal{C}_n \setminus \alpha_{i^*}) = P(d, \mathcal{C}_n)$.

Otherwise, $f_{ij}^* > 0$ for all $(i, j) \in \{\alpha_1, \ldots, \alpha_n\}$. Let $g$ be the vector that

$$ g_{ij} = \begin{cases} -1 & \text{if } (i, j) \in \{\alpha_1, \ldots, \alpha_n\} \\ 1 & \text{if } (i, j) \in \{(1, 1), (2, 2), \ldots, (n, n)\} \\ 0 & \text{otherwise.} \end{cases} $$

In network flow theory, $g$ is known as an augmenting cycle for $f^*$. Let $\delta^* = \min_{i,j \in \{1, \ldots, n\}} f_{ij}$, $(i, j) \in \{\alpha_1, \ldots, \alpha_n\}$. Note that $f^* + \delta^* g$ is a feasible and optimal solution of $P(d, \mathcal{C}_n)$. Moreover, $f_{ij}^* + \delta^* g_{ij} = 0$ for some $(i, j) \in \{\alpha_1, \ldots, \alpha_n\}$. Thus, there is some $i^*$ such that $f_{i^*j}^* + \delta^* g_{i^*j} = 0$ and this implies that $P(d, \mathcal{C}_n \setminus \alpha_{i^*}) = P(d, \mathcal{C}_n)$.

**Lemma 10.** Let $\alpha_i = (i, i+1)$ for $i = 1, \ldots, n-1$, and $\alpha_n = (n, 1)$. For nonnegative vectors $p \in \mathbb{R}^2$, $c \in \mathbb{R}^n$, $d \in \mathbb{R}^n$, and $u \in \mathbb{R}^2$, if $\sum_{i=1}^n p_{ij} \leq \sum_{i=1}^n c_{ij}$, and $u_{ik} = c_{il} \forall 1 \leq k \leq n$, there exists some $1 \leq i^* \leq n$ such that $P(p, c, d, u, \mathcal{C}_n \setminus \alpha_{i^*}) = P(p, c, d, u, \mathcal{C}_n)$.

**Proof of Lemma 10.** From the assumption on $p$, $\sum_{i,j \in \mathcal{C}_n} p_{ij} g_{ij} \geq 0$, where $g$ is the augmenting cycle vector defined in the proof of Lemma 9. Thus, we can follow the proof from Lemma 9 to prove Lemma 10. Let $\delta^* = \min_{i,j \in \{1, \ldots, n\}} f_{ij}$, $(i, j) \in \{\alpha_1, \ldots, \alpha_n\}$. Then again, $f^* + \delta^* g$ is a feasible and optimal solution of $P(p, c, d, u, \mathcal{C}_n)$ and $f_{ij}^* + \delta^* g_{ij} = 0$ for some $(i, j) \in \{\alpha_1, \ldots, \alpha_n\}$. This implies that $P(p, c, d, u, \mathcal{C}_n \setminus \alpha_{i^*}) = P(p, c, d, u, \mathcal{C}_n)$.

**Acknowledgments**

The authors thank the anonymous referees and the associate editor for their suggestions, which have greatly improved the quality of the paper. This research was supported by the National Science Foundation [Contract CMMI-0758069], Masdar Institute of Science and Technology, the Ford-MIT Alliance, and the Natural Sciences and Engineering Research Council of Canada Postgraduate Scholarship.

**References**


**David Simchi-Levi** is a professor of engineering systems at Massachusetts Institute of Technology. The work described in this paper is part of a larger research project that deals with effective supply chain and procurement strategies that improve supply chain performance.

**Yehua Wei** is a Ph.D. candidate at the Operations Research Center of the Massachusetts Institute of Technology. His primary research interest lies in the field of supply chain management. Currently, he is studying supply chain strategies that enable firms to cope with uncertainties in demand and supply.