Flexible system design: A perspective from service levels

Lifei Sheng a,*, Huan Zheng b, Ying Rong b, Woonghee Tim Huh a

a Sauder School of Business, University of British Columbia, Canada
b Antai College of Economics and Management, Shanghai Jiaotong University, China

1. Introduction and literature review

This paper studies the problem of capacity portfolio investment under demand uncertainty. For example, when a firm manufactures several products, which flexible machines should be used and how much capacity should be invested?

The studies of process flexibility stem from the seminal work of Jordan and Graves [12], who have shown that limited flexibility can reap most of the benefits from the fully flexible machines. In the limited flexible system that they have considered, machines produce only 2 products each and form a "2-chain" structure (a long circle connecting all the machines and products). Subsequently, desirable properties of the 2-chain have been shown. Chou et al. [6] have shown that the 2-chain structure can achieve more than 90% benefit of full flexibility if demand follows a certain demand distribution. Simchi-Levi and Wei [15] have argued that the 2-chain is the best flexibility structure in a balanced system where all machines are 2-flexible and each product is produced from 2 machines. However, Desir et al. [11] have shown that 2-chain may not be optimal even with independent and identically distributed demand distributions if the design limits the number of possible "edges" to 2N, where N is the total number of products. Wang and Zhang [17] have provided a closed-form distribution-free bound on the performance of general k-chain. Deng and Shen [10] have characterized the distribution-free performance of any given flexibility configuration. Chen et al. [5] have analyzed the problem of how to design an optimal sparse flexibility structure which uses the smallest number of flexibility links but achieves at least 1 − ε fraction of the sales under full flexibility. Chou et al. [7] have studied the worst case performance of generalized sparse flexibility structure. Other works that calibrate and design sparse flexibility structure with fixed capacity investment include [8,13,14,9].

In comparison, capacity investment involving both dedicated and flexible machines has been investigated by Van Mieghem [16]. His 2-product model does not readily generalize to an arbitrary number of products since the number of possible configurations is exponential in the number of products. Bassamboo et al. [2] have studied a similar problem using newsvendor networks, where only two "adjacent levels of flexibility" is needed in capacity investment when system is symmetric, and Bassamboo et al. [1] have extended to the case with a submodular set of product sets that a machine can process. A similar investigation in [3] with a parallel queuing system has shown that only dedicated machines and flexible machines capable of producing only 2 products is needed asymptotically.

In a flexible resource selection setting similar to [2], we consider the problem of maximizing the joint service level (the probability of having no stock-out) subject to a budget constraint. (In comparison, the objective of [2] is to minimize the penalty cost that is linear in shortage amounts.) Instead of working directly with this objective function, we use a lower bound, which we obtain by working with the notion of the largest inscribed sphere, and consequently we build up a linear program formulation for approximating the flexible system design problem.
The main contributions of this work are summarized as below: (1) We study the flexible system design problem from a perspective of service levels. By approximating the objective with its largest inscribed sphere, we rewrite the problem into a linear program (LP) as a heuristic. To the best of our knowledge, we are the first one to apply this idea to the flexible system design problem. (2) We show that our LP heuristic has a good performance under uniform or normal demands. For general settings, where the demand support can be characterized or approximated by a simple polytope, our LP heuristic will still work well. In addition, while our objective is to maximize the joint service level, our LP heuristic solution also leads to a good performance of fill rate (the fraction of demand satisfied). (3) While we do not impose any a priori structure on flexibility (such as the chaining structure of [15]), we can show the following property for the approximate problem with the above-mentioned lower bound as the objective—that its solution can be characterized or approximated by a simple LP as a heuristic. To the best of our knowledge, we are the first to consider this approximation model. This model is based on a novel idea of an inscribed sphere, and it is easy to solve since it can be transformed to an LP of a reasonable size.

2. Problem description

Let the products be indexed by $i$, where $i = 1, \ldots, N$, and $N$ be the total number of products. Let $D = (D_1, \ldots, D_N)$ denote the random demand vector. A machine can produce one or more of the products, and we say a machine has level-$k$ flexibility if it is designed to manufacture $k$ different products. We assume that the production rate of a machine is constant and independent of the product and machine type. Let $S \subseteq \{1, \ldots, N\}$ denote the capacity type, and let $x = (x_S)_{S \subseteq \{1, \ldots, N\}}$ be the decision variable associated with how much capacity would be invested on each machine. Let $c_S$ denote the capacity installation cost for machine type $S$. Also, let $B$ be the investment budget.

Our objective is to maximize the joint service level, the probability that all demands are satisfied under the given capacity decision $x$. Given the demand realization $(d_1, \ldots, d_N)$, it can be satisfied if there is a feasible allocation $(u_S)$ satisfying the following inequalities:

$$\sum_{S \subseteq \{1, \ldots, N\} : \Omega(x) \cap S^c \neq \emptyset} u_S \geq d_i \quad \forall i,$$

$$\sum_{i \in \{1, \ldots, N\} : \Omega(x) \cap (0)^c \neq \emptyset} u_S \leq x_S \quad \forall S,$$

and $u_{S_i} \geq 0 \quad \forall S, i$.

It can be shown, using the max-flow min-cut theorem, that $(d_1, \ldots, d_N)$ can be satisfied under capacity investment $x = (x_S)_{S \subseteq \{1, \ldots, N\}}$ if and only if $(d_1, \ldots, d_N) \in \Omega(x)$, where

$$\Omega(x) = \left\{(d_1, \ldots, d_N) \mid \sum_{j \in J} d_j' \leq \sum_{(S^c) \neq J} x_S \right\}$$

for each $J \subseteq \{1, \ldots, N\}$. (1)

Thus, the problem of maximizing the joint service level under a budget constraint can be formulated as:

$$\max \left\{ P(D \in \Omega(x)) \mid \sum_{S \subseteq \{1, \ldots, N\}} c_S x_S \leq B, \ x_S \geq 0 \right\}$$

for each $S \subseteq \{1, \ldots, N\}$. (2)

Note that the number of decision variables is exponential in $N$, and thus may be difficult to solve when $N$ is large. Thus, we first analyze a symmetric and uniformly distributed demand case (Section 3), and then discuss possible extensions (Section 4).

3. Lower bound model: symmetric uniform demand

Since the exact formulation in (2) can be difficult to analyze, we consider instead an approximate model where its objective is replaced with a lower bound. In this section, we are to establish a lower bound approximation model under symmetric and uniformly distributed demand. We will show that the lower bound model is equivalent to a linear program problem (LP) and therefore leads to attractive properties in its optimal solution. Further, we will discuss the effectiveness of our approximation, compared to the exact problem (2).

We start by introducing a few assumptions.

**Assumption 1** (Symmetric Cost). All level-$k$ flexibility has the same unit cost denoted by $s_k$, i.e., $c_S = s_k$ for each $S$ satisfying $|S| = k$.

Without loss of generality, we normalize $s_1 = 1$. Since $c_S$ is increasing in $S$, it follows that $1 = s_1 \leq s_2 \leq \cdots \leq s_N$.

**Assumption 2** (Uniform and Symmetric Demand). Demand $D_i$, $i = 1, \ldots, N$, is independently and uniformly distributed on $[0, 1]$.

Under Assumption 2, the joint service level is equivalent to the volume of the intersection between the feasible production region $\Omega(x)$ and demand support $[0, 1]^N$, but the objective function in (2) may not behave as well. For example, if $N = 2$, it can be shown that $P(D \in \Omega(x)) = (x_1 + x_2)(x_2 + x_2) - (x_2)^2/2$, given that $0 \leq x_1 + x_2 \leq 1$ (i = 1, 2), which is neither convex nor concave in $(x_1, x_2, x_3)$. (See Fig. 1(a) for an illustration.) To avoid the non-convexity of the objective function, we propose a lower bound approximation model. This model is based on a novel idea of an inscribed sphere, and it is easy to solve since it can be transformed to an LP of a reasonable size.

3.1. The lower bound model: description and analysis

Instead of computing the volume of the intersection between $\Omega(x)$ and $[0, 1]^N$, we compute the volume of the largest inscribed sphere within it. Clearly, the sphere’s volume provides a lower bound to the joint service level. (See Fig. 1(b) for an illustration.)

Consider a sphere centered at $p = (p_1, \ldots, p_N)$ and radius $r$. The volume of this $N$-dimensional sphere is

$$\frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2} + 1\right)} r^N,$$

where $\Gamma$ indicates the Gamma function $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Note that the volume of the sphere is strictly increasing in $r$, thus the maximization of the volume is equivalent to maximizing $r$. If the sphere is inside the intersection between the feasible production region $\Omega(x)$ and the support of the uniform demand distribution, we must have

$$p \in \Omega(x),$$

$$r \leq \text{dist}(p, \partial_0 \Omega(x)),$$

$$r \leq p_i \quad \text{and} \quad r \leq 1 - p_i \quad \forall i,$$

where the first condition states that the center $p$ should be in $\Omega(x)$, and the second condition requires that the radius of the sphere should not exceed the distance from the center $p$ to the boundary of $\Omega(x)$ denoted by $\partial_0 \Omega(x)$. The third condition ensures that the sphere lies within the support of demand distribution, i.e. the unit cube $[0, 1]^N$.

From (1), the boundaries of $\Omega(x)$ are the hyperplanes obtained by fixing each inequality to equality. By referring to the formula of the Euclidean distance from a point to a hyperplane, one can show
that the constraint (5) is equivalent to the following collection of inequalities:

$$r \leq \sum_{j_1, \ldots, j_k \in S \neq \emptyset} x_S - \sum_{t=1}^k p_t \sqrt{k}$$

for any $1 \leq j_1 \cdots < j_k \leq N$ and $k \leq N$.  

(7)

Note that the constraint (7) and the nonnegativity of $r$ will imply (4). Therefore, based on the discussion, if we approximate $r \leq \sum_{i=1}^m r_i$ the exact capacity investment decision (2) can be approximated with the following linear program:

$$\max \left\{ r \left| \sum_{S \subseteq \{1, \ldots, N\}} c_S x_S \leq B, \ r \geq 0, \ x_S \geq 0 \right. \right\}.$$  

for each $S \subseteq \{1, \ldots, N\}$, subject to (6) and (7) 

(8)

The decision variables above are $p = (p_1, \ldots, p_N)$, $r$ and $x_S$ for each subset $S$. Let $p^r = (p^r_1, \ldots, p^r_N)$, $r^r$ and $c_S$ denote the optimal solution.

While we have established a linear program which gives a lower bound approximation on the exact problem (2), we now turn to analyze its optimal solution. Since all demand and cost parameters are symmetric across products (Assumptions 1 and 2), we expect the optimal solution of (8) to exhibit a symmetric property.

**Lemma 1 (Persistence of Symmetry).** Under Assumptions 1 and 2, there exists an optimal solution to the lower bound model (8), $p^r = (p^r_1, \ldots, p^r_N)$, $r^r$ and $c_S$ denote the symmetric property whenever $|S| = |S'|$, and (ii) $p^r_1 = \cdots = p^r_N$.

This lemma substantially simplifies the lower bound model (8) due to symmetry. Let $y_k$ denote $x_S^r$ for any $S$ satisfying $|S| = k$, and let $p_i = p_i$ for $i \in \{1, \ldots, N\}$. Also, for a set $J \subseteq \{1, \ldots, N\}$ of given size $|J| = k$, let $v_{k,m}$ denote the number of size-$m$ sets $S \subseteq \{1, \ldots, N\}$ that intersect with $J$, i.e., $J \cap S \neq \emptyset$. Thus, $v_{k,m} = \binom{N}{m} - \binom{N-k}{m}$ if $k < N$, and $v_{k,m} = \binom{N}{m}$ if $k = N$. Then, we can solve (8) by the following formulation:

$$\max r$$

s.t. $$r - p \leq 0$$

$$r - (1 - p) \leq 0$$

$$\frac{\sum_{m=1}^N v_{k,m} y_m - kp}{\sqrt{k}} \leq 0$$ for each $k = 1, \ldots, N$ 

(10)

We denote the optimal solution to this formulation by $p^r, r^r$ and $(y^r_m)$. We remark that (9) identifies only a symmetric solution (8), and does not rule out the possibility that (8) admits an asymmetric optimal solution.

Below, using the strong duality of linear programs, we identify a close-form solution to (9) under certain conditions.

**Theorem 1.** Suppose that Assumptions 1 and 2 hold, and the budget satisfies $B \leq \frac{\sqrt{N}}{2} s_2 + \sqrt{N(2 - 2\delta)}$. Let $\delta = s_2 - s_1 = s_2 - 1 \geq 0$. Then, there exists an optimal solution to (9) satisfying $y_j^r = y_j^r = \cdots = y_N^r = 0$ if the unit costs satisfy $s_j \geq 1 + (j - 1)\delta$ for each $j = 3, \ldots, N$. In this case,

$$r^r = \frac{B}{\sqrt{N(2 + (\sqrt{N} - N - 1)s_2)}}$$

$$y_j^r = \frac{2B}{s_j N \sqrt{N + (2 - s_2)N}}$$

$$y_N^r = \frac{2B}{N(2 + 2\sqrt{N} - s_2 + Ns_2)}$$  

(11)

The conclusion of Theorem 1 is that the manager needs to activate only the first two levels of flexibility (i.e., each machine is capable of producing only one or two products). This is reasonable since it avoids investment in high-level flexibility, which is more costly. This corroborates the findings in the literature attesting to the benefits of the limited amount of flexibility. Furthermore, Theorem 1 provides an easily-computable close-form expression for the capacity investment.

There are two main assumptions in Theorem 1. The first assumption is the lower bound on the cost of flexible machines, given by $s_j \geq 1 + (j - 1)\delta$, which is satisfied, for example, when the cost
parameter $s_j$ is convex in $j$ (see Fig. 2). The second condition requires that the budget is not too large, i.e., $B \leq \frac{\sqrt{N}}{2} s_2 + \sqrt{N} \frac{s_2^2}{4}$. (If this condition is violated, then the solution given in (11) is no longer feasible for the second constraint in (9).) This condition, it turns out, is not severely restrictive, especially when the number of products $N$ is large. Note that $\frac{\sqrt{N}}{2} s_2 + \sqrt{N} \frac{s_2^2}{4}$ as the upper bound of $B$ is higher than $\frac{N}{2} + \frac{\sqrt{N}}{2}$, where the inequality holds tight if $s_2 = 1$. Since each demand is independent and uniformly distributed in $[0, 1]$, when $N$ is sufficiently large, by the central limit theorem, the aggregate demand approximately follows a normal distribution with $\frac{N}{2}$ and standard deviation $\sqrt{\frac{N}{4}}$. This implies that the budget capacity of $\frac{N}{2} s_2 + \sqrt{N} \frac{s_2^2}{4}$ corresponds to a demand that is at least $\sqrt{3}$ (approximately 1.73) standard deviation higher than the aggregate mean demand.

When the budget condition is violated, it is possible that other levels of flexibility are activated. It is reasonable to use machines with more flexibility when more budget is available. Fig. 3 illustrates the capacity investment generated by (9) in a 4-product case as a function of the budget $B$. We observe that full flexibility (machines capable of producing all 4 products) will be activated after the budget exceeds the threshold discussed in the theorem.

### 3.2. Effectiveness of the lower bound approximation

In this section, we investigate the quality of the solution that we obtain from our lower bound model (9) by considering how it performs for the original exact model (2).

Let $x^*$ be the optimal solution of the exact model (2), and let $SL^*$ be the corresponding optimal joint service level. Let $y^*$ be the optimal solution to the lower bound model (9). From our discussion in Section 3.1, $y^*$ corresponds to a feasible capacity investment solution that is feasible to the exact model (2), in which we denote its service level by $SL^*$. Notice that $SL^*$ represents the probability evaluated using (2) instead of the volume of the inscribed sphere based on (3). We characterize the effectiveness of our lower bound model by comparing $SL^*$ and $SL^*$.

We first consider a 2-product case, i.e., $N = 2$. In this case, we show that the closed-form expression that we obtain in (9) attains a service level that is provably 66% as good as the optimal service level. The following proposition depends on the value of $s_2$. (We do not consider the scenario where $s_2 \geq 2$ since it would then be too expensive to invest in a flexible machine.)

**Proposition 1.** Suppose $N = 2$ and Assumptions 1 and 2 hold.

(a) Case $1 \leq s_2 < 1 + \frac{1}{\sqrt{2}}$:

\[
SL^* = \begin{cases} 
\frac{(1 + 2\sqrt{2})(4s_2^2 - s_2^2 - 2)}{(2 + \sqrt{2 - 1}s_2)^2} & \text{if } B \leq 4s_2 - 2 - s_2^2 \\
\frac{(1 + 2\sqrt{2})B^2(2 - s_2)^2}{(2 + \sqrt{2 - 1}s_2)^2} & \text{if } 4s_2 - 2 - s_2^2 < B < \sqrt{2} + \left(1 - \frac{1}{\sqrt{2}}\right)s_2 \\
1 & \text{otherwise.}
\end{cases}
\]

(b) Case $1 + \frac{1}{\sqrt{2}} \leq s_2 < 2$:

\[
SL^* = \begin{cases} 
\frac{(1 + 2\sqrt{2})(4s_2^2 - s_2^2 - 2)}{(2 + \sqrt{2 - 1}s_2)^2} & \text{if } B \leq \sqrt{2} + \left(1 - \frac{1}{\sqrt{2}}\right)s_2 \\
\frac{(4 + 4B - B^2 - 8s_2 + 4s_2^2)(4s_2^2 - s_2^2 - 2)}{B^2(2 - s_2)^2} & \text{if } \sqrt{2} + \left(1 - \frac{1}{\sqrt{2}}\right)s_2 < B \leq 4s_2 - 2 - s_2^2 \\
1 & \text{otherwise.}
\end{cases}
\]

Moreover, the ratio $SL^*/SL^*$ reaches the minimum of 66% at $s_2 = 1$ and $B \leq 1$.

The above proposition guarantees that the joint service level under our approximate capacity investment is no less than 66% of that under the exact optimal capacity investment, implying that the performance of the lower bound model is not too bad. Note that it is the worst case bound, and the ratio can be much larger, and indeed attains 100% especially when the budget is sufficiently large.

Now, we turn our attention to general $N$, in particular, $N > 2$. In this case, we do not have an analytic solution to the exact model (2), and thus we use the sample average approximation (SAA) method to find the optimal solution to (2) as well as to evaluate the performance of the solution through the lower bound problem (9).

There are five methods of finding capacities that we have considered: (i) Ball: solution to the lower bound model (9); (ii) Ball-2: closed-form solution stated in Theorem 1; (iii) SAA-JSL: solution to the SAA version of the exact model (2); (iv) SAA-FR: same as SAA-JSL except that the objective function is the fill rate, the fraction of demand satisfied. (v) All-in-2: solution to the SAA version of the exact model (2), which evenly allocates capacity to all level-2 flexibilities.

We are interested in two performance measures: (a) JSL: the joint-service-level, given in the objective function of (2); (b) FR: the fill rate. We have used 1000 demand samples to generate the SAA solutions and have evaluated these five solutions under another 10,000 samples. Table 1 shows that our approximation solutions (Ball and Ball-2) provide decent performance not only in joint service level but also in fill rate, even though the original objective under our investigation is the joint service level. Also, the running time associated with SAA problems grew in a manner exponential in $N$, but the running time associated with Ball and Ball-2 was negligible.

While both Ball and Ball-2 perform well, it is interesting to note that Ball-2 performs better than Ball, making the simpler close-form approximation more attractive. Under linear cost case, though Ball-2 is an optimal solution to the approximate formulation (9) according to Theorem 1, the existence of multiple optimal solutions of Ball leads to the different performances between Ball and Ball-2. More importantly, even though Ball is optimal under (9), such approximation might be so conservative that it is too far...
Joint service levels (JSL) and fill rates (FR) with uniform distribution $U[0,1], B = N/2+2\sqrt{N/T}$. Linear cost means $s_j = 1+(j-1)/2$; Convex cost means $s_j = 1+(j-1)/2^{1/3}$.

<table>
<thead>
<tr>
<th>Cost type</th>
<th>JSL</th>
<th>$N=2$</th>
<th>$N=3$</th>
<th>$N=4$</th>
<th>$N=5$</th>
<th>FR</th>
<th>$N=2$</th>
<th>$N=3$</th>
<th>$N=4$</th>
<th>$N=5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>SAA-JSL</td>
<td>93.2%</td>
<td>79.5%</td>
<td>66.2%</td>
<td>54.7%</td>
<td>SAA-FR</td>
<td>100.0%</td>
<td>100.0%</td>
<td>99.9%</td>
<td>99.6%</td>
</tr>
<tr>
<td></td>
<td>Ball</td>
<td>91.4%</td>
<td>74.4%</td>
<td>52.9%</td>
<td>37.4%</td>
<td>Ball</td>
<td>100.0%</td>
<td>99.9%</td>
<td>99.5%</td>
<td>99.4%</td>
</tr>
<tr>
<td></td>
<td>Ball-2</td>
<td>91.4%</td>
<td>74.4%</td>
<td>59.3%</td>
<td>46.2%</td>
<td>Ball-2</td>
<td>99.2%</td>
<td>99.1%</td>
<td>99.1%</td>
<td>99.1%</td>
</tr>
<tr>
<td></td>
<td>All-in-2</td>
<td>68.8%</td>
<td>62.5%</td>
<td>56.3%</td>
<td>50.7%</td>
<td>All-in-2</td>
<td>99.2%</td>
<td>99.1%</td>
<td>99.1%</td>
<td>99.1%</td>
</tr>
<tr>
<td>Convex</td>
<td>SAA-JSL</td>
<td>94.4%</td>
<td>83.8%</td>
<td>74.6%</td>
<td>65.1%</td>
<td>SAA-FR</td>
<td>100.0%</td>
<td>100.0%</td>
<td>99.9%</td>
<td>99.6%</td>
</tr>
<tr>
<td></td>
<td>Ball</td>
<td>93.4%</td>
<td>78.1%</td>
<td>63.9%</td>
<td>51.4%</td>
<td>Ball</td>
<td>100.0%</td>
<td>99.9%</td>
<td>99.5%</td>
<td>99.5%</td>
</tr>
<tr>
<td></td>
<td>Ball-2</td>
<td>93.4%</td>
<td>78.1%</td>
<td>63.9%</td>
<td>51.4%</td>
<td>Ball-2</td>
<td>100.0%</td>
<td>99.9%</td>
<td>99.5%</td>
<td>99.5%</td>
</tr>
<tr>
<td></td>
<td>All-in-2</td>
<td>72.4%</td>
<td>67.7%</td>
<td>63.2%</td>
<td>58.0%</td>
<td>All-in-2</td>
<td>99.2%</td>
<td>99.1%</td>
<td>99.1%</td>
<td>99.1%</td>
</tr>
<tr>
<td>Concave</td>
<td>SAA-JSL</td>
<td>90.3%</td>
<td>73.5%</td>
<td>58.3%</td>
<td>45.5%</td>
<td>SAA-FR</td>
<td>100.0%</td>
<td>99.9%</td>
<td>99.8%</td>
<td>99.6%</td>
</tr>
<tr>
<td></td>
<td>Ball</td>
<td>89.6%</td>
<td>70.8%</td>
<td>53.5%</td>
<td>39.4%</td>
<td>Ball</td>
<td>100.0%</td>
<td>99.6%</td>
<td>99.4%</td>
<td>99.4%</td>
</tr>
<tr>
<td></td>
<td>Ball-2</td>
<td>89.6%</td>
<td>70.9%</td>
<td>54.2%</td>
<td>41.1%</td>
<td>Ball-2</td>
<td>100.0%</td>
<td>99.6%</td>
<td>99.4%</td>
<td>99.4%</td>
</tr>
<tr>
<td></td>
<td>All-in-2</td>
<td>64.0%</td>
<td>56.5%</td>
<td>49.6%</td>
<td>40.3%</td>
<td>All-in-2</td>
<td>99.2%</td>
<td>99.1%</td>
<td>99.1%</td>
<td>99.1%</td>
</tr>
</tbody>
</table>

Table 1
Joint service levels (JSL) and fill rates (FR) with uniform distribution $U[0,1], B = N/2+2\sqrt{N/T}$. Linear cost means $s_j = 1+(j-1)/2$; Convex cost means $s_j = 1+(j-1)/2^{1/3}$; Concave cost means $s_j = 1+(j-1)/2^{1/3}$. Sample size for generating SAA solutions: 1000; Sample size for evaluation: 10,000.

4. Extensions

While the idea of inscribed sphere has been carried out for the symmetric uniform distribution, we would like to show that it can be applied to other demand distributions. We will consider asymmetric uniform distribution, normal distribution as well as general independent distribution.

4.1. Asymmetric uniform distribution

First of all, we investigate the asymmetric uniform demand case where each $D_j$ independently follows uniform distribution $U[0, M_j]$ for some $M_j$, $j = 1, \ldots, N$. One approach involves linearly shrinking each axis by a factor of $M_j$, such that the resulting distribution becomes uniform $[0, 1]$ for each product. Define $\tilde{D}_j = D_j/M_j$, $j = 1, \ldots, N$, therefore $\tilde{D}_j \sim U[0, 1]$. The feasible production region $\Omega(x)$ defined in (1) is equivalent to the following $\tilde{\Omega}(x)$:

$$\tilde{\Omega}(x) = \left\{ (\tilde{d}_1, \ldots, \tilde{d}_n) \mid \sum_{j \in J} M_j \tilde{d}_j \leq \sum_{(S \subseteq \{1, \ldots, N\}, S \neq \emptyset)} x_S \right\}$$

for each $J \subseteq \{1, \ldots, N\}$.

As discussed in Section 3.1, the problem of finding the largest inscribed sphere inside the intersection between the transformed feasible production region $\tilde{\Omega}(x)$ and the transformed demand support $[0, 1]^N$ can be formulated as the following LP:

$$\begin{align*}
\text{max} & \quad r \\
\text{s.t.} & \quad r \leq p_i \quad \text{and} \quad r \leq 1 - p_i \quad \forall i = 1, \ldots, N, \\
& \quad \sum_{(j \in J)} x_S - \sum_{i=1}^m M_i p_i \\
& \quad \sum_{i=1}^m M_i \\
& \quad \text{for any } 1 \leq j_1, \ldots, j_m \leq N \text{ and } m \leq N. \\
& \quad \sum_{S \subseteq \{1, \ldots, N\}} c_S x_S \leq B \\
& \quad x_S, r \geq 0
\end{align*}$$

(12)

where the right-hand-side of constraint (12), modified from (7), indicates the Euclidean distance from the center $p$ to the boundary away from the true optimal solution of (2). In fact, the optimal solutions of SAA-JSL activate far less higher flexible capacity compared to the Ball solutions. On the other hand, the Ball-2 solutions are, though maybe not optimal under the approximate formulation (9), closer to the true optimal solution.

To further evaluate our approximation solution, we compare it with All-in-2, a simple heuristic that invests only in level-2 flexibility. According to Table 1, Ball-2 seems to have a better performance though may be not optimal under the approximate formulation (9), to the Ball solutions. On the other hand, the Ball-2 solutions are, less than that of All-in-2.

Fig. 3. Capacity investment under varying budget: cost parameters $(s_1, s_2, s_3, s_4) = (1, 1.1, 1.21, 1.33)$.

Joint service levels (JSL) and fill rates (FR) with normal distribution \( N(4, 1) \), \( B = N \mu + 2 \sigma \sqrt{N} \) and \( W = 6 \). Linear cost means \( s_j = 1 + (j - 1)/2 \); Convex cost means \( s_j = 1 + ((j - 1)/2)^2 \). Concave cost means \( s_j = 1 + (j - 1)/2 \)^4. Sample size for generating SAA solutions: 1000; Sample size for evaluation: 10,000.

<table>
<thead>
<tr>
<th>Cost type</th>
<th>JSL</th>
<th>N = 2</th>
<th>N = 3</th>
<th>N = 4</th>
<th>N = 5</th>
<th>FR</th>
<th>N = 2</th>
<th>N = 3</th>
<th>N = 4</th>
<th>N = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAA-JSL</td>
<td>89.8%</td>
<td>79.8%</td>
<td>69.3%</td>
<td>58.9%</td>
<td>99.0%</td>
<td>FR</td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.8%</td>
<td>99.7%</td>
</tr>
<tr>
<td>CONV</td>
<td>89.2%</td>
<td>77.2%</td>
<td>68.3%</td>
<td>49.4%</td>
<td>99.8%</td>
<td></td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.8%</td>
<td>99.6%</td>
</tr>
<tr>
<td>Fix-C</td>
<td>89.1%</td>
<td>78.4%</td>
<td>62.7%</td>
<td>48.3%</td>
<td>Fix-C</td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.7%</td>
<td>99.6%</td>
<td></td>
</tr>
<tr>
<td>All-in-2</td>
<td>29.9%</td>
<td>16.5%</td>
<td>9.3%</td>
<td>5.1%</td>
<td>All-in-2</td>
<td>99.1%</td>
<td>99.0%</td>
<td>99.0%</td>
<td>99.0%</td>
<td></td>
</tr>
<tr>
<td>Convex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAA-JSL</td>
<td>90.2%</td>
<td>82.1%</td>
<td>73.2%</td>
<td>64.6%</td>
<td>99.0%</td>
<td>FR</td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.8%</td>
<td>99.7%</td>
</tr>
<tr>
<td>CONV</td>
<td>90.0%</td>
<td>80.9%</td>
<td>71.0%</td>
<td>61.9%</td>
<td>99.8%</td>
<td></td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.8%</td>
<td>99.6%</td>
</tr>
<tr>
<td>Fix-C</td>
<td>90.0%</td>
<td>80.2%</td>
<td>69.5%</td>
<td>58.2%</td>
<td>Fix-C</td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.7%</td>
<td>99.6%</td>
<td></td>
</tr>
<tr>
<td>All-in-2</td>
<td>38.0%</td>
<td>24.0%</td>
<td>15.9%</td>
<td>9.5%</td>
<td>All-in-2</td>
<td>99.1%</td>
<td>99.0%</td>
<td>99.0%</td>
<td>99.0%</td>
<td></td>
</tr>
<tr>
<td>Concave</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAA-JSL</td>
<td>88.3%</td>
<td>76.7%</td>
<td>65.3%</td>
<td>53.6%</td>
<td>99.0%</td>
<td>FR</td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.8%</td>
<td>99.7%</td>
</tr>
<tr>
<td>CONV</td>
<td>88.2%</td>
<td>76.4%</td>
<td>63.9%</td>
<td>50.7%</td>
<td>CONV</td>
<td>100.0%</td>
<td>99.8%</td>
<td>99.7%</td>
<td>99.6%</td>
<td></td>
</tr>
<tr>
<td>Fix-C</td>
<td>88.2%</td>
<td>76.3%</td>
<td>63.3%</td>
<td>50.2%</td>
<td>Fix-C</td>
<td>100.0%</td>
<td>99.7%</td>
<td>99.7%</td>
<td>99.6%</td>
<td></td>
</tr>
<tr>
<td>All-in-2</td>
<td>22.0%</td>
<td>10.8%</td>
<td>5.1%</td>
<td>2.4%</td>
<td>All-in-2</td>
<td>99.1%</td>
<td>99.0%</td>
<td>99.0%</td>
<td>99.0%</td>
<td></td>
</tr>
</tbody>
</table>

The above LP shares almost the same structure as our lower bound model (9), which implies that a similar result as Theorem 1 can be shown. Considering the space limitation, we omit the detailed discussion here.

4.3. General independent distribution

For general independent demand case, we will directly modify our lower bound model so that it becomes a heuristic. The method is motivated by the fact that for i.i.d. normal distributions, with high probability, any demand sample should fall within the sphere centered at the expectation \( \mu \) of radius \( 3 \sigma \sqrt{N} \). Similar result is also true for general i.i.d. distributions given that the radius is large enough.

We fix the center \( p \) as the demand expectation denoted by \( \mu \). In addition, there may not exist an upper bound on demand, therefore we remove the constraint \( r - (1-p) \leq 0 \). The corresponding linear program should be:

\[
\begin{align*}
\text{max} & \quad r \\
\text{s.t.} & \quad r \leq p - \lambda_1, \quad \lambda_2 \leq p, \\
& \quad r - \frac{\sum_{m=1}^{N} v_{k,m} y_m - k \mu}{\sqrt{k}} \leq 0 \quad \text{for each } k = 1, \ldots, N, \\
& \quad \sum_{m=1}^{N} \left( \frac{N}{m} \right) s_m y_m \leq B, \\
& \quad y_m, r \geq 0.
\end{align*}
\]

Call the heuristic in Section 4.2 as Convolution (CONV) and the one in Section 4.3 as Fix-Center (Fix-C). As defined, SAA-JSL indicates the optimal solution to the SAA version of the exact model (2) and All-in-2 indicates a solution that evenly allocates capacity to all level-2 flexibilities. Table 2 tests the performance of the two heuristics under normal distribution in terms of joint-service level and fill rate. As we can see, Convolution and Fix-Center provide a good performance. Moreover, both of them greatly outperform All-in-2 in the normal distribution case.

Acknowledgments

The authors gratefully acknowledge editor-in-chief Jan Karel Lenstra, the area editor Patrice Marcotte, and an anonymous referee for their valuable comments that helped improve this paper considerably. The research of Huan Zheng was supported by the National Science Foundation of China (Grant 71371119); the research of Ying Rong was supported by Shanghai Pujiang Program and the National Science Foundation of China (Grant 71202068); the research of Woonghee Tim Huh was supported by The Natural Sciences and Engineering Research Council of Canada.
Appendix A


Appendix B. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.orl.2015.01.009.

References