An Approximation Scheme for Stochastic Integer Programs Arising in Capacity Expansion*

Shabbir Ahmed

School of Industrial & Systems Engineering

Georgia Institute of Technology

765 Ferst Drive, Atlanta, GA 30332

Nikolaos V. Sahinidis†

Department of Chemical Engineering

University of Illinois

600 South Mathews Avenue, Urbana, IL 61801

January 11, 2001. Revised: November 26, 2001

*Subject Classifications: Analysis of algorithms: asymptotically optimal heuristics. Facilities/equipment planning: capacity expansion. Programming: stochastic, integer.

†Corresponding author (email: nikos@uiuc.edu)

Abstract

Planning for capacity expansion forms a crucial part of the strategic level decision making in many applications. Consequently, quantitative models for economic capacity expansion planning have been the subject of intense research. However, much of the work in this area has been restricted to linear cost models and/or limited degree of uncertainty to make the problems analytically tractable. This paper addresses a stochastic capacity expansion problem where the economies-of-scale in expansion costs are handled via fixed-charge cost functions, and forecast uncertainties in the problem parameters are explicitly considered by specifying a set of scenarios. The resulting formulation is a multi-stage stochastic integer program. We develop a fast, linear programming based, approximation scheme that exploits the decomposable structure and is guaranteed to produce feasible solutions for this problem. Through probabilistic analysis tools, we prove that the optimality gap of the heuristic solution almost surely vanishes asymptotically as the problem size increases.

1 Introduction

Planning for capacity expansion consists, primarily, of determining future expansion times, sizes, and locations to support anticipated demand growth. This activity forms a crucial part of the strategic level decision making in many applications. Examples can be found in heavy process industries (Sahinidis & Grossmann 1992), communication networks (Chang & Gavish 1993, Saniee 1995, Laguna 1998), electric utilities (Murphy, Sen & Soyster 1987, Murphy & Weiss 1990), automobile industries (Eppen, Martin & Schrage 1989), service industries (Berman, Ganz & Wagner 1994, Berman & Ganz 1994), and, more recently, in electronic goods and semiconductor industries (Rajagopalan, Singh & Morton 1998, Bermon & Hood 1999, Swaminathan 2000). In all of these applications, the expansion of production capacity requires the commitment of substantial capital resources over long periods of time. Furthermore, the economies-of-scale in the expansion costs, as well as the uncertainties in the long range forecasts for costs and demands, make these decision problems very complex. Consequently, quantitative models for economic capacity expansion planning have been the subject of intense research since the early 1960s. However, much of the work in this area has been restricted to linear cost models and/or limited degree of uncertainty to make the problems analytically tractable.

Early approaches for solving stochastic capacity expansion problems were based on stochastic control theory (Manne 1961, Freidenfelds 1980, David, Dempster, Sethi & Vermes 1987, Bean, Higle & Smith 1992). In these models, the demands were assumed to be simple stochastic processes to render analytical tractability. With the advent of stochastic programming (cf. Kall & Wallace (1994) or Birge & Louveaux (1997)) and increased computational power, the use of scenarios to model uncertainties in planning models has become increasingly popular. In two-stage stochastic programming approaches for capacity planning (Eppen et al. 1989, Fine & Freund 1990, Berman et al. 1994, Swaminathan 2000), it is assumed that the entire capacity expansion schedule is decided before the uncertainty is realized, and only some recourse actions can be taken in order to correct any infeasibilities. Since all capacity expansion decisions, and hence any fixed-charge expansion costs, are restricted to the first stage of the problem, standard stochastic programming decomposition methods can be used to solve these models.

Multi-stage models extend the two-stage stochastic programming models by allowing revised decisions in each time stage based upon the uncertainty realized so far. The uncertainty information in a multi-stage stochastic program is modeled as a multi-layered scenario tree, and the optimization problem consists of determining an expansion schedule that hedges against this scenario tree. Multi-stage stochastic linear programming has been extensively treated in the literature (cf. Birge & Louveaux (1997)). Solution approaches for multi-stage stochastic linear programs include nested Benders decomposition (Birge 1985) and progressive hedging (Rockafellar & Wets 1991). However, these methods are inapplicable to stochastic capacity expansion problems with fixed-charge costs owing to the non-convexities caused by the presence of integer variables in later stages.

In the context of capacity planning, Rajagopalan et al. (1998) proposed a multi-stage capacity expansion and replacement model where capacity becomes available only at certain time periods. The demand is assumed to be non-decreasing and the available capacity of a technology, when it appears, is assumed to be sufficient. The authors exploited the structure of the optimal solution to develop a dynamic programming strategy. More recently, Chen, Li & Tirupati (2001) addressed a multi-stage stochastic capacity expansion model for technology selection. The authors assumed linear expansion cost functions and used Lagrangian decomposition to solve the problem.

In this paper, we consider a stochastic capacity expansion problem where the economies-of-scale in expansion costs are handled via fixed-charge cost functions and forecast uncertainties in the problem parameters are explicitly considered by specifying a set of scenarios.

The resulting formulation is a multi-stage stochastic mixed-integer program with binary variables in all stages. The proposed model is fairly general in terms of the existing literature. It allows for multiple production facilities, does not require non-decreasing demand patterns, and allows for limited availability of procurable capacity of the various facility types. However, this generalization makes the problem $\mathcal{NP}$-hard even in a deterministic setting. This computational complexity motivates the need for efficient heuristic methods. Heuristics for solving deterministic capacity expansion problems are prevalent in the literature (Fong & Srinivasan 1981a, Klincewicz, Luss & Yu 1988, Li & Tirupati 1994). However, theoretical analyses of the performance of these heuristics have not been reported. In a recent line of work, Liu & Sahinidis (1997), Ahmed & Sahinidis (2000a), and Ahmed & Sahinidis (2000b) proposed LP-based heuristics for deterministic capacity expansion problems in the chemical process industries and in manufacturing technology selection. Using probabilistic analysis tools, these schemes were proven to be asymptotically optimal in the number of time periods. In this paper, we extend this approach to the stochastic case.

The remainder of this paper is organized as follows. Section 2 presents the multi-stage integer programming formulation. A heuristic scheme for the problem is presented in Section 3. The main idea is to decompose the problem into a sequence of smaller, deterministic problems each of which is solved by an efficient heuristic. A probabilistic analysis of the approach is carried out in Section 4. For a standard probability model, we show that the heuristic is asymptotically optimal as the planning horizon increases. Finally, in Section 5, we present computational results in the context of capacity expansion of chemical processing networks under uncertainty.

2 Problem Statement

In this section, we present a multi-stage stochastic programming formulation for capacity expansion under uncertainty with fixed-charge expansion costs.

We consider the problem of determining the timing and the level of capacity acquisitions for a set of production facilities $\mathcal{I}$, along with a policy for allocating the available capacity to satisfy the demand of a set of product families $\mathcal{J}$, while minimizing the expected total discounted investment and allocation cost for a planning horizon of n periods. The product demands (d), variable and fixed costs of capacity acquisition (α and β), and the costs for allocating capacity to products (δ) are assumed to be stochastic. We model uncertainty as a multi-layered tree. Each node layer in the tree corresponds to a time period t. A scenario s corresponds to a single path from the root to a unique leaf of this tree, representing a joint realization of the uncertain parameters over all time periods, *i.e.*, $\{d_{jt}^s, \alpha_{it}^s, \beta_{it}^s, \delta_{ijt}^s\}_{t=1}^n$, where $i \in \mathcal{I}$ and $j \in \mathcal{J}$. Figure 1 presents an example of a scenario tree. At any time stage, a node may be identified with the "bundle" B_t of scenarios passing through it. We denote the collection of nodes or bundles at a time stage t by $\mathcal{B}_t$. The collection at the root node $\mathcal{B}_1$ contains a single bundle B_1 consisting of all scenarios $s = 1, \ldots, S$. Similarly, the collection $\mathcal{B}_n$ at the leaf nodes consists of S members each of which contains a single scenario. The probability associated with a scenario (path) s is denoted by p^s.

With the uncertainty information structure specified as above, we can now state a formulation for the problem. The following notation will be used to describe the model:

Sets and indices:

- i index for a production facility ($i \in \mathcal{I}$);
- j index for a product family ($j \in \mathcal{J}$);
- $\mathcal{I}_j$ set of facilities that are capable of producing product family j;
- $\mathcal{J}_i$ set of product families that can be produced by facility i;
- t index for time periods ($t = 1, \dots, n$);
- s index for scenarios ($s = 1, \dots, S$).

Parameters:

- α_{it}^s variable cost associated with acquiring unit capacity of facility i in period t under scenario s;
- β_{it}^s fixed cost associated with acquiring capacity of facility i in period t under scenario s;
- δ_{ijt}^s cost associated with allocating a unit of capacity of facility i to product family j in period t under scenario s;
- μ_{ij} (deterministic) yield rate of product family j per unit capacity of facility i;
- d_{jt}^s demand of product family j in period t under scenario s;
- U_{it} (deterministic) capacity of facility i available for acquisition in time period t;
- X_{i0} initial capacity of facility i;
- p^s probability of scenario s ($\sum_{s=1}^{S} p^s = 1$).

Variables:

- W_{ijt}^s amount of capacity of facility i allocated to product family j in period t under scenario s;
- X_{it}^s amount of capacity addition to facility i in period t under scenario s;
- Y_{it}^s binary variables equal to 1 if capacity of facility i is acquired in period t under scenario s and equal to 0 otherwise.

We assume that $\cup_{i \in \mathcal{I}} \mathcal{J}_i = \mathcal{J}$ and $\cup_{j \in \mathcal{J}} \mathcal{I}_j = \mathcal{I}$, so that there is a facility available for each product family and $\mathcal{I}$ consists of only those facilities that can produce one or more of the product families in $\mathcal{J}$. We also assume that all cost parameters are appropriately discounted to their present values. Following Rajagopalan (1994) and Li & Tirupati (1994), we ignore inventory fluctuations and do not permit disposal of excess capacity. With the above assumptions, the stochastic capacity expansion problem is formulated as follows.

$$\text{(SCAP)}: \quad \min \textstyle\sum_{s=1}^{S} p^s \left\{ \sum_{t=1}^{n} \sum_{i\in\mathcal{I}} \left(\alpha_{it}^s X_{it}^s + \beta_{it}^s Y_{it}^s + \sum_{j\in J_i} \delta_{ijt}^s W_{ijt}^s \right) \right\} \tag{1}$$

subject to

$$X_{it}^s \le U_{it} Y_{it}^s \qquad t=1,\ldots,n; \quad i\in\mathcal{I}; \quad s=1,\ldots,S \tag{2}$$

$$\sum_{j\in\mathcal{J}_i} W_{ijt}^s \le X_{i0} + \sum_{\tau=1}^{t} X_{i\tau}^s \qquad t=1,\ldots,n; \quad i\in\mathcal{I}; \quad s=1,\ldots,S \tag{3}$$

$$\sum_{i\in\mathcal{I}_j} \mu_{ij} W_{ijt}^s = d_{jt}^s \qquad t=1,\ldots,n; \quad j\in\mathcal{J}; \quad s=1,\ldots,S \tag{4}$$

$$W_{ijt}^s, X_{it}^s \ge 0 \qquad t=1,\ldots,n; \quad j\in\mathcal{J}_i; \quad i\in\mathcal{I}; \quad s=1,\ldots,S \tag{5}$$

$$Y_{it}^s \in \{0,1\} \qquad t=1,\ldots,n; \quad i\in\mathcal{I}; \quad s=1,\ldots,S \tag{6}$$

$$X_{it}^{s_1} = X_{it}^{s_2} \qquad \forall (s_1,s_2)\in B_t, \forall B_t \in \mathcal{B}_t, t=1,\ldots,n; \quad i\in\mathcal{I} \tag{7}$$

$$Y_{it}^{s_1} = Y_{it}^{s_2} \qquad \forall (s_1,s_2)\in B_t, \forall B_t \in \mathcal{B}_t, t=1,\ldots,n; \quad i\in\mathcal{I} \tag{8}$$

$$W_{ijt}^{s_1} = W_{ijt}^{s_2} \qquad \forall (s_1,s_2)\in B_t, \forall B_t \in \mathcal{B}_t, t=1,\ldots,n; \quad j\in\mathcal{J}_i; \quad i\in\mathcal{I} \tag{9}$$

In the formulation above, the objective (1) minimizes the expected total investment and allocation costs over the planning horizon. Constraint (2) ensures that the capacity acquired in any period and any scenario does not exceed the upper bound on the acquirable capacity. Constraint (3) enforces the condition that, for any facility, the total capacity allocated to the product families does not exceed the installed capacity. Constraint (4) links the allocated capacities to the product demand. Non-negativity and binary restrictions of the variables are enforced through (5) and (6). Notice that, at time stage t, the decision maker cannot distinguish between two scenarios s_1 and s_2 that belong to the same node B_t of the scenario tree. Consequently, the decisions corresponding to scenarios s_1 and s_2 have to be identical. These *non-anticipativity* restrictions (cf. Birge & Louveaux (1997)) are enforced by constraints (7), (8), and (9).

Ahmed & Sahinidis (2000b) proved that, owing to the presence of finite bounds on the capacity additions, the deterministic capacity expansion problem ($S = 1$) is $\mathcal{NP}$-hard with respect to the number of time periods even for a single facility. Since the deterministic problem is a special single-scenario version of (SCAP), (SCAP) is $\mathcal{NP}$-hard. Currently, no practicable general-purpose solution methodology exists for the exact solution of multi-stage stochastic integer programs. In principle, with the scenario tree specified, the problem is a large-scale deterministic mixed-integer program and can be solved by standard integer programming techniques. However, such a scheme will be very expensive computationally. In the next section, we describe an efficient decomposition-based heuristic strategy to construct good quality solutions to (SCAP).

3 An Approximation Scheme

In this section, we develop an approximation scheme to construct solutions to (SCAP). The approach is motivated by the following observations: relaxing the integrality restrictions reduces the problem to a stochastic linear program which can be solved by standard decomposition methods; and relaxing the non-anticipativity constraints decomposes the problem into S instances of the deterministic capacity expansion problem which can be solved independently.

For notational ease, let us denote a joint realization of the uncertain parameters (or a scenario) by $\omega^s := (\omega_1^s, \ldots, \omega_n^s)$ where $\omega_t^s := (\alpha_t^s, \beta_t^s, \delta_t^s, d_t^s)$. Note that the subscripts i and j have been omitted. The technological constraints (2)-(6) in (SCAP) corresponding to scenario s will be concisely denoted by $\mathcal{X}(\omega^s)$. The decision variables corresponding to scenario s will be denoted by $x^s := (x_1^s, \ldots, x_n^s)$ with $x_t^s := (X_t^s, Y_t^s, W_t^s)$. The objective function (1) corresponding to scenario s for an $n-$period problem will be denoted by $f_n^s(\cdot)$. The non-anticipativity constraints (7)-(9) will collectively be denoted by $\mathcal{N}$. Using this notation, we can concisely represent the problem as:

$$\text{(SCAP)}: \quad \min \quad z_n = \sum_{s=1}^{S} p^s f_n^s(x^s) \tag{10}$$

$$\text{s.t.} \quad x^s \in \mathcal{X}(\omega^s) \cap \mathcal{N} \quad \forall s = 1, \ldots, S \tag{11}$$

From now on, we shall refer to the above representation of (SCAP) for convenience.

Note that, for a solution $(x^1, \ldots, x^S)$ to be feasible, it needs to satisfy both the technological constraints $\mathcal{X}(\omega^s)$ and the non-anticipativity constraints $\mathcal{N}$, *i.e.*, $x^s \in \mathcal{X}(\omega^s) \cap \mathcal{N}$ for $s = 1, \ldots, S$. A solution that satisfies only the non-anticipativity constraints, but not necessarily the technological constraints, is called *implementable*; while a solution that satisfies only the technological constraints, but not necessarily the non-anticipativity constraints, is called *admissible* (Rockafellar & Wets 1991). Thus, a solution is *feasible* if it is both implementable and admissible. We propose to construct a feasible solution to (SCAP) in the following three phases:

I. Relax the integrality requirement in the technological constraints $\mathcal{X}(\omega^s)$ and construct an implementable solution $(\overline{x}^1, \ldots, \overline{x}^S) \in \mathcal{N}$. If the current solution is also admissible, *i.e.*, $\overline{x}^s \in \mathcal{X}(\omega^s)$ for $s = 1, \ldots, S$, then stop. Otherwise, go to Phase II.

II. Relax the non-anticipativity constraints $\mathcal{N}$ and perturb $(\overline{x}^1, \ldots, \overline{x}^S)$ to construct an admissible solution $(\underline{x}^1, \ldots, \underline{x}^S)$ such that $\underline{x}^s \in \mathcal{X}(\omega^s)$ for $s = 1, \ldots, S$. If such a solution is also implementable, *i.e.*, $(\underline{x}^1, \ldots, \underline{x}^S) \in \mathcal{N}$, then stop. Otherwise, go to

Phase III.

III. Re-enforce the non-anticipativity constraints on $(\underline{x}^1, \ldots, \underline{x}^S)$ to construct a feasible solution $(\widehat{x}^1, \ldots, \widehat{x}^S)$ such that $\widehat{x}^s \in \mathcal{X}(\omega^s) \cap \mathcal{N}$ for $s = 1, \ldots, S$.

Details of each of the above steps are described next.

3.1 Phase I: Constructing an implementable solution

Relaxing the integrality requirement in the constraint set $\mathcal{X}(\omega^s)$ transforms (SCAP) into a multi-stage stochastic *linear* program, which can then be solved using LP technology or specialized decomposition techniques such as the nested Benders decomposition (Birge 1985) or progressive hedging (Rockafellar & Wets 1991). Performance of solution approaches to stochastic linear programs largely depends on problem size and structure and a problem-specific implementation may be required for very large-scale problems. Now matter how this LP is solved, its solution clearly obeys the non-anticipativity restrictions.

3.2 Phase II: Constructing an admissible solution

Relaxing the non-anticipativity constraints $\mathcal{N}$ decomposes (SCAP) into S instances of the deterministic capacity expansion problem (for $s = 1, \ldots, S$):

$$\text{(CAP)}: \quad \min \textstyle\sum_{t=1}^{n} \sum_{i \in \mathcal{I}} \left(\alpha_{it}^s X_{it}^s + \beta_{it}^s Y_{it}^s + \sum_{j \in J_i} \delta_{ijt}^s W_{ijt}^s \right)$$

subject to

$$\begin{aligned} X_{it}^s &\le U_{it}^s Y_{it}^s & t = 1, \ldots, n; \;\; i \in \mathcal{I} \\ \sum_{j \in \mathcal{J}_i} W_{ijt}^s &\le X_{i0} + \sum_{\tau=1}^{t} X_{i\tau}^s & t = 1, \ldots, n; \;\; i \in \mathcal{I} \end{aligned}$$

$$\sum_{i \in \mathcal{I}_j} \mu_{ij} W_{ijt}^s = d_{jt}^s \qquad t = 1, \ldots, n; \quad j \in \mathcal{J}$$

$$W_{ijt}^s, X_{it}^s \geq 0 \qquad t = 1, \ldots, n; \quad i \in \mathcal{I}; \quad j \in \mathcal{J}$$

$$Y_{it}^s \in \{0, 1\} \qquad t = 1, \ldots, n; \quad i \in \mathcal{I}$$

Liu & Sahinidis (1997), Ahmed & Sahinidis (2000a) and Ahmed & Sahinidis (2000b) proposed temporal capacity shifting heuristics based upon perturbing the LP relaxation solution to construct integral solutions to (CAP). The motivation for the development of these heuristics comes from the empirical results of Chang & Gavish (1995) and Liu & Sahinidis (1995), who observed decreasing relaxation gaps of deterministic capacity expansion problems with respect to the planning horizon length. This empirical evidence suggests the possibility of construction of good quality solutions from the LP relaxation solution for instances of (CAP) with large planning horizons. Note that simply rounding up the values of the binary variables (Y_{it}) in the LP relaxation of (CAP) results in a feasible solution. However, such a naive strategy might result in very poor solutions, possibly requiring capacity expansion to be carried out in all periods. It is important to perturb the LP relaxation solution in a way so as to keep the number of expansion decisions small. Note that, since fixed costs of capacity addition are typically high, we wish to acquire as much capacity as possible whenever the decision to expand is taken. It is easy to see that, if the investment costs are constant across all periods, there is an optimal solution to (CAP) where capacity acquisitions are made only in the earliest periods. Furthermore, in this case, there is an optimal solution where the capacity addition equals the availability bound (U_{it}) in all periods except perhaps the last one in which capacity was added. Using these observations, Ahmed & Sahinidis (2000a) and Ahmed & Sahinidis (2000b) proposed to perturb the LP relaxation solution by shifting capacity additions from later periods to the earlier periods if capacity is available.

Having decomposed (SCAP) by relaxing the non-anticipativity constraints, we can apply the above temporal capacity shifting heuristic to the implementable (LP relaxation) solution from Phase I to construct admissible (integral) solutions for each of the scenario subproblems of (SCAP). A formal statement of this is presented below:

1. Start with the implementable (LP relaxation) solution from Phase I $(\overline{x}^1, \ldots, \overline{x}^S)$ where $\overline{x}^s := (\overline{x}_1^s, \ldots, \overline{x}_n^s)$, and $\overline{x}_t^s := (\overline{X}_{it}^s, \overline{Y}_{it}^s, \overline{W}_{ijt}^s)$.

2. Denote the solution obtained in this phase by $(\underline{x}^1, \ldots, \underline{x}^S)$ where $\underline{x}^s := (\underline{x}_1^s, \ldots, \underline{x}_n^s)$, and $\underline{x}_t^s := (\underline{X}_{it}^s, \underline{Y}_{it}^s, \underline{W}_{ijt}^s)$. Set $\underline{W}_{ijt}^s = \overline{W}_{ijt}^s$.

3. Repeat the following temporal capacity shifting heuristic for all $s = 1, \ldots, S$. The superscripts s have been eliminated for brevity.

   (a) For each $i \in \mathcal{I}$, let $T_i := \{t | \overline{Y}_{it} > 0\}$, *i.e.*, the set of time periods when capacity is added in the LP solution of Phase I. Let $T_i = \{t_1, t_2, \ldots, t_{p_i}\}$.

   (b) Repeat the following step for all $i \in \mathcal{I}$:

Do for $h = 1, \ldots, p_i$

Set $\underline{X}_{it_h} \leftarrow \overline{X}_{it_h}$, $\underline{Y}_{it_h} \leftarrow 0$, and $k \leftarrow h + 1$.

While $\underline{X}_{it_h} < U_{it_h}$ and $k \leq p_i$ do,

Let $\delta = \min\{U_{it_h} - \underline{X}_{it_h}, \overline{X}_{it_k}\}$.

Set $\underline{X}_{it_h} \leftarrow \underline{X}_{it_h} + \delta$ and $\overline{X}_{it_k} \leftarrow \overline{X}_{it_k} - \delta$.

Set $k \leftarrow k + 1$.

End While.

If $\underline{X}_{it_h} > 0$, set $\underline{Y}_{it_h} \leftarrow 1$.

End Do.

The following two properties of the solution obtained by the above heuristic are obvious from the construction:

Lemma 3.1 *For any production facility* $i \in \mathcal{I}$, $\sum_{\tau=1}^{t} \underline{X}_{it}^{s} \geq \sum_{\tau=1}^{t} \overline{X}_{it}^{s}$ *for all* $t = 1, \ldots, n$ *and* $s = 1, \ldots, S$.

Lemma 3.2 *For any production facility* $i \in \mathcal{I}$, $\sum_{t=1}^{n} \underline{X}_{it}^{s} = \sum_{t=1}^{n} \overline{X}_{it}^{s}$ *for all* $s = 1, \ldots, S$.

Since the implementable solution $\overline{x}^s$ satisfies all constraints in $\mathcal{X}(\omega^s)$, except perhaps the integrality requirements, by the above results and the construction the solution $\underline{x}^s$ satisfies all constraints in $\mathcal{X}(\omega^s)$ including the integrality requirements.

Proposition 3.3 *The solution* $(\underline{x}^1, \ldots, \underline{x}^S)$ *obtained in Phase II is admissible.*

We shall now establish a crucial property of the temporal capacity shifting heuristic that is needed for the probabilistic analysis in the next section.

Lemma 3.4 *For any facility* $i \in \mathcal{I}$, $\sum_{t=1}^{n} U_{it}\underline{Y}_{it}^{s} - \sum_{t=1}^{n} U_{it}\overline{Y}_{it}^{s} \leq U_i^{max}$, *where* $U_i^{max} = \max_{t=1,\ldots,n}\{U_{it}\}$. *If* U_{it} *is constant across time periods, the above reduces to* $\sum_{t=1}^{n} \underline{Y}_{it}^{s} - \sum_{t=1}^{n} \overline{Y}_{it}^{s} \leq 1$.

Proof: Recall that, by construction, the solution in Phase II consists of capacity expansions up to the available capacity level for all periods in which the capacity is acquired except perhaps the last. Let t' be the last period in which capacity is acquired in the heuristic solution. Then, for any facility i, $\sum_{t=1}^{n} \underline{X}_{it}^{s} = \sum_{t=1}^{n} U_{it}\underline{Y}_{it}^{s} - \epsilon_{it'}^{s}$, where $\epsilon_{it'}^{s} = U_{it'} - \underline{X}_{it'}^{s}$. The LP relaxation solution satisfies $\sum_{t=1}^{n} \overline{X}_{it}^{s} = \sum_{t=1}^{n} U_{it}\overline{Y}_{it}^{s}$. Noting that $\epsilon_{it'}^{s} \leq U_i^{max}$, the result follows from Lemma 3.2. □

Note that the above property will not be satisfied by a naive round-up strategy, since such a scheme could potentially lead to rounding up the binary variables in all periods. Let us illustrate this fact with an example.

Example

Consider the following deterministic single-facility capacity expansion problem.

$$\begin{array}{ll} \min & \sum_{t=1}^{n} \alpha_t X_t + \beta_t Y_t \\ \text{s.t.} & 0 \le X_t \le U_t Y_t \quad t = 1, \ldots, n \\ & \sum_{\tau=1}^{t} X_\tau \ge d_t \quad t = 1, \ldots, n \\ & Y_t \in \{0, 1\} \quad t = 1, \ldots, n \end{array}$$

Let us assume that $\alpha_t > \alpha_{t+1}$ and $\beta_t > \beta_{t+1}$ for all $t = 1, \ldots, n$, *i.e.*, it is cheaper to postpone capacity addition. We also assume that $U_t \ge (d_t - \max_{1\le\tau\le t-1} d_\tau)^+$, where $(\cdot)^+ = \max(0, \cdot)$.

Consider now the class of problem instances for which the demand parameters are given by $d_t = \sum_{\tau=1}^{t} \frac{C}{\tau^2}$ and $U_t = C$ for all $t = 1, \ldots, n$, for some $C > 0$. Note that the demand d_t is bounded above by $\frac{C\pi^2}{6}$.

It is then clear that an optimal solution to the LP-relaxation of the above problem is

$$\overline{X}_t = (d_t - \max_{1\le\tau\le t-1} d_\tau)^+ \text{ and } \overline{Y}_t = \frac{(d_t - \max_{1\le\tau\le t-1} d_\tau)^+}{U_t} \quad \text{for all } t = 1, \ldots, n.$$

Specifically, the LP-relaxation solution is $\overline{Y}_t = \frac{1}{t^2}$ for all $t = 1, \ldots, n$. If we use a naive round-up strategy to construct a feasible integer solution, then the resulting solution is $Y_t^R = \lceil \frac{1}{t^2} \rceil = 1$ yielding a total number of expansion set-ups of $\sum_{t=1}^{n} Y_t^R = n$. On the other hand, by Lemma 3.4, it is easily seen that the proposed shifting heuristic guarantees $\sum_{t=1}^{n} \underline{Y}_t^s \le \sum_{t=1}^{n} \overline{Y}_t^s + 1 \le \lfloor \frac{\pi^2}{6} + 1 \rfloor = 2$. □

Phase III: Constructing a feasible solution

In this final phase of the heuristic, we construct a solution that is both implementable and admissible, hence feasible. Let $\widehat{x}^s$ denote this solution. Note that the capacity shifting step might destroy the non-anticipativity structure of the capacity expansion variables (X_t^s, Y_t^s). We recover non-anticipativity by a procedure we call *capacity bundling* where we set $\widehat{X}_{it}^s = \max_{s \in B_t} \{\underline{X}_{it}^s\}$ for all $s \in B_t$ for all $B_t \in \mathcal{B}_t$. This guarantees that the capacity acquired in any period is the same in all scenarios of a scenario bundle. Finally, the values of the binary variables are rounded up accordingly.

Figure 2 illustrates the heuristic strategy for a simple 3-period, 4-scenario example. The solutions obtained in each of the three phases of the heuristic are plotted. The heights of the rectangular blocks represent the capacity expansion bounds, and the heights of fillings in the block represent the amounts of capacity added in the corresponding solution. Note that the LP relaxation solution satisfies the non-anticipativity constraints. For example, the capacity additions in scenarios 2 and 3 in time period 2 are the same since these scenarios belong to the same bundle. However, after the temporal capacity shifting, the non-anticipativity structure is destroyed. The capacity bundling phase restores the non-anticipativity structure.

From construction of the heuristic, it can be easily verified that:

Theorem 3.5 *The solution obtained by the proposed heuristic is feasible to* (SCAP)*, i.e.,* $\widehat{x}^s \in \mathcal{X}(\omega) \cap \mathcal{N}$.

The proposed heuristic has a running time of $O(T_{LP} + mn^2 + Sn)$, where T_{LP} is the effort required to solve the LP relaxation, m is the number of facilities, n is the number of time periods, and S is the total number of scenarios in the scenario tree.

The heuristic can be easily improved by shifting capacity only to periods that offer an expected cost benefit in Phase II. Furthermore, the proposed strategy can potentially be integrated with other heuristic methods such as those proposed by Fong & Srinivasan (1981b) and Li & Tirupati (1994). Such improvements will only produce better quality solutions. However, in the next section, we show that the proposed heuristic even in its simple form is asymptotically optimal in the number of planning periods.

4 Probabilistic Analysis

In this section, we carry out a probabilistic analysis to characterize the probable performance of the heuristic on "typical" problem instances. We consider a fixed set of product families $\mathcal{J}$, a fixed set of production facilities $\mathcal{I}$, and a fixed set of yield rates μ_{ij}. The probabilistic analysis will be carried out on instances of (SCAP) consisting of increasingly more time periods where the problem parameters such as costs, product demands, and technological availability are drawn from the following probabilistic model.

- The capacity expansion bounds, $\{U_{it}\}_{t=1}^n$, of a facility i are drawn from distributions with bounded support, *i.e.*, $U_{it} \in [\underline{U}_i, \overline{U}_i]$, with $\underline{U}_i > 0$ for all $i \in \mathcal{I}$.
- The demands of the various product families and the cost parameters for the various technologies are assumed to follow one of the following distributions:

(D1) For each product family j, the demand, $\{d_{jt}\}_{t=1}^n$, is a sequence of i.i.d random variables with finite first and second moments. For a given technology i, the cost parameters $\{\alpha_{it}\}_{t=1}^n$ and $\{\beta_{it}\}_{t=1}^n$ are either sequences of i.i.d random variables with finite first and second moments, or are sequences of random variables (not necessarily i.i.d) with bounded supports.

(D2) For each product family j, the demand, $\{d_{jt}\}_{t=1}^n$, is a sequence of independent random variables (not necessarily identically distributed) with bounded second moments. For a given technology i, the cost parameters $\{\alpha_{it}\}_{t=1}^n$ and $\{\beta_{it}\}_{t=1}^n$ are sequences of random variables with bounded supports.

(D3) For each product family j, the demand, $\{d_{jt}\}_{t=1}^{n}$, is a sequence of random variables with bounded support. For a given technology i, the cost parameters $\{\alpha_{it}\}_{t=1}^{n}$ and $\{\beta_{it}\}_{t=1}^{n}$ are either sequences of i.i.d random variables with finite first and second moments or are sequences of random variables with bounded supports.

- The problem parameters are such that the random instance is feasible. This can be easily ensured by including an expensive artificial facility with infinite capacity that is capable of producing all product families.

To specify an instance of (SCAP), we need to construct a scenario tree of realizations for the stochastic cost and demand parameters. The generation of scenario trees for multi-stage stochastic programming is an active field of research. See Dupačová, Consigli & Wallace (2000) for a survey. In the current analysis, we assume that the scenario tree is generated from data sample paths. In this scheme, the first step is to delineate the initial structure of the scenario tree, *i.e.*, the number of stages and the branching scheme. Independent sample paths of the stochastic problems parameters are generated by simulation. The sample paths are then "fitted" onto the scenario tree by first discretizing the range of sample points according to the number of nodes of the tree at a particular stage. The weights of the scenario paths are then computed by collecting the sample paths that pass through the data ranges in the nodes of that scenario path. Each scenario in the tree is then either a sample path of data realizations or a collection of such sample paths. This scheme is also followed, for example, in building scenario trees from simulations in IBM's commercial stochastic programming software (IBM Corporation 1998).

Using the above probability model, we shall now prove that the optimality gap of the proposed heuristic almost surely vanishes asymptotically as the problem size increases. The main tool in our probabilistic analysis is the asymptotic convergence properties of extreme order statistics (Galambos 1987). For a sequence of random variables $\{x_1, x_2, \ldots, x_n\}$, consider the random variables $x_n^{max} = \max_{j=1,\ldots,n}\{x_j\}$ and $x_n^{min} = \min_{j=1,\ldots,n}\{x_j\}$. The asymptotic theory of extreme order statistics concerns with the study of the limiting distributions of the statistics x_n^{max} and x_n^{min} as n approaches infinity. We shall make use of the following two asymptotic properties of x_n^{max}:

Lemma 4.1 *Suppose $\{x_j\}_{j=1}^n$ is a sequence of non-negative i.i.d random variables with finite second moment. Let $x_n^{max} = \max_{j=1,n}\{x_j\}$. Then, $\lim_{n\to\infty} \frac{x_n^{max}}{\sqrt{n}} = 0$ with probability 1 (w.p. 1).*

Lemma 4.2 *Suppose $\{x_j\}_{j=1}^n$ is a sequence of non-negative independent random variables with bounded first and second moments. Let $x_n^{max} = \max_{j=1,n}\{x_j\}$. Then, $\lim_{n\to\infty} \frac{x_n^{max}}{n} = 0$ w.p. 1.*

Lemmas 4.1 and 4.2 can be easily shown using classical results in extreme value theory (cf. Leadbetter, Lindgren & Rootźen (1983) or Galambos (1987)).

Lemma 4.3 *Consider instances of* (SCAP) *generated from distributions* (D1) *or* (D3) *for the problem parameters. Then, the solution corresponding to any scenario s obtained at the end of Phase II of the heuristic satisfies the following:*

$$\text{For any } i \in \mathcal{I}, \qquad \lim_{n\to\infty} \frac{\sum_{t=1}^n Y_{it}^s}{\sqrt{n}} = 0 \quad \text{w.p. 1.}$$

If the instances are generated from distribution (D2) *for the problem parameters, we have:*

$$\text{For any } i \in \mathcal{I}, \qquad \lim_{n\to\infty} \frac{\sum_{t=1}^n Y_{it}^s}{n} = 0 \quad \text{w.p. 1.}$$

Proof: For a random instance of (SCAP) with n time periods, let the maximum demand of product family j in scenario s be $d_j^s(n) = \max_{t=1,n}\{d_{jt}^s\}$. Clearly, in the optimal solution of the LP relaxation of (SCAP) obtained in Phase I, the final capacity of any facility i in scenario

s will satisfy $X_{i0} + \sum_{t=1}^{n} \overline{X}_{it}^{s} \leq \sum_{j \in \mathcal{J}_i} d_j^s(n)$ while, additionally, $\sum_{t=1}^{n} \overline{X}_{it}^{s} = \sum_{t=1}^{n} U_{it} \overline{Y}_{it}^{s}$. From Lemma 3.4, we then have

$$\sum_{t=1}^{n} U_{it} \underline{Y}_{it}^{s} \leq \sum_{j \in \mathcal{J}_i} (d_j^s(n) - X_{i0})^+ + \overline{U}_i,$$

or,

$$\sum_{t=1}^{n} \underline{Y}_{it}^{s} \leq \frac{\sum_{j \in \mathcal{J}_i} (d_j^s(n) - X_{i0})^+}{\underline{U}_i} + \frac{\overline{U}_i}{\underline{U}_i}, \tag{12}$$

where $(\cdot)^+ = \max(\cdot, 0)$. Recall that our scenario tree is constructed by collecting sample paths $\{d_{jt}^p\}$ generated according to one of the distributions. Let for each sample path p, $d_j^p(n) = \max_{t=1,n}\{d_{jt}^p\}$. Let $\mathcal{P}_s$ be the set of sample paths that correspond to scenario s. Since the scenario parameters are obtained by discretizing the range of the path parameters, then $\min_{p \in \mathcal{P}_s}\{d_j^p(n)\} \leq d_j^s(n) \leq \max_{p \in \mathcal{P}_s}\{d_j^p(n)\}$. If the demand sequence satisfies distribution (D1), then by Lemma 4.1, for each sample path $p \in \mathcal{P}_s$, $\lim_{n \to \infty} \frac{d_j^p(n)}{\sqrt{n}} = 0$ w.p. 1 for all $j \in \mathcal{J}_i$. If the demands satisfy distribution (D2), then by Lemma 4.2, for each sample path $p \in \mathcal{P}_s$, $\lim_{n \to \infty} \frac{d_j^p(n)}{n} = 0$ w.p. 1 for all $j \in \mathcal{J}_i$. If the demands satisfy distribution (D3), then for each $j \in \mathcal{J}_i$ there exists a finite upper bound $\overline{d}_j$ on the demand in all time periods. Then, for each sample path $p \in \mathcal{P}_s$, $d_j^p(n) \leq \overline{d}_j$ w.p. 1 and $\lim_{n \to \infty} \frac{d_j^p(n)}{\sqrt{n}} = 0$ w.p. 1. Thus, under distributions (D1) and (D3), $\lim_{n \to \infty} \frac{d_j^s(n)}{\sqrt{n}} = 0$ w.p. 1, and under distribution (D2), $\lim_{n \to \infty} \frac{d_j^s(n)}{n} = 0$ w.p. 1 for all $j \in \mathcal{J}$. Using these limits in (12) and the fact $\overline{U}_i$ and $\underline{U}_i$ are independent of n, we have the desired results. □

Proposition 4.4 *For a given scenario s,*

$$\lim_{n \to \infty} \frac{f_n^s(\underline{x}^s) - f_n^s(\overline{x}^s)}{n} = 0 \qquad \textit{w.p. 1.}$$

Proof: From the construction of the Phase II solution, we have:

$$f_n^s(\underline{x}^s) - f_n^s(\overline{x}^s) \leq \sum_{t=1}^{n} \sum_{i \in \mathcal{I}} \left[\alpha_{it}^s \underline{X}_{it}^s - \alpha_{it}^s \overline{X}_{it}^s + \beta_{it}^s \underline{Y}_{it}^s - \beta_{it}^s \overline{Y}_{it}^s \right].$$

Let $\alpha_i^{max} = \max_{t=1,n}\{\alpha_{it}^s\}$, $\alpha_i^{min} = \min_{t=1,n}\{\alpha_{it}^s\}$, $\beta_i^{max} = \max_{t=1,n}\{\beta_{it}^s\}$, and $\beta_i^{min} = \min_{t=1,n}\{\beta_{it}^s\}$. Using Lemma 3.2 and Lemma 3.4, we have

$$\begin{aligned} f_n^s(\underline{x}^s) - f_n^s(\overline{x}^s) &\leq \sum_{i \in \mathcal{I}} \left[\left(\alpha_i^{max} - \alpha_i^{min}\right) \sum_{t=1}^{n} \underline{X}_{it}^s + \beta_i^{max} \sum_{t=1}^{n} \underline{Y}_{it}^s - \beta_i^{min} \left(\sum_{t=1}^{n} \underline{Y}_{it}^s - 1 \right) \right] \\ &\leq \sum_{i \in \mathcal{I}} \left[\beta_i^{min} + \left\{ \overline{U}_i (\alpha_i^{max} - \alpha_i^{min}) + (\beta_i^{max} - \beta_i^{min}) \right\} \sum_{t=1}^{n} \underline{Y}_{it}^s \right]. \end{aligned}$$

Since $0 \leq \alpha_i^{min}$ and $0 \leq \beta_i^{min} \leq \beta_i^{max}$ for all $i \in \mathcal{I}$, we have:

$$f_n^s(\underline{x}^s) - f_n^s(\overline{x}^s) \leq \sum_{i \in \mathcal{I}} \left[\beta_i^{max} + (\alpha_i^{max} \overline{U}_i + \beta_i^{max}) \sum_{t=1}^{n} \underline{Y}_{it}^s \right].$$

Dividing by n, we obtain:

$$\frac{f_n^s(\underline{x}^s) - f_n^s(\overline{x}^s)}{n} \leq \sum_{i \in \mathcal{I}} \left[\frac{\beta_i^{max}}{n} + \left(\overline{U}_i \frac{\alpha_i^{max}}{\sqrt{n}} + \frac{\beta_i^{max}}{\sqrt{n}} \right) \frac{\sum_{t=1}^{n} \underline{Y}_{it}^s}{\sqrt{n}} \right].$$

Under distributions (D1) and (D3), the cost parameters are either i.i.d or have bounded support. In either case, $\lim_{n\to\infty} \frac{\alpha_i^{max}}{\sqrt{n}} = 0$ and $\lim_{n\to\infty} \frac{\beta_i^{max}}{\sqrt{n}} = 0$ w.p. 1. Thus, the result follows from Lemma 4.3. Similarly, we also have,

$$\frac{f_n^s(\underline{x}^s) - f_n^s(\overline{x}^s)}{n} \leq \sum_{i \in \mathcal{I}} \left[\frac{\beta_i^{max}}{n} + \left(\alpha_i^{max} \overline{U}_i + \beta_i^{max} \right) \frac{\sum_{t=1}^{n} \underline{Y}_{it}^s}{n} \right].$$

Under distribution (D2), the cost parameters have bounded support. Hence, $(\overline{U}_i \alpha_i^{max} + \beta_i^{max}) < +\infty$ and, for all i, $\beta_i^{max} < +\infty$ and are independent of n. Thus, the result follows from Lemma 4.3. $\square$

Proposition 4.5 *For any scenario s,*

$$\lim_{n\to\infty} \frac{f_n^s(\widehat{x}^s) - f_n^s(\underline{x}^s)}{n} = 0 \qquad w.p.\ 1.$$

Proof: Recall that in the capacity bundling phase we increase the capacity installed for time periods that do not satisfy the non-anticipativity restriction. For a facility i, let s_i' be the scenario that requires the largest number of capacity expansion sequences, *i.e.*, the worst-case scenario. Then, clearly, the increase in capacity for other scenarios in the capacity bundling phase will at most be that of the worst-case scenario. Thus

$$\frac{f_n^s(\widehat{x}^s) - f_n^s(\underline{x}^{(s)})}{n} \leq \sum_{i\in\mathcal{I}} \left[\left(\alpha_i^{max}\overline{U}_i + \beta_i^{max}\right) \frac{\sum_{t=1}^n \underline{Y}_{it}^{s_i'}}{n} \right].$$

Taking the limit and invoking Lemma 4.3 for s_i' completes the proof. □

For an instance of (SCAP) with n time periods, let z_n^{LP}, z_n^{IP}, and z_n^H denote the optimal value of LP relaxation solution obtained in Phase I, the optimal value of the integer program (SCAP), and the value of the heuristic solution obtained in Phase III, respectively. We now state the main results of this section.

Theorem 4.6

$$\lim_{n\to\infty} \frac{z_n^H - z_n^{IP}}{n} = 0 \qquad w.p.\ 1.$$

Proof: Note that $0 \leq z_n^H - z_n^{IP} \leq z_n^H - z_n^{LP}$. Thus,

$$\begin{aligned} \frac{z_n^H - z_n^{IP}}{n} &\leq \frac{\sum_{s=1}^S p^s[f_n^s(\widehat{x}^s) - f_n^s(\overline{x}^s)]}{n} \\ &= \sum_{s=1}^S p^s \left[\frac{f_n^s(\widehat{x}^s) - f_n^s(\underline{x}^s)}{n} + \frac{f_n^s(\underline{x}^s) - f_n^s(\overline{x}^s)}{n} \right]. \end{aligned}$$

Invoking Propositions 4.4 and 4.5 completes the proof. □

To characterize the asymptotic properties of the relative error of the heuristic solution, we make the following assumptions:

Assumption 4.7 *For any* s *and* t, $\delta^s_{ijt}/\mu_{ij} \geq 1$ *for all* $j \in \mathcal{J}_i$ *and* $i \in \mathcal{I}$.

Assumption 4.8 *For any* s *and* t, $d^s_{jt} \geq 1$ *for at least one* $j \in \mathcal{J}$.

The quantity δ^s_{ijt}/μ_{ij} can be interpreted as the unit production cost of family j from technology i in period t and scenario s. Then, Assumption 4.7 states that, in each scenario, the unit production cost of a product family in any period is at least 1. Similarly, Assumption 4.8 states that in each scenario and in each period there is unit demand of at least one of the product families. These assumptions are satisfied by appropriate scaling for positive lower bounds on the production costs and demands.

Corollary 4.9 *Under Assumptions 4.7 and 4.8 and the specified probability model,*

$$\lim_{n\to\infty} \frac{z^H_n - z^{IP}_n}{z^{IP}_n} = 0 \quad \textit{w.p. 1.}$$

Proof: In light of Theorem 4.6, it suffices to show that z^{IP}_n is $\Omega(n)$, *i.e.*, there exist positive constants C and n_0 such that $z^{IP}_n \geq Cn$ for all $n \geq n_0$.

First, let us rewrite the non-anticipativity constraints (7), (8), and (9) in (SCAP) as follows:

$$\textstyle\sum_{s=1}^S L^s_{it} X^s_{it} = 0 \quad t = 1, \ldots, n; \quad i \in \mathcal{I} \tag{13}$$

$$\textstyle\sum_{s=1}^S M^s_{it} Y^s_{it} = 0 \quad t = 1, \ldots, n; \quad i \in \mathcal{I} \tag{14}$$

$$\textstyle\sum_{s=1}^S N^s_{ijt} W^s_{ijt} = 0 \quad t = 1, \ldots, n; \quad i \in \mathcal{I}; \quad j \in \mathcal{J} \tag{15}$$

where the matrices L^s_{it}, M^s_{it}, and N^s_{ijt} appropriately defined. Now consider the dual to the LP relaxation of (SCAP):

$$
\begin{array}{lrl}
z_n^D = & \max & \sum_{s=1}^S \sum_{t=1}^n \left[\left(\sum_{i \in \mathcal{I}} X_{i0} \gamma^s_{it} + \xi^s_{it} \right) + \sum_{j \in \mathcal{J}} d^s_{jt} \eta^s_{jt} \right] \\
\text{s.t.} & \lambda^s_{it} + L^s_{it} u_{it} - \sum_{\tau=t}^{n} \gamma_{i\tau} \le p^s \alpha^s_{it} & t = 1, \ldots, n; \quad s = 1, \ldots, S; \quad i \in \mathcal{I} \\
& -U_i \lambda^s_{it} + M^s_{it} v_{it} + \xi^s_{it} \le p^s \beta^s_{it} & t = 1, \ldots, n; \quad s = 1, \ldots, S; \quad i \in \mathcal{I} \\
& \gamma^s_{it} + N^s_{ijt} w_{ijt} + \mu_{ij} \eta_{jt} \le p^s \delta^s_{ijt} & t = 1, \ldots, n; \quad s = 1, \ldots, S; \quad i \in \mathcal{I}; \quad j \in \mathcal{J}_i \\
& \lambda^s_{it}, \gamma^s_{it}, \xi^s_{it} \le 0 & t = 1, \ldots, n; \quad s = 1, \ldots, S; \quad i \in \mathcal{I} \\
& \eta^s_{jt}, u_{it}, v_{it}, w_{ijt} \text{unrestricted} & t = 1, \ldots, n; \quad s = 1, \ldots, S; \quad i \in \mathcal{I}; \quad j \in \mathcal{J}
\end{array}
$$

where λ^s_{it}, γ^s_{it}, and η^s_{jt} are the dual variables corresponding to constraints (2), (3), and (4), respectively; ξ^s_{it} are the dual variables corresponding to the LP relaxation of (6); and u_{it}, v_{it}, and w_{ijt} are the dual variables corresponding to the non-anticipativity constraints (13), (14), and (15), respectively. Consider the following feasible solution to the above dual problem: $\lambda^s_{it} = 0$, $\xi^s_{it} = 0$, $\gamma^s_{it} = 0$, $u_{it} = 0$, $v_{it} = 0$, $w_{ijt} = 0$, and $\eta^s_{jt} = \min_{i \in \mathcal{I}_j} \{ p^s \delta^s_{ijt} / \mu_{ij} \}$. Then, under Assumptions 4.7 and 4.8, clearly z_n^D is $\Omega(n)$. Since the problem instances are assumed to be feasible, we also have that z_n^{IP} is $\Omega(n)$. □

5 Capacity Expansion of Chemical Processing Networks under Uncertainty

In this section, we demonstrate the asymptotic convergence behavior of the proposed heuristic in the context of capacity expansion of chemical processing networks under uncertainty. Given a potential network consisting of a set of processes interconnected by a set of chemicals, the problem consists of (i) selecting processes from among competing technologies, (ii) timing and sizing process expansions, (iii) determining the optimal production levels for the installed processes. The objective is to minimize the discounted cost of the entire processing network over a long-range horizon. A detailed description of the deterministic version of this problem appears in Sahinidis, Grossmann, Fornari & Chathrathi (1989). Here, we use (SCAP) to model the stochastic version.

We applied the proposed heuristic on randomly generated problem instances from a set of four basic process networks. The first of these networks (see Figure 3) is from Ahmed & Sahinidis (1998) while the next two are from Liu (1995). These networks involve three, four, and four processes (square nodes in the figures), and four, five, and five chemicals (circular nodes). The fourth basic network is that of an industrial petrochemical processing chain with 38 processes and 24 chemicals and is described in Sahinidis et al. (1989). For the first three networks, we report results with planning horizons ranging from 2 to 10 time periods in unit increments. For the fourth network, we present results with up to 8 time periods. Parameters for the problem instances were generated randomly according to the probability model described in Section 4. The uncertainty was modeled as a binary tree with a total of 2^{n-1} scenarios for a problem with n time periods. The sizes of the deterministic equivalent of the instances with n time periods are presented in Table 1. For each network, 5 instances were randomly generated corresponding to each planning horizon. Thus, the entire problem set consisted of 170 problem instances.

For all networks, the proposed heuristic was compared against solving the deterministic equivalent integer program using state-of-the-art integer programming techniques. CPLEX 7.0 (CPLEX 2000) was used with default strategies to solve the linear and integer programs on an IBM RISC System/6000 Model-43P machine with 128MB of memory. Most of the generated integer problems are not solvable within reasonable computing times with CPLEX unless cutting planes and extensive problem preprocessing are used.

Figure 3 presents the network structure and the per-period heuristic error as a function of the number of time periods for the first three examples. The asymptotic convergence of the normalized error as proved in Theorem 4.6 is clearly observed in all cases. This is more profound in Figure 4 where the percentage relative error of the heuristic solution is plotted against the problem size. It can be observed that, for large problems in this set, the heuristic provided solutions which are within an average of 1% of optimality. The largest problem in this set involved $36,858$ binary variables, $184,291$ continuous variables, and $368,641$ constraints. This problem was solved by the heuristic to within 1% of optimality in 13 CPU minutes. Exact solution of these problems required up to 14 times more than the time required by the heuristic. The only exception was Network 1 with $n = 10$ for which the IP was not solved after 500 CPU minutes and 186,000 branch and bound nodes. On the other hand, the heuristic took only 3.5 minutes to solve this problem with a gap from the LP relaxation of only 0.57%.

The per-period error of the heuristic solution for the industrial scale processing Network 4 is presented in Figure 5. The asymptotic convergence of the optimality gap is observed in this case as well. The largest problem in this set that was solved by the IP solver involved $n = 7$ time periods. The problem with $n = 8$ required $48,602$ binary variables, $217,431$ continuous variables, and $415,745$ constraints. This problem was not solved by the IP solver even after 12 CPU hours and 64,500 branch and bound nodes. On the other hand, the problem was solved by the heuristic with a 10.7% gap from the LP relaxation value in less than 16 CPU minutes – the major portion of which was the time required for solving the LP relaxation. It should be noted that the gap of 10.7% is with respect to the LP relaxation value. Our experience from the smaller ($n \leq 7$) problem instances suggests that the true optimality gap is significantly smaller.

Finally, we compare the quality of simple LP rounding solution to that of the heuristic solutions. Recall that simple rounding up of the values of the binary variables in the corresponding LP relaxation solution produces an integer feasible solution. Figure 7 presents the normalized errors for simple LP rounding and that of the proposed heuristic for Network 3. Even though the normalized errors for LP rounding do show a decreasing trend with the number of time periods, these errors are several orders of magnitude higher than the errors for the heuristic solutions, and therefore of little practical value. Also, recall that the example in Section 3.2 demonstrates that the normalized error of LP rounding will not decrease in general.

Acknowledgments

Partial financial support was received from the IBM Research Division, the Mitsubishi Chemical Corporation, and the National Science Foundation under awards DMI 95-02722 and CTS 97-04643. The authors also thank two anonymous referees and the associate editor for helpful comments.

References

Ahmed, S. & Sahinidis, N. (1998). Robust process planning under uncertainty. *Industrial & Engineering Chemistry Research*, Vol. 37, pp. 1883–1892.

Ahmed, S. & Sahinidis, N. V. (2000a). Analytical investigations of the process planning problem. *Computers & Chemical Engineering*, Vol. 5, pp. 1605–1621.

Ahmed, S. & Sahinidis, N. V. (2000b). Selection, acquisition, and allocation of manufacturing technology in a multi-product environment. Submitted to *Management Science*.

Bean, J. C., Higle, J. L. & Smith, R. L. (1992). Capacity expansion under stochastic demands. *Operations Research*, Vol. 40, pp. S210–S216.

Berman, O. & Ganz, Z. (1994). The capacity expansion problem in the service industry. *Computers & Operations Research*, Vol. 21, pp. 557–572.

Berman, O., Ganz, Z. & Wagner, J. M. (1994). A stochastic optimization model for planning capacity expansion in a service industry under uncertian demand. *Naval Research Logistics*, Vol. 41, pp. 545–564.

Bermon, S. & Hood, S. (1999). Capacity optimization planning system (CAPS). *Interfaces*, Vol. 29, pp. 31–50.

Birge, J. R. (1985). Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, Vol. 33, pp. 989–1007.

Birge, J. R. & Louveaux, F. (1997). *Introduction to Stochastic Programming*. Springer, New York, NY.

Chang, S.-G. & Gavish, B. (1993). Telecommunications network topological design and capacity expansion: Formulations and algorithms. *Telecommunication Systems*, Vol. 1, pp. 99–131.

Chang, S.-G. & Gavish, B. (1995). Lower bounding procedures for multiperiod telecommunications network expansion problems. *Operations Research*, Vol. 43, pp. 43–57.

Chen, Z.-L., Li, S. & Tirupati, D. (2001). A scenario based stochastic programming approach for technology and capacity planning. To appear in *Computers & Operations Research*.

CPLEX (2000). *CPLEX 7.0 User's Manual*. ILOG CPLEX Division. Incline Village, NV.

David, M. H. A., Dempster, M. A. H., Sethi, S. P. & Vermes, D. (1987). Optimal capacity expansion under uncertainty. *Advances in Applied Probability*, Vol. 19, pp. 156–176.

Dupačová, J., Consigli, G. & Wallace, S. W. (2000). Scenarios for multi-stage stochastic programs. *Annals of Operations Research*, Vol. 100, pp. 25–53.

Eppen, G. D., Martin, R. K. & Schrage, L. (1989). A scenario approach to capacity planning. *Operations Research*, Vol. 37, pp. 517–527.

Fine, C. H. & Freund, R. M. (1990). Optimal investment in product-flexible manufacturing capacity. *Management Science*, Vol. 36, pp. 449–466.

Fong, C. O. & Srinivasan, V. (1981a). The multiregion dynamic capacity expansion problem: Part I. *Operations Research*, Vol. 29, pp. 787–799.

Fong, C. O. & Srinivasan, V. (1981b). The multiregion dynamic capacity expansion problem: Part II. *Operations Research*, Vol. 29, pp. 800–816.

Freidenfelds, J. (1980). Capacity expansion when demand is a birth-death random process. *Operations Research*, Vol. 28, pp. 712–721.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*. Robert E. Krieger Publishing Co., Florida.

IBM Corporation (1998). IBM Stochastic Solutions.
URL: `http://www6.software.ibm.com/es/oslv2/features/stoch.htm`.

Kall, P. & Wallace, S. W. (1994). *Stochastic Programming*. John Wiley and Sons, Chichester, England.

Klincewicz, J. G., Luss, H. & Yu, C.-S. (1988). A large-scale multilocation capacity planning model. *European Journal of Operational Research*, Vol. 34, pp. 178–190.

Laguna, M. (1998). Applying robust optimization to capacity expansion of one location in telecommunications with demand uncertainty. *Management Science*, Vol. 44, pp. S101–S110.

Leadbetter, M. R., Lindgren, G. & Rootźen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer.

Li, S. & Tirupati, D. (1994). Dynamic capacity expansion problem with multiple products: technology selection and timing of capacity additions. *Operations Research*, Vol. 42, pp. 958–976.

Liu, M. L. (1995). *Optimization Tools for Process Planning*. PhD thesis. University of Illinois at Urbana-Champaign.

Liu, M. L. & Sahinidis, N. V. (1995). Computational trends and effects of approximations in an MILP model for process planning. *Industrial & Engineering Chemistry Research*, Vol. 34, pp. 1662–1673.

Liu, M. L. & Sahinidis, N. V. (1997). Bridging the gap between heuristics and optimization: The capacity expansion case. *AIChE Journal*, Vol. 43, pp. 2289–2299.

Manne, A. S. (1961). Capacity expansion and probabilistic growth. *Econometrica*, Vol. 29, pp. 632–649.

Murphy, F. H. & Weiss, H. J. (1990). An approach to modeling electric utility capacity expansion planning. *Naval Research Logistics*, Vol. 37, pp. 827–845.

Murphy, F. H., Sen, S. & Soyster, A. L. (1987). Electric utility expansion planning in the presence of existing capacity: A nondifferentiable, convex programming approach. *Computers & Operations Research*, Vol. 14, pp. 19–31.

Rajagopalan, S. (1994). Capacity expansion with alternative technology choices. *European Journal of Operational Research*, Vol. 77, pp. 392–403.

Rajagopalan, S., Singh, M. R. & Morton, T. E. (1998). Capacity expansion and replacement in growing markets with uncertain technological breakthroughs. *Management Science*, Vol. 44, pp. 12–30.

Rockafellar, R. T. & Wets, R. J.-B. (1991). Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, Vol. 16,(1) pp. 119–147.

Sahinidis, N. V. & Grossmann, I. E. (1992). Reformulation of the multiperiod MILP model for capacity expansion of chemical processes. *Operations Research*, Vol. 40, Supp. No. 1, pp. S127–S144.

Sahinidis, N. V., Grossmann, I. E., Fornari, R. E. & Chathrathi, M. (1989). Optimization model for long range planning in the chemical industry. *Computers & Chemical Engineering*, Vol. 13, pp. 1049–1063.

Saniee, I. (1995). An efficient algorithm for the multiperiod capacity expansion of one location in telecommunications. *Operations Research*, Vol. 43, pp. 187–190.

Swaminathan, J. M. (2000). Tool capacity planning for semiconductor fabrication facilities under demand uncertainty. *European Journal of Operational Research*, Vol. 120, pp. 545–558.

Problem	Number of Processes	Number of Chemicals	Binary Variables	Continuous Variables	Constraints
Network 1	3	4	$3n2^{n-1}$	$14n2^{n-1}$	$10n2^{n-1} + 14(n2^{n-1} - 2^n + 1)$
Network 2	4	5	$4n2^{n-1}$	$18n2^{n-1}$	$13n2^{n-1} + 18(n2^{n-1} - 2^n + 1)$
Network 3	5	6	$5n2^{n-1}$	$22n2^{n-1}$	$16n2^{n-1} + 22(n2^{n-1} - 2^n + 1)$
Network 4	38	24	$38n2^{n-1}$	$132n2^{n-1}$	$104n2^{n-1} + 132(n2^{n-1} - 2^n + 1)$

Table 1: Dimensions of the problem instances

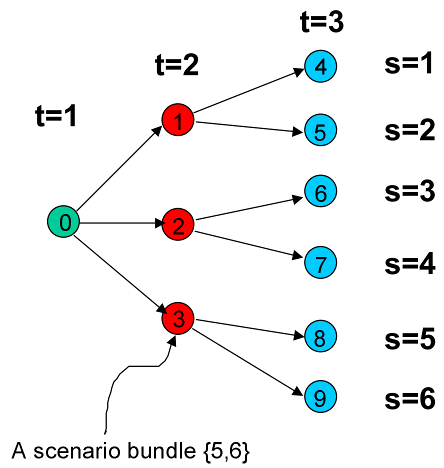

Figure 1: Scenario tree used to model uncertainty

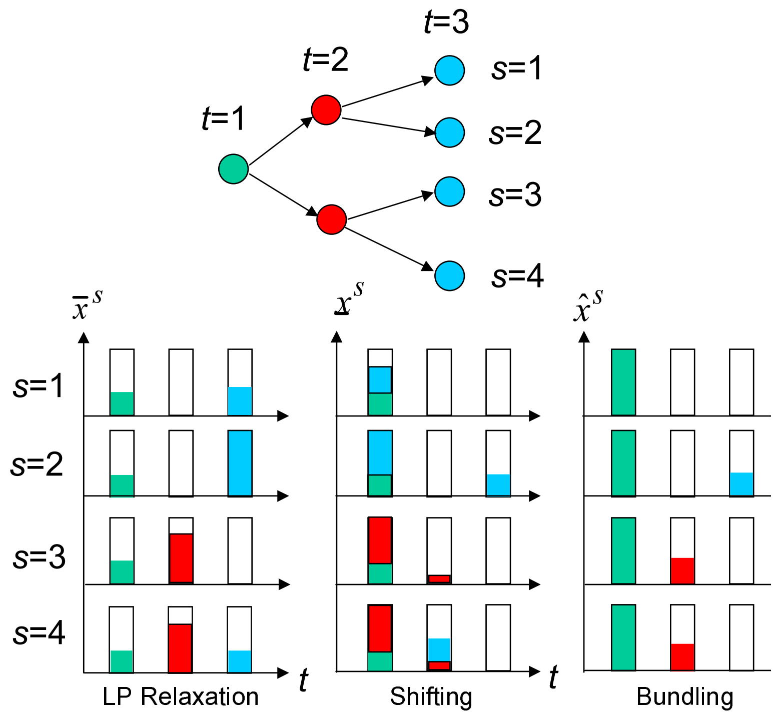

Figure 2: The heuristic strategy

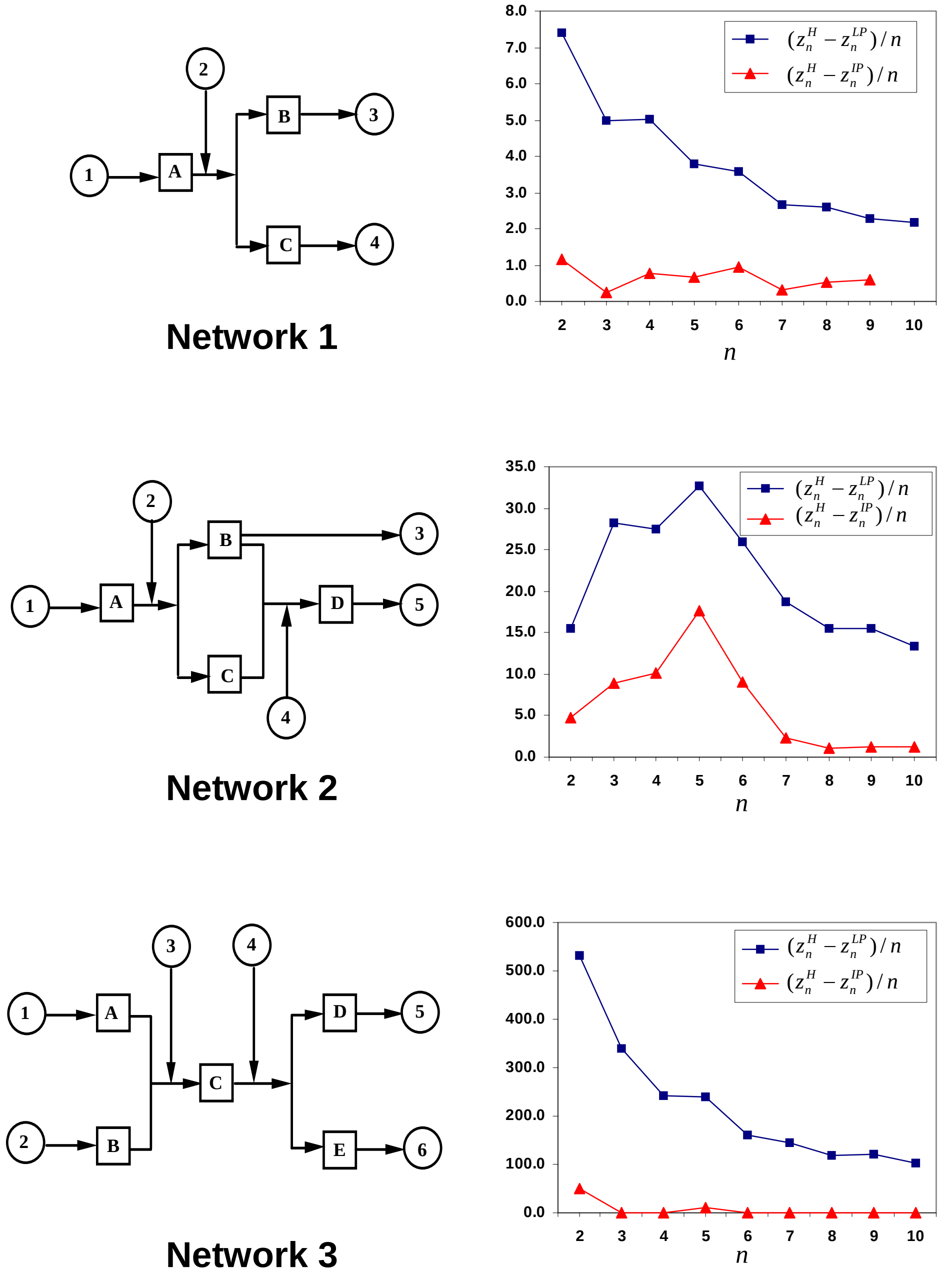

Figure 3: Results for Networks 1, 2, and 3

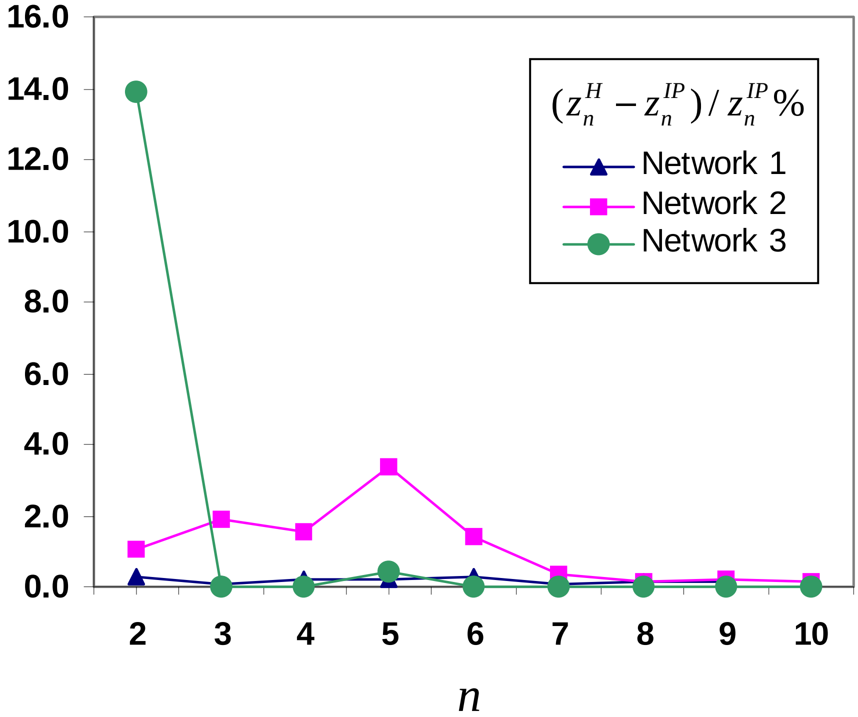

Figure 4: Relative errors for Networks 1, 2, and 3

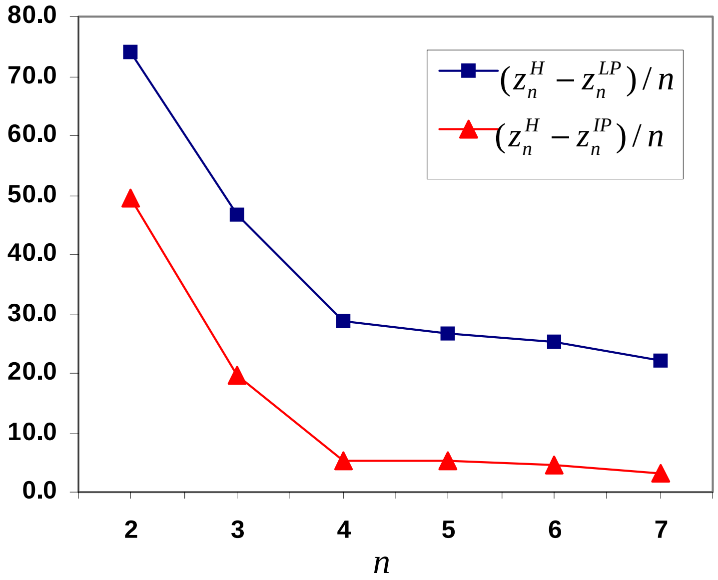

Figure 5: Heuristic error bounds for Network 4

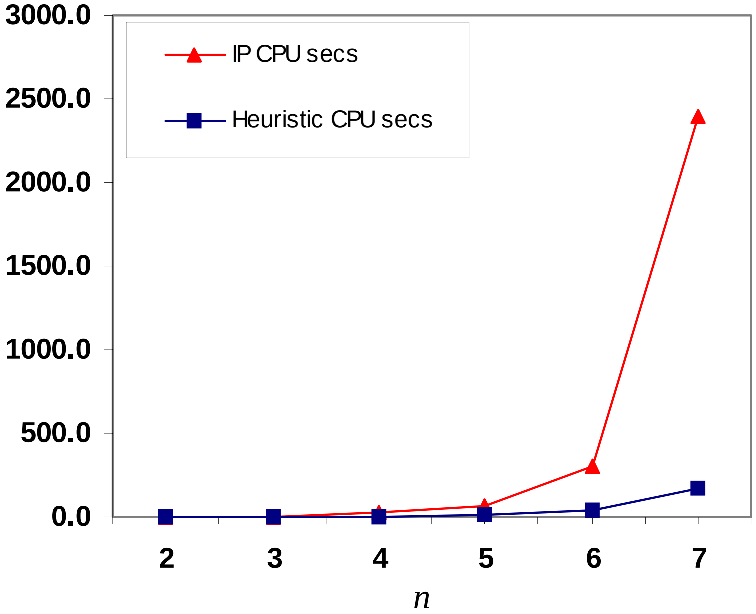

Figure 6: Solution times for Network 4

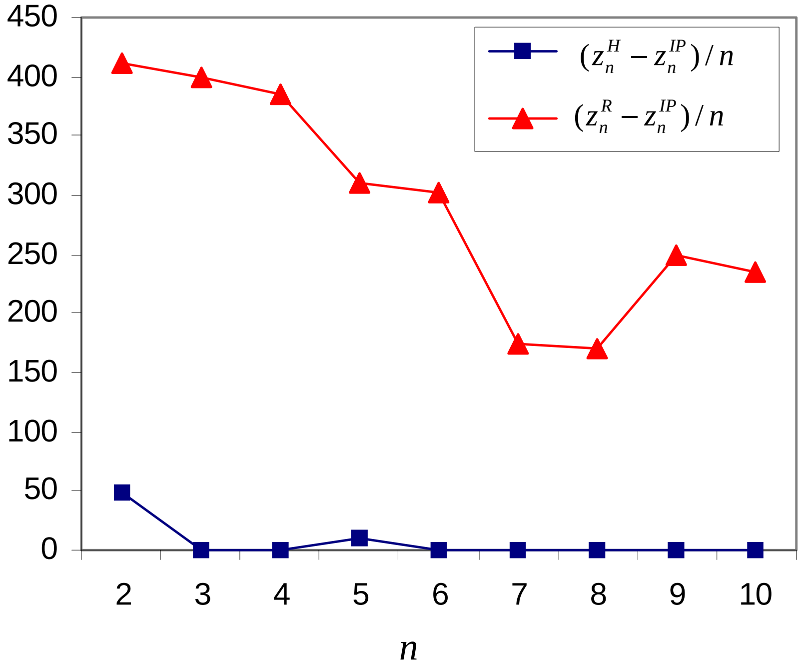

Figure 7: Rounding vs. Heuristic Errors for Network 3