# New Valid Inequalities and Formulation for the Static Chance-constrained Lot-Sizing Problem *

Zeyang Zhang, Chuanhou Gao
School of Mathematical Sciences, Zhejiang University, Hangzhou, China
zy_zhang@zju.edu.cn, gaochou@zju.edu.cn

James Luedtke
Department of Industrial & Systems Engineering
Wisconsin Institute for Discovery
University of Wisconsin-Madison, Madison, WI, USA
jim.luedtke@wisc.edu

June 17, 2021

**Abstract**

We study the static chance-constrained lot sizing problem, in which production decisions over a planning horizon are made before knowing random future demands, and the backlog and inventory variables are then determined by the demand realizations. The chance constraint imposes a service level constraint requiring that the probability that any backlogging is required should be below a given threshold. We model uncertain outcomes with a finite set of scenarios, and begin by applying existing results about chance-constrained programming to obtain an initial extended mixed-integer programming formulation. We further strengthen this formulation with a new class of valid inequalities that generalizes the classical $(\ell, S)$ inequalities for the deterministic uncapacitated lot sizing problem. In addition, we prove an optimality condition of the solutions under a modified Wagner-Whitin condition, and based on this derive a new extended mixed-integer programming formulation. We also discuss how our model and methods can be extended to a model in which the time horizon is split into two parts, where demands are known in the first part and random in the latter part. We conduct a thorough computational study demonstrating the effectiveness of the new valid inequalities and extended formulation.

## 1 Introduction

The lot-sizing problem is a classic production planning problem in which production and inventory levels are planned over a finite set of discrete time periods. In the deterministic uncapacitated lot-sizing problem (ULS) (without backlogging) [22], the problem is to determine a production plan for a product to satisfy demands over a finite time horizon while minimizing the sum of setup, production, and inventory holding costs. In the ULS problem, the demand in every period is assumed to be known (deterministic). However, in many realistic settings future demand is predicted by a forecast which inevitably has errors, and the actual realized demands may therefore be modeled as random variables. We consider the static stochastic lot-sizing (SLS) problem, in which it is assumed that all the production amounts are chosen before observing any demand realizations, and the inventory levels over time adjust to random demands. We use a chance constraint to require that the probability that the chosen production schedule meets all demands over time is at least $1 - \epsilon$, where $\epsilon \in (0, 1)$ is a given risk tolerance.

Chance-constrained programming (CCP) dates back to [8, 9]. The CCP formulation that arises in the SLS problem is a special case of the CCP problem with stochastic right-hand side under a finite discrete

*Z. Zhang and C. Gao thank the financial support from the National Nature Science Foundation of China under Grant No. 11671418 and 12071428, and Zhejiang Provincial Natural Science Foundation of China under Grant No. LZ20A010002.

distribution. This special case has been studied extensively in the literature on CCP. Given the discrete scenarios, a deterministic equivalent formulation using binary variables to indicate for which scenarios the constraints are satisfied can be constructed. In [18] it was observed that this formulation can be strengthened using mixing inequalities [5, 12]. Further valid inequalities were investigated in [1, 15, 24].

For the deterministic ULS problem, an explicit convex hull description is given by [6] utilizing the so-called $(\ell, S)$ inequalities. The first polyhedral study of the deterministic ULS problem with backlogging (ULSB) is performed by [19], in which the authors reformulate the structure of the original formulation by several methods to obtain extended formulations. The complete linear description of the convex hull of ULSB is provided by [16] by generalizing the valid inequalities of [19]. In addition, [20] conduct a polyhedral study of the lot-sizing problem under several different conditions with Wagner-Whitin costs.

When considering the uncertainty of demands, and penalizing the expected cost of shortages, a stochastic uncapacitated lot-sizing problem (SULS) is proposed. In [2] and [3] the stochastic capacity expansion problem is studied, which includes SULS as a submodel. A polyhedral study of the SULS problem based on a scenario tree is conducted in [11]. They provide several kinds of valid inequalities, and give a sufficient condition under which those inequalities are facet-defining. Afterwards, [10] propose an efficient dynamic programming algorithm for SULS, and similar algorithms can be generalized to SULS with random lead times ([13] and [14]). None of these papers consider chance constraints.

The static chance-constrained lot-sizing problem was first proposed in [7] as an application of general CCP. However, they did not consider the inventory cost in the objective function. In [1, 15, 24] a model that includes inventory cost is solved by a branch-and-cut algorithm, but the model in these works is used as a test case for general-purpose methods without investigating the particular structure of the SLS. A dynamic variant of this problem which updates the production schedule after the scenario realization of the former time periods is studied in [23]. The work [17] is the most closely related to our work. They provide the first polyhedral study exploiting the lot-sizing structure to identify valid inequalities for the SLS.

In our paper, we investigate valid inequalities for SLS that exploit the structure of both CCP and the lot-sizing problem. The model we investigate is a slight extension of the standard SLS model, in which the time horizon is split into two phases. In the first phase, the demands are assumed to be known, whereas the demands are random in the second phase. The SLS is a special case of our model in which the first phase has no periods. We derive an initial strong extended formulation using the CCP results in [18], which we call E-SLS. Next, we propose a new class of valid inequalities, the CC-$(\ell, S)$ inequalities, to strengthen the formulation E-SLS by exploiting the characterizations of lot-sizing problem. We also derive a property of optimal solutions under a modified Wagner-Whitin condition, and use this to construct a new extended formulation NE-SLS, which is valid under this condition. We conduct a computational study to compare the performance between the valid inequalities in [17] and our CC-$(\ell, S)$ inequalities, and find that our inequalities lead to significantly better relaxations which translates into faster solve times. Additional experiments demonstrate the potential value of the new extended formulation in some cases.

This paper is organized as follows. In Sect. 2.1, we write the original mathematical formulation for SLS. In Sect. 2.2, we construct a stronger extended formulation E-SLS. In Sect. 2.3, we present our new CC-$(\ell, S)$ inequalities. In Sect. 3, we derive the new extended formulation NE-SLS. For simplicity the results in Sect. 2.3 and 3 are presented for the classical SLS problem in which demands are uncertain in all time periods. In Sect. 4, we present the generalization of these results to our model in which some demands are known in an initial set of time periods. We present results of our computational study in Sect. 5 and make concluding remarks in Sect. 6.

**Notation.** For integers $a \leq b$, we define $[a, b] = \{a, a+1, \ldots, b-1, b\}$. For a set $Y \subseteq \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p\}$ we define $\text{Proj}_x(Y) = \{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^p, (x, y) \in Y\}$.

# 2 Formulations

## 2.1 SLS Formulation

We consider a planning horizon with length $T$. We assume that the demands in periods $i \in [1, p]$ are known values $d_i$, where $p \in [0, T]$. We assume the demand in period $i \in [p+1, T]$ is a random variable $\xi_i$ and the joint distribution of $\xi = (\xi_{p+1}, \ldots, \xi_T)$ has known finite support. The main decision variables are the production levels $x_t$ for $t \in [1, T]$ and the production setup variables $y_t$ for $t \in [1, T]$, where $y_t = 1$ indicates a setup is done in period $t$ and $y_t = 0$ otherwise. We assume all these decisions must be made at the beginning of the planning horizon before observing the values of the random demands. In addition, the model keeps track of the inventory levels in the initial time periods. For periods $t \in [1,p]$, $s_t$ represents the inventory level at the end of period $t$, and let $s^1 = (s_1, \ldots, s_p)$. Let $\alpha = (\alpha_1, \ldots, \alpha_T)$ and $\beta = (\beta_1, \ldots, \beta_T)$ be the unit production cost vector and fixed setup cost vector, respectively; $h^1 = (h_1, \ldots, h_p)$ be the holding cost vector for the first $p$ periods; and $\varepsilon \in (0,1)$ be the given risk tolerance. With initial inventory $s_0 = 0$, the corresponding chance-constrained stochastic lot-sizing (SLS) problem is formulated as follows:

$$\min \quad \alpha^\top x + \beta^\top y + {h^1}^\top s^1 + \mathbb{E}_\xi(\Theta(x^2, \xi)) \tag{1}$$

$$x_i + s_{i-1} = d_i + s_i, \quad i \in [1,p] \tag{2}$$

$$\mathbb{P}\Big(s_p + \sum_{t=p+1}^{i} x_t \geq \sum_{t=p+1}^{i} \xi_t,\ i \in [p+1,T]\Big) \geq 1 - \varepsilon \tag{3}$$

$$x_i \leq M_i y_i, \quad i \in [1,T] \tag{4}$$

$$x \in \mathbb{R}_+^T, \quad y \in \{0,1\}^T, \quad s^1 \in \mathbb{R}_+^p, \tag{5}$$

Constraints (2) are the relations among production, inventory and demand in the first $p$ periods. Constraints (3) ensure that the probability of meeting all demands from periods $p+1$ to $T$ is at least $1-\epsilon$. The constraints (4) ensure that production is zero in periods where no setup is done; in these constraints, for each $i \in [1,T]$ the constant $M_i$ is large enough so that when $y_i = 1$ the corresponding constraint is redundant. For a given $x^2 = (x_{p+1}, \ldots, x_T)$ and observation of the random demands $\xi$, $\Theta(x^2, \xi)$ calculates the holding costs over the periods $p+1, \ldots, T$:

$$\Theta(x^2, \xi) = \min \quad (h^2)^\top s^2$$

$$s_i \geq s_p + \sum_{t=p+1}^{i} (x_t - \xi_t), \quad i \in [p+1,T] \tag{6}$$

$$s^2 \in \mathbb{R}_+^{T-p}. \tag{7}$$

Here $s^2 = (s_{p+1}, \ldots, s_T)$ is the vector of inventory variables in periods $[p+1,T]$ and $h^2 = (h_{p+1}, \cdots, h_T)$ is the vector of (nonnegative) holding costs. Constraints (6) and (7) ensure that $s_i$ is equal to the inventory level in period $i$, if it is positive, and $s_i$ is equal to zero in period $i$, otherwise.

Let $\Omega = \{1, \ldots, m\}$ be the index set of demand scenarios and let $\pi_j$ be the probability of scenario $j$, for all $j \in \Omega$. In addition, let $d_{ji}$ be the demand for period $i$ under scenario $j$, for all $i \in [p+1,T]$ and $j \in \Omega$ (i.e., $\xi_i = d_{ji}$ under scenario $j \in \Omega$).

Under the given finite scenario model of the random demands, problem (1)-(5) can be reformulated as an explicit deterministic mixed-integer program (refer to [17]). Specifically, for each scenario $j \in \Omega$ let $z_j$ be a binary variable, which equals to 0 if the demand in all time periods is satisfied under scenario $j$, and 1 otherwise. In addition, let $s_{ji}$ be the inventory at the end of time period $i \in [p+1,T]$ in scenario $j \in \Omega$. Then the deterministic equivalent formulation is:

$$\min \quad \alpha^\top x + \beta^\top y + (h^1)^\top s^1 + \sum_{j=1}^{m} \pi_j (h^2)^\top s_j^2 \tag{8}$$

$$x_i + s_{i-1} = d_i + s_i, \quad i \in [1,p] \tag{9}$$

$$s_p + \sum_{t=p+1}^{i} x_t \geq \sum_{t=p+1}^{i} d_{jt}(1 - z_j), \quad i \in [p+1,T], \quad j \in \Omega \tag{10}$$

$$\sum_{j=1}^{m} \pi_j z_j \leq \varepsilon \tag{11}$$

$$s_{ji} \geq \sum_{t=p+1}^{i} (x_t - d_{jt}) + s_p, \quad i \in [p+1,T], \quad j \in \Omega \tag{12}$$

$$x_i \le M_i y_i, \quad i \in [1,T] \tag{13}$$

$$x \in \mathbb{R}^T_+, \quad y \in \{0,1\}^T, \quad s^1 \in \mathbb{R}^p_+, \quad z \in \{0,1\}^m \tag{14}$$

$$s^2_j \in \mathbb{R}^{T-p}_+, \quad j \in \Omega \tag{15}$$

With the given scenario model on demands, a large enough value for $M_i$ can be calculated as $M_i = \sum_{t=i}^{p} d_t + \max_{j\in\Omega}\{\sum_{t=p+1}^{T} d_{jt}\}$ for $i \in [1,p]$, and $M_i = \max_{j\in\Omega}\{\sum_{t=i}^{T} d_{jt}\}$, for $i \in [p+1,T]$.

In the following sections we explore how this formulation can be improved so that the linear programming (LP) relaxation is a closer approximation to the optimal value. In Sect. 2.2, we exploit the chance constraint structure to obtain a stronger extended formulation. In Sect. 2.3, we further use the lot-sizing structure of the problem to propose a class of valid inequalities.

## 2.2 An Extended Formulation of SLS

Let $P$ be the feasible region of SLS, i.e. $P = \{(x, s^1, s^2, y, z) \in \mathbb{R}^{T+p+m(T-p)}_+ \times \{0,1\}^{T+m} : (9)-(15)\}$, where $s^2 = (s^2_1, \ldots, s^2_m)$. Define $D_{ji} = \sum_{t=p+1}^{i} d_{jt}$, for every $i \in [p+1,T]$ and $j \in \Omega$, i.e., $D_{ji}$ is the cumulative demand from period $(p+1)$ to period $i$ in scenario $j$. For each $i \in [p+1,T]$, let $\{\sigma^i_1, \sigma^i_2 \ldots, \sigma^i_m\}$ be a permutation of index set $\Omega$, which satisfies to $D_{\sigma^i_1 i} \ge D_{\sigma^i_2 i} \ge \cdots \ge D_{\sigma^i_m i}$. As a notation simplification, we define $D_{\sigma^i_j} = D_{\sigma^i_j i}$, for each $i \in [p+1,T]$ and $j \in \Omega$. Furthermore, let $q^*_i = \min\{q \in \Omega : \sum_{j=1}^{q} \pi_{\sigma^i_j} > \epsilon\}$. Next, introducing additional binary variables $w^i_j$ for $j = 1,\ldots,q^*_i$ and $i \in [p+1,T]$ we construct an extended formulation for SLS by the methods in [18] as the following:

$$\min \quad \alpha^\top x + \beta^\top y + (h^1)^\top s^1 + \sum_{j=1}^{m} \pi_j (h^2)^\top s^2_j$$

$$x_i + s_{i-1} = d_i + s_i, \quad i \in [1,p] \tag{16}$$

$$s_p + \sum_{t=p+1}^{i} x_t + \sum_{j=1}^{q^*_i-1} (D_{\sigma^i_j} - D_{\sigma^i_{j+1}}) w^i_j \ge D_{\sigma^i_1}, \quad i \in [p+1,T] \tag{17}$$

$$w^i_j - w^i_{j+1} \ge 0, \quad j \in [1, q^*_i - 2], \quad i \in [p+1,T] \tag{18}$$

$$z_{\sigma^i_j} - w^i_j \ge 0, \quad j \in [1, q^*_i - 1], \quad i \in [p+1,T] \tag{19}$$

$$\sum_{j=1}^{m} \pi_j z_j \le \varepsilon \tag{20}$$

$$s_{ji} \ge s_p + \sum_{t=p+1}^{i} x_t - D_{ji}, \quad i \in [p+1,T], \quad j \in \Omega \tag{21}$$

$$x_i \le M_i y_i, \quad i \in [1,T] \tag{22}$$

$$x \in \mathbb{R}^T_+,\ y \in \{0,1\}^T,\ s^1 \in \mathbb{R}^p_+,\ z \in \{0,1\}^m, \tag{23}$$

$$s^2_j \in \mathbb{R}^{T-p}_+,\ j \in \Omega,\ w^i \in \{0,1\}^{q^*_i-1},\ i \in [p+1,T] \tag{24}$$

We denote the new extended formulation as E-SLS, in which constraints (17)-(19) play the same role as constraint (10) in SLS. Let $Q$ be the feasible region of E-SLS, i.e.

$$Q = \{(x, s^1, s^2, y, z, w) : (16) - (24)\},$$

then the results in [18] imply $\mathrm{Proj}_{(x,s^1,s^2,y,z)} Q = P$, and therefore, E-SLS is an valid model for SLS. In [18] it is demonstrated that for chance-constrained *linear* programs the LP relaxation of the extended formulation derived this way can be significantly closer to the optimal value than the LP relaxation of the original formulation. We find similar improvement from using E-SLS in our results, but given the presence of the binary setup variables, we explore how to further strengthen the LP relaxation of E-SLS in the next subsection.

### 2.3 CC-$(\ell, S)$ Inequalities for E-SLS

In the deterministic uncapacitated lot-sizing problem, the class of inequalities known as the $(\ell, S)$ inequalities are sufficient to give the complete linear description of convex hull. In this section, we propose an extension of the $(\ell, S)$ inequalities to the static chance-constrained lot-sizing problem. For simplicity of exposition, we assume in this section that $p = 0$ (i.e., the initial time phase where demands are known with certainty is empty), and remind the reader of our assumption that $s_0 = 0$. (Extension of all results to $s_0 > 0$ is trivial.) We present the generalization of our results to the case $p > 0$ in Sect. 4.

Let $D^j_{i\ell} = \sum_{t=i}^{\ell} d_{jt}$ for $1 \le i \le \ell \le T$, $j \in \Omega$ denote the cumulative demand from period $i$ to period $\ell$ in scenario $j$, and let $\bar{D}_{i\ell} = \max_{j\in\Omega}\{D^j_{i\ell}\}$ be the maximum cumulative demand from period $i$ to period $\ell$ over all scenarios. Then we define the CC-$(\ell, S)$ inequalities as the following.

**Definition 1.** *For each $\ell \in [1, T]$ and $S \subseteq [1, \ell]$, the inequality*

$$\sum_{t\in\bar{S}} x_t + \sum_{t\in S} \bar{D}_{t\ell} y_t + \sum_{j=1}^{q^*_\ell - 1} (D_{\sigma^\ell_j} - D_{\sigma^\ell_{j+1}}) w^\ell_j \ge D_{\sigma^\ell_1} \tag{25}$$

*is called a CC-$(\ell, S)$ inequality for E-SLS, where $\bar{S} = [1, \ell] \setminus S$.*

We define E-SLS+CC-$(\ell, S)$ to be the formulation defined by the E-SLS model with the addition of the CC-$(\ell, S)$ inequalities for all $\ell \in [1, T]$ and $S \subseteq [1, \ell]$. We note that the proposed CC-$(\ell, S)$ inequalities are not *valid inequalities* for the formulation E-SLS in the classical sense, as it is possible for there to be solutions to E-SLS that are not feasible to E-SLS+CC-$(\ell, S)$. However, we next argue that the formulation E-SLS+CC-$(\ell, S)$ is a valid formulation of SLS, which is the underlying model we wish to solve, and thus we may safely add teh CC-$(\ell, S)$ inequalities to formulation E-SLS.

One direction of this argument is straightforward. If $(x, s, y, z, w)$ is a feasible solution of E-SLS+CC-$(\ell, S)$, then it is also a feasible solution of E-SLS, and because E-SLS is a valid extended formulation of SLS, this implies $(x, s, y, z)$ is a feasible solution of SLS. Hence, we only need to show that for any feasible solution $(x, s, y, z)$ of SLS, there exists $w$ such that $(x, s, y, z, w)$ is a feasible solution of E-SLS+CC-$(\ell, S)$.

**Theorem 1.** *For any feasible solution $(x, s, y, z)$ of SLS, there exists $w \in \{0, 1\}^{\sum_{i=1}^T (q^*_i - 1)}$ such that $(x, s, y, z, w)$ is also a feasible solution of E-SLS+CC-$(\ell, S)$, and therefore, E-SLS+CC-$(\ell, S)$ is a valid model for SLS.*

*Proof.* For each $i \in [1, T]$, define $\bar{j}(i) = \min\{j \in [1, q^*_i] : z_{\sigma^i_j} = 0\}$, and observe that by definition $z_{\sigma^i_1} = z_{\sigma^i_2} = \cdots = z_{\sigma^i_{\bar{j}(i)-1}} = 1$. Now, for $i \in [1, T]$ define

$$w^i_j = \begin{cases} 1 & \text{for } j \in [1, \bar{j}(i) - 1] \\ 0 & \text{for } j \in [\bar{j}(i), q^*_i]. \end{cases}$$

Then, by construction, $w$ and $z$ satisfy (18) and (19). Next, for any $i \in [1, T]$, using (10) and $z_{\sigma^i_{\bar{j}(i)}} = 0$, we have

$$\sum_{t=1}^{i} x_t \ge \sum_{t=1}^{i} d_{\sigma^i_{\bar{j}(i)},t} = D_{\sigma^i_{\bar{j}(i)}}. \tag{26}$$

Using the definition of $w^i_j$ we have

$$\sum_{j=1}^{q^*_i - 1} (D_{\sigma^i_j} - D_{\sigma^i_{j+1}}) w^i_j = \sum_{j=1}^{\bar{j}(i)-1} (D_{\sigma^i_j} - D_{\sigma^i_{j+1}}) = D_{\sigma^i_1} - D_{\sigma^i_{\bar{j}(i)}}. \tag{27}$$

Combining (26) and (27) yields

$$\sum_{t=1}^{i} x_t + \sum_{j=1}^{q^*_i - 1} (D_{\sigma^i_j} - D_{\sigma^i_{j+1}}) w^i_j \ge D_{\sigma^i_1}$$

thus showing that (17) is satisfied and hence $(x,s,y,z,w)$ is a feasible solution of E-SLS.

It remains to show that $(x,s,y,z,w)$ satisfies (25) for each $\ell \in [1,T]$ and $S \subseteq [1,\ell]$. Thus, fix $\ell \in [1,T]$ and $S \subseteq [1,\ell]$.

Suppose first $y_t = 0$ for all $t \in S$. This implies $x_t = 0$ for $t \in S$. Then,

$$\begin{aligned}\sum_{t\in\bar{S}} x_t + \sum_{t\in S} \bar{D}_{t\ell} y_t + \sum_{j=1}^{q^*_\ell - 1} (D_{\sigma^\ell_j} - D_{\sigma^\ell_{j+1}}) w^\ell_j \\ = \sum_{t=1}^{\ell} x_t + \sum_{j=1}^{q^*_\ell - 1} (D_{\sigma^\ell_j} - D_{\sigma^\ell_{j+1}}) w_j \geq D_{\sigma^\ell_1}\end{aligned}$$

where the inequality holds because we have already shown (16) holds.

Next, assume $y_t = 1$ for some $t \in S$, then let $t^* = \min\{t \in S : y_t = 1\}$. We then have

$$\sum_{t\in\bar{S}} x_t + \sum_{t\in S} \bar{D}_{t\ell} y_t + \sum_{j=1}^{q^*_\ell - 1} (D_{\sigma^\ell_j} - D_{\sigma^\ell_{j+1}}) w^\ell_j$$

$$\geq \sum_{t=1}^{t^*-1} x_t + \bar{D}_{t^*\ell} + \sum_{j=1}^{q^*_\ell - 1} (D_{\sigma^\ell_j} - D_{\sigma^\ell_{j+1}}) w^\ell_j \tag{28}$$

$$= \sum_{t=1}^{t^*-1} x_t + \bar{D}_{t^*\ell} + D_{\sigma^\ell_1} - D_{\sigma^\ell_{\bar{j}(\ell)}} \tag{29}$$

$$\geq D_{\sigma^\ell_{\bar{j}(\ell)}, t^*-1} + \bar{D}_{t^*\ell} + D_{\sigma^\ell_1} - D_{\sigma^\ell_{\bar{j}(\ell)}} \tag{30}$$

$$\geq D_{\sigma^\ell_1} \tag{31}$$

where (28) follows from the definition of $t^*$, (29) follows from (27), (30) follows from (26) because $z_{\sigma_{\bar{j}(\ell)}} = 0$, and (31) follows because

$$\begin{aligned}D_{\sigma^\ell_{\bar{j}(\ell)}} = \sum_{t=1}^{\ell} d_{\sigma^\ell_{\bar{j}(\ell)},t} \leq \sum_{t=1}^{t^*-1} d_{\sigma^\ell_{\bar{j}(\ell)},t} + \max_{j\in\Omega} \sum_{t=t^*}^{\ell} d_{jt} \\ = D_{\sigma^\ell_{\bar{j}(\ell)}, t^*-1} + \bar{D}_{t^*,\ell}.\end{aligned}$$

□

As the number of CC-$(\ell, S)$ inequalities grows exponentially with $T$, we require a separation algorithm for identifying violated inequalities from this class. Let $(\bar{x}, \bar{s}, \bar{y}, \bar{z}, \bar{w}) \in \mathbb{R}_+^{(m+1)T} \times [0,1]^{T+m+\sum_{i=1}^{T}(q^*_i - 1)}$ be a given relaxation solution. Algorithm 1 provides a natural generalization of the separation algorithm for the traditional $(\ell, S)$ inequalities to the inequalities of the form (25). The running time of this algorithm is $\mathcal{O}(T \log T)$.

At each iteration $\ell$ where an $(\ell,S)$ pair is output in line 8, the corresponding inequality (25) is violated by the current solution. In case that there are no such outputs, there are no violated inequalities. The proof of correctness of this algorithm follows exactly that of the separation algorithm for the traditional $(\ell, S)$ inequalities (see [21, Page 219]).

# 3 A New Extended Formulation of SLS under a Modified Wagner-Whitin Condition

In this section we derive a potentially stronger formulation in a lifted variable space, which is valid under a modified Wagner-Whitin cost condition.

**Assumption 1** (Modified Wagner-Whitin condition)**.** *The SLS problem satisfies the following conditions:*

**Algorithm 1** Separation Algorithm for CC-$(\ell, S)$ Inequalities

**for** $t = 1, \ldots, T$ **do**
  Determine $\ell(t) \in [t, T]$ such that $\bar{D}_{t,\ell(t)-1}\bar{y}_t < \bar{x}_t \leq \bar{D}_{t,\ell(t)}\bar{y}_t$ by bisection.
**end for**
Let $\Delta_0 = 0$,
**for** $\ell = 1, \ldots, T$ **do**
  Select sets $Y_\ell = \{t \in [1, \ell] : \ \ell(t) > \ell\}$, $X_\ell = \{t \in [1, \ell] : \ \ell(t) = \ell\}$.
  Calculate $\Delta_\ell = \Delta_{\ell-1} + (\bar{D}_{t,\ell} - \bar{D}_{t,\ell-1})(\sum_{t \in Y_\ell} \bar{y}_t) + \sum_{t \in X_\ell} (\bar{x}_t - \bar{D}_{t,\ell-1}\bar{y}_t)$.
  If $\Delta_\ell < D_{\sigma_1^\ell} - \sum_{j=1}^{q_\ell^*-1} (D_{\sigma_j^\ell} - D_{\sigma_{j+1}^\ell})\bar{w}_j^\ell$, output $\ell$ and $S = Y_\ell$.
**end for**

$$\alpha_i + (1 - \varepsilon)h_i \geq \alpha_{i+1}, \quad \textit{for } i \in [1, T-1].$$

Assumption 1 yields a special property of the optimal solution of the SLS, which we subsequently use to derive the new formulation.

**Lemma 2.** *Suppose Assumption 1 holds, let* $(x, y, s, z)$ *be a feasible solution to the SLS problem* (9) - (15)*, and let* $I = \{i_1, i_2, \ldots, i_r\} = \{i \in [1, T] : y_i = 1\}$ *where* $1 = i_1 < i_2 < \ldots < i_r \leq T$*. Then, there exists* $\bar{s}$ *such that the solution* $(\bar{x}, y, \bar{s}, z)$ *with production levels* $\bar{x}$ *defined as*

$$\bar{x}_{i_k} = \delta_{i_{k+1}}(z) - \delta_{i_k}(z), \quad k \in [1, r] \tag{32}$$

*and* $\bar{x}_i = 0$ *for* $i \in [1, T] \setminus I$ *is feasible to* (9) - (15) *and has cost not more than the cost of* $(x, y, s, z)$*, where* $i_{r+1} := T + 1$, $\delta_i(z) := \max_{j \in J_z}\{\sum_{t=1}^{i-1} d_{jt}\}$*, and* $J_z = \{j \in \Omega : z_j = 0\}$.

*Proof.* Let

$$f(x, y, s, z) = \alpha^\top x + \beta^\top y + (h^1)^\top s^1 + \sum_{j=1}^{m} \pi_j (h^2)^\top s_j^2$$

be the cost of solution $(x, y, s, z)$ in (8).

If there is some $k \in [1, r]$ such that $x_{i_k} \neq \delta_{i_{k+1}}(z) - \delta_{i_k}(z)$, then let $k^* = \min\{k \in [1, r] : \ x_{i_k} \neq \delta_{i_{k+1}}(z) - \delta_{i_k}(z)\}$. If $x_{i_{k^*}} < \delta_{i_{k^*+1}}(z) - \delta_{i_{k^*}}(z)$, then there must be some scenario $j \in J_z$ whose demand in period $(i_{k^*+1} - 1)$ can not be satisfied, so we may assume $x_{i_{k^*}} > \delta_{i_{k^*+1}}(z) - \delta_{i_{k^*}}(z)$.

Define

$$g(x, y, s, z) = \sum_{i \in [1,T] \setminus [i_{k^*}, i_{k^*+1}]} (\alpha_i x_i + \beta_i y_i) + \sum_{j=1}^{m} \Big(\pi_j \sum_{i \in [1,T] \setminus [i_{k^*}, i_{k^*+1}-1]} h_i s_{ji}\Big),$$

then

$$\begin{aligned} f(x, y, s, z) = {} & g(x, y, s, z) + (\alpha_{i_{k^*}} x_{i_{k^*}} + \beta_{i_{k^*}} y_{i_{k^*}}) \\ & + \sum_{j=1}^{m} \Big(\pi_j \sum_{i=i_{k^*}}^{i_{k^*+1}-1} h_i s_{ji}\Big) + (\alpha_{i_{k^*+1}} x_{i_{k^*+1}} + \beta_{i_{k^*+1}} y_{i_{k^*+1}}). \end{aligned}$$

Let $\bar{x}$ and $\bar{s}$ be defined by $\bar{x}_{i_{k^*}} = x_{i_{k^*}} - \eta$, $\bar{s}_{ji} = s_{ji} - \eta$, for $i \in [i_{k^*}, i_{k^*+1} - 1]$ and $j \in J_z$, $\bar{s}_{ji} = [s_{ji} - \eta]^+$, for $i \in [i_{k^*}, i_{k^*+1} - 1]$ and $j \in \Omega \setminus J_z$, $\bar{x}_{i_{k^*+1}} = x_{i_{k^*+1}} + \eta$, and other components are the same as $x$ and $s$.

Then

$$
\begin{aligned}
f(\bar{x}, y, \bar{s}, z) = g(x,y,s,z) + (\alpha_{i_{k^*}}(x_{i_{k^*}} - \eta) + \beta_{i_{k^*}} y_{i_{k^*}}) \\
+ \sum_{j \in J_z} \big(\pi_j \sum_{i=i_{k^*}}^{i_{k^*+1}-1} h_i(s_{ji} - \eta)\big) + \sum_{j \in \Omega \setminus J_z} \big(\pi_j \sum_{i=i_{k^*}}^{i_{k^*+1}-1} h_i[s_{ji} - \eta]^+\big) \\
+ (\alpha_{i_{k^*+1}}(x_{i_{k^*+1}} + \eta) + \beta_{i_{k^*+1}} y_{i_{k^*+1}}) \\
&= f(x,y,s,z) - \Big(\alpha_{i_{k^*}} + \sum_{i=i_{k^*}}^{i_{k^*+1}-1} (\sum_{j \in J_z} \pi_j) h_i - \alpha_{i_{k^*}+1}\Big)\eta \\
&\quad + \sum_{j \in \Omega \setminus J_z} \big(\pi_j \sum_{i=i_{k^*}}^{i_{k^*+1}-1} h_i([s_{ji} - \eta]^+ - s_{ji})\big) \\
&\le f(x,y,s,z) - \Big(\alpha_{i_{k^*}} + \sum_{i=i_{k^*}}^{i_{k^*+1}-1} (\sum_{j \in J_z} \pi_j) h_i - \alpha_{i_{k^*}+1}\Big)\eta \\
&= f(x,y,s,z) - \sum_{i=i_{k^*}}^{i_{k^*+1}-1} (\alpha_i + (\sum_{j \in J_z} \pi_j) h_i - \alpha_{i+1})\eta.
\end{aligned}
$$

By Assumption 1, $\sum_{i=i_{k^*}}^{i_{k^*+1}-1}(\alpha_i + (\sum_{j \in J_z} \pi_j) h_i - \alpha_{i+1}) \ge \sum_{i=i_{k^*}}^{i_{k^*+1}-1}(\alpha_i + (1-\varepsilon) h_i - \alpha_{i+1}) \ge 0$, then setting $\eta = x_{i_{k^*}} - (\delta_{i_{k^*+1}}(z) - \delta_{i_{k^*}}(z))$, $(\bar{x}, y, \bar{s}, z)$ is also a feasible solution of SLS with $f(\bar{x}, y, \bar{s}, z) \le f(x,s,y,z)$, and satisfies $\bar{x}_{i_{k^*}} = \delta_{i_{k^*+1}}(z) - \delta_{i_{k^*}}(z)$.

This process can now be repeated with the solution $(\bar{x}, y, \bar{s}, z)$ as long as there is an index $i$ that does not satisfy (32), eventually yielding a solution that does satisfy (32) and has cost no worse than the cost of $(x,y,s,z)$. □

The new formulation introduces new decision variables that take advantage of the structure we can restrict the production variables to follow, without loss of optimality, according to Lemma 2. Specifically, we introduce the following new set of variables:

- $\phi_{it}^{jk} = 1$ if an amount $D_{\sigma_k^t} - D_{\sigma_j^{i-1}}$ is produced in period $i$, for $1 \le i \le t \le T$, $j \in [1, q_{i-1}^*]$, and $k \in [1, q_t^*]$,

where we define $q_0^* := 1$ and $D_{\sigma_1^0} := 0$. We call the production plan determined by setting $\phi_{it}^{jk} = 1$ (i.e., producing $D_{\sigma_k^t} - D_{\sigma_j^{i-1}}$ in period $i$ to meet demand in periods $i+1, \ldots, t$) a subplan.

We then obtain the following new extended formulation(NE-SLS):

$$
\min \quad \alpha^\top x + \beta^\top y + \sum_{j=1}^{m} \pi_j h^\top s_j
$$

$$
\sum_{\tau=1}^{T} \sum_{k=1}^{q_\tau^*} \phi_{1\tau}^{1k} = 1 \tag{33}
$$

$$
\sum_{i=1}^{t-1} \sum_{j=1}^{q_{i-1}^*} \sum_{k=1}^{q_{t-1}^*} \phi_{i,t-1}^{jk} - \sum_{\tau=t}^{T} \sum_{j=1}^{q_{t-1}^*} \sum_{k=1}^{q_\tau^*} \phi_{t\tau}^{jk} = 0, \quad t \in [2,T] \tag{34}
$$

$$
\sum_{i=1}^{T} \sum_{j=1}^{q_{i-1}^*} \sum_{k=1}^{q_T^*} \phi_{iT}^{jk} = 1 \tag{35}
$$

$$
\sum_{\tau=t}^{T} \sum_{j=1}^{q_{t-1}^*} \sum_{k=1}^{q_\tau^*} \phi_{t\tau}^{jk} \le y_t, \quad t \in [1,T] \tag{36}
$$

$$\sum_{\tau=t}^{T}\sum_{j=1}^{q^*_{t-1}}\sum_{k=1}^{q^*_\tau}(D_{\sigma^\tau_k} - D_{\sigma^{t-1}_j})\phi^{jk}_{t\tau} = x_t, \quad t \in [1,T] \tag{37}$$

$$\sum_{t=i}^{T}\sum_{k=1}^{q^*_t}\phi^{jk}_{it} = w^{i-1}_{j-1} - w^{i-1}_j, \quad j \in [1, q^*_{i-1}-1], \quad i \in [2,T] \tag{38}$$

$$\sum_{i=1}^{t}\sum_{j=1}^{q^*_{i-1}}\phi^{jk}_{it} = w^t_{k-1} - w^t_k, \quad k \in [1, q^*_t - 1], \quad t \in [1,T] \tag{39}$$

$$w^i_0 = y_{i+1}, \quad i \in [1,T] \tag{40}$$

$$w^i_0 \geq w^i_1, \quad i \in [1,T] \tag{41}$$

$$\phi^{jk}_{it} \in \{0,1\}, \quad j \in [1, q^*_{i-1}], \quad k \in [1, q^*_t], \quad 1 \leq i \leq t \leq T$$

$$w^i_0 \in \{0,1\}, \quad i \in [1,T]$$

$$(x, y, s, z, w) \text{ satisfy } (18) - (21), (23) - (24)$$

where we define $y_{T+1} := 1$. Constraints (33)-(35) are the flow conservation constraints that model a production plan as a sequence of subplans. Constraints (36) ensure that if $y_t = 0$ then no subplan can start in period $t$, for $t \in [1,T]$. Constraints (37) calculate the production levels in each period. Constraints (38)-(41) together with (18)-(20) ensure that the production plan determined by the subplans satisfies the chance constraint. The auxiliary variables $(w^i_0, w^i_1, \ldots, w^i_{q^*_i - 1})$ for $i \in [1,T]$ play a similar role as in E-SLS, except that here they are extended to include $w^i_0$, which by (40) is equal to $y_{i+1}$. (Although $w^i_0$ can be eliminated we use it because it simplifies presentation of constraints (38) and (39).) This modification ensures that if production is not done in period $i+1$ ($y_{i+1} = 0$) then $w^i_j = 0$ for all $j \in [1, q^*_j - 1]$, so that in this case these variables to not impact the $z_j$ variables determining which scenarios are satisfied via (19). For each $i \in [1,T]$, constraints (41) and (18) ensure that there is at most one $j \in [1, q^*_i - 1]$ such that $w^i_{j-1} - w^i_j = 1$, and $w^i_{j-1} - w^i_j = 0$ otherwise. As in E-SLS, $w^i_{j-1} - w^i_j = 1$ is an indication that there is sufficient inventory available in period $i$ to meet scenarios with demand up to $D_{\sigma^i_j}$. Thus, constraints (38) enforce consistency between this in period $i-1$ and the subplan determined by the $\phi^{jk}_{it}$ variables. Likewise, constraints (39) enforce consistency between the indication of the available inventory in period $t$ (as determined by the expression $w^t_{k-1} - w^t_k$) and the subplan determined by the $\phi^{jk}_{it}$ variables. As in the E-SLS formulation constraints (41), (18) and (19) enforce consistency between the $w^i_j$ variables and the $z_j$ variables used in the chance constraint (20).

**Proposition 3.** *Under Assumption 1, the new extended formulation NE-SLS is a valid model for SLS.*

*Proof.* Let $(x, y, s, z)$ be a feasible solution of SLS where $x$ satisfies (32). Let $I = \{i_1, i_2, \ldots, i_r\} = \{i \in [1,T] : y_i = 1\}$ where $1 = i_1 < i_2 < \ldots < i_r \leq T$. For each $i \in [1,T]$, define $\bar{j}(i) = \min\{j \in [1, q^*_i] : z_{\sigma^i_j} = 0\}$. Now, for $t \in [1,T]$, $k \in [1, q^*_t]$, define

$$\phi^{1k}_{1t} = \begin{cases} 1 & \text{if } t = i_1 - 1, \quad k = \bar{j}(i_1 - 1) \\ 0 & \text{otherwise ,} \end{cases}$$

and for $2 \leq i \leq t \leq T$, $j \in [1, q^*_{i-1}]$, $k \in [1, q^*_t]$, define

$$\phi^{jk}_{it} = \begin{cases} 1 & \text{if } i = i_u,\ t = i_{u+1} - 1,\ j = \bar{j}(i_u - 1),\ k = \bar{j}(i_{u+1} - 1),\ u \in [1,r] \\ 0 & \text{otherwise ,} \end{cases}$$

and for $i \in [1,T]$, $j \in [0, q^*_i]$ define

$$w^i_j = \begin{cases} 1 & \text{if } i+1 \in I \cup \{i_{r+1}\},\ j \in [0, \bar{j}(i) - 1] \\ 0 & \text{otherwise .} \end{cases}$$

Direct calculations verify that $(x, y, s, z, w, \phi)$ satisfies all the constraints of NE-SLS.

Let $(x, y, s, z, w, \phi)$ be a feasible solution of the NE-SLS. For $i \in [1, T]$, if $y_i = 0$, then (36)-(37) imply $x_i = 0$, which indicates that $(x, y, s, z)$ satisfies constraints (13). Obviously, $(x, y, s, z)$ satisfies (11) and (12) as they are contained in the formulation of NE-SLS. Let $\hat{I} = \{1\} \cup \{i \in [1, T] : \exists t \in [i, T], j \in [1, q^*_{i-1}], k \in [1, q^*_t],$ s.t. $\phi^{jk}_{it} = 1\} = \{\hat{i}_1, \hat{i}_2, \ldots, \hat{i}_r\}$ where $1 = \hat{i}_1 < \hat{i}_2 < \ldots < \hat{i}_r \leq T < \hat{i}_{r+1} := T + 1$. Then by (33)-(35) there exists exactly one $j_u \in [1, q^*_{\hat{i}_u - 1}]$ and one $k_u \in [1, q^*_{\hat{i}_{u+1}-1}]$, such that $\phi^{j_u k_u}_{\hat{i}_u, \hat{i}_{u+1}-1} = 1$ for $u \in [1, r]$. For $u \in [1, r-1]$, if $\phi^{j_u k_u}_{\hat{i}_u, \hat{i}_{u+1}-1} = \phi^{j_{u+1} k_{u+1}}_{\hat{i}_{u+1}, \hat{i}_{u+2}-1} = 1$, then we conclude that $k_u = j_{u+1}$. If $k_u \neq q^*_{\hat{i}_{u+1}}$, then by (39) there is $w^{\hat{i}_{u+1}-1}_{k_u - 1} - w^{\hat{i}_{u+1}-1}_{k_u} = \phi^{j_u k_u}_{\hat{i}_u, \hat{i}_{u+1}-1} = 1 = \phi^{j_{u+1} k_{u+1}}_{\hat{i}_{u+1}, \hat{i}_{u+2}-1}$, therefore, by (38) and the flow conservation constraints (33)-(35) we can obtain $k_u = j_{u+1}$; if $k_u = q^*_{\hat{i}_{u+1}}$, then by (39)-(40) we have $w^{\hat{i}_{u+1}-1}_0 = w^{\hat{i}_{u+1}-1}_1 = \cdots = w^{\hat{i}_{u+1}-1}_{q^*_{\hat{i}_{u+1}-1}} = 1$, also, by (38) and the flow conservation constraints (33)-(35) we can obtain $j_{u+1} = q^*_{\hat{i}_{u+1}} = k_u$. Thus, for any $i \in [\bar{i}_u, \bar{i}_{u+1}]$, $u \in [1, r]$, we have

$$
\begin{aligned}
\sum_{t=1}^{i} x_t &= \sum_{v=1}^{u} (D_{\sigma^{\bar{i}_{v+1}-1}_{k_v}} - D_{\sigma^{\bar{i}_v - 1}_{j_v}}) \phi^{j_v k_v}_{\bar{i}_v \bar{i}_{v+1}} \\
&= \sum_{v=1}^{u} (D_{\sigma^{\bar{i}_{v+1}-1}_{k_v}} - D_{\sigma^{\bar{i}_v - 1}_{j_v}}) \\
&= D_{\sigma^{\bar{i}_{u+1}-1}_{k_u}} \\
&\geq \begin{cases} D_{\sigma^{\bar{i}_{u+1}-1}_j} (1 - z_{\sigma^{\bar{i}_{u+1}-1}_j}) & \text{for } j \in [1, k_u - 1] \\ D_{\sigma^{\bar{i}_{u+1}-1}_j} & \text{for } j \in [k_u, m] \end{cases} \\
&\geq \begin{cases} D_{\sigma^{\bar{i}_{u+1}-1}_j i} (1 - z_{\sigma^{\bar{i}_{u+1}-1}_j}) & \text{for } j \in [1, k_u - 1] \\ D_{\sigma^{\bar{i}_{u+1}-1}_j i} & \text{for } j \in [k_u, m] \end{cases} \\
&\geq D_{ji}(1 - z_j) \quad \text{for } j \in \Omega.
\end{aligned}
$$

Thus, $(x, y, s, z)$ satisfies (10) and hence is a feasible solution of SLS. □

The new model NE-SLS potentially has a much tighter linear programming relaxation than other models, which can lead to better root relaxation gaps and fewer nodes explored nodes when used withing a branch-and-bound algorithm.

**Remark 1.** *Although NE-SLS is only proved to be valid under the modified Wagner-Whitin condition, in our computational experience we have observed that the optimal solution of NE-SLS is usually equal to the optimal solution of SLS even when this condition does not hold. Thus, in cases where solving NE-SLS is more efficient, it may be a useful formulation computationally as an inner approximation even when the modified Wagner-Whitin condition does not hold.*

# 4 The Results for General SLS

In this section we adapt the results for SLS with $p = 0$ to the general case with $p \geq 0$. First, we define the CC-$(\ell, S)$ inequalities for general SLS.

Let $d^\ell_i = \sum_{t=i}^{\ell} d_t$, for $1 \leq i \leq \ell \leq p$ denote the cumulative demand from period $i$ to period $\ell$, let $D^j_{i\ell} = \sum_{t=i}^{\ell} d_{jt}$ for $p + 1 \leq i \leq \ell \leq T$, $j \in \Omega$ be the cumulative demand from period $i$ to period $\ell$ in scenario $j$, and let $\bar{D}_{i\ell} = \max_{j \in \Omega}\{D^j_{i\ell}\}$, for $p + 1 \leq i \leq \ell \leq T$ be the maximum cumulative demand from period $i$ to period $\ell$ over all scenarios. Then we can describe the general CC-$(\ell, S)$ inequalities as the following.

**Definition 2.** *The following two kinds of inequalities are called the CC-$(\ell, S)$ inequalities for general E-SLS:*
*1. For each $\ell \in [1, p]$ and $S \subseteq [1, \ell]$:*

$$
\sum_{t \in \bar{S}} x_t + \sum_{t \in S} d^\ell_t y_t \geq d^\ell_1, \tag{42}
$$

*where* $\bar{S} = [1,\ell] \setminus S$

*2. For each* $\ell \in [p+1,T]$, $S_1 \subseteq [1,p]$, *and* $S_2 \subseteq [p+1,\ell]$*:*

$$\sum_{t\in \bar{S}_1} x_t + \sum_{t\in \bar{S}_2} x_t + \sum_{t\in S_1} (d_t^p + \bar{D}_{p+1,\ell}) y_t + \sum_{t\in S_2} \bar{D}_{t\ell} y_t + \sum_{j=1}^{q_\ell^*-1} (D_{\sigma_j^\ell} - D_{\sigma_{j+1}^\ell}) w_j^\ell \geq d_1^p + D_{\sigma_1^\ell}, \tag{43}$$

*where* $\bar{S}_1 = [1,p] \setminus S_1$, $\bar{S}_2 = [p+1,T] \setminus S_2$, *and* $S = S_1 \cup S_2$.

Similarly, we define general E-SLS+CC-$(\ell,S)$ to be the formulation defined by the general E-SLS model with the addition of the general CC-$(\ell,S)$ inequalities for all $\ell \in [1,T]$, $S \subseteq [1,\ell]$. Then we have the similar theorem as Theorem 1.

**Theorem 4.** *For any feasible solution* $(x,s^1,s^2,y,z)$ *of general SLS, there exists* $w \in \{0,1\}^{\sum_{i=p+1}^T (q_i^*-1)}$ *such that* $(x,s^1,s^2,y,z,w)$ *is also a feasible solution of general E-SLS+CC-*$(\ell,S)$*, and therefore, general E-SLS+CC-*$(\ell,S)$ *is a valid model for general SLS.*

Let $(\bar{x},\bar{s^1},\bar{s^2},\bar{y},\bar{z},\bar{w}) \in \mathbb{R}_+^{T+p+m(T-p)} \times [0,1]^{T+m+\sum_{i=p+1}^T (q_i^*-1)}$ be a given relaxation solution. Algorithm 2 provides the separation algorithm for the general CC-$(\ell,S)$ inequalities (42) and (43). The running time of this algorithm is also $\mathcal{O}(T\log T)$.

**Algorithm 2** Separation Algorithm for General CC-$(\ell,S)$ Inequalities

```
for t = 1,...,T do
    Determine ℓ(t) ∈ [t,T] such that d̄_{t,ℓ(t)−1} ȳ_t < x̄_i ≤ d̄_{t,ℓ(t)} ȳ_t by bisection, where d̄_{tk} = d_t^k for 1 ≤ t ≤
  k ≤ p, d̄_{tk} = d_t^p + D̄_{p+1,k} for 1 ≤ t ≤ p < k ≤ T, and d̄_{tk} = D̄_{tk} for p+1 ≤ t ≤ k ≤ T.
end for
Let Δ_0 = 0,
for ℓ = 1,...,p do
    Select sets Y_ℓ = {t ∈ [1,ℓ] : ℓ(t) > ℓ}, X_ℓ = {t ∈ [1,ℓ] : ℓ(t) = ℓ}.
    Calculate Δ_ℓ = Δ_{ℓ−1} + d_ℓ( Σ_{t∈Y_ℓ} ȳ_t) + Σ_{t∈X_ℓ} (x̄_t − d̄_{t,ℓ−1} ȳ_t).
    If Δ_ℓ < d̄_{1ℓ}, output ℓ and S = Y_ℓ.
end for
for ℓ = p+1,...,T do
    Select sets Y_ℓ = {t ∈ [1,ℓ] : ℓ(t) > ℓ}, X_ℓ = {t ∈ [1,ℓ] : ℓ(t) = ℓ}.
    Calculate Δ_ℓ = Δ_{ℓ−1} + (d̄_{p+1,ℓ} − d̄_{p+1,ℓ−1})(Σ_{t∈Y_ℓ∩[1,p]} ȳ_t) + (d̄_{tℓ} − d̄_{t,ℓ−1})(Σ_{t∈Y_ℓ∩[p+1,ℓ]} ȳ_t) + Σ_{t∈X_ℓ} (x̄_t −
  d̄_{t,ℓ−1} ȳ_t).
    If Δ_ℓ < d̄_{1ℓ} − Σ_{j=1}^{q_ℓ^*−1} (D_{σ_j^ℓ} − D_{σ_{j+1}^ℓ}) w̄_j^ℓ, output ℓ and S = Y_ℓ.
end for
```

Like in Sect. 3, we derive a potentially stronger formulation for the general SLS in a lifted variable space, which is valid under a modified Wagner-Whitin cost condition.

**Assumption 2** (Modified Wagner-Whitin condition)**.** *The general SLS problem satisfies the following conditions:*

$$\alpha_i + h_i \geq \alpha_{i+1}, \quad \textit{for } i \in [1,p],$$

*and*

$$\alpha_i + (1-\varepsilon)h_i \geq \alpha_{i+1}, \quad \textit{for } i \in [p+1,T-1].$$

**Lemma 5.** *Suppose Assumption 2 holds, let $(x, y, s^1, s^2, z)$ be a feasible solution to the general SLS problem (9)-(15), and let $I = \{i_1, i_2, \ldots, i_r\} = \{i \in [1, T] : y_i = 1\}$ where $1 = i_1 < i_2 < \ldots < i_r \leq T$. Then, there exists $\bar{s}^1$, $\bar{s}^2$ such that the solution $(x, y, \bar{s}^1, \bar{s}^2, z)$ with production levels $\bar{x}$ defined as*

$$\bar{x}_{i_k} = \delta_{i_{k+1}}(z) - \delta_{i_k}(z), \quad k \in [1, r].$$

*and $\bar{x}_i = 0$ for $i \in [1, T] \setminus I$ is feasible to (9)-(15) and has cost not more than the cost of $(x, y, s^1, s^2, z)$, where $i_r := T + 1$, $\delta_i(z) = \sum_{t=1}^{i-1} d_t$, for $i \in [1, p]$, $\delta_i(z) = \sum_{t=1}^{p} d_t + \max_{\tau \in J_z} \{\sum_{t=p+1}^{i-1} d_{\tau t}\}$, for $i \in [p+1, T]$, and $J_z = \{j \in \Omega : z_j = 0\}$.*

Similarly, we introduce some new sets of variables to construct the new formulation:

- $\psi_{it} = 1$ if an amount $d_i^t$ is produced in period $i$, for $1 \leq i \leq t \leq p$,
- $\varphi_{it}^j = 1$ if an amount $d_i^p + D_{\sigma_j^t}$ is produced in period $i$, for $1 \leq i \leq p$, $p + 1 \leq t \leq T$, and $j \in [1, q_t^*]$,
- $\phi_{it}^{jk} = 1$ if an amount $D_{\sigma_k^t} - D_{\sigma_j^{i-1}}$ is produced in period $i$, for $p + 1 \leq i \leq t \leq T$, $j \in [1, q_{i-1}^*]$, and $k \in [1, q_t^*]$,

where $q_p^* := 1$ and $D_{\sigma_1^p} := 0$.

We then obtain the following new extended formulation (general NE-SLS):

$$\min \quad \alpha^\top x + \beta^\top y + (h^1)^\top s^1 + \sum_{j=1}^{m} \pi_j (h^2)^\top s_j^2$$

$$\sum_{\tau=1}^{p} \psi_{1\tau} + \sum_{t=p+1}^{T} \sum_{j=1}^{q_\tau^*} \varphi_{1\tau}^j = 1$$

$$\sum_{i=1}^{t-1} \psi_{i,t-1} - (\sum_{\tau=t}^{p} \psi_{t\tau} + \sum_{\tau=p+1}^{T} \sum_{j=1}^{q_\tau^*} \varphi_{t\tau}^j) = 0, \quad t \in [2, p]$$

$$\sum_{i=1}^{p} \psi_{ip} - \sum_{\tau=p+1}^{T} \sum_{k=1}^{q_\tau^*} \phi_{p+1,\tau}^{1k} = 0$$

$$\sum_{i=1}^{p} \sum_{j=1}^{q_{t-1}^*} \varphi_{i,t-1}^j + \sum_{i=p+1}^{t-1} \sum_{j=1}^{q_{i-1}^*} \sum_{k=1}^{q_{t-1}^*} \phi_{i,t-1}^{jk} - \sum_{\tau=t}^{T} \sum_{j=1}^{q_{t-1}^*} \sum_{k=1}^{q_\tau^*} \phi_{t\tau}^{jk} = 0, \quad t \in [p+2, T]$$

$$\sum_{i=1}^{p} \sum_{j=1}^{q_T^*} \varphi_{iT}^j + \sum_{i=p+1}^{T} \sum_{j=1}^{q_{i-1}^*} \sum_{k=1}^{q_T^*} \phi_{iT}^{jk} = 1$$

$$\sum_{\tau=t}^{p} \psi_{t\tau} + \sum_{\tau=p+1}^{T} \sum_{j=1}^{q_\tau^*} \varphi_{t\tau}^j \leq y_t, \quad t \in [1, p]$$

$$\sum_{\tau=t}^{T} \sum_{j=1}^{q_{t-1}^*} \sum_{k=1}^{q_\tau^*} \phi_{t\tau}^{jk} \leq y_t, \quad t \in [p+1, T]$$

$$\sum_{\tau=t}^{p} d_t^\tau \psi_{t\tau} + \sum_{\tau=p+1}^{T} \sum_{j=1}^{q_\tau^*} (d_t^p + D_{\sigma_j^\tau}) \varphi_{t\tau}^j = x_t, \quad t \in [1, p]$$

$$\sum_{\tau=t}^{T} \sum_{j=1}^{q_{t-1}^*} \sum_{k=1}^{q_\tau^*} (D_{\sigma_k^\tau} - D_{\sigma_j^{t-1}}) \phi_{t\tau}^{jk} = x_t, \quad t \in [p+1, T]$$

$$\sum_{t=i}^{T}\sum_{k=1}^{q_t^*}\phi_{it}^{jk} = w_{j-1}^{i-1} - w_j^{i-1}, \quad j \in [1, q_{i-1}^* - 1], \quad i \in [p+2, T]$$

$$\sum_{i=1}^{p}\varphi_{it}^{k} + \sum_{i=p+1}^{t}\sum_{j=1}^{q_{i-1}^*}\phi_{it}^{jk} = w_{k-1}^{t} - w_k^{t}, \; k \in [1, q_t^* - 1], \quad t \in [p+1, T]$$

$$w_0^i = y_{i+1}, \quad i \in [p+1, T]$$

$$w_0^i \geq w_1^i, \quad i \in [p+1, T]$$

$$\psi_{it} \in \{0,1\}, \quad 1 \leq i \leq t \leq p,$$

$$\varphi_{it}^{j} \in \{0,1\}, \quad j \in [1, q_t^*], \quad i \in [1, p], \quad t \in [p+1, T]$$

$$\phi_{it}^{jk} \in \{0,1\}, \quad j \in [1, q_i^*], \quad k \in [1, q_t^*], \quad p+1 \leq i \leq t \leq T$$

$$w_0^i \in \{0,1\}, \quad i \in [p+1, T]$$

$$(x, y, s^1, s^2, z, w) \text{ satisfy } (16), (18) - (21), (23) - (24),$$

where $y_{T+1} := 1$.

**Proposition 6.** *Under Assumption 2, the new extended formulation general NE-SLS is a valid model for general SLS.*

# 5 Computational experiments

In this section we report results comparing the computational performance of the four formulations: SLS, E-SLS, E-SLS+CC-$(\ell, S)$ and NE-SLS. All the experiments were executed on a Windows 10 Home workstation with 1.80GHz Intel(R) Core(TM) i7-8550U CPU and 16.0 GB RAM. The algorithms tested in the computational experiment were implemented using Python programming language, with Python 3.7 and Gurobi 8.1.1. A time limit of one hour were enforced.

For formulation E-SLS+CC-$(\ell, S)$ we use Algorithm 1 to separate the CC-$(\ell, S)$ inequalities and add them to the LP relaxation iteratively in a cutting plane loop. We keep adding cuts until the LP relaxation is not improved for five consecutive iterations. After that, the formulation with these cuts included is given the the solver Gurobi and solved.

As an additional comparison, we investigate the use of the valid inequalities of Proposition 3.2 in [17] (which we refer to as LK-Cuts) added to formulation E-SLS, which we refer to as E-SLS+LK. We separate LK-Cuts using the heuristic separation algorithm proposed in [17]. We found that adding these cuts within a cutting plane loop led to a large number LK-Cuts, which after a while had very marginal improvement on the LP relaxation. To achieve the most benefit from these cuts while also limiting the number added we implemented a stopping condition for the cut generation loop. Let $objval1$ be the LP relaxation objective value of E-SLS after adding the LK-Cuts for some iterations, and $objval2$ be the LP relaxation objective value of E-SLS after one more iteration. Then, we stop the process of generating LK-Cuts once $\frac{objval2 - objval1}{objval1} \leq \beta$ occurs in five consecutive iterations of the cutting plane loop. We tested the performance of E-SLS+LK with $\beta = 0.0001$ and $\beta = 0.02$, and found that generally E-SLS+LK with $\beta = 0.02$ performs better in terms of total computation time. Thus, we use $\beta = 0.02$ in our experiments using the E-SLS+LK method. Results of E-SLS+LK with these two parameter values are reported in the Appendix.

In our test instances we set $p = 0$ and $\pi_j = \frac{1}{m}$ for all $j \in \Omega$. The data we used is adapted from [4], in which the integrality gap of the lot-sizing problem instances is influenced by the ratio between the setup cost and inventory holding cost. Therefore, the instances are generated for varying setup/holding cost ratios $f \in \{100,\ 200,\ 500,\ 1000\}$. For all instances, holding cost $h_i$ equals to 10, and unit production cost $\alpha_i$ and setup cost $\beta_i$ are drawn from discrete uniform distribution over $[81, 119]$ and $[9f,\ 11f]$, respectively, for all $i \in [1, T]$. In addition, the demand $d_{ji}$ follows discrete uniform distribution $[1, 19]$, for each period $i \in [1, T]$ and scenario $j \in \Omega$. In all tables, each row reports the average of three randomly generated instances.

Before analyzing the computational results among different formulations, we introduce the meaning for every column in the following tables. The "Time" column reports the average solution time, in seconds, for the instances that are solved to optimality within the time limit, the number in the brackets "[ ]"under the "Time" column indicates how many instances among the three are solved to optimality within the time limit, and the " * " sign under the "Time" column indicates that none of the three instances is solved to optimality within the time limit. The "Nodes" column reports the average number of nodes explored during the branch-and-bound process, and the " > " sign under the "Nodes" column indicates that the reported number of nodes is a lower bound, because not all the instances were solved within the time limit and the average is computed using the number of nodes processed up to the limit. The "Gap" column reports the average percentage optimality gap at the time limit, i.e., $Gap = \frac{ubval - lbval}{ubval}$ where $ubval$ is the value of the best feasible solution found by that method within the time limit and $lbval$ is the lower bound obtained within the time limit. The "LP.Gap" and "R.Gap" columns report the average percentage gap of the LP relaxation and root node, respectively, i.e., $LP.Gap = \frac{bestubval - lpval}{bestubval} \times 100$, and $R.Gap = \frac{bestubval - rootval}{bestubval} \times 100$, where $bestubval$, $lpval$ and $rootval$ are the objective function values of the best feasible solution (the best solution found among all the comparing formulations), the initial LP relaxation, and LP relaxation after all cuts are added before branching. The "Cuts" column reports the average number of cuts added by the MIP solver Gurobi in the search tree. The "LK-Cuts" and "CC-($\ell$, $S$) Cuts" columns, which only appear under E-SLS+LK and CC-($\ell, S$)(indicating formulation E-SLS+CC-($\ell, S$)), respectively, are the average number of valid inequalities of Proposition 3.2 in [17] and our proposed CC-($\ell$, $S$) inequalities added to E-SLS.

First, we compare the performance between the original formulation SLS and the extended formulation E-SLS in Tables 1 and 2. The size of the instances we used here are small enough so that both of the formulations can solve all the instances to optimality. These results indicate that these instances can be solved significantly faster using formulation E-SLS than with SLS. Therefore, we do not report results with SLS in further experiments.

Table 1: Comparison between SLS and E-SLS

| Instances | | SLS | | E-SLS | |
|---|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | Time | Nodes | Time | Nodes |
| (0.1, 30, 100) | 100 | 506.7 | 70721 | 0.9 | 387 |
| | 200 | 159.1 | 22076 | 0.7 | 338 |
| | 500 | 103.8 | 15097 | 0.7 | 217 |
| | 1000 | 46.3 | 5968 | 0.5 | 100 |

Table 2: Comparison between SLS and E-SLS

| Instances | | SLS | | | E-SLS | | |
|---|---|---|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | LP Gap (%) | Root Gap (%) | Cuts | LP Gap (%) | Root Gap (%) | Cuts |
| (0.1, 30, 100) | 100 | 41.06 | 14.28 | 1497 | 12.26 | 2.38 | 189 |
| | 200 | 41.30 | 14.63 | 915 | 14.70 | 2.54 | 130 |
| | 500 | 41.01 | 19.01 | 807 | 15.69 | 3.30 | 124 |
| | 1000 | 36.30 | 16.40 | 532 | 14.65 | 0.72 | 73 |

In Tables 3-5, we present the computational results comparing formulation E-SLS, E-SLS+LK and E-SLS+CC-($\ell, S$) on large-scale instances. We use a relatively large value of $\varepsilon = 0.1$ in all these instances, as smaller values of $\varepsilon$ lead to easier to solve instances. The results in Table 3 indicate that E-SLS+CC-($\ell, S$) leads to much smaller computation time than E-SLS and E-SLS+LK. Most of the instances with $T = 60$ are not solved within the time limit using E-SLS and E-SLS+LK, while E-SLS+CC-($\ell, S$) solves all these instances within 1000 seconds. From Table 4 we observe that E-SLS-CC-($\ell, S$) is the most effective at closing the LP and root gap. E-SLS+LK does close significant gap compared to E-SLS, but as observed in Table 3, this does not generally translate to improved computation time or ending gap. Finally, from Table 5 we observe that the number of nodes explored is reduced significantly by adding CC-($\ell, S$) inequalities. Therefore, we can conclude that our proposed CC-($\ell, S$) inequalities are very effective. As the LK-Cuts do not improve the computational performance in our test instances, we do not display the results of E-SLS+LK in the next experiments.

Table 3: Comparison among E-SLS, E-SLS+LK and E-SLS+CC-($\ell, S$)

| Instances | | E-SLS | | E-SLS+LK | | CC-($\ell, S$) | |
|---|---|---|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | Time | Gap | Time | Gap | Time | Gap |
| (0.1, 30, 500) | 100 | 8.1 | 0.00 | 7.9 | 0.00 | 2.7 | 0.00 |
| | 200 | 25.7 | 0.00 | 29.9 | 0.00 | 8.9 | 0.00 |
| | 500 | 7.7 | 0.00 | 8.4 | 0.00 | 5.7 | 0.00 |
| | 1000 | 2.0 | 0.00 | 2.4 | 0.00 | 2.1 | 0.00 |
| (0.1, 30, 1000) | 100 | 15.6 | 0.00 | 24.1 | 0.00 | 5.4 | 0.00 |
| | 200 | 41.6 | 0.00 | 67.4 | 0.00 | 12.6 | 0.00 |
| | 500 | 12.1 | 0.00 | 18.3 | 0.00 | 8.9 | 0.00 |
| | 1000 | 3.6 | 0.00 | 7.1 | 0.00 | 3.5 | 0.00 |
| (0.1, 30, 2000) | 100 | 158.4 | 0.00 | 212.7 | 0.00 | 21.7 | 0.00 |
| | 200 | 93.7 | 0.00 | 141.1 | 0.00 | 19.2 | 0.00 |
| | 500 | 29.2 | 0.00 | 63.5 | 0.00 | 9.7 | 0.00 |
| | 1000 | 9.7 | 0.00 | 14.4 | 0.00 | 7.4 | 0.00 |
| (0.1, 60, 500) | 100 | [1]3327.9 | 1.17 | * | 1.74 | 590.7 | 0.00 |
| | 200 | * | 4.00 | * | 5.40 | 1775.4 | 0.00 |
| | 500 | [2]2244.3 | 0.89 | [1]1711.3 | 1.71 | 497.4 | 0.00 |
| | 1000 | 488.6 | 0.00 | 729.6 | 0.00 | 217.0 | 0.00 |
| (0.1, 60, 1000) | 100 | * | 1.58 | * | 2.66 | 487.4 | 0.00 |
| | 200 | * | 2.93 | * | 3.10 | 526.3 | 0.00 |
| | 500 | 1990.6 | 0.00 | * | 2.83 | 457.9 | 0.00 |
| | 1000 | 371.8 | 0.00 | 821.2 | 0.00 | 193.8 | 0.00 |
| (0.1, 60, 2000) | 100 | * | 2.36 | * | 2.63 | 675.7 | 0.00 |
| | 200 | * | 3.68 | * | 4.35 | 997.2 | 0.00 |
| | 500 | [2]2522.3 | 0.82 | [1]3352.3 | 1.89 | 684.1 | 0.00 |
| | 1000 | 1223.4 | 0.00 | [2]2252.4 | 0.70 | 598.5 | 0.00 |

Tables 6-8 show the comparison between E-SLS and E-SLS+CC-($\ell, S$) for varying values of $\varepsilon$. As in Table 3 we see that the instances with setup/holding cost ratio $f = 200$ are the hardest to solve, we only do experiments on instances with $f = 200$ here. Although not all instances are solved using E-SLS+CC-($\ell, S$) for the instances with $\varepsilon = 0.2$, the ending gap is still significantly smaller than that with E-SLS. These results further demonstrate the effectiveness of the CC-($\ell, S$) inequalities.

In Tables 9-11, we present results comparing the performance of E-SLS, E-SLS+CC-($\ell, S$) and NE-SLS on instances having more time periods but fewer scenarios, for instances which satisfy the modified Wagner-Whitin condition. We observe that both E-SLS+CC-($\ell, S$) and NE-SLS perform much better than E-SLS, so we do not display the detailed results for E-SLS in Table 11. The comparison between E-SLS+CC-($\ell, S$) and NE-SLS is interesting. NE-SLS performs better when $\varepsilon$ equals to 0.05 or 0.1. In Table 11, we can see that NE-SLS has smaller root gap and explores many fewer nodes to reach optimality than E-SLS+CC-($\ell, S$) when $\epsilon$ is smaller and when $f$ is larger. Despite having better root relaxation gap for larger $f$, when $\epsilon = 0.2$ NE-SLS still has worse ending optimality gap than E-SLS+CC-($\ell, S$) due to the increased size of the formulation.

Due to its size, formulation NE-SLS can be less effective on instances with more scenarios. To demonstrate the limits of NE-SLS, in Tables 12-13, we compare the results of E-SLS+CC-($\ell, S$) and NE-SLS on instances with more scenarios and fixed $\varepsilon = 0.05$. Generally, NE-SLS continues to have better LP and root gaps, particularly when $f$ is larger, but its performance relative to E-SLS+CC-($\ell, S$) degrades as the number of scenarios increases. Thus, we conclude that NE-SLS and E-SLS+CC-($\ell, S$) are complementary: NE-SLS

Table 4: Comparison among E-SLS, E-SLS+LK and E-SLS+CC-($\ell, S$)

| Instances | | E-SLS | | E-SLS+LK | | CC-($\ell, S$) | |
|---|---|---|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | LP Gap (%) | Root Gap (%) | LP Gap (%) | Root Gap (%) | LP Gap (%) | Root Gap (%) |
| (0.1, 30, 500) | 100 | 12.08 | 5.07 | 7.32 | 3.94 | 3.67 | 1.48 |
| | 200 | 17.97 | 7.17 | 10.59 | 5.42 | 4.94 | 2.36 |
| | 500 | 20.54 | 8.17 | 10.07 | 6.82 | 5.53 | 2.33 |
| | 1000 | 15.29 | 0.03 | 5.83 | 1.76 | 4.27 | 0.03 |
| (0.1, 30, 1000) | 100 | 11.61 | 4.78 | 7.95 | 4.69 | 3.38 | 1.45 |
| | 200 | 17.91 | 7.95 | 10.79 | 6.46 | 5.14 | 2.68 |
| | 500 | 19.73 | 8.24 | 9.63 | 6.55 | 5.14 | 1.75 |
| | 1000 | 14.27 | 0.08 | 5.63 | 0.23 | 3.40 | 0.03 |
| (0.1, 30, 2000) | 100 | 14.37 | 6.90 | 10.83 | 6.60 | 3.83 | 2.00 |
| | 200 | 17.60 | 8.10 | 11.99 | 7.75 | 4.73 | 2.38 |
| | 500 | 19.30 | 9.24 | 10.95 | 7.30 | 4.93 | 0.74 |
| | 1000 | 16.30 | 0.03 | 9.64 | 1.87 | 4.69 | 0.04 |
| (0.1, 60, 500) | 100 | 14.26 | 7.11 | 11.40 | 7.01 | 4.09 | 2.95 |
| | 200 | 20.56 | 12.23 | 16.84 | 11.28 | 4.99 | 3.66 |
| | 500 | 25.25 | 14.38 | 17.23 | 10.72 | 5.64 | 3.69 |
| | 1000 | 27.52 | 15.58 | 19.19 | 12.86 | 6.29 | 4.11 |
| (0.1, 60, 1000) | 100 | 13.56 | 6.49 | 10.98 | 6.88 | 3.81 | 2.41 |
| | 200 | 17.81 | 9.34 | 14.05 | 9.56 | 4.38 | 3.02 |
| | 500 | 23.61 | 13.13 | 16.47 | 10.94 | 5.69 | 3.41 |
| | 1000 | 26.01 | 13.12 | 16.77 | 11.46 | 5.95 | 3.75 |
| (0.1, 60, 2000) | 100 | 13.12 | 6.82 | 11.04 | 6.48 | 3.42 | 2.00 |
| | 200 | 17.60 | 10.13 | 14.16 | 9.49 | 4.75 | 3.03 |
| | 500 | 22.66 | 11.86 | 17.11 | 11.15 | 5.18 | 3.43 |
| | 1000 | 25.70 | 13.50 | 18.28 | 11.06 | 6.37 | 3.97 |

Table 5: Comparison among E-SLS, E-SLS+LK and E-SLS+CC-$(\ell, S)$

| Instances | | E-SLS | E-SLS+LK | | CC-$(\ell, S)$ | |
|---|---|---|---|---|---|---|
| $(\varepsilon, T, m)$ | $f$ | Nodes | LK Cuts | Nodes | CC-$(\ell, S)$ Cuts | Nodes |
| (0.1, 30, 500) | 100 | 2437 | 246 | 1025 | 352 | 367 |
| | 200 | 5903 | 309 | 2847 | 486 | 835 |
| | 500 | 1224 | 325 | 862 | 511 | 467 |
| | 1000 | 263 | 281 | 298 | 339 | 137 |
| (0.1, 30, 1000) | 100 | 1936 | 203 | 2720 | 308 | 381 |
| | 200 | 5015 | 316 | 4103 | 420 | 684 |
| | 500 | 1023 | 368 | 898 | 482 | 317 |
| | 1000 | 290 | 320 | 453 | 336 | 54 |
| (0.1, 30, 2000) | 100 | 5598 | 200 | 7983 | 485 | 607 |
| | 200 | 6253 | 287 | 4700 | 475 | 639 |
| | 500 | 1170 | 377 | 973 | 445 | 119 |
| | 1000 | 266 | 232 | 279 | 325 | 234 |
| (0.1, 60, 500) | 100 | >144338 | 415 | >43377 | 1505 | 23324 |
| | 200 | >85451 | 476 | >22267 | 2720 | 26140 |
| | 500 | >101084 | 805 | >47486 | 3064 | 4984 |
| | 1000 | 29255 | 634 | 14411 | 2864 | 2698 |
| (0.1, 60, 1000) | 100 | >60912 | 433 | >26711 | 1590 | 6876 |
| | 200 | >45426 | 495 | >24349 | 2016 | 5042 |
| | 500 | 47938 | 688 | >20537 | 2714 | 2311 |
| | 1000 | 10803 | 722 | 7624 | 2696 | 997 |
| (0.1, 60, 2000) | 100 | >23062 | 368 | >16805 | 1538 | 4591 |
| | 200 | >22749 | 497 | >11080 | 2027 | 4617 |
| | 500 | >26921 | 532 | >14056 | 2416 | 2468 |
| | 1000 | 13575 | 580 | >10040 | 2613 | 2287 |

Table 6: Comparison between E-SLS and E-SLS+CC-$(\ell, S)$

| Instances | | E-SLS | | CC-$(\ell, S)$ | |
|---|---|---|---|---|---|
| $(T, m, f)$ | $\varepsilon$ | Time | Gap | Time | Gap |
| (60, 1000, 200) | 0.05 | 1953.4 | 0.00 | 111.0 | 0.00 |
| | 0.1 | * | 2.93 | 526.3 | 0.00 |
| | 0.2 | * | 5.41 | [1]377.5 | 0.71 |

Table 7: Comparison between E-SLS and E-SLS+CC-$(\ell, S)$

| Instances | | E-SLS | | CC-$(\ell, S)$ | | |
|---|---|---|---|---|---|---|
| $(T, m, f)$ | $\varepsilon$ | Nodes | Cuts | Nodes | CC-$(\ell, S)$ Cuts | Cuts |
| (60, 1000, 200) | 0.05 | 97697 | 1559 | 2427 | 1809 | 756 |
| | 0.1 | >45426 | 1693 | 5042 | 2016 | 1068 |
| | 0.2 | >22597 | 1464 | >7185 | 2122 | 1771 |

Table 8: Comparison between E-SLS and E-SLS+CC-($\ell, S$)

| Instances | | E-SLS | | CC-($\ell, S$) | |
|---|---|---|---|---|---|
| ($T$, $m$, $f$) | $\varepsilon$ | LP.Gap (%) | R.Gap (%) | LP.Gap (%) | R.Gap (%) |
| (60, 1000, 200) | 0.05 | 17.55 | 7.05 | 4.21 | 2.01 |
| | 0.1 | 17.81 | 9.34 | 4.38 | 3.02 |
| | 0.2 | 18.67 | 11.44 | 4.91 | 3.62 |

Table 9: Comparison of the average time among E-SLS, E-SLS+CC-($\ell, S$) and NE-SLS

| Instances | | E-SLS | CC-($\ell, S$) | NE-SLS |
|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | | | |
| (0.05, 90, 100) | 100 | 1370.2 | 151.7 | 95.6 |
| | 200 | [1]1824.0 | 1503.5 | 184.3 |
| | 500 | [2]1900.3 | 951.6 | 70.7 |
| | 1000 | 1133.1 | 775.7 | 34.8 |
| (0.1, 90, 100) | 100 | * | [2]1054.0 | [2]1205.5 |
| | 200 | * | * | 2472.8 |
| | 500 | * | [1]1784.5 | 1375.2 |
| | 1000 | * | 1116.8 | 388.2 |
| (0.2, 90, 100) | 100 | * | * | * |
| | 200 | * | * | * |
| | 500 | * | * | * |
| | 1000 | * | [2]3072.6 | * |

Table 10: Comparison of the average ending optimality gap among E-SLS, E-SLS+CC-($\ell, S$) and NE-SLS

| Instances | | E-SLS | CC-($\ell, S$) | NE-SLS |
|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | | | |
| (0.05, 90, 100) | 100 | 0.00 | 0.00 | 0.00 |
| | 200 | 0.46 | 0.00 | 0.00 |
| | 500 | 0.21 | 0.00 | 0.00 |
| | 1000 | 0.00 | 0.00 | 0.00 |
| (0.1, 90, 100) | 100 | 2.11 | 0.19 | 0.70 |
| | 200 | 4.04 | 1.10 | 0.00 |
| | 500 | 5.33 | 0.65 | 0.00 |
| | 1000 | 2.26 | 0.00 | 0.00 |
| (0.2, 90, 100) | 100 | 5.45 | 1.46 | 6.92 |
| | 200 | 8.85 | 2.46 | 7.20 |
| | 500 | 7.86 | 1.22 | 2.55 |
| | 1000 | 7.49 | 0.24 | 2.31 |

Table 11: Comparison between E-SLS+CC-($\ell, S$) and NE-SLS

| Instances | | CC-($\ell, S$) | | | NE-SLS | | |
|---|---|---|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | Nodes | LP Gap (%) | Root Gap (%) | Nodes | LP Gap (%) | Root Gap (%) |
| (0.05, 90, 100) | 100 | 7072 | 3.24 | 1.19 | 296 | 2.01 | 0.71 |
| | 200 | 31613 | 4.00 | 1.92 | 1308 | 2.60 | 1.15 |
| | 500 | 4400 | 4.30 | 2.24 | 164 | 1.62 | 0.51 |
| | 1000 | 3204 | 4.31 | 2.02 | 1 | 1.01 | 0.00 |
| (0.1, 90, 100) | 100 | >79882 | 3.82 | 2.25 | >7029 | 4.44 | 3.01 |
| | 200 | >33572 | 4.51 | 2.94 | 7669 | 4.43 | 2.79 |
| | 500 | >25015 | 5.21 | 3.61 | 2714 | 3.63 | 1.93 |
| | 1000 | 4671 | 5.05 | 3.16 | 1 | 2.03 | 0.00 |
| (0.2, 90, 100) | 100 | >45203 | 4.24 | 3.18 | >1 | 5.88 | 4.84 |
| | 200 | >20914 | 5.15 | 4.24 | >2 | 6.85 | 5.50 |
| | 500 | >20551 | 5.77 | 3.78 | >465 | 3.95 | 2.59 |
| | 1000 | >31086 | 5.94 | 4.78 | >394 | 3.84 | 2.47 |

generally performs better if the number of scenarios and $\varepsilon$ are small, and $f$ is large, and E-SLS+CC-($\ell, S$) generally performs better otherwise.

Table 12: Comparison between E-SLS+CC-($\ell, S$) and NE-SLS

| Instances | | CC-($\ell, S$) | | NE-SLS | |
|---|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | Time | Gap | Time | Gap |
| (0.05, 90, 200) | 100 | 1423.7 | 0.00 | [2]1397.8 | 0.41 |
| | 200 | [1]2521.1 | 0.33 | 2341.9 | 0.00 |
| | 500 | [1]1946.4 | 0.95 | 1750.1 | 0.00 |
| | 1000 | [1]1345.6 | 1.09 | 1007.4 | 0.00 |
| (0.05, 90, 300) | 100 | [2]1347.2 | 0.41 | * | 2.34 |
| | 200 | * | 1.27 | * | 2.54 |
| | 500 | [1]2104.9 | 0.75 | * | 1.65 |
| | 1000 | 1785.3 | 0.00 | [2]1886.8 | 0.46 |

# 6 Conclusions

We explore the static chance-constrained lot-sizing problem in this paper. We first apply an existing chance constrained formulation to reformulate the original formulation into an extended formulation E-SLS. We then propose the CC-($\ell, S$) inequalities that further strengthen E-SLS. Finally, under a modified Wagner-Whitin condition, we derive an optimality condition which forms the basis of a new extended formulation NE-SLS. The computational experiments show that our CC-($\ell, S$) inequalities and the new formulation NE-SLS are quite effective in solving the stochastic lot-sizing problem.

There are a couple directions that could be explored related to this work. Rather than adding the CC-($\ell, S$) inequalities to formulation E-SLS, is it possible to derive equivalently strong valid inequalities for the original formulation E-SLS? In our experience, we found that for most of the instances we tested, formulation NE-SLS has the same optimal solutions as SLS, even when Assumption 1 does not hold. Thus another interesting direction to explore is to see if it can be shown that NE-SLS is a valid formulation under weaker assumptions.

Table 13: Comparison between E-SLS+CC-$(\ell, S)$ and NE-SLS

| Instances | | CC-$(\ell, S)$ | | | NE-SLS | | |
|---|---|---|---|---|---|---|---|
| $(\varepsilon, T, m)$ | $f$ | Nodes | LP Gap (%) | Root Gap (%) | Nodes | LP Gap (%) | Root Gap (%) |
| (0.05, 90, 200) | 100 | 20034 | 3.57 | 1.85 | 7916 | 3.62 | 2.26 |
| | 200 | 51377 | 4.28 | 2.54 | 4196 | 3.68 | 2.26 |
| | 500 | 26374 | 5.39 | 3.49 | 3862 | 3.01 | 1.65 |
| | 1000 | 16883 | 5.87 | 4.17 | 1568 | 2.97 | 1.34 |
| (0.05, 90, 300) | 100 | 20600 | 3.80 | 1.99 | 752 | 3.89 | 2.78 |
| | 200 | 20823 | 4.44 | 2.77 | 1028 | 4.13 | 2.89 |
| | 500 | 21982 | 5.15 | 3.23 | 609 | 3.28 | 2.16 |
| | 1000 | 11451 | 5.46 | 3.38 | 416 | 2.57 | 1.22 |

# References


[1] Abdi A, Fukasawa R (2016) On the mixing set with a knapsack constraint. Mathematical Programming 157(1):191–217

[2] Ahmed S, Sahinidis NV (2003) An approximation scheme for stochastic integer programs arising in capacity expansion. Operations Research 51(3):461–471

[3] Ahmed S, King AJ, Parija G (2003) A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. Journal of Global Optimization 26(1):3–24

[4] Atamtürk A, Muñoz JC (2004) A study of the lot-sizing polytope. Mathematical Programming 99(3):443–465, DOI 10.1007/s10107-003-0465-8

[5] Atamtürk A, Nemhauser GL, Savelsbergh MW (2000) The mixed vertex packing problem. Mathematical Programming 89(1):35–53

[6] Barany I, Van Roy T, Wolsey LA (1984) Uncapacitated lot-sizing: The convex hull of solutions, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 32–43

[7] Beraldi P, Ruszczyński A (2002) A branch and bound method for stochastic integer problems under probabilistic constraints. Optimization Methods and Software 17(3):359–382

[8] Charnes A, Cooper WW (1963) Deterministic equivalents for optimizing and satisficing under chance constraints. Operations Research 11(1):18–39

[9] Charnes A, Cooper WW, Symonds GH (1958) Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil. Management Science 4(3):235–263

[10] Guan Y, Miller AJ (2008) Polynomial-time algorithms for stochastic uncapacitated lot-sizing problems. Operations Research 56(5):1172–1183

[11] Guan Y, Ahmed S, Nemhauser GL, Miller AJ (2006) A branch-and-cut algorithm for the stochastic uncapacitated lot-sizing problem. Mathematical Programming 105(1):55–84

[12] Günlük O, Pochet Y (2001) Mixing mixed-integer inequalities. Mathematical Programming 90(3):429–457

[13] Huang K, Küçükyavuz S (2008) On stochastic lot-sizing problems with random lead times. Operations Research Letters 36(3):303 – 308

[14] Jiang R, Guan Y (2011) An o(n2)-time algorithm for the stochastic uncapacitated lot-sizing problem with random lead times. Operations Research Letters 39(1):74 – 77

[15] Küçükyavuz S (2012) On mixing sets arising in chance-constrained programming. Mathematical Programming 132(1):31–56

[16] Küçükyavuz S, Pochet Y (2009) Uncapacitated lot sizing with backlogging: the convex hull. Mathematical Programming 118(1):151–175

[17] Liu X, Küçükyavuz S (2018) A polyhedral study of the static probabilistic lot-sizing problem. Annals of Operations Research 261(1):233–254

[18] Luedtke J, Ahmed S, Nemhauser GL (2010) An integer programming approach for linear programs with probabilistic constraints. Mathematical Programming 122(2):247–272

[19] Pochet Y, Wolsey LA (1988) Lot-size models with backlogging: Strong reformulations and cutting planes. Mathematical Programming 40(1):317–335

[20] Pochet Y, Wolsey LA (1994) Polyhedra for lot-sizing with wagner—whitin costs. Mathematical Programming 67(1):297–323

[21] Pochet Y, Wolsey LA (2006) Production Planning by Mixed Integer Programming, New York: Springer, pp 207–234

[22] Wagner HM, Whitin TM (1958) Dynamic version of the economic lot size model. Management Science 5(1):89–96

[23] Zhang M, Küçükyavuz S, Goel S (2014) A branch-and-cut method for dynamic decision making under joint chance constraints. Management Science 60(5):1317–1333

[24] Zhao M, Huang K, Zeng B (2017) A polyhedral study on chance constrained program with random right-hand side. Mathematical Programming 166(1):19–64

# Appendix

Tables 14-16 present results comparing the performance of ELS+LK using $\beta = 0.0001$ and $\beta = 0.02$ as the stopping condition when generating the LK-Cuts. We can see that for most instances, when $\beta = 0.02$ E-SLS+LK performs better in terms of solution time and ending optimality gap, and using $\beta = 0.0001$ only performs better on these metrics in some particular cases. However, when $\beta = 0.0001$ ELS+LK yields a better LP gap and root gap. We find that the challenge in effectively using the LK-Cuts is that reducing the gap requires adding a large number of LK-Cuts, so that the overall solution time is increased if they are added too aggressively. On the other hand, although the performance using $\beta = 0.02$ is better, the impact of the cuts in that case becomes less significant because the gap reduction is not so great.

Table 14: Comparison between different criterion for E-SLS+LK

| Instances | | criterion = 0.0001 | | criterion = 0.02 | |
|---|---|---|---|---|---|
| $(\varepsilon, T, m)$ | $f$ | Time | Gap | Time | Gap |
| (0.1, 30, 500) | 100 | 15.7 | 0.00 | 7.9 | 0.00 |
| | 200 | 67.7 | 0.00 | 29.9 | 0.00 |
| | 500 | 49.8 | 0.00 | 8.4 | 0.00 |
| | 1000 | 12.8 | 0.00 | 2.4 | 0.00 |
| (0.1, 30, 1000) | 100 | 49.2 | 0.00 | 24.1 | 0.00 |
| | 200 | 352.9 | 0.00 | 67.4 | 0.00 |
| | 500 | 207.6 | 0.00 | 18.3 | 0.00 |
| | 1000 | 46.0 | 0.00 | 7.1 | 0.00 |
| (0.1, 30, 2000) | 100 | 1520.5 | 0.00 | 212.7 | 0.00 |
| | 200 | 1104.1 | 0.00 | 141.1 | 0.00 |
| | 500 | 566.3 | 0.00 | 63.5 | 0.00 |
| | 1000 | 318.7 | 0.00 | 14.4 | 0.00 |
| (0.1, 60, 500) | 100 | * | 1.35 | * | 1.74 |
| | 200 | * | 3.31 | * | 5.40 |
| | 500 | [1]1932.7 | 2.22 | [1]1711.3 | 1.71 |
| | 1000 | 2752.0 | 0.00 | 729.6 | 0.00 |
| (0.1, 60, 1000) | 100 | * | 2.44 | * | 2.66 |
| | 200 | * | 3.27 | * | 3.10 |
| | 500 | * | 4.74 | * | 2.83 |
| | 1000 | * | 3.46 | 821.2 | 0.00 |
| (0.1, 60, 2000) | 100 | * | 3.78 | * | 2.63 |
| | 200 | * | 5.93 | * | 4.35 |
| | 500 | * | 8.21 | [1]3352.3 | 1.89 |
| | 1000 | * | 9.21 | [2]2252.4 | 0.70 |

Table 15: Comparison between different values of $\beta$ for E-SLS+LK

| Instances | | $\beta = 0.0001$ | | | $\beta = 0.02$ | | |
|---|---|---|---|---|---|---|---|
| ($\varepsilon$, $T$, $m$) | $f$ | Nodes | LK Cuts | Cuts | Nodes | LK Cuts | Cuts |
| (0.1, 30, 500) | 100 | 938 | 1133 | 331 | 1025 | 246 | 340 |
| | 200 | 2117 | 1694 | 528 | 2847 | 309 | 515 |
| | 500 | 1438 | 1956 | 395 | 862 | 325 | 211 |
| | 1000 | 542 | 1676 | 159 | 298 | 281 | 96 |
| (0.1, 30, 1000) | 100 | 1093 | 1330 | 398 | 2720 | 203 | 423 |
| | 200 | 2984 | 2543 | 931 | 4103 | 316 | 572 |
| | 500 | 2187 | 3358 | 514 | 898 | 368 | 250 |
| | 1000 | 183 | 4063 | 114 | 453 | 320 | 175 |
| (0.1, 30, 2000) | 100 | 6263 | 2473 | 671 | 7983 | 200 | 739 |
| | 200 | 3009 | 4154 | 942 | 4700 | 287 | 672 |
| | 500 | 544 | 5757 | 394 | 973 | 377 | 454 |
| | 1000 | 420 | 6615 | 202 | 279 | 232 | 161 |
| (0.1, 60, 500) | 100 | >20628 | 3709 | 2181 | >43377 | 415 | 2650 |
| | 200 | >8492 | 5195 | 2262 | >22267 | 476 | 1527 |
| | 500 | >5867 | 6274 | 1354 | >47486 | 805 | 2285 |
| | 1000 | 7345 | 6563 | 579 | 14411 | 634 | 1421 |
| (0.1, 60, 1000) | 100 | >7635 | 4196 | 1147 | >26711 | 433 | 2139 |
| | 200 | >2848 | 5400 | 1495 | >24349 | 495 | 2757 |
| | 500 | >1824 | 7376 | 662 | >20537 | 688 | 2808 |
| | 1000 | >2386 | 7911 | 742 | 7624 | 722 | 1193 |
| (0.1, 60, 2000) | 100 | >1718 | 4846 | 922 | >16805 | 368 | 2382 |
| | 200 | >927 | 7063 | 1360 | >11080 | 497 | 2756 |
| | 500 | >554 | 10533 | 1469 | >14056 | 532 | 1956 |
| | 1000 | >572 | 14114 | 1074 | >10040 | 580 | 1560 |

Table 16: Comparison between different values of $\beta$ for E-SLS+LK

| Instances | | $\beta = 0.0001$ | | $\beta = 0.02$ | |
|---|---|---|---|---|---|
| $(\varepsilon, T, m)$ | $f$ | LP.Gap (%) | R.Gap (%) | LP.Gap (%) | R.Gap (%) |
| (0.1, 30, 500) | 100 | 4.28 | 2.60 | 7.32 | 3.94 |
| | 200 | 5.70 | 4.35 | 10.59 | 5.42 |
| | 500 | 6.00 | 4.72 | 10.07 | 6.82 |
| | 1000 | 4.62 | 0.03 | 5.83 | 1.76 |
| (0.1, 30, 1000) | 100 | 4.70 | 3.16 | 7.95 | 4.69 |
| | 200 | 6.02 | 4.69 | 10.79 | 6.46 |
| | 500 | 5.31 | 4.13 | 9.63 | 6.55 |
| | 1000 | 3.81 | 0.04 | 5.63 | 0.23 |
| (0.1, 30, 2000) | 100 | 5.91 | 4.89 | 10.83 | 6.60 |
| | 200 | 6.15 | 5.27 | 11.99 | 7.75 |
| | 500 | 6.16 | 4.66 | 10.95 | 7.30 |
| | 1000 | 5.78 | 1.83 | 9.64 | 1.87 |
| (0.1, 60, 500) | 100 | 6.17 | 4.66 | 11.40 | 7.01 |
| | 200 | 7.65 | 6.22 | 16.84 | 11.28 |
| | 500 | 7.43 | 6.34 | 17.23 | 10.72 |
| | 1000 | 7.96 | 7.06 | 19.19 | 12.86 |
| (0.1, 60, 1000) | 100 | 6.19 | 5.25 | 10.98 | 6.88 |
| | 200 | 7.00 | 5.69 | 14.05 | 9.56 |
| | 500 | 8.06 | 7.11 | 16.47 | 10.94 |
| | 1000 | 8.50 | 7.34 | 16.77 | 11.46 |
| (0.1, 60, 2000) | 100 | 6.40 | 5.34 | 11.04 | 6.48 |
| | 200 | 7.43 | 6.56 | 14.16 | 9.49 |
| | 500 | 7.06 | 6.45 | 17.11 | 11.15 |
| | 1000 | 8.14 | 7.43 | 18.28 | 11.06 |