Flow-Based Stochastic Guaranteed Service Models for the Outsourcing-Aware Base-Stock Policy

Dominik Kamp and Jörg Rambau

May 2, 2023

Abstract

The Stochastic Guaranteed Service Model with Demand Propagation SGSM-DP is an MILP model that can pro-actively decide on safety stocks and outsourced quantities in a multi-echelon supply chain distribution network in the presence of a discrete demand random distribution and immediate outsourcing options. However, on the one hand, it can only handle small networks, on the other hand its objective does not precisely capture the holding costs for the expected inventory levels. In this paper, new models are presented that overcome both short-comings: The F-SGSM-DP is essentially equivalent to the SGSM-DP but can handle large networks fast. The F-SGSM-DF is an improved model compared to the F-SGSM-DP with a similar scalability but with a more accurate objective. Both new models are based on a representation of a logical material flow with a flow-design in a certain time-expanded network. Numerical results on randomized data confirm that the F-SGSM-DF is faster to compute, predicts better the realized cost, and lead to substantial cost savings in dynamic simulations.

1 Introduction

This paper studies MILP (mixed-integer linear programming) formulations for the Stochastic Guaranteed Service Model with Demand Propagation (SGSM-DP), introduced by Löhnert and Rambau (2018), an extension of the earlier Stochastic Guaranteed Service Model (SGSM), introduced by Rambau and Schade (2014). In contrast to the original Guaranteed Service Model (GSM) (Graves and Willems 2000, Lesnaia 2004, Eruguz et al. 2016), the SGSM and the SGSM-DP support stochastic demand in terms of a discrete demand distribution and an immediate outsourcing option at each node. Thus, the SGSM and the SGSM-DP can suggest to reduce safety stock at the cost of occasional outsourcing in high-demand scenarios. Excess demand must be outsourced, since backlogging is not allowed, as in the original GSM. In contrast to the SGSM, the SGSM-DP handles the demand propagation inside the network by endogenous calculations. This is important, since outsourcing decisions reduce the internal demands upstream. A first MILP formulation of the SGSM-DP was given by Löhnert and Rambau (2018). However, the best models so far were only able to deal with distribution networks up to at most 20 nodes for few scenarios. Moreover, its objective function measures the holding cost for the whole base-stock levels, not for the expected inventory levels.

This is a special topic in the field of Multi Echelon Inventory Optimization (MEIO). MEIO, in general, asks for a supply chain network which policies should be used in each node for ordering, producing, and supplying material so as to optimize an objective function. A survey with a classification of MEIO problem settings and solution methods was provided by de Kok et al. (2018). Graves

Research is supported by the German Research Foundation (DFG), project "The Stochastic Guaranteed Service Model with Demand Propagation" (RA 1033/3-1) and the Bayreuth research center for modeling and simulation (MODUS).

and Willems (2000) compare the two basic modelling paradigms of the Stochastic Service Model (SSM) and the Guaranteed Service Model (GSM). It is safe to say that the SSM in its exact form is only tractable for highly restricted network structures and/or parameter settings. For more complicated structures, one has to resort to approximations. The GSM as such is already an approximation of the original stochastic MEIO problem. It has successfully been used for complicated network structures in practical applications (Eruguz et al. 2016).

This paper deals with the GSM-branch of research. Its contributions are twofold: first, a new MILP formulation for the SGSM-DP, logically equivalent to the one by Löhnert and Rambau (2018), is presented that can solve instances up to 50 nodes in reasonable time (details below); second, in a modified model, a new objective function is incorporated that better reflects the expected inventory holding cost over time, so that in model-independent simulations on random data the average holding and outsourcing cost is significantly reduced (details below). Parts of the first contribution have been presented without detailed proofs in an extended abstract of the MATHMOD2022 conference already by Kamp and Rambau (2022).

The paper is organized as follows. In Section 2 notations are introduced and a formal description of an outsourcing-aware base-stock policy is provided. In Section 3.1 an all new flow-design formulation F-SGSM-DP is presented together with a formal description under which conditions there is a logical equivalence to the SGSM-DP. Section 3.2 introduces a non-equivalent model F-SGSM-DF in which the objective function better reflects the holding cost of the expected inventory level rather than the base-stock level. Additionally, in this modification, the evaluation of recourse costs is adjusted to better describe the actual outsourcing cost. Computational results are presented in Section 4. The first part on solving the models shows strong evidence for the fact that the new F-SGSM-DP is a much tighter MILP than the known one by Löhnert and Rambau (2018). While with a time limit of 1000 seconds the SGSM-DP can only be solved for instances of up to 20 nodes, the F-SGSM-DP can handle networks with as many as 400 nodes. The second part on the dynamic performance of the model-optimal decision shows by simulation that the long-term average cost is better predicted and reduced by the modified F-SGSM-DF. Finally, Section 5 presents the conclusions.

2 Problem Statement

We consider the task of allocating safety stocks in a multi-echelon distribution network for one type of product with a stochastic end customer demand represented by finite scenarios. Operations follow a periodic-review base-stock policy with guaranteed service times in the presence of expensive but immediate and unbounded outsourcing options at all inventories. The goal is to find constant base-stock levels, guaranteed outbound service times, and upper order limits that generate the minimum expected total of holding and outsourcing costs. The motivation to specify upper order limits is that this way the endogenous demand propagation can be reduced deliberately in anticipation of possible future outsourcing. In contrast to this, directly deciding the outsourced rate in each node could occasionally lead to unnecessary outsourcing in the presence of sufficient stock.

Formally, we investigate an inventory problem of type

$$\mathtt{n}^{\mathrm{ech}},\mathtt{D}^{\mathrm{net}}|\mathtt{I}^{\mathrm{cap}},\mathtt{C}^{\mathrm{del}}|\mathtt{U}^{\mathrm{dem}},\mathtt{G}^{\mathrm{cus}}|\mathtt{P}^{\mathrm{res}}|\mathtt{C}^{\mathrm{obj}}$$

in the condensed notation by de Kok et al. (2018).

More specifically, let a supply chain network be represented by a finite graph $G := (N, A)$. As in the traditional GSM-framework by Graves and Willems (2000) each node $i \in N$ represents a stock point where a unique type of product can be stored at a holding cost rate of $H_i \in \mathbb{Q}_{\geq 0}$. Every arc $(j, i) \in A \subseteq N \times N$ represents the assembly dependency that a product from j is required to generate a product at i. This operation is assumed to take a deterministic number of periods given by the lead time $T_i \in \mathbb{Z}_{>0}$. Nodes without predecessors are supplied externally. The demand set $D \subseteq N$ consists of the nodes that face non-trivial random demand by external customers. Each customer order must be satisfied within a time window defined by the outbound service time $S_i \in \mathbb{Z}_{\geq 0}$. In contrast to the original assumption by Simpson Jr. (1958) of exogenously defined service levels for indeterminate demand distributions, we are interested in an endogenous determination of optimal service levels for an explicit demand distribution. Therefore, a discrete demand distribution is assumed, which is joint among the demand nodes. This non-trivial demand distribution is determined by a finite set of scenarios Ω, probabilities $P^\omega \in \mathbb{Q}_{<1}^{>0}$ for all scenarios $\omega \in \Omega$, and conditional demand rates $D_i^\omega \in \mathbb{Z}_{\geq 0}$ so that for all $i \in D$ there is some $\omega \in \Omega$ with $D_i^\omega > 0$. Stochastic dependencies over time are neglected. Instead, in each scenario the joint demand sequence is assumed to be constant. Every node $i \in N$ is additionally in touch with an emergency supplier, which can immediately serve a product at compensation cost $C_i \in \mathbb{Q}_{\geq 0}$. Based on this information, the task is to find for every node scenario-independent guaranteed inbound service times $s_i \in \mathbb{Z}_{\geq 0}$ and base-stock levels $y_i \in \mathbb{Z}_{\geq 0}$ as well as scenario-dependent propagated demand rates $n_i^\omega \in \mathbb{Q}_{\geq 0}$ satisfying the service and conservation conditions in order to minimize the expected total of holding and compensation costs. In the following, for a node $i \in N$ in graph G we denote the set of predecessors by $\delta_G^-(i)$ and the set of successors by $\delta_G^+(i)$. This investigation is restricted to distribution networks.

Definition 1 (Distribution Network)**.** A *distribution network* is a finite directed graph with the properties:

(i) [Connected] In G between all nodes there are undirected paths.

(ii) [Acyclic] In G there is no closed directed path.

(iii) [Divergent] For every node $i \in N$ it is $|\delta_G^-(i)| \leq 1$.

Assumption 1(i) can always be obtained by considering each connected sub-network separately. Assumption 1(ii) is a common restriction to avoid self-dependencies by propagated production requirements. In total, assumption 1 asserts that there is a unique node without predecessors called the root node $1 \in N$ and any other node has a unique predecessor. For that reason, every assembly process can be interpreted as either an internal transportation between two nodes or an external procurement towards the root node. Let for every $i \in N$ the set of reachable nodes via directed paths be denoted by the downstream set $\varrho_G^+(i) \subseteq N$ including node i. Then for this node in each scenario $\omega \in \Omega$ the *full demand rate* is defined by

$$\Delta_i^\omega := \sum_{j \in \varrho_G^+(i) \cap D} D_j^\omega \tag{1}$$

as the sum of customer demand rates in the downstream network of node i. Furthermore, let the unique backward path from i to the root node 1 be given by the upstream set $P_i \subseteq N$ including initial node i. Then, the *root process duration* is defined by

$$\tau_i := \sum_{j \in P_i} T_j \tag{2}$$

as the sum of upstream lead times to node i. The corresponding *relevant time set* is defined by

$$K_i := \{0, \ldots, \tau_i\} \tag{3}$$

as in Kamp and Rambau (2022). Since we want to be able to optimize base-stocks *in anticipation of outsourcing*, we need to extend the base-stock policy that is usually applied in the GSM-research. If stock-outs can be compensated by immediate outsourcing, then the decision between storage and outsourcing can be based on cost considerations only.

Definition 2 (Outsourcing-Aware Base-Stock Policy)**.** The *Outsourcing-Aware Base-Stock Policy* is a periodic-review replenishment policy with the following non-negative integral parameters for each node: a base-stock level, an outbound service time, and an upper order limit per period. The outsourcing-aware base-stock policy works as follows: On an incoming order at a node, the order is registered for delivery with a delay of the outbound service time. Then, the outsourcing order is given in each period by the part of the sent quantity that exceeds the inventory level after regular deliveries have arrived. For each node in reverse topological order, the full replenishment order is determined to lift the current inventory position to the base-stock level. Here, the expected outsourcing orders up to the outbound service time are taken into account. Pending orders will arrive with a delay given by the outbound service time of the regular supplier plus the expected lead time. Finally, the propagated replenishment order is the full replenishment order bounded from below by zero and from above by the upper order limit.

The base-stock levels, the outbound service times, and the upper order limits can be determined either from optimization results (if a model provides these data like the new model in this paper) or heuristically (if a model does not provide them like the GSM with a fixed service level). In the SGSM by Rambau and Schade (2014), the full replenishment order quantities are propagated immediately since an outsourcing-implied demand reduction is not considered there. In contrast, the SGSM-DP by Löhnert and Rambau (2018) determines all required information. In the objective function of this MILP, the holding cost term does not represent the expected holding cost but the holding cost for the base-stocks. In a sense, it estimates the cost for providing enough capacity for the base-stocks and not the capital bound in the stored materials. Thus, the higher the expected propagated demand relative to the maximum propagated demand, the more the SGSM-DP overestimates the expected holding costs.

Recall, that the following model by Kamp and Rambau (2022) is a non-linear formulation of the SGSM-DP:

Problem 1 (SGSM-DP)**.** The following problem is called the *Stochastic Guaranteed Service Model with Demand Propagation*:

$$\min \sum_{i \in N} \Big(H_i \cdot y_i + C_i \cdot \sum_{\omega \in \Omega} P^{\omega} \cdot q_i^{\omega} \Big) \qquad \text{(estimated cost)} \qquad (4)$$

$$x_i - s_i + s_i^{+} = T_i \qquad \forall i \in N \qquad \text{(replenishment delay)} \qquad (5)$$

$$-s_j + s_i^{+} \le 0 \qquad \forall (i,j) \in A \qquad \text{(internal service)} \qquad (6)$$

$$s_i^{+} \le S_i \qquad \forall i \in D \qquad \text{(external service)} \qquad (7)$$

$$y_i - x_i \cdot n_i^{\omega} \ge 0 \qquad \forall i \in N, \omega \in \Omega \qquad \text{(base-stock level)} \qquad (8)$$

$$q_i^{\omega} - x_i \cdot m_i^{\omega} = 0 \qquad \forall i \in N, \omega \in \Omega \qquad \text{(demand propagation)} \qquad (9)$$

$$q_i^{\omega} - m_i^{\omega} \ge 0 \qquad \forall i \in N, \omega \in \Omega \qquad \text{(demand passage)} \qquad (10)$$

$$n_i^{\omega} + m_i^{\omega} - \sum_{j \in \delta_G^{+}(i)} n_j^{\omega} = \begin{cases} D_i^{\omega} & \text{if } i \in D \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in N, \omega \in \Omega \qquad \text{(flow conservation)} \qquad (11)$$

$$s_i, s_i^{+}, x_i, y_i \in \mathbb{Z}_{\ge 0} \qquad \forall i \in N \qquad \text{(natural anticipation)} \qquad (12)$$

$$q_i^{\omega} \in \mathbb{Z}_{\ge 0}, n_i^{\omega}, m_i^{\omega} \in \mathbb{Q}_{\ge 0} \qquad \forall i \in N, \omega \in \Omega \qquad \text{(mixed reaction)} \qquad (13)$$

For every node $i \in N$ the inbound service time s_i is required to quantify the delay between an order is propagated and the corresponding delivery depart from the regular supplier. This makes products arrive with a delay of $s_i + T_i$. Like in the traditional description by Graves and Willems (2000) for simplicity announced demand of all successors are served with a common delay given by the outbound service time s_i^{+}. These times must comply with internal service guarantees (6) and external service guarantees (7). The common time difference between departing and arriving deliveries is given by the replenishment delay x_i determined in (5). Propagated demand during this delay has to be covered from the base-stock level y_i. To keep the demand reduction by outsourcing into account, in every scenario $\omega \in \Omega$ the propagated demand rate n_i^ω needs to be determined endogenously. This is done by introducing the outsourced demand rate m_i^ω of due orders served by the local emergency supplier. It is used in (11) to balance out the demand conservation condition. For feasible propagated demand rates the tightest upper order limit is given by $\lceil \max_{\omega \in \Omega} n_i^\omega \rceil$. The base-stock must withstand the demand load, which is equal to the replenishment delay times the propagated demand rate. This is enforced by the non-linear condition (8). Finally, additional outsourced quantities q_i^ω are calculated by (9) and (10) to evaluate the objective function (4).

We are interested in special solutions of the SGSM-DP that avoid artificial service delays which are neither evaluated in the objective (4) nor implemented in the policy given by definition 2.

Definition 3 (Slack-Free Solution)**.** For distribution networks, a feasible solution of the SGSM-DP is called slack-free, iff $s_1 = 0$ for the unique root node $1 \in N$ and $s_i = s_j^+$ for every arc $(j,i) \in A$.

For the GSM and the SGSM on distribution networks with non-decreasing safety stock functions there is always a slack-free optimal solution as shown by Schade (2012, Theorem 3.3.1). However, this is no longer valid in general for the SGSM-DP. Nevertheless, these solutions are preferable because proper service time slacks in constraint (6) of the original SGSM-DP can spoil the cost evaluation. This comes from the fact, that either additional early arrival stocks will result from slacks, or some base-stock levels are unnecessarily large if slacks are resolved by delaying internal services.

The following model is a suitable restriction of the SGSM-DP. Moreover, it will establish a relation between the non-linear SGSM-DP and the flow-linearized F-SGSM-DP.

Problem 2 (I-SGSM-DP)**.** The following problem is called the *Instant Stochastic Guaranteed Service Model with Demand Propagation*:

$$\min \sum_{i \in N} \Big(H_i \cdot y_i + C_i \cdot \sum_{\omega \in \Omega} P^\omega \cdot q_i^\omega \Big) \qquad \text{(estimated cost)} \qquad (14)$$

$$x_i - s_i + s_i^+ = T_i \qquad \forall i \in N \qquad \text{(replenishment delay)} \qquad (15)$$

$$-s_j + s_i^+ = 0 \qquad \forall (i,j) \in A \qquad \text{(internal service)} \qquad (16)$$

$$s_i^+ \le S_i \qquad \forall i \in D \qquad \text{(external service)} \qquad (17)$$

$$s_i \le \tau_i - T_i \qquad \forall i \in N \qquad \text{(relevance bound)} \qquad (18)$$

$$y_i - x_i \cdot n_i^\omega \ge 0 \qquad \forall i \in N, \omega \in \Omega \qquad \text{(base-stock level)} \qquad (19)$$

$$q_i^\omega - x_i \cdot m_i^\omega = 0 \qquad \forall i \in N, \omega \in \Omega \qquad \text{(demand propagation)} \qquad (20)$$

$$q_i^\omega - \quad m_i^\omega \ge 0 \qquad \forall i \in N, \omega \in \Omega \qquad \text{(demand passage)} \qquad (21)$$

$$n_i^\omega + m_i^\omega - \sum_{j \in \delta_G^+(i)} n_j^\omega = \begin{cases} D_i^\omega & \text{if } i \in D \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in N, \omega \in \Omega \qquad \text{(flow conservation)} \qquad (22)$$

$$s_i, s_i^+, x_i, y_i \in \mathbb{Z}_{\ge 0} \qquad \forall i \in N \qquad \text{(natural anticipation)} \qquad (23)$$

$$q_i^\omega \in \mathbb{Z}_{\ge 0}, n_i^\omega, m_i^\omega \in \mathbb{Q}_{\ge 0} \qquad \forall i \in N, \omega \in \Omega \qquad \text{(mixed reaction)} \qquad (24)$$

In contrast to problem 1, condition (6) is restricted to equality (16) and constraint (18) is added to fix $s_1 = 0$ since $\tau_1 = T_1$.

The big-M-formulation of the SGSM-DP provided by Löhnert and Rambau (2018) has some serious drawbacks: first, it has a very large integrality gap and optimal results could so far only be obtained for networks with up to around 20 nodes (for mild parameters). Second, its objective estimates the holding cost of the full base-stock levels rather than the expected inventories. This can be suitable if holding costs are driven by stock-point capacities, but it is not suitable if holding costs are mainly driven by capital investment into stored material. Third, the compensation cost is based on the outsourced quantity given by the number of outsourced products during the replenishment delay, and therefore does not match the holding cost, which is evaluated per period.

The problem studied in this paper is to find a model formulation for an optimized computation of parameters to the outsourcing-aware base-stock policy so that also networks with, say, 50 nodes are tractable. A further task is to find a model modification to estimate the expected inventory levels and outsourced rates more faithfully. For the new and old models, the corresponding parameters for the outsourcing-aware base-stock policy shall be compared in simulations.

3 New Models for the Parameter Optimization for the Outsourcing-Aware Base-Stock Policy

In this section, we present new models for the parameters of the outsourcing-aware base-stock policy based on an MILP for a flow-design in a special underlying graph. The benefit of these models is that they have small integrality gaps so that networks with about 50 nodes (depending on the other data) are tractable.

We present two new model types, both based on flow-design ideas that are completely new in this context: the first model strives to find the optimal solution of the SGSM-DP faster so that larger networks become tractable. It is (with slight modifications) equivalent to the known SGSM-DP, i.e., it can find the same parameter values for the outsourcing-aware base-stock policy. The second model strives to improve the solution performance in dynamic simulations, i.e., in general, it determines different parameter values. The core idea of both models is to consider a *logical material flow in a time-expanded network*, where inventory withdrawals correspond to accelerations and are, therefore, represented by flows *backwards in time*.

3.1 A Flow-Design Model Equivalent to the SGSM-DP

The following describes a generalization of the main theoretic result in Kamp and Rambau (2022) corresponding to the flow-based formulation of the SGSM-DP with integral demand rates and includes an extensive proof. To represent the logic of the I-SGSM-DP, each node $i \in N$ in the inventory network G is split up into an inventory node i^{inv} where products are stored, a dispatch node i^{dis} where demand is propagated, and an outsourcing node i^{out} where outsourced supplies are received. Material is provided by the additional source node 0^{dis} without any intermediate stock (cross-docking). The resulting *split graph* G^{spl} is formally defined as follows.

$$G^{\text{spl}} := (N^{\text{spl}}, A^{\text{spl}}) \qquad \text{(split graph)} \qquad (25)$$

with

$$N^{\text{spl}} := M^{\text{spl}} \dot\cup O^{\text{spl}} \qquad \text{(split node set)} \qquad (26)$$

$$A^{\text{spl}} := E^{\text{spl}} \dot\cup F^{\text{spl}} \dot\cup V^{\text{spl}} \dot\cup R^{\text{spl}} \qquad \text{(split arc set)} \qquad (27)$$

$$D^{\text{spl}} := \{ i^{\text{dis}} | i \in D \} \qquad \text{(split demand set)} \qquad (28)$$

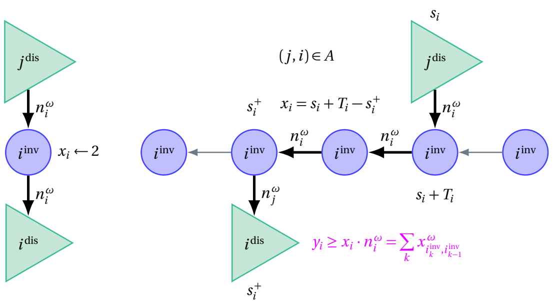

Figure 1: Structure of Flow Linearization: Service acceleration imposes demand load via logical flow backwards in time through inventory node, while deliveries via dispatch node arrive at replenishment time $s_i + T_i$ and depart at service time s_i^+.

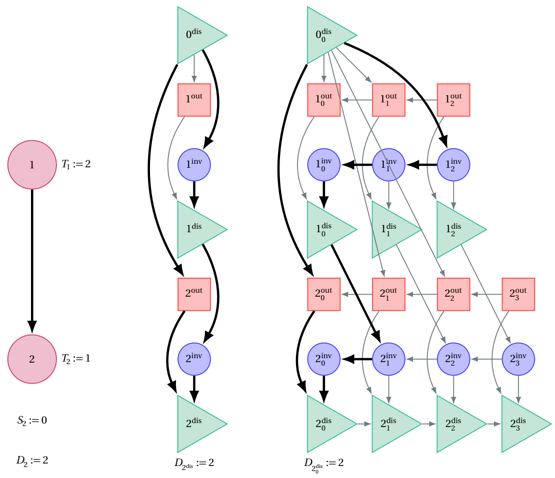

Figure 2: F-SGSM-DP-Network with a Feasible Flow: At the left hand side original network, in the middle split network where original nodes are split into inventory, outsourcing, and dispatch nodes, at the right hand side flow network leading a scenario flow with immediate service.

where

$$
\begin{aligned}
&M^{\text{spl}} := \{i^{\text{inv}} | i \in N\}, && \tau_{i*} := \tau_i && \text{(material node set)} && (29)\\
&O^{\text{spl}} := \{0^{\text{dis}}\} \dot{\cup} \{i^{\text{dis}}, i^{\text{out}} | i \in N\}, && \tau_{0^{\text{dis}}} := 0 && \text{(operational node set)} && (30)\\
&E^{\text{spl}} := \{(j^{\text{dis}}, i^{\text{inv}}) | (j,i) \in \{(0,1)\} \dot{\cup} A\}, && T_{j^{\text{dis}}, i^{\text{inv}}} := T_i && \text{(entering arc set)} && (31)\\
&F^{\text{spl}} := \{(i^{\text{inv}}, i^{\text{dis}}), (i^{\text{out}}, i^{\text{dis}}) | i \in N\}, && T_{i*, i^{\text{dis}}} := 0 && \text{(forwarding arc set)} && (32)\\
&V^{\text{spl}} := \{(i^{\text{inv}}, i^{\text{inv}}), (i^{\text{out}}, i^{\text{out}}) | i \in N\}, && T_{i*, i*} := -1 && \text{(accelerating arc set)} && (33)\\
&R^{\text{spl}} := \{(i^{\text{dis}}, i^{\text{dis}}) | i \in D\}, && T_{i^{\text{dis}}, i^{\text{dis}}} := 1 && \text{(receiving arc set)} && (34)
\end{aligned}
$$

Here, τ_{i*} denotes the time horizon of any node $i* \in N^{\text{spl}}$ and $T_{i*,j*}$ the corresponding time delay on any arc $(i*, j*)$ from the respective arc set. Accordingly, the relevant time sets are defined by $K_i := \{0, \ldots, \tau_i\}$ for all split nodes $i \in N^{\text{spl}}$ as well as the relevant guaranteed service times by $S_{i^{\text{dis}}} := \min(S_i, \tau_i)$ for all split demand nodes $i^{\text{dis}} \in D^{\text{spl}}$. This split graph is time-expanded up to the time horizons. Additionally, certain outsourcing arcs are added, which are required to enforce the special implications of the outsourced quantities prescribed by the SGSM-DP. These emanate from 0_0^{dis} and terminate in each instance of i^{out} for which there is an entering arc terminating in the corresponding instance of i^{inv} one period later. Formally, the resulting *t-graph* G^{exp} is defined by

$$G^{\text{exp}} := (N^{\text{exp}}, A^{\text{exp}}) \qquad \text{(t-graph)} \qquad (35)$$

with

$$
\begin{aligned}
&N^{\text{exp}} := \{i_k | i \in N^{\text{spl}}, k \in K_i\} && \text{(t-nodes)} && (36)\\
&A^{\text{exp}} := \{(i_k, j_{k+T_{i,j}}) | (i,j) \in A^{\text{spl}}, k \in K_i, k + T_{i,j} \in K_j\} \dot{\cup} S^{\text{exp}} && \text{(t-arcs)} && (37)\\
&D^{\text{exp}} := \{i_{S_i} | i \in D^{\text{spl}}\} && \text{(t-demands)} && (38)
\end{aligned}
$$

where

$$S^{\text{exp}} := \{(0_0^{\text{dis}}, i^{\text{out}}_{k+T_{j^{\text{dis}}, i^{\text{inv}}}-1}) | (j^{\text{dis}}, i^{\text{inv}}) \in E^{\text{spl}}, k \in K_{j^{\text{dis}}}\} \qquad \text{(t-supplies)} \qquad (39)$$

like depicted in figure 2. In the t-graph the customer requirements are imposed in time and value by the flow demands

$$D_i^{\omega} := \begin{cases} -\Delta_1^{\omega} & \text{if } i = 0_0^{\text{dis}} \\ D_j^{\omega} & \text{if } i = j_k^{\text{dis}} \in D^{\text{exp}} \\ 0 & \text{otherwise} \end{cases} \qquad (40)$$

for all t-nodes $i \in N^{\text{exp}}$ in every scenario $\omega \in \Omega$. On this structure we are able to construct the following model.

Problem 3 (F-SGSM-DP)**.** The following problem is called the *Flow-based Stochastic Guaranteed*

Service Model with Demand Propagation:

$$\min \sum_{i\in N}\Big(H_i \cdot y_i + C_i \cdot \sum_{\omega\in\Omega} P^{\omega} \cdot q_i^{\omega}\Big) \qquad \text{(cost)} \qquad (41)$$

$$\sum_{k\in K_i} z_{i,k} = 1 \qquad \forall i \in N \qquad \text{(choice)} \qquad (42)$$

$$x^{\omega}_{j_k^{\text{dis}}, i^{\text{inv}}_{k+T_i}} + x^{\omega}_{0_0^{\text{dis}}, i^{\text{out}}_{k+T_i-1}} \le \Delta_i^{\omega} \cdot z_{j,k} \qquad \forall (j,i) \in A, \forall k \in K_j, \omega \in \Omega \qquad \text{(inbound)} \qquad (43)$$

$$x^{\omega}_{i_k^{\text{inv}}, i_k^{\text{dis}}} + x^{\omega}_{i_k^{\text{out}}, i_k^{\text{dis}}} \le \Delta_i^{\omega} \cdot z_{i,k} \qquad \forall i \in N, \forall k \in K_i, \omega \in \Omega \qquad \text{(outbound)} \qquad (44)$$

$$\sum_{k\in K_i\setminus\{0\}} x^{\omega}_{i_k^{\text{inv}}, i_{k-1}^{\text{inv}}} \le y_i \qquad \forall i \in N, \omega \in \Omega \qquad \text{(load)} \qquad (45)$$

$$\sum_{k\in K_i\setminus\{0\}} (x^{\omega}_{i_k^{\text{out}}, i_{k-1}^{\text{out}}} + x^{\omega}_{i_{k-1}^{\text{out}}, i_{k-1}^{\text{dis}}}) = q_i^{\omega} \qquad \forall i \in N, \omega \in \Omega \qquad \text{(quantity)} \qquad (46)$$

$$\sum_{j\in\delta^-_{G^{\text{exp}}}(i)} x^{\omega}_{j,i} - \sum_{j\in\delta^+_{G^{\text{exp}}}(i)} x^{\omega}_{i,j} = D_i^{\omega} \qquad \forall i \in N^{\text{exp}}, \omega \in \Omega \qquad \text{(flow)} \qquad (47)$$

$$z_{i,k} \in \mathbf{B} \qquad \forall i \in N, \forall k \in K_i \qquad \text{(service)} \qquad (48)$$

$$y_i \in \mathbb{Z}_{\ge 0} \qquad \forall i \in N \qquad \text{(base)} \qquad (49)$$

$$q_i^{\omega} \in \mathbb{Z}_{\ge 0} \qquad \forall i \in N, \omega \in \Omega \qquad \text{(emergency)} \qquad (50)$$

$$x^{\omega}_{i,j} \in \mathbb{Q}_{\ge 0} \qquad \forall (i,j) \in A^{\text{exp}}, \omega \in \Omega \qquad \text{(direction)} \qquad (51)$$

The multiple-choice variable $z_{i,k}$ (48) unified by constraints (42) indicates the selection of the outbound service time $s_i^+ = k$. Variable $x^{\omega}_{i,j}$ (51) describes the flow rate on (i,j) in the t-graph in scenario ω conserved by conditions (47). The flow-design constraints (43) and (44) enforce the selected outbound service times in the t-graph flow. The acceleration arcs along the expansion of i^{inv} and i^{out} backwards in time are utilized to evaluate the non-linear demand loads on the base-stock level y_i (49) in (19) and outsourced quantities q_i^{ω} (50) in (20) by linear constraints (45) and (46) as illustrated in figure 1. In the objective (41) holding costs for the base-stocks and compensation costs for the outsourced quantities are evaluated. Given a feasible solution of the F-SGSM-DP, we can derive parameters for the outsourcing-aware base-stock policy as follows: for each $i \in N$, the base-stock level is given by y_i, the guaranteed outbound service time is given by $\sum_{k\in K_i} k \cdot z_{i,k}$, and the upper order limit is given by $\lceil \max_{\omega\in\Omega} \sum_{k\in K_i} x^{\omega}_{i_k^{\text{inv}}, i_k^{\text{dis}}} \rceil$.

The following theorem shows that, essentially, we may solve the F-SGSM-DP instead of the I-SGSM-DP to obtain the same result.

Theorem 1. *The F-SGSM-DP and the I-SGSM-DP are equivalent if one of the following modifications is made.*

(i) Integrality constraints on outsourced quantities are removed.

(ii) Integrality constraints on flow values are added.

(iii) Linear constraints $\sum_{k\in K_i} k \cdot z_{i,k} - \sum_{k\in K_j} k \cdot z_{j,k} \le T_i$ for all $(j,i) \in A$ are added to the F-SGSM-DP.

Furthermore, the I-SGSM-DP and the SGSM-DP are equivalent if either modification (i) or (ii) is made.

Before we formally prove this theorem, we provide some preliminary insights. An advantage of a slack-free formulation is, that only finite sets of service times have to be considered as already indicated by Kamp and Rambau (2022).

Lemma 1 (Relevant Time Horizon)**.** *For distribution networks, every slack-free solution of the SGSM-DP satisfies*

$$s_i^+ + x_i \in K_i \tag{52}$$

for each node $i \in N$.

Proof. Proof: For an arbitrary $i \in N$ summing up equations (5) on the upstream set P_i results in

$$s_{*i}^+ + \sum_{j\in P_i} x_{*j} = \underbrace{s_{*1}}_{=0} + \underbrace{\sum_{j\in P_i} T_j}_{=\tau_i} = \tau_i \tag{53}$$

where $i \in P_i$ and the non-negativity of the feasible replenishment delays imply (52). □

Furthermore, for every node $i \in N$, condition (5) implies $T_i + s_i \in K_i$ resulting in $s_i \le \tau_i - T_i$. Therefore, the I-SGSM-DP restricts the SGSM-DP to slack-free solutions. Next, it is shown that the full demand rates are valid upper demand bounds for the SGSM-DP and the I-SGSM-DP.

Lemma 2 (Node Supply Bound)**.** *For distribution networks, every feasible solution of the SGSM-DP or the I-SGSM-DP satisfies*

$$n_i^\omega + m_i^\omega \le \Delta_i^\omega \tag{54}$$

for each node $i \in N$ in every scenario $\omega \in \Omega$.

Proof. Proof: For arbitrary $i \in N$ and $\omega \in \Omega$ summing up equations (11) or (22) of the downstream set $\varrho_G^+(i)$ results in

$$\sum_{j\in\varrho_G^+(i)} \Big(n_j^\omega - \sum_{k\in\delta_G^+(j)} n_k^\omega + m_j^\omega\Big) = \sum_{j\in\varrho_G^+(i)\cap D} D_j^\omega = \Delta_i^\omega \tag{55}$$

where $i \in \varrho_G^+(i)$ and the non-negativity of the feasible outsourced demand rates imply (54) via assumptions 1(ii) and (iii). □

Let in the following the largest lower bound of a model objective be indicated by a trailing *. Indeed, SGSM-DP* is upper-bounded by F-SGSM-DP*.

Lemma 3 (Primal Bound for SGSM-DP by F-SGSM-DP)**.** *For distribution networks, every feasible solution of the F-SGSM-DP induces a feasible solution of the SGSM-DP with identical objective value.*

Proof. Proof: Let an arbitrary feasible solution of the F-SGSM-DP have outbound service indicators $z_{i,k} \in \mathbf{B}$ for all $i \in N$ with $k \in K_i$ and flow values $x_{i,j}^\omega \in \mathbb{Q}_{\ge 0}$ for all t-arcs $(i,j) \in A^{\text{exp}}$ in every scenario $\omega \in \Omega$. The selected outbound service times are assigned by

$$s_i^+ \leftarrow \sum_{k\in K_i} k \cdot z_{i,k} \in K_i \subset \mathbb{Z}_{\ge 0} \tag{56}$$

for all $i \in N$ as well as $z_{0,0} := 1$ for $s_0^+ := 0$. Furthermore, the supply values are defined by

$$n_i^\omega \leftarrow \sum_{k\in K_i} x^\omega_{i_k^{\text{inv}}, i_k^{\text{dis}}} \in \mathbb{Q}_{\ge 0} \tag{57}$$

$$m_i^\omega \leftarrow \sum_{k\in K_i} x^\omega_{i_k^{\text{out}}, i_k^{\text{dis}}} \in \mathbb{Q}_{\ge 0} \tag{58}$$

for all $i \in N$ in every $\omega \in \Omega$, while satisfying constraints (11) by summing the conservation conditions (47) for all instances of i^{dis} in each scenario separately. The remaining times are set to

$$x_i \leftarrow \max(s_j^+ + T_i - s_i^+, 0) \in \mathbb{Z}_{\geq 0} \tag{59}$$

$$s_i \leftarrow x_i + s_i^+ - T_i = \max(s_j^+, s_i^+ - T_i) \in \mathbb{Z}_{\geq 0} \tag{60}$$

for all $(j,i) \in \{(0,1)\} \cup A$, while satisfying constraints (5) and (6). For conditions (7), consider an arbitrary demand node $i \in D$ and a scenario $\omega \in \Omega$ in which $D_i^\omega > 0$ is satisfied. This customer demand applies to i^{dis} at time $S_{i^{\text{dis}}}$ and therefore requires a positive flow into the node expansion of i^{dis} at a time not after $S_{i^{\text{dis}}}$ because the connecting receiving arcs point forwards in time. Therefore, conditions (44) imply that $s_i^+ \leq S_{i^{\text{dis}}} \leq S_i$ showing the validity of conditions (7). Using the same base-stock levels and outsourced quantities leads to the same objective value, while constraints (8) and (9) are validated by showing

$$\sum_{k \in K_i \setminus \{0\}} x^\omega_{i_k^{\text{inv}}, i_{k-1}^{\text{inv}}} = x_i \cdot n_i^\omega \tag{61}$$

$$\sum_{k \in K_i \setminus \{0\}} (x^\omega_{i_k^{\text{out}}, i_{k-1}^{\text{out}}} + x^\omega_{i_{k-1}^{\text{out}}, i_{k-1}^{\text{dis}}}) = x_i \cdot m_i^\omega \tag{62}$$

for all $i \in N$ in every $\omega \in \Omega$ based on constraints (45) and (46). So let $(j,i) \in \{(0,1)\} \cup A$ and $\omega \in \Omega$ be arbitrary.

1. Case: $n_i^\omega > 0 \lor m_i^\omega > 0$
   Conditions (44) imply that a flow of value n_i^ω leaves the node expansion of i^{inv} and a flow of value m_i^ω leaves the node expansion of i^{out} both only at time s_i^+. Since at least one of these flows is positive, this requires a positive flow to enter the node expansion of either i^{inv} or i^{out} at a time not before s_i^+ because the connecting accelerating arcs point backwards in time. By conditions (43) a positive flow can only enter the node expansion of i^{inv} at time $s_j^+ + T_i$ and the node expansion of i^{out} at time $s_j^+ + T_i - 1$, which implies $s_j^+ + T_i \geq s_i^+$. Therefore, by definition it is $x_i = s_j^+ + T_i - s_i^+$ being the length of the flow line passing the node expansion of i^{inv} with value n_i^ω and if $x_i > 0$ the by one extended length of the line passing the node expansion of i^{out} with flow m_i^ω. Otherwise, it is $m_i^\omega = 0$ because in this case $s_j^+ + T_i - 1 < s_i^+$, while satisfying constraint (10). This shows (61) and (62).

2. Case: $n_i^\omega = 0 \land m_i^\omega = 0$
   Since $m_i^\omega = 0$, constraint (10) is fulfilled. Moreover, there is no flow leaving the node expansions of i^{inv} and i^{out}. Therefore, no flow passes the corresponding accelerating arcs showing (61) and (62).

Altogether the defined solution is feasible for the SGSM-DP and has the same objective value as the initial solution of the F-SGSM-DP. □

It can be seen that a non-trivial demand distribution is merely required to infer the customer service guarantees from the flow conservation conditions. The reason is that the corresponding indicator variables are forced to be active only if the underlying flow is non-zero. Moreover, this also shows the statement for the alternative assumption in Kamp and Rambau (2022) that no demand node has successors. Under this condition, demand nodes without customers face no demand at all. In this case, the customer service requirement does not have to be enforced since it can be satisfied by reducing the outbound service time at no cost.

On the other hand, I-SGSM-DP* is lower-bounded by F-SGSM-DP*.

Lemma 4 (Dual Bound for I-SGSM-DP by F-SGSM-DP)**.** *For distribution networks, every feasible solution of the I-SGSM-DP induces a feasible solution of the F-SGSM-DP with identical objective value.*

Proof. Proof: Let an arbitrary feasible solution of the I-SGSM-DP have for all nodes $i \in N$ inbound service times $s_i \in \mathbb{Z}_{\geq 0}$, outbound service times $s_i^+ \in \mathbb{Z}_{\geq 0}$, replenishment delays $x_i \in \mathbb{Z}_{\geq 0}$, as well as additionally in every scenario $\omega \in \Omega$ propagated rates $n_i^\omega \in \mathbb{Q}_{\geq 0}$, and outsourced rates $m_i^\omega \in \mathbb{Q}_{\geq 0}$. By conditions (15) and (18) it is $s_i^+ \in K_i$ for all $i \in N$. Therefore, we can select these outbound service times by assigning

$$z_{i,k} \leftarrow \mathbf{1}(k = s_i^+) \in \mathbf{B} \tag{63}$$

for all $i \in N$ and $k \in K_i$ as well as $s_0^+ := s_1 = 0$ for $z_{0,0} := 1$, while satisfying constraints (42). Furthermore, in every $\omega \in \Omega$ the underlying flow values are set to

$$x^\omega_{j_k^{\text{dis}}, i^{\text{inv}}_{k+T_i}} \leftarrow n_i^\omega \cdot z_{j,k} \in \mathbb{Q}_{\geq 0} \tag{64}$$

$$x^\omega_{0_0^{\text{dis}}, i^{\text{out}}_{k+T_i-1}} \leftarrow m_i^\omega \cdot z_{j,k} \in \mathbb{Q}_{\geq 0} \tag{65}$$

for all $(j,i) \in \{(0,1)\} \cup A$ and $k \in K_j$ since $T_i > 0$, while satisfying constraints (43), as well as

$$x^\omega_{i_k^{\text{inv}}, i_k^{\text{dis}}} \leftarrow n_i^\omega \cdot z_{i,k} \in \mathbb{Q}_{\geq 0} \tag{66}$$

$$x^\omega_{i_k^{\text{out}}, i_k^{\text{dis}}} \leftarrow m_i^\omega \cdot z_{i,k} \in \mathbb{Q}_{\geq 0} \tag{67}$$

for all $i \in N$ and $k \in K_i$, while satisfying constraints (44), because of the valid supply bounds (54). Because of the time conditions (18), (16), and (15) as well as the artificially defined outbound service time of the source node, it is $s_j^+ + T_i = s_i + T_i = s_i^+ + x_i$ for every $(j,i) \in \{(0,1)\} \cup A$. This is why the node expansion to i^{inv} is supplied at time $s_i^+ + x_i$ as well as demanded at time $s_i^+ \leq s_i^+ + x_i$ by conditional flows n_i^ω, the node expansion to i^{out} is supplied at time $s_i^+ + x_i - 1$ as well as demanded at time s_i^+ by conditional flows m_i^ω being non-zero only for $x_i > 0$ due to propagation conditions (20) together with passage conditions (21), and the node expansion to i^{dis} is only for $i \in D$ internally supplied at time s_i^+ as well as externally demanded at time $S_{i^{\text{dis}}} = \min(S_i, \tau_i) \geq s_i^+$ by conditional flows D_i^ω due to external service conditions (17) and conservation conditions (22). For these physical flows the conservation conditions (47) are balanced by assigning in each scenario $\omega \in \Omega$ for every node $i \in N$ the logical flow $n_i^\omega \in \mathbb{Q}_{\geq 0}$ to every accelerating arc on the line from $i^{\text{inv}}_{s_i^+ + x_i}$ to $i^{\text{inv}}_{s_i^+}$, if $x_i > 0$ the logical flow $m_i^\omega \in \mathbb{Q}_{\geq 0}$ to every accelerating arc on the line from $i^{\text{out}}_{s_i^+ + x_i - 1}$ to $i^{\text{out}}_{s_i^+}$, and if $i \in D$ the logical flow $D_i^\omega \in \mathbb{Q}_{\geq 0}$ to every receiving arc on the line from $i^{\text{dis}}_{s_i^+}$ to $i^{\text{dis}}_{S_{i^{\text{dis}}}}$, while annulling all other flow values. Consequently, the sums in (45) and (46) are equal to the conditional propagated and outsourced rates respectively multiplied by the associated replenishment delay. Thus, using the same base-stock levels and outsourced quantities validates constraints (45) and (46) by the corresponding constraints (19) and (20), which shows that the constructed solution is feasible for the F-SGSM-DP and has the same objective value as the initial solution of the I-SGSM-DP. □ □

In combination, the preceding lemmas show that

$$\text{SGSM-DP}^* \leq \text{F-SGSM-DP}^* \leq \text{I-SGSM-DP}^* \tag{68}$$

under the given assumptions.

Now we are in a position to complete the proof of our main theoretical result.

Proof. Proof of Theorem 1: For the equivalence to the F-SGSM-DP at first consider modification (i) or (ii). Let an arbitrary feasible solution of the SGSM-DP have for all nodes $i \in N$ inbound service times

$s_i \in \mathbb{Z}_{\geq 0}$, outbound service times $s_i^+ \in \mathbb{Z}_{\geq 0}$, replenishment delays $x_i \in \mathbb{Z}_{\geq 0}$, as well as additionally in every scenario $\omega \in \Omega$ outsourced rates m_i^ω and outsourced quantities q_i^ω. It can be converted to a feasible solution of the I-SGSM-DP by defining $s_0^+ := 0$ and subsequently adapting

$$s_i \leftarrow s_j^+ \in \mathbb{Z}_{\geq 0} \tag{69}$$

$$x_i \leftarrow \min(x_i, s_i + T_i) \in \mathbb{Z}_{\geq 0} \tag{70}$$

$$s_i^+ \leftarrow s_i + T_i - x_i \in \mathbb{Z}_{\geq 0} \tag{71}$$

for every $(j,i) \in \{(0,1)\} \cup A$ in topological order starting at the artificial source node 0. Since in each iteration the local inbound service time is set to the outbound service time of the unique predecessor, this validates conditions (15), (16) and (18) by $s_1 = 0$. Furthermore, it can be seen by induction that no time is increased. This is obviously true for the root arc $(0,1)$ since the resulting setting, expressed in the original variables, is given by inbound service time $0 \leq s_1$, replenishment delay $\min(x_1, T_1) \leq x_1$, and outbound service time $\max(T_1 - x_1, 0) \leq s_1^+$ because of condition (5) and the sign constraints on service times. Let for an arbitrary $(j,i) \in A$ the topologically preceding times be not increased. Then the resulting setting, expressed in the variable values before the corresponding iteration, is given by inbound service time $s_j^+ = s_j + T_j - x_j \leq \min(\tau_i - T_i, s_i)$ because of condition (6), replenishment delay $\min(x_i, s_j^+ + T_i) \leq x_i$, and outbound service time $\max(s_j^+ + T_i - x_i, 0) \leq s_i + T_i - x_i = s_i^+$ because of condition (6) and (5), which completes the inductive argument. Therefore, also the outbound service times of demand nodes are not increased and conditions (17) remain valid. As the replenishment delays are not increased, the base-stock level conditions (19) and flow conservation conditions (22) are validated by using the same base-stock levels and propagated rates based on the corresponding conditions (8) and (11). To fulfill the propagation constraints (20) the outsourced quantities are reduced to

$$q_i^\omega \leftarrow x_i \cdot m_i^\omega \geq 0 \tag{72}$$

for all $i \in N$ in every $\omega \in \Omega$, which is valid for modification (i) or (ii) because by the integrality of replenishment delays the resulting outsourced quantities are integral if the outsourced rates are. Above, a replenishment delay is never strictly reduced to zero since the corresponding lead time is positive by definition, which is why passage conditions (21) are validated for resulting positive replenishment delays by the valid conditions (20) and for resulting zero replenishment delays by the corresponding conditions (10) for unchanged zero outsourced quantities. This shows that the resulting solution with the same base-stock levels and propagated rates is feasible for the I-SGSM-DP and since the outsourced quantities are not increased, the objective value compared to the initial solution of the SGSM-DP is neither. The solution mappings from the proofs of lemmas 3 and 4 are also valid for fractional outsourced quantities since they remain untouched and the integrality constraints are not required otherwise. Also for integral flow values these remain valid because the integrality constraints on supply rates and physical flows are actually conserved, which induces integral logical flows on accelerating and receiving arcs in the t-graph. So there is a cycle of objective-protecting solution mappings from the SGSM-DP over the I-SGSM-DP to the F-SGSM-DP and back, which establishes the equivalence of these models for modification (i) or (ii).

Next, consider modification (iii). Then, the solution mapping from the F-SGSM-DP to the SGSM-DP in the proof of lemma 3 indeed results in a solution of the I-SGSM-DP because the additional constraints together with $s_0^+ := 0$ for all $(j,i) \in \{(0,1)\} \cup A$ imply

$$s_i^+ - s_j^+ \leq T_i \tag{73}$$

since $s_1^+ \leq T_1 = \tau_1$ by $s_1^+ \in K_1$ and therefore

$$x_i = s_j^+ + T_i - s_i^+ \tag{74}$$

$$s_i = s_j^+ \tag{75}$$

in the resulting setting, while satisfying the restricting constraints (16) and (18) because of $s_i = 0$. On the other hand, the solution mapping from the I-SGSM-DP to the F-SGSM-DP in the proof of lemma 4 results in a solution which already satisfies the additional constraints because conditions (15) and (16) imply

$$T_i = x_i + s_i^+ - s_j^+ \geq s_i^+ - s_j^+ = \sum_{k \in K_i} k \cdot z_{i,k} - \sum_{k \in K_j} k \cdot z_{j,k} \tag{76}$$

for all $(j,i) \in A$ by the sign restrictions on replenishment delays as well as $z_{i,k} = \mathbf{1}(k = s_i^+)$ for all $i \in N$ and $k \in K_i$ in the resulting setting. So there is a cycle of objective-protecting solution mappings from the I-SGSM-DP to the F-SGSM-DP and back, which establishes the equivalence of these models for modification (iii). □ □

If integrality constraints on base-stock levels are removed, the previous statements remain valid, since the constructed solution mappings do not rely on these integrality constraints and leave the base-stock levels unchanged. In combination with modification (i) this shows that the continuous-commodity variants of the SGSM-DP, the I-SGSM-DP, and the F-SGSM-DP are equivalent.

We have shown that under mild conditions the F-SGSM-DP can compute an optimal solution for the I-SGSM-DP. This is most interesting if the F-SGSM-DP can be solved for larger instances. We compare the solution times and the integrality gaps of the new F-SGSM-DP with the known big-M-linearization of the SGSM-DP in numerical experiments in Section 4.

3.2 An Improved Flow-Design Model

This section provides the semantic model improvement suggested by Kamp (2021). The improved model is different in essentially two ways: first, in the objective the inventory levels are reduced by the average demand load to estimate better the expected inventory levels; second, unified outsourced rates instead of accumulated outsourced quantities impose compensation costs. The transition from

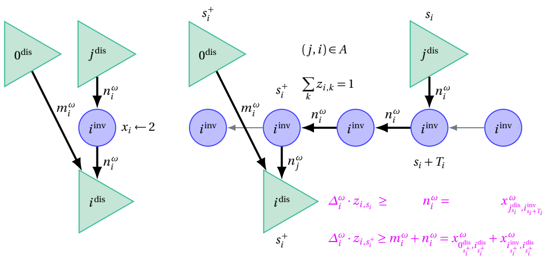

Figure 3: Structure of Flow Design: Deliveries at inventory node i must obey the guaranteed service rules, which is enforced by big-M-conditions with respect to binary selector variables $z_{i,k}$ restricting each scenario flow through dispatch node i^{dis} to service time s_i^+.

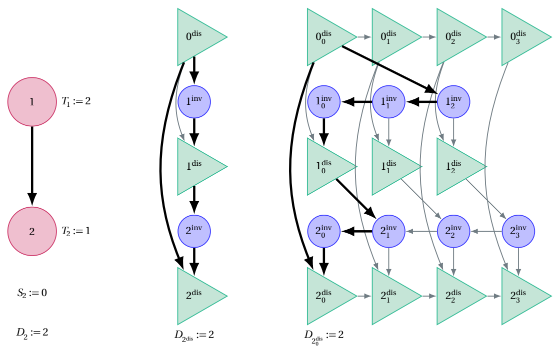

Figure 4: F-SGSM-DF-Network with a Feasible Flow: At the left hand side original network, in the middle split network where original nodes are split into inventory, and dispatch nodes, at the right hand side flow network leading a scenario flow with immediate service.

outsourced quantities to outsourced rates allows for a simplification in the flow-design network. This time, each node $i \in N$ in the inventory network G only gives rise to split nodes $i^{\text{inv}}, i^{\text{dis}}$ supplied by a source node 0^{dis} in another *split graph* G^{spl} in the following way: The inventory node i^{inv} holds the regularly replenished stock. The dispatch node i^{dis} assigns the order quantities to deliver from stock and from outsourcing. The source node supplies the root inventory node 1^{inv} with the regular delay T_1 and every dispatch node i^{dis} as a zero-delay emergency supplier. The resulting split graph is defined as follows.

$$G^{\text{spl}} := (N^{\text{spl}}, A^{\text{spl}}) \qquad \text{(split graph)} \qquad (77)$$

with

$$N^{\text{spl}} := M^{\text{spl}} \dot\cup O^{\text{spl}} \qquad \text{(split node set)} \qquad (78)$$

$$A^{\text{spl}} := E^{\text{spl}} \dot\cup F^{\text{spl}} \dot\cup S^{\text{spl}} \dot\cup V^{\text{spl}} \dot\cup R^{\text{spl}} \qquad \text{(split arc set)} \qquad (79)$$

$$D^{\text{spl}} := \{i^{\text{dis}} | i \in D\} \qquad \text{(split demand set)} \qquad (80)$$

where

$$M^{\text{spl}} := \{i^{\text{inv}} | i \in N\}, \qquad \tau_{i^*} := \tau_i \qquad \text{(material node set)} \qquad (81)$$

$$O^{\text{spl}} := \{i^{\text{dis}} | i \in \{0\} \dot\cup N\}, \qquad \tau_{0^{\text{dis}}} := \max_{i \in N} \tau_i \qquad \text{(operational node set)} \qquad (82)$$

$$E^{\text{spl}} := \{(j^{\text{dis}}, i^{\text{inv}}) | (j,i) \in \{(0,1)\} \dot\cup A\}, \qquad T_{j^{\text{dis}}, i^{\text{inv}}} := T_i \qquad \text{(entering arc set)} \qquad (83)$$

$$F^{\text{spl}} := \{(i^{\text{inv}}, i^{\text{dis}}) | i \in N\}, \qquad T_{i^{\text{inv}}, i^{\text{dis}}} := 0 \qquad \text{(forwarding arc set)} \qquad (84)$$

$$S^{\text{spl}} := \{(0^{\text{dis}}, i^{\text{dis}}) | i \in N\}, \qquad T_{0^{\text{dis}}, i^{\text{dis}}} := 0 \qquad \text{(outsourcing arc set)} \qquad (85)$$

$$V^{\text{spl}} := \{(i^{\text{inv}}, i^{\text{inv}}) | i \in N\}, \qquad T_{i^{\text{inv}}, i^{\text{inv}}} := -1 \qquad \text{(accelerating arc set)} \qquad (86)$$

$$R^{\text{spl}} := \{(i^{\text{dis}}, i^{\text{dis}}) | i \in D\}, \qquad T_{i^{\text{dis}}, i^{\text{dis}}} := 1 \qquad \text{(receiving arc set)} \qquad (87)$$

Again, τ_{i^*} denotes the time horizon of any node $i^* \in N^{\text{spl}}$ and T_{i^*,j^*} the corresponding time delay on any arc (i^*, j^*) from the respective arc set with relevant time sets $K_i := \{0, \ldots, \tau_i\}$ for all split nodes $i \in N^{\text{spl}}$ as well as relevant guaranteed service times $S_{i^{\text{dis}}} := \min(S_i, \tau_i)$ for all split demand nodes $i^{\text{dis}} \in D^{\text{spl}}$. This split graph is time-expanded up to the time horizons without the need of further extensions, which results in the *t-graph* G^{exp} defined by

$$G^{\text{exp}} := (N^{\text{exp}}, A^{\text{exp}}) \qquad \text{(t-graph)} \qquad (88)$$

with

$$N^{\text{exp}} := \{i_k | i \in N^{\text{spl}}, k \in K_i\} \qquad \text{(t-nodes)} \qquad (89)$$

$$A^{\text{exp}} := \{(i_k, j_{k+T_{i,j}}) | (i,j) \in A^{\text{spl}}, k \in K_i, k + T_{i,j} \in K_j\} \qquad \text{(t-arcs)} \qquad (90)$$

$$D^{\text{exp}} := \{i_{S_i} | i \in D^{\text{spl}}\} \qquad \text{(t-demands)} \qquad (91)$$

as depicted in figure 4 with flow demands (40). This enables us to construct the following model with improved objective and outsourcing rates instead of outsourcing quantities.

Problem 4 (F-SGSM-DF)**.** The following problem is called the *Flow-based Stochastic Guaranteed*

Service Model with Demand Flow:

$$\min \sum_{\omega\in\Omega} P^{\omega}\cdot\sum_{i\in N}\Big(H_i\cdot a_i^{\omega}+C_i\cdot\sum_{k\in K_i} x^{\omega}_{0_k^{\text{dis}},i_k^{\text{dis}}}\Big) \qquad \text{(cost)} \qquad (92)$$

$$\sum_{k\in K_i} z_{i,k}=1 \qquad \forall i\in N \qquad \text{(choice)} \qquad (93)$$

$$x^{\omega}_{j_k^{\text{dis}},i_{k+T_i}^{\text{inv}}}\le \Delta_i^{\omega}\cdot z_{j,k} \qquad \forall (j,i)\in A,\forall k\in K_j,\omega\in\Omega \qquad \text{(inbound)} \qquad (94)$$

$$x^{\omega}_{i_k^{\text{inv}},i_k^{\text{dis}}}+x^{\omega}_{0_k^{\text{dis}},i_k^{\text{dis}}}\le \Delta_i^{\omega}\cdot z_{i,k} \qquad \forall i\in N,\forall k\in K_i,\omega\in\Omega \qquad \text{(outbound)} \qquad (95)$$

$$\sum_{k\in K_i\setminus\{0\}} x^{\omega}_{i_k^{\text{inv}},i_{k-1}^{\text{inv}}}=y_i-a_i^{\omega} \qquad \forall i\in N,\omega\in\Omega \qquad \text{(load)} \qquad (96)$$

$$\sum_{j\in\delta^-_{G^{\text{exp}}}(i)} x^{\omega}_{j,i}-\sum_{j\in\delta^+_{G^{\text{exp}}}(i)} x^{\omega}_{i,j}=D_i^{\omega} \qquad \forall i\in N^{\text{exp}},\omega\in\Omega \qquad \text{(flow)} \qquad (97)$$

$$z_{i,k}\in\mathbf{B} \qquad \forall i\in N,\forall k\in K_i \qquad \text{(service)} \qquad (98)$$

$$y_i\in\mathbb{Z}_{\ge 0} \qquad \forall i\in N \qquad \text{(base)} \qquad (99)$$

$$x^{\omega}_{i,j}\in\mathbb{Q}_{\ge 0} \qquad \forall (i,j)\in A^{\text{exp}},\omega\in\Omega \qquad \text{(direction)} \qquad (100)$$

$$a_i^{\omega}\in\mathbb{Q}_{\ge 0} \qquad \forall i\in N,\omega\in\Omega \qquad \text{(stock)} \qquad (101)$$

The parameters for the outsourcing-aware base-stock policy are still given for each $i\in N$ by the base-stock level y_i (99), the guaranteed outbound service time $\sum_{k\in K_i} k\cdot z_{i,k}$ inferred from the multiple-choice variables $z_{i,k}$ (98) specified by the set-partitioning constraints (93), and the upper order limit $\lceil\max_{\omega\in\Omega}\sum_{k\in K_i} x^{\omega}_{i_k^{\text{inv}},i_k^{\text{dis}}}\rceil$ determined smallest possible to permit the propagated demand values $x^{\omega}_{i_k^{\text{inv}},i_k^{\text{dis}}}$ (100) validated by the demand conservation conditions (97). Variables a_i^{ω} (101) are determined in condition (96) to evaluate the expected inventory level along with the expected outsourced rates in objective (92). These inventory levels are given by the slacks of constraints (45) and represent the parts of the base-stocks which remain in the inventories for the current demand loads. Flow-design conditions (94) and (95) impose unique outbound service times across scenarios in the flow network, i.e., positive flow values are only allowed at consistent time slots as illustrated in figure 3.

The adaptions made in the F-SGSM-DF are supported by dynamic simulation results on random instances provided in the next section.

4 Computational Experiments

To asses the run-time complexities of the models' solution processes and performances of the respective model-optimal solutions in the dynamic context, a computational study on randomized data is performed. Optimal solutions in the following are computed by a standard MacBook Air (11 Inch, Mid 2012, macOS Catalina 10.15.7, 2.6 GHz Core i5-3317U, 4 GB DDR3-RAM) with a patched variant of `SCIP 7.0.3` embedding LP-solver `SoPlex 5.0.2` with default parameters apart from tolerances "epsilon" of 10^{-10} and "feastol" of 10^{-7}. (Gamrath et al. 2020). Networks are generated by successively adding nodes and choosing unique predecessors at random.

At first the computational performance of the SGSM-DP by Löhnert and Rambau (2018), the T-SGSM-DP by Kamp and Rambau (2022) (a purely technical tightening of the SGSM-DP), and the F-SGSM-DP given by problem 3 (all with continuous demand rates) is compared. For this, the same randomization parameters as for the mild instance set by Kamp and Rambau (2022) are used with

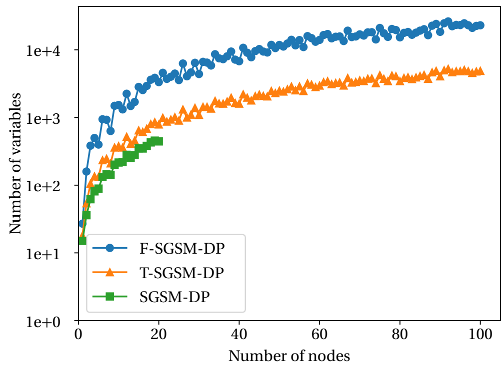

Figure 5: Comparison of Model Sizes for Mild Parameters: The flow-based formulation has almost 10 times as many variables as the tightened formulation.

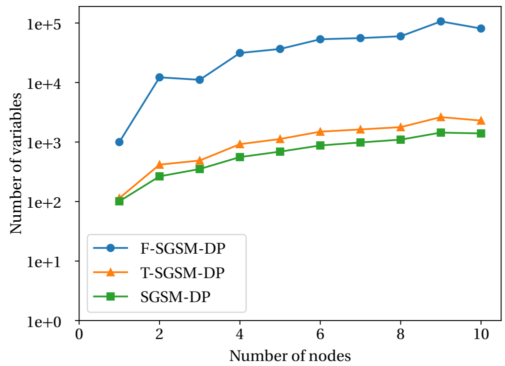

Figure 6: Comparison of Model Sizes for Heavy Parameters: The flow-based formulation has almost 100 times as many variables as the tightened formulation.

$|N| \in \{1,\ldots,100\}$. That is, for $|\Omega| = 3$ the data is independently sampled from the ranges

$$T_i \in \{1,\ldots,4\} \tag{102}$$
$$H_i \in \{1,2\} \tag{103}$$
$$C_i \in \{H_i+1,\ldots,H_i+8\} \tag{104}$$

for node $i \in N$ and

$$S_i \in \{0,1\} \tag{105}$$
$$P^\omega \in \{1,\ldots,100\} \text{ normalized} \tag{106}$$
$$D_i^\omega \in \{1,\ldots,\lceil 4\cdot \tfrac{\omega}{|\Omega|}\rceil\} \tag{107}$$

for demand node $i \in D$ in scenario $\omega \in \Omega$ respectively. Figures 5, 7, and 8 display the number of variables, the resulting total computation times and the relative integrality gaps defined by $1-\frac{\text{LP-Opt}}{\text{MILP-Opt}}$ for a time limit of 1000 seconds.

The F-SGSM-DP is by a factor of 100 faster than the T-SGSM-DP, which in turn is 10 times faster than the original SGSM-DP for more than 10 nodes. With the F-SGSM-DP, an optimal solution can still be computed within the time limit for even up to 400 nodes. Hence, the larger model size of the F-SGSM-DP is in this case overcompensated by its low integrality gap.

Another heavy instance set is generated for $|N| \in \{1,\ldots,10\}$ and $|\Omega| = 30$ with ranges $T_i \in \{1,\ldots,31\}$, $S_i \in \{0,\ldots,30\}$, and $D_i^\omega \in \{1,\ldots,\lceil 62\cdot \frac{\omega}{|\Omega|}\rceil\}$ in the same way as before. These instances are equipped with a practical cost structure by at first generating the compensation marginals in topological order with mean $16\tau_i$ and then the holding marginals in reverse topological order with mean 16, while the range widths are adaptively chosen, so that both marginal sets are topologically increasing as well as every compensation marginal is larger than the mean and the realization of the corresponding holding marginal.

For these heavy instances, figure 9 shows that there is no longer a consistent difference between the F-SGSM-DP and the T-SGSM-DP. A possible explanation for this is the substantial presolving effort when processing the huge number of variables in the F-SGSM-DP, which might be counteracted

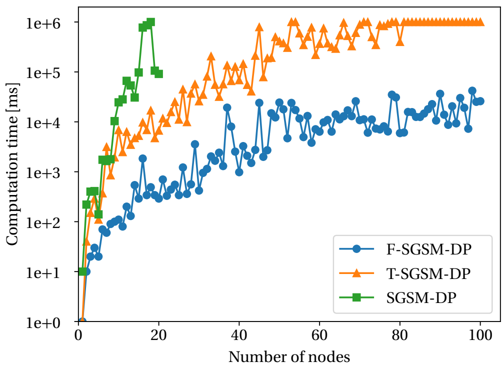

Figure 7: Comparison of Computation Times for Mild Parameters: The tightened formulation requires almost 100 times as much time to be solved as the flow-based formulation.

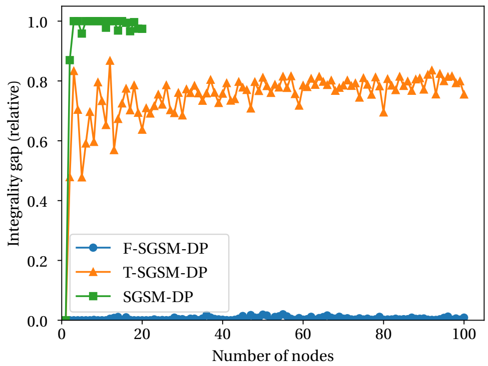

Figure 8: Comparison of Integrality Gaps for Mild Parameters: The flow-based formulation is almost tight (close to zero) whereas the big-M formulations rarely goes below 70%.

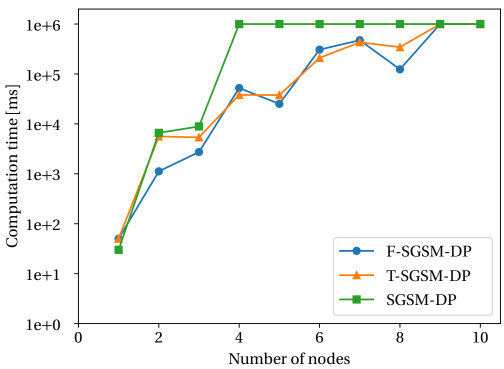

Figure 9: Comparison of Computation Times for Heavy Parameters: The tightened formulation and the flow-based formulation require roughly the same time to be solved.

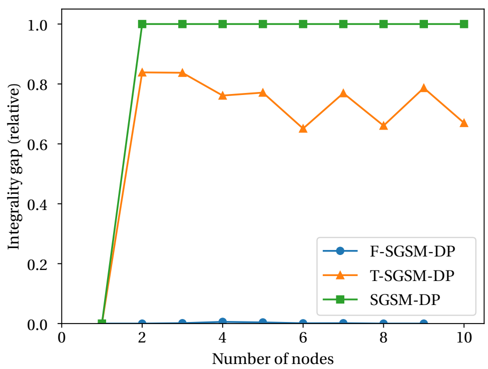

Figure 10: Comparison of Integrality Gaps for Heavy Parameters: The flow-based formulation is almost tight (close to zero) whereas the big-M formulations remains above 60%.

by a tailored decomposition approach. Nevertheless, figure 10 demonstrates again that the static LP-relaxation of the F-SGSM-DP often turns out to be tight and is still computable consistently faster than a proven optimal solution of the SGSM-DP (here by a factor of around 10).

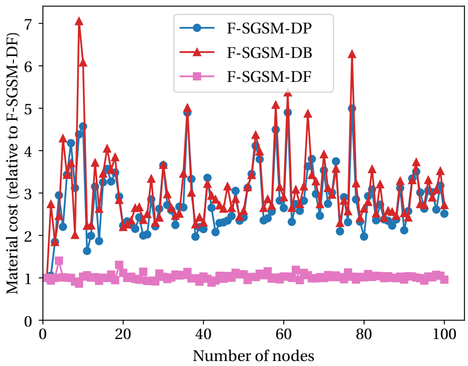

Figure 11: Comparison of Realized Costs Relative to F-SGSM-DF-objective for Mild Parameters: The F-SGSM-DF achieves lowest cost and predicts well (close to one).

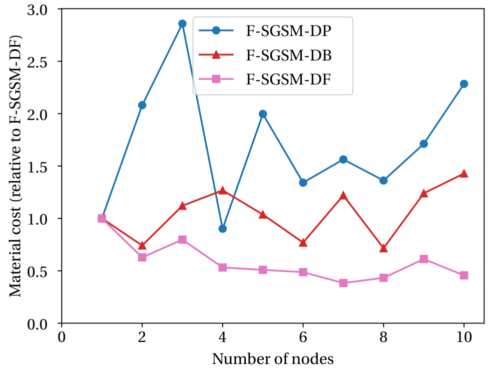

Figure 12: Comparison of Realized Costs Relative to F-SGSM-DF-objective for Heavy Parameters: The F-SGSM-DF achieves lowest cost but overestimates (lower than one).

For the evaluation of the realized cost in the dynamic application of the outsourcing-aware base-stock policy with parameter settings from the various models, the simulation is set up as follows. In every period, customer demands are generated independently over time and jointly for all demand nodes. Simulated are the same demand distributions as presented to the optimization models. Lead times remain deterministic. All nodes are driven by the outsourcing-aware base-stock policy by definition 2 for the best solutions found in the respective models within the time limit. Costs are evaluated over 1000 periods after a settle time of 100 periods starting with empty inventories. To estimate the expected cost, an average over 10 simulation runs is computed for each problem and model instance. The measured cost is the holding cost for the inventory storage over time (without pipeline storage) plus the compensation cost for the outsourcing actions. Figures 11 and 12 show the resulting total cost relative to the objective value of the F-SGSM-DF. There, the F-SGSM-DB refers to an intermediate model given by the F-SGSM-DF with holding costs for the base-stocks (as in the F-SGSM-DP). It turns out, that on the mild instance set the cost predicted by the F-SGSM-DF approximates well the realized cost in the dynamic simulation and performs consistently better than the other approaches. However, on the heavy instance set the F-SGSM-DF overestimates costs, which might be due to the longer lead times making dynamical shortcomings in the representation of the demand loads evident. Nevertheless, even for these instances the F-SGSM-DF dominates the F-SGSM-DB and the F-SGSM-DP.

Summarized, computational efficiency, dynamic cost prediction quality, and dynamic cost efficiency of the new F-SGSM-DF are a huge leap forward compared to the former models.

5 Conclusions

We have presented the outsourcing-aware base-stock policy that allows to determine smaller base-stocks in anticipation of outsourcing options. Its parameters can be optimized by the SGSM-DP-MILP by Löhnert and Rambau (2018), but only for small networks and with imprecise dynamic cost predictions. In this work, we have presented an all new MILP-formulation F-SGSM-DP for the SGSM-DP based on flow-design that has a very small integrality gap and can handle distribution networks with up to 400 nodes for mild parameters. Moreover, by modification of the objective we obtained the F-SGSM-DF. Simulations on randomized instances for demand distributions varying mildly over time show a much more precise dynamic cost prediction, where the F-SGSM-DF can still handle networks with 50 nodes in reasonable time for mild parameters.

Future research concerns demand distributions that vary strongly over time (like many days without any demand and few days with a substantial demand). Furthermore, an effective decomposition scheme for dynamic column generation would be helpful to tackle instances with hard parameters like long lead times. The F-SGSM-DP/F-SGSM-DF are still tight for such instances, but the model size grows very large, which leads to a longer computation time and a higher memory consumption with standard solvers.

References

de Kok T, Grob C, Laumanns M, Minner S, Rambau J, Schade K (2018) A typology and literature review on stochastic multi-echelon inventory models. *European Journal of Operational Research* 269(3):955–983, ISSN 0377-2217, URL `http://dx.doi.org/10.1016/j.ejor.2018.02.047`.

Eruguz AS, Sahin E, Jemai Z, Dallery Y (2016) A comprehensive survey of guaranteed-service models for multi-echelon inventory optimization. *International Journal of Production Economics* 172:110–125, ISSN 0925-5273.

Gamrath G, Anderson D, Bestuzheva K, Chen WK, Eifler L, Gasse M, Gemander P, Gleixner A, Gottwald L, Halbig K, Hendel G, Hojny C, Koch T, Le Bodic P, Maher SJ, Matter F, Miltenberger M, Mühmer E, Müller B, Pfetsch ME, Schlösser F, Serrano F, Shinano Y, Tawfik C, Vigerske S, Wegscheider F, Weninger D, Witzig J (2020) The scip optimization suite 7.0. technical report, Zuse Institut Berlin, Berlin, URL `http://www.optimization-online.org/DB_HTML/2020/03/7705.html`.

Graves SC, Willems SP (2000) Optimizing strategic safety stock placement in supply chains. *Manufacturing & Service Operations Management* 2(1):68–83, ISSN 1523-4614, URL `https://pubsonline.informs.org/doi/abs/10.1287/msom.2.1.68.23267`.

Kamp D (2021) *The Guaranteed Service Model with Demand Propagation – New Models based on Flow Networks*. Master's thesis, University of Bayreuth.

Kamp D, Rambau J (2022) Flow formulation of demand propagation in guaranteed service models. *IFAC - PapersOnLine* 55(20):505–510, ISSN 2405-8963, URL `https://www.sciencedirect.com/science/article/pii/S2405896322013465`.

Lesnaia E (2004) *Optimizing Safety Stock Placement in General Network Supply Chains*. doctoral thesis, Massachusetts Institute of Technology, Massachusetts, URL `https://dspace.mit.edu/handle/1721.1/28856`.

Löhnert M, Rambau J (2018) Modeling demand propagation in guaranteed service models. *IFAC - PapersOnLine* 51(2):284–289, ISSN 2405-8963, URL `https://www.sciencedirect.com/science/article/pii/S2405896318300533`.

Rambau J, Schade K (2014) The stochastic guaranteed service model with recourse for multi-echelon warehouse management. *Mathematical Methods of Operations Research* 79(3):293–326, ISSN 1432-2994, URL `https://link.springer.com/article/10.1007%2Fs00186-014-0464-y`.

Schade K (2012) *Stochastische Optimierung*. Stochastic Programming (Wiesbaden: Vieweg+Teubner Verlag), 1 edition, ISBN 978-3-8348-1821-8.

Simpson Jr KF (1958) In-process inventories. *Operations Research* 6(6):863–873, ISSN 0030-364X, URL `https://www.jstor.org/stable/167080`.