



HAL
open science

Modeling and solving complex multi-item lot-sizing problems with inventory constraints

Mehdi Charles

► **To cite this version:**

Mehdi Charles. Modeling and solving complex multi-item lot-sizing problems with inventory constraints. Other. Université de Lyon, 2021. English. NNT : 2021LYSEM039 . tel-03867260

HAL Id: tel-03867260

<https://theses.hal.science/tel-03867260>

Submitted on 23 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT : 2021LYSEM039

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
l'Ecole des Mines de Saint-Etienne

Ecole Doctorale N° 488
Sciences, Ingénierie, Santé

Spécialité de doctorat : Génie Industriel

Soutenue publiquement le 16/12/2021, par :
Mehdi Charles

Modeling and solving capacitated lot-sizing problems with setup times and inventory constraints

Devant le jury composé de :

Elloumi, Sourour	Professeur, ENSTA	Présidente
Gicquel, Céline	Maître de conférences, IUT Orsay	Rapporteure
Yalaoui, Alice	Maître de conférences, Université de Technologie de Troyes	Rapporteure
Absi, Nabil	Professeur, Mines Saint-Etienne	Examineur
Penz, Bernard	Professeur, Grenoble INP	Examineur
Dauzère-Pérès, Stéphane	Professeur, Mines Saint-Etienne	Directeur de thèse
Kedad-Sidhoum, Safia	Professeur, CNAM	Directrice de thèse
Mazhoud, Issam	Ingénieur, DecisionBrain	Co-encadrant
Focacci, Filippo	Ingénieur, DecisionBrain	Invité
Godard, Daniel	Ingénieur, DecisionBrain	Invité

Spécialités doctorales
 SCIENCES ET GENIE DES MATERIAUX
 MECANIQUE ET INGENIERIE
 GENIE DES PROCEDES
 SCIENCES DE LA TERRE
 SCIENCES ET GENIE DE L'ENVIRONNEMENT

Responsables :
 K. Wolski Directeur de recherche
 S. Drapier, professeur
 F. Gruy, Maître de recherche
 B. Guy, Directeur de recherche
 V.Laforest, Directeur de recherche

Spécialités doctorales
 MATHEMATIQUES APPLIQUEES
 INFORMATIQUE
 SCIENCES DES IMAGES ET DES FORMES
 GENIE INDUSTRIEL
 MICROELECTRONIQUE

Responsables
 M. Batton-Hubert
 O. Boissier, Professeur
 J.C. Pinoli, Professeur
 N. Absi, Maître de recherche
 Ph. Lalevé, Professeur

EMSE : Enseignants-chercheurs et chercheurs autorisés à diriger des thèses de doctorat (titulaires d'un doctorat d'Etat ou d'une HDR)

ABSI	Nabil	MR	Génie industriel	CMP
AUGUSTO	Vincent	MR	Génie industriel	CIS
AVRIL	Stéphane	PR	Mécanique et ingénierie	CIS
BADEL	Pierre	PR	Mécanique et ingénierie	CIS
BALBO	Flavien	PR	Informatique	FAYOL
BASSEREAU	Jean-François	PR	Sciences et génie des matériaux	SMS
BATTON-HUBERT	Mireille	PR	Mathématiques appliquées	FAYOL
BEIGBEDER	Michel	MA	Informatique	FAYOL
BILAL	Blayac	DR	Sciences et génie de l'environnement	SPIN
BLAYAC	Sylvain	PR	Microélectronique	CMP
BOISSIER	Olivier	PR	Informatique	FAYOL
BONNEFOY	Olivier	PR	Génie des Procédés	SPIN
BORBELY	Andras	DR	Sciences et génie des matériaux	SMS
BOUCHER	Xavier	PR	Génie Industriel	FAYOL
BRUCHON	Julien	PR	Mécanique et ingénierie	SMS
CAMEIRAO	Ana	PR	Génie des Procédés	SPIN
CHRISTIEN	Frédéric	PR	Science et génie des matériaux	SMS
DAUZERE-PERES	Stéphane	PR	Génie Industriel	CMP
DEBAYLE	Johan	MR	Sciences des Images et des Formes	SPIN
DEGEORGE	Jean-Michel	MA	Génie industriel	Fayol
DELAFOSSE	David	PR	Sciences et génie des matériaux	SMS
DELORME	Xavier	PR	Génie industriel	FAYOL
DESRAYAUD	Christophe	PR	Mécanique et ingénierie	SMS
DJENIZIAN	Thierry	PR	Science et génie des matériaux	CMP
BERGER-DOUCE	Sandrine	PR	Sciences de gestion	FAYOL
DRAPIER	Sylvain	PR	Mécanique et ingénierie	SMS
DUTERTRE	Jean-Max	PR	Microélectronique	CMP
EL MRABET	Nadia	MA	Microélectronique	CMP
FAUCHEU	Jenny	MA	Sciences et génie des matériaux	SMS
FAVERGEON	Loïc	MR	Génie des Procédés	SPIN
FEILLET	Dominique	PR	Génie Industriel	CMP
FOREST	Valérie	PR	Génie des Procédés	CIS
FRACZKIEWICZ	Anna	DR	Sciences et génie des matériaux	SMS
GAVET	Yann	MA	Sciences des Images et des Formes	SPIN
GERINGER	Jean	MA	Sciences et génie des matériaux	CIS
GONDRAN	Natacha	MA	Sciences et génie de l'environnement	FAYOL
GONZALEZ FELIU	Jesus	MA	Sciences économiques	FAYOL
GRAILLOT	Didier	DR	Sciences et génie de l'environnement	SPIN
GRIMAUD	Frederic	EC	Génie mathématiques et industriel	FAYOL
GROSSEAU	Philippe	DR	Génie des Procédés	SPIN
GRUY	Frédéric	PR	Génie des Procédés	SPIN
HAN	Woo-Suck	MR	Mécanique et ingénierie	SMS
HERRI	Jean Michel	PR	Génie des Procédés	SPIN
ISMAILOVA	Esma	MC	Microélectronique	CMP
KERMOUCHE	Guillaume	PR	Mécanique et Ingénierie	SMS
KLOCKER	Helmut	DR	Sciences et génie des matériaux	SMS
LAFOREST	Valérie	DR	Sciences et génie de l'environnement	FAYOL
LERICHE	Rodolphe	DR	Mécanique et ingénierie	FAYOL
LIOTIER	Pierre-Jacques	MA	Mécanique et ingénierie	SMS
MEDINI	Khaled	EC	Sciences et génie de l'environnement	FAYOL
MOLIMARD	Jérôme	PR	Mécanique et ingénierie	CIS
MOULIN	Nicolas	MA	Mécanique et ingénierie	SMS
MOUTTE	Jacques	MR	Génie des Procédés	SPIN
NAVARRO	Laurent	MR	Mécanique et ingénierie	CIS
NEUBERT	Gilles	PR	Génie industriel	FAYOL
NIKOLOVSKI	Jean-Pierre	Ingénieur de recherche	Mécanique et ingénierie	CMP
O CONNOR	Rodney Philip	PR	Microélectronique	CMP
PICARD	Gauthier	PR	Informatique	FAYOL
PINOLI	Jean Charles	PR	Sciences des Images et des Formes	SPIN
POURCHEZ	Jérémy	DR	Génie des Procédés	CIS
ROUSSY	Agnès	MA	Microélectronique	CMP
SANAUR	Sébastien	MA	Microélectronique	CMP
SERRIS	Eric	IRD	Génie des Procédés	FAYOL
STOLARZ	Jacques	CR	Sciences et génie des matériaux	SMS
VALDIVIESO	François	PR	Sciences et génie des matériaux	SMS
VIRICELLE	Jean Paul	DR	Génie des Procédés	SPIN
WOLSKI	Krzystof	DR	Sciences et génie des matériaux	SMS
XIE	Xiaolan	PR	Génie industriel	CIS
YUGMA	Gallian	MR	Génie industriel	CMP

Acknowledgments

These past years were full of meaningful encounters. Doing a PhD is not the easiest task, and there are many people I would like to thank.

First of all, I would like to thank my academic and industrial supervisors. Many thanks to both my academic supervisors, Stéphane Dauzère-Pérès and Safia Kedad-Sidhoum, for always being there to guide and help me with my research, but still letting me have flexibility in my work and believing in me. I have never felt left out during these past years and that is something I am grateful for.

I would like to thank Filippo Focacci, Daniel Godard and Issam Mazhoud for enabling me to work in the best conditions possible at DecisionBrain and for trusting me.

I would also like to thank the jury of my PhD defense for the interest they showed for my work and for the relevant questions they raised during the defense. I would like to give a special thank to the reviewers of my manuscript, Alice Yalaoui and Céline Gicquel, who provided great feedbacks and very insightful remarks.

I thank all my colleagues at DecisionBrain and at the Centre Microélectronique de Provence in Gardanne for the various exchanges and for the nice working environment. I have not been able to travel to Gardanne as much as I would have like to, but I have always spend a great time there. In both Paris and Gardanne I have met great people that helped me both in and out of the scope of my PhD.

Lastly, I thank my family and my friends and for always being there for me during these past three years.

Contents

General introduction	1
1 Industrial context and a short state of the art	5
1.1 Production planning	5
1.1.1 Lot-sizing in the industry	5
1.1.2 Production planning at DecisionBrain	6
1.2 Industrial application: Spare part manufacturer	8
1.3 State of the art	9
1.3.1 Modeling lot-sizing problems	9
1.3.2 Solution approaches	12
1.4 Conclusion	14
2 Capacitated lot-sizing problem with minimum and maximum ending inventories and setup times	17
2.1 Related works	18
2.2 Motivations	19
2.2.1 Problem formulation	20
2.2.2 Impact of global ending inventory	21
2.2.3 Planning on a rolling horizon	22
2.2.4 Initial inventories	25
2.3 Inventory levels for the capacitated lot-sizing problem with setup times on a rolling horizon	25
2.3.1 Multi-item cyclical production planning with bounded average capacity consumption	26
2.3.2 Similarities between (P_C) and the optimization problem of Trigeiro et al. (1989)	27
2.3.3 Analysis of the continuous relaxation of (P_C)	28
2.3.4 Lost sales costs	30
2.4 Extensions to more general cases	34
2.4.1 Average demand per item	34
2.4.2 Average demand per item and correlated setup costs and times	34
2.5 New instance generation scheme	35
2.5.1 Original generation scheme	35
2.5.2 New generation scheme	36
2.5.3 Fitting I^{inf} and setting initial inventories	37
2.6 Computational experiments	39
2.6.1 Comparison of original and new instances	39
2.6.2 Analysis on a rolling horizon	45

2.7	Conclusions	48
3	Capacitated lot-sizing problem with inventory constraints within periods	51
3.1	Related works	52
3.2	Motivations	52
3.2.1	Original problem formulation	52
3.2.2	Inventory evolution within periods	53
3.3	Uniform production and demand within a cone of uncertainty	54
3.3.1	Demand approximation within a cone of uncertainty	54
3.3.2	Modeling inventory constraints	56
3.3.3	Worst case errors	59
3.4	Production at maximum rate and instantaneous demand	61
3.4.1	Modeling inventory constraints	61
3.4.2	The uncapacitated lot-sizing problem with constant inventory bounds and inventory constraints within periods	62
3.5	Computational experiments	65
3.5.1	Planning reconstruction	65
3.5.2	Analysis	66
3.6	Additional studies	68
3.6.1	Backlogs	68
3.6.2	Production and demand at maximum rates	69
3.6.3	Joint storage capacity	71
3.7	Conclusions	73
4	Parallelized decomposition approaches for capacitated lot-sizing problems	77
4.1	Related works	78
4.2	Problem formulations	78
4.2.1	General problem formulation	78
4.2.2	Disaggregated formulation	80
4.3	Parallelized relax-and-fix	81
4.3.1	Principle	82
4.3.2	Parallelized relax-and-fix	83
4.3.3	Reconstruction method	84
4.4	A Lagrangian relaxation based heuristic	86
4.4.1	Relaxation of the capacity constraints	86
4.4.2	Lagrangian heuristic	87
4.4.3	Forward and backward passes	89
4.4.4	Vertical pass	90
4.4.5	Deriving a feasible solution for (P)	91
4.5	Uncapacitated lot-sizing problem without lost sales and with ending inventory constraints	91
4.5.1	Problem analysis	91
4.5.2	Heuristic approach	93
4.5.3	Polynomial case	94
4.6	Computational experiments	96
4.6.1	Instance generation scheme	96
4.6.2	Analysis	97
4.7	Conclusions	100

5 Industrial application	103
5.1 Introduction	103
5.2 Complex lot-sizing constraints and industrial application	104
5.2.1 Minimum production quantities over several periods	104
5.2.2 Shared resources and setup states	104
5.2.3 Problem formulation for a single resource	106
5.2.4 Multi-objective optimization	107
5.3 Extending the relax-and-fix heuristic	107
5.3.1 Infeasibility issues	107
5.3.2 Extending the parallelized relax-and-fix to multi-objective optimization	109
5.3.3 Other types of decompositions	109
5.3.4 Updated parallelized relax-and-fix	109
5.4 Computational experiments	111
5.5 Conclusion and perspectives	113
6 Conclusions and perspectives	117
6.1 Conclusions	117
6.2 Perspectives	118
6.2.1 Short-term perspectives	118
6.2.2 Mid-term perspectives	119
6.2.3 Long-term perspectives	120
A Extended summary in French	123
List of Figures	vi
List of Tables	ix
List of Algorithms	xi
Bibliography	xiii

General introduction

Lot-sizing problems aim at determining a production or distribution plan that satisfies demands over a planning horizon discretized into periods. The objective is to find a production plan that minimizes the total production, inventory and setup costs. Lot-sizing problems are very common in manufacturing and logistics as there can be important economical incentives in making rational production and distribution decisions. The complexity of the problems varies depending on the constraints that are considered. The first objective of the thesis is to study lot-sizing problems that are closer to industrial reality and with better management of inventory constraints. To solve lot-sizing problems of the complexity and size of the problems that are faced at DecisionBrain, using an industrial solver is not sufficient to provide good solutions in reasonable times. Consequently, the second objective of this thesis is to develop generic heuristics to solve the addressed lot-sizing problems. We have focused our work on extending methods that are already widely used to solve lot-sizing problems.

Contributions

- The first contribution of this thesis consists in analyzing literature instances commonly used as a benchmark for studies on the capacitated lot-sizing problem and point out biases and issues with regards to the optimal production plans for these instances compared to industrial reality. We propose a method to modify these instances in order to make them more realistic.
- Lot-sizing problems are oftentimes optimized on a rolling horizon to fix production decisions. As described later in the manuscript, these decisions are affected by the end-of-horizon effect. To mitigate the end-of-horizon effect and its impact on a rolling horizon, we introduce an extension of the capacitated lot-sizing problem with setup times and lost sales by adding minimum and maximum ending inventory constraints. We show in numerical experiments that these new constraints can reduce the end-of-horizon effect on a rolling horizon. We extend the notion of Time Between Order to take into account the impact of capacity when considering the problem with setup times. We also define relevant lost sales costs based on the analysis of a static and cyclical lot-sizing model.
- In order to better take inventory constraints into account when solving lot-sizing problems, we study the evolution of inventory within periods and propose two new models to better model the dynamic nature of the inventory evolution. These two models make different assumptions on the shapes of production and demand evolution that are relevant in different contexts. In the first model, we consider a uniform production rate within the period and a demand within a cone of uncertainty. Under these

assumptions, we bound the inventory by adding inventory constraints within each period. In the second model, we consider an instantaneous demand and a production at a maximum production rate. We add inventory constraints within each period as well. We also add new capacity constraints before and after each demand.

- Finally, the last contribution of this thesis consists in extending commonly used heuristics to solve lot-sizing problems by using parallelization. We have defined a parallelized version of the relax-and-fix heuristic and extended a Lagrangian relaxation based heuristic to tackle the CLSP with minimum and maximum ending inventory constraints. We introduce strategies to pick intervals to fix at each iteration in the parallelized relax-and-fix and compare the results with the relax-and-fix heuristic. The Lagrangian relaxation based heuristic is based on a relaxation of the capacity constraints combined with a reconstruction heuristic to obtain good feasible solutions. We show that the related uncapacitated problem is polynomial when the maximum ending inventory is identical for all items and NP-hard otherwise. We also propose another version of the parallelized relax-and-fix algorithm that we implemented in an optimization tool. We test this heuristic on a real industrial problem with production constraints that extend over several periods and multiple machines working in parallel.

Thesis structure

In Chapter [1](#), we present the industrial context and a short state of the art. We first focus on production planning and complex lot-sizing problems faced by DecisionBrain. We provide a state of the art on some classes of lot-sizing problems related to the ones considered in this thesis.

In Chapter [2](#), we define adequate ending inventory levels and relevant indicators to analyze the shape of an optimal solution and mitigate the end-of-horizon effect. We perform an analysis on a static and cyclical model to define relevant indicators and inventory levels on a rolling horizon.

In Chapter [3](#), we point out the importance of inventory management at each period when solving manufacturing problems. In this chapter, we propose two ways to model the demand and production rates and identify the minimum and maximum inventory levels reached within each period. Numerical experiments to define the relevance and use of each model are performed.

In Chapter [4](#), we introduce new decomposition methods to solve generic lot-sizing problems. These heuristics can be fastened by using parallelization. Two heuristics are proposed to solve the capacitated lot-sizing problem with ending inventory constraints: (1) A parallelized version of the relax-and-fix algorithm, and (2) A Lagrangian relaxation based heuristic. We perform numerical experiments to assess the efficiency of these new heuristics.

In Chapter [5](#), we discuss the integration of the parallelized relax-and-fix heuristic in the planning engine developed by DecisionBrain. We discuss the limits of this method applied to generic lot-sizing problems, and propose ways to cope with these issues. The adapted heuristic is tested on an industrial capacitated lot-sizing problem.

Chapter [6](#) sums up the contributions of the manuscript and suggests perspectives and future works.

Chapter 1

Industrial context and a short state of the art

Since the middle of the 20th century, lot-sizing has been a very active research field, in particular because of its numerous applications in manufacturing and logistics. Yet, the first results on lot-sizing come from the beginning of the 20th century, where [Harris \(1913\)](#) developed the notion of Economic Order Quantity (EOQ) for a cyclical and stationary single-item production planning problem. A global overview on the various types of lot-sizing problems can be found in [Pochet and Wolsey \(2006\)](#). In this section, we point out the complexity of modeling industrial production planning problems and the efforts in the literature to better model and solve these problems. In order to emphasize the correlation between industrial problems and lot-sizing problems, we introduce an industrial application and the corresponding lot-sizing model. We point out that, for some of the constraints of this lot-sizing model, there are little or no literature references.

In [Section 1.1](#), we describe the role and evolution of production planning in the industry and, in [Section 1.2](#), an industrial application in production planning. This industrial application will be further analyzed in [Chapter 5](#). [Section 1.3](#) provides a state of the art on the modeling of lot-sizing problems and on solutions approaches.

1.1 Production planning

1.1.1 Lot-sizing in the industry

Production management problems in the industry usually can be divided into three hierarchical levels of decisions:

- **Strategic:** This level corresponds to long range decision making, which spans over several years. The strategies to pursue are defined at this level, as well as the most critical resources to acquire and allocate.
- **Tactical:** This level corresponds to middle range decision making, which usually applies on several weeks or months. Based on the capacity allocated at the strategic level, this is typically at this level that lot-sizing decisions are taken based on customer orders and demand forecasts. Lot-sizing problems usually consist in determining the lots to produce for each item and in each period of a planning horizon.

- **Operational:** This level corresponds to short range decision making, which is taken for several hours or days, and can include real time decisions, that are sometimes considered in a fourth level. The lots decided at the tactical level are typically sequenced and scheduled at the operational level on the resources of the factory.

Being able to make good decisions at each level is crucial because each level is linked to the others. Lot-sizing problems are very important in the industry. They are mainly used for tactical planning in order to make general decisions about the quantities to produce for multiple items on large time buckets (weeks or months), i.e. periods, in a planning horizon. Lot-sizing problems can also be used on smaller time buckets, i.e. days or even hours. Lot-sizing problems adapted to the industry need to better model the industrial reality, with models becoming more and more complex (new constraints and objectives, integration of other problems, stochastic parameters, ...). Usually industrial problems are solved on a rolling horizon, meaning that, in order to fix the production plan for the next period, decisions over a longer planning horizon are optimized. Demands are approximated, and the longer the planning horizon the riskier it is to approximate future demands. A review of lot-sizing in the industry can be found in [Jans and Degraeve \(2008\)](#). We can also cite [Vofsi and Woodruff \(2006\)](#) for a review of optimization models for production planning problems. Production planning problems often combine a lot-sizing problem with a scheduling problem to sequence the productions ([Elmaghraby \(1978\)](#), [Dauzère-Pérès and Lasserre \(2002\)](#), [Seeanner and Meyr \(2013\)](#)). The main feature of a lot-sizing problem is that the time is divided into discrete periods, while in practice the schedule occurs in continuous time ([Dobson \(1987\)](#), [Holmbom and Segerstedt \(2014\)](#)).

1.1.2 Production planning at DecisionBrain

DecisionBrain [\[1\]](#) is a software company that specializes in providing optimization solutions primarily in the fields of production planning and workforce management. DecisionBrain is an IBM business partner and thus often uses in its solutions the IBM's Mixed Integer Programming solver CPLEX and Constraint Programming solver CPO. The range of problems varies from routing problems to scheduling problems to very complex multi-level lot-sizing problems. DecisionBrain develops an optimization engine (called the Planning Engine) that uses IBM ILOG CPLEX and that aims at solving a large variety of lot-sizing problems. In this section, we give insights on issues that might occur when trying to solve industrial problems.

Complex constraints

The constraints that are modeled in industrial lot-sizing problems are more complex than the problems that are usually studied in the academic literature. An industrial problem oftentimes has specific rules. For instance, many industrial problems have multiple levels or bounds on the production and inventory quantities. The planning engine developed by DecisionBrain can handle several types of lot-sizing problems such as multi-level problems with parallel machines, bounds on the decision variables.

Some of the constraints that are modeled in the Planning Engine have been studied in the literature. For instance, in the Planning Engine, demands have time windows to be satisfied instead of being defined per period ([Lee et al. \(2001\)](#), [Dauzère-Pérès et al. \(2002\)](#)). Another

¹<https://www.decisionbrain.com/>

feature concerns the perishability of items that cannot remain in the inventory for more than a given number of periods (Friedman and Hoch (1978)). It is also possible to model lead times, meaning that a production for an item will not be available in the inventory before a given number of periods (Billington et al. (1986), Berretta et al. (2005), Almeder et al. (2015)). Other features implemented in the engine have not been extensively studied in the literature, such as a minimum production quantity that spans over several periods.

Because industrial problems come from many different sectors with different planning horizons and period lengths, it is a challenging task to develop a generic solver for lot-sizing problems. Some constraints of problems faced at DecisionBrain also model the link between production planning and other types of problem, such as routing or scheduling problems. These constraints are modeled in the Planning Engine. For instance it is possible to model transfer processes to move inventory from one location to another.

Integrating lot-sizing and scheduling problems

Decisions taken after solving a lot-sizing problem are aggregated, meaning that for each period only the production quantities are known and not the sequence of the related production operations. Many production management problems combine both a lot-sizing problem (*tactical*) and a scheduling problem (*operational*). Once the quantities to produce have been decided, the best production sequence that satisfies the operational constraints in continuous time needs to be found. There can be a decoupling between lot-sizing and scheduling decisions. For instance, in a two-level production planning problem where an intermediate product needs to be processed on one machine to execute another item on another machine, then the production sequence is very important because the decisions taken after solving the lot-sizing problem may lead to an infeasible schedule. Another example is the dynamic evolution of the inventory within each period. If the inventory has a maximum (or minimum) level, the classical lot-sizing constraints only ensure that the maximum (or minimum) inventory level is satisfied at the very beginning and at the very end of each period. If a period has a duration of one month and the demand only occurs in a single day, the exact time of the production within the period can affect the feasibility of the production plan. The combination of lot-sizing and scheduling decisions has been extensively studied in the literature and models with small time buckets were defined, such as the Discrete Lot-Sizing and Scheduling Problem (Fleischmann (1990)), where at most one item is produced at each period and this production has to last for the entire period. The Continuous Setup Lot-Sizing Problem (Karmarkar et al. (1987)) is similar but the production does not have to last for the entire period. Some lot-sizing problems mix big and small buckets models such as the General Lot-Sizing and scheduling problem (Fleischmann and Meyr (1997), Copil et al. (2017)) where macro and micro-periods are defined such that each macro-period is composed of a set of micro-periods and at most one setup can occur for each micro-period. There is still progress to be made, especially in developing heuristics to quickly find good solutions for very large instances. Small time bucket models often lead to a higher number of decision variables because periods are usually smaller. A survey on the combination of lot-sizing and scheduling can be found in Drexl and Kimms (1997).

Complex structure and instance size

An additional complexity of industrial instances is the size of the problem instances that need to be solved. Industrial instances can have thousands of demands for hundreds of items. By

comparison, the literature instances proposed by [Trigeiro et al. \(1989\)](#) and often used to benchmark solution approaches to solve the capacitated lot-sizing problem with setup times have at most 30 items. In most industrial production planning problems, items are produced by a set of machines with limited capacity, thus decisions must be taken not only on the quantities to produce but also on the capacity allocation on each machine. Oftentimes, the final demands concern end items that require several production steps, requiring the production of intermediate items at each step.

1.2 Industrial application: Spare part manufacturer

The industrial application introduced in this section concerns a manufacturer of automobile spare parts. This example will be used in [Chapter 5](#) to adapt the parallelized relax-and-fix heuristic proposed in [Chapter 4](#) to more generic lot-sizing problems. The finished products are spare parts that are first produced and painted, and then assembled by customer. Thus, for the manufacturer, the demands are on the painted spare parts. The manufacturing process is performed sequentially as follows:

1. A first workshop produces the spare parts using molds to obtain the correct shapes.
2. A second workshop paints the spare parts coming from the first workshop in order to satisfy the demands for the finished products.



Figure 1.1: Painting spare parts ²

The two workshops being linked, decisions taken at the first workshop deeply affect the quality of the decisions in the second workshop. Indeed, the painting line schedule highly depends on the available inventory of non-painted spare parts. Yet, the available inventory at each period is a consequence of the decisions taken for the whole production line. The decisions of the second workshop also affect the decisions in the first workshop because the demands for non-painted spare parts are deduced from the forecasted schedule of the painting line. The planning problem in the first workshop has been modeled as a lot-sizing problem. The problem is solved using the Planning Engine of DecisionBrain. Because the schedule of the painting line needs to be synchronized with the production decisions on the first

²Attribution: [RoboGuru](#), [CC BY-SA 3.0](#), via Wikimedia Commons

workshop, demands for non-painted spare parts are defined every hour. The time buckets are small (two hours) to match the demands. The planning horizon is set to one day and about one hundred different spare parts need to be produced.

Inventory management and constraints are critical in this problem, which was one of the motivations to tackle this problem in this thesis. The problem includes many complex key features of lot-sizing problems. The spare parts are produced on a set of parallel machines. Each machine needs to use a specific mold to produce each spare part, and a mold can be shared by several machines. Inventory constraints are considered to limit the inventory of non-painted spare parts.

Because the periods are short and the setup times are a large fraction of the period length, the production of one spare part usually lasts for more than one period. To model this, the lot-sizing model allows what are called setup carry-overs, and the minimum order quantities can exceed the capacity available in a single period. The consideration of the inventory evolution within periods, which is the topic of Chapter 3 and especially the model proposed in Section 3.3, is relevant for this type of problems.

1.3 State of the art

In this section, we give a general overview of the literature regarding both the modeling of lot-sizing problems and the solution approaches proposed to solve these problems. This state of the art focuses on the problems we are analyzing throughout the manuscript. In each chapter, a specific literature review related to the problems that are studied in this work is given in the introduction sections.

1.3.1 Modeling lot-sizing problems

The literature on lot-sizing started as early as the beginning of the 1900's in the economic literature with the definition by Harris of the Economic Order Quantity for a single-item lot-sizing problem with a continuous production evolution (Harris (1913)). We provide in this section a short review of the modeling of dynamic lot-sizing problems, first for single-item problems and then for multi-item capacitated problems.

Single-item lot-sizing problems

The single-item uncapacitated lot-sizing problem was first introduced by Wagner and Whitin in the late 1950' (Wagner and Whitin (1958)). This problem is polynomial and can be solved using dynamic programming. It is a guideline for many lot-sizing problems, yet it is among the few lot-sizing problems that can be polynomially solved. A survey on single-item lot-sizing problems can be found in Brahim, Dauzère-Pérès, Najid and Nordli (2006) and was updated in Brahim et al. (2017). The problem is an uncapacitated single-item lot-sizing problem where production, inventory and setup decisions need to be taken over a planning horizon of T periods to meet the demand of a single item at each period. In this manuscript, as it is most often the case in lot-sizing papers, we assume that the costs are linear. Each period $t \in \llbracket 1, T \rrbracket$ induces a fixed setup cost f_t if there is production in t , a variable processing cost per unit p_t and a holding cost per unit h_t .

Let us define for each $t \in \llbracket 1, T \rrbracket$ the decision variables as follows:

- $X_t \geq 0$: Quantity produced at period t ,
 $Y_t \in \{0, 1\}$: Setup variable equals to 1 if there is an order at period t , and 0 otherwise,
 $I_t \geq 0$: Inventory at the end of period t .

The formulation of the single-item uncapacitated lot-sizing problem is given below:

$$\min \sum_{t=1}^T (f_t Y_t + p_t X_t + h_t I_t) \quad (1.1)$$

$$I_{t-1} + X_t = d_t + I_t, \quad \forall t \in 1, \dots, T \quad (1.2)$$

$$X_t \leq M_t Y_t, \quad \forall t \in 1, \dots, T \quad (1.3)$$

$$Y_t \in \{0, 1\}, \quad \forall t \in 1, \dots, T \quad (1.4)$$

$$X_t, I_t \geq 0, \quad \forall t \in 1, \dots, T \quad (1.5)$$

The objective function (1.1) minimizes the total production, setup and inventory costs over the planning horizon. Constraints (1.2) are the flow conservation constraints that balance the inventory at period $t-1$ and the production quantities at period t with the inventory and the demand at period t . Constraints (1.3) link the continuous production variables with the binary setup variables, M_t being an upper bound on the optimal production quantity (for instance $M_t = \sum_{k=t}^T d_k$). Constraints (1.4) and (1.5) define the domain of the variables.

Because periods are defined by discretizing time and thus have fixed durations, a natural extension to this problem is to limit the production at each period. A maximum capacity c_t^{max} can be consumed by production at each period t (Florian and Klein (1971)).

$$X_t \leq c_t^{max}, \quad \forall t \in 1, \dots, T \quad (1.6)$$

Constraints (1.6) can then be extended by adding a fixed processing time s_t and a processing time per unit b_t in period t . The capacity consumed in each period by both the fixed and the variable processing times must not exceed the maximum capacity.

$$s_t Y_t + b_t X_t \leq c_t^{max}, \quad \forall t \in 1, \dots, T \quad (1.7)$$

Constraints (1.7) ensure that the capacity consumed by setup and production times does not exceed the maximum production capacity c_t^{max} .

In case the capacity at a given period is not sufficient to satisfy all the demands, lost sales can be considered to allow some of the demands not to be fulfilled (Sandbothe and Thompson (1990), Aksen et al. (2003), Absi and Kedad-Sidhoum (2009)). Let us introduce the variable L_t that corresponds to the number of unsatisfied units of demand d_t , and the parameter l_t that represents the unit lost sales cost. The formulation of the single-item lot-sizing problem with lost sales is given below:

$$\min \sum_{t=1}^T (f_t Y_t + p_t X_t + h_t I_t + l_t L_t) \quad (1.8)$$

$$I_{t-1} + X_t + L_t = d_t + I_t, \quad \forall t \in 1, \dots, T \quad (1.9)$$

$$L_t \leq d_t, \quad \forall t \in 1, \dots, T \quad (1.10)$$

$$L_t \geq 0, \quad \forall t \in 1, \dots, T \quad (1.11)$$

$$(1.3), (1.4), (1.5)$$

Lost sales costs are added to the objective function (1.8). The material flow constraints (1.9) are modified to consider only the satisfied demand at each period that is $d_t - L_t$. Constraints (1.10) bound the number of lost sales by the demand at each period.

This problem was then extended to the possibility for some demands to be backlogged, hence not satisfied at the period they occur but later on the planning horizon (Zangwill (1966), Millar and Yang (1994)). Let us introduce the variable B_t that corresponds to the number of units of demand that are backlogged at the end of period t , and the parameter b_t that represents the unit backlog cost. The formulation of the single-item lot-sizing problem with backlogs is given below:

$$\min \sum_{t=1}^T (f_t Y_t + p_t X_t + h_t I_t + b_t B_t) \quad (1.12)$$

$$I_{t-1} + B_t + X_t = d_t + I_t + B_{t-1}, \quad \forall t \in 1, \dots, T \quad (1.13)$$

$$X_t \leq \left(\sum_{t=1}^T d_t \right) Y_t, \quad \forall t \in 1, \dots, T \quad (1.14)$$

$$B_t \geq 0, \quad \forall t \in 1, \dots, T \quad (1.15)$$

$$(1.4), (1.5)$$

Backlog costs are added to the objective function (1.12). The material flow constraints (1.13) are modified to consider only the satisfied demand at period t or the previous demands that were backlogged. The satisfied demand at t , taking into account the previous backlog, is $d_t - B_t + B_{t-1}$.

Multi-item lot-sizing problems

Throughout the years, lot-sizing problems have become more and more complex because of the need to model industrial constraints as precisely as possible.

In the 1980's, the single-item lot-sizing problem was extended to take into account the production of multiple items, which is a more realistic representation of industrial problems. In the single machine problem, the decisions on the production quantities of N items are interrelated because the machine is shared by all the items. Let us extend the production, inventory and setup variables by adding an index $i \in \llbracket 1, N \rrbracket$ that corresponds to item i . Let us define a fixed processing time s_{it} and a variable processing time b_{it} for each item $i \in \llbracket 1, N \rrbracket$ at each period $t \in \llbracket 1, T \rrbracket$.

$$\sum_{i=1}^N s_{it} Y_{it} + b_{it} X_{it} \leq c_t^{max}, \quad \forall t \in 1, \dots, T \quad (1.16)$$

Multi-item capacitated lot-sizing problems are more complex and harder to solve than single-item uncapacitated lot-sizing problems. For the capacitated lot-sizing problem with setup times, even the feasibility problem is NP-complete (Trigeiro et al. (1989)). The capacitated lot-sizing problem with setup times (CLSP) has been extensively studied in the literature (Manne (1958), Bitran and Matsuo (1986), Barany et al. (1984)) and remains a very active field of study.

Additional constraints on the production quantities can be considered. Indeed, in some cases, a production might only be possible or beneficial if the quantity produced exceeds a given quantity. This can be because of hard constraints (for instance production in batches),

or because it is not profitable to setup a machine for small production quantities, even if it leads to the best solution when only relying on the costs and on the capacity constraints. To that intent, many articles study the CLSP with a Minimum Order Quantity (Anderson and Cheah (1993), Hellion et al. (2012), Park and Klabjan (2015)). Let us introduce q_{it}^{min} , the minimum production quantity for item i at period t .

$$X_{it} \geq q_{it}^{min} Y_{it}, \quad \forall t \in 1, \dots, T, \forall i \in 1, \dots, N \quad (1.17)$$

Inventory management can also be important to take into account when solving a lot-sizing problem (Liu and Tu (2008)). It might correspond to a guideline as to how the production is supposed to be shaped (lean manufacturing, Time Between Order, ...). It can also take the form of hard inventory constraints (Love (1973), Gutiérrez et al. (2003), Liu (2008)) that come from warehouse capacities or minimum inventory quantities to provide robustness to the production plan. Soft constraints are also considered, mainly in the form of safety stocks (Absi and Kedad-Sidhoum (2009)), which are also used to provide robustness to the production plan, penalizing the inventory below a certain level. However, safety stocks are less restrictive than minimum inventory levels. If I_{it} (resp. \bar{I}_{it}) denotes the minimum (resp. maximum) inventory level for each item at each period, we have:

$$I_{it} \geq \underline{I}_{it}, \quad \forall t \in 1, \dots, T, \forall i \in 1, \dots, N \quad (1.18)$$

$$I_{it} \leq \bar{I}_{it}, \quad \forall t \in 1, \dots, T, \forall i \in 1, \dots, N \quad (1.19)$$

Because it is not realistic to consider a problem with only a single machine producing all items, a natural extension to the lot-sizing problem is to consider multiple machines that can perform operations in parallel. Each item can be produced by a subset of machines. A dimension is added to single-machine lot-sizing problems, where not only the quantities produced at each period must be optimized, but also the distribution of the production quantities between machines (Fiorotto et al. (2015)). Because of uncertainties in the exact period demands are collected, models with different modeling of the demands are considered. For instance, the lot-sizing problem with time windows considers that each demand needs to be fulfilled within a specific period interval (Lee et al. (2001)). Time windows have also been defined for production quantities in Dautère-Pères et al. (2002) and Brahimi, Dautère-Pères and Najid (2006). Both types of time windows are considered in Absi et al. (2011).

Recently, the focus has been on the study of lot-sizing problems with more than one level which have many applications. For instance, the production of one item can require an intermediate item (Maes et al. (1991), Stadtler (2003), Helber and Sahling (2010)). Several other industrial aspects that go beyond the scope of this thesis have been modeled as lot-sizing problems. A thorough review of lot-sizing models to solve production planning problems can be found in Díaz-Madroño et al. (2014).

1.3.2 Solution approaches

Several directions can be taken when designing approaches to solve a lot-sizing problem. As the problems become more and more complex and the size of the instances to solve grows, the need to find solution approaches that can apply to lot-sizing problems with complex constraints and provide good solutions becomes critical.

Solution approaches for uncapacitated lot-sizing problems

For some lot-sizing problems, dynamic programming methods can be implemented using the structural properties an optimal solutions. For the uncapacitated single-item lot-sizing problem for instance, Wagner and Whitin proved in the late 50's the so-called zero-inventory property and developed an algorithm to solve this problem using dynamic programming (Wagner and Whitin (1958)) in $O(T^2)$ under concavity assumptions on the costs. This complexity was later improved to $O(T \log T)$ (Federgruen and Tzur (1991), Wagelmans et al. (1992)). Heuristics were also proposed, such as the simple one in Silver and Meal (1973) based on a linear forward method that computes the average cost per period.

The single-item uncapacitated problem with lost sales can be solved in $O(T^2)$ (Aksen et al. (2003)). The uncapacitated problem with backlogs was proven to be polynomially solvable in $O(T^3)$ when costs are concave (Zangwill (1966)). The same author later showed that the problem can be solved in $O(T^2)$ with linear production and setup costs (Zangwill (1969)). Aggarwal and Park (1993) extended this result to the problem with concave costs. Other extensions of this problem were proven to be polynomial, such as the single-item lot-sizing problem with minimum order quantities (Hellion et al. (2012), Okhrin and Richter (2011)) or with inventory bounds (Love (1973), Atamtürk and Küçükyavuz (2008)). All of these problems were solved using dynamic programming. These problems have the similar property that they can be decomposed based on discrete optimal values for some of the decision variables.

Solution approaches for capacitated lot-sizing problems

Most capacitated lot-sizing problems are NP-hard, except when strong assumptions are made on the nature of the costs (Florian et al. (1980), Bitran and Yanasse (1982)). For instance, in Florian and Klein (1971) the authors prove that the single-item capacitated lot-sizing problem is polynomial under static assumptions on the capacity and concavity assumptions on the costs. Without setup times, the feasibility problem for the multi-item capacitated lot-sizing problem is polynomial and specific heuristics can be defined to find feasible solutions (Dixon and Silver (1981)). Various approaches were introduced and studied to solve lot-sizing problems with capacity constraints.

A first approach consists in relaxing the integrality constraints of some of the boolean variables to fasten the optimization process. A feasible production plan can be found by solving different subproblems and sometimes performing a reconstruction heuristic (Chen and Thizy (1990), Helber and Sahling (2010), Maes et al. (1991), Federgruen et al. (2007), Absi and Kedad-Sidhoum (2007)). In Pochet and Van Vyve (2004), the authors compare different approaches based on linear relaxations of a production planning problem, including the relax-and-fix algorithm (Absi and Kedad-Sidhoum (2007)), an iterative heuristic that consists in performing iterative linear relaxations of the binary decision variables. The authors show that this heuristic can perform the best for specific types of instances.

Another approach consists in using a Lagrangian relaxation based heuristic where some of the complex constraints are relaxed, enabling the decomposition of the problem into subproblems oftentimes polynomial and that can be solved in parallel. This method has been extensively used in the literature to find good solutions for the CLSP (Thizy and Van Wassenhove (1985)), also with the addition of setup times (Süral et al. (2009), Diaby et al. (1992), Absi et al. (2013), Millar and Yang (1994)). Usually for capacitated lot-sizing problems, the capacity constraints are relaxed because they are often the only linking constraints between the different items. For instance, the relaxed problem for the CLSP with

setup times becomes separable into single-item uncapacitated lot-sizing problems that can be solved polynomially using dynamic programs (Trigeiro et al. (1989)). Approaches based on a Lagrangian decomposition can also be used to solve capacitated lot-sizing problems. These methods consist in duplicating some of the variables and then relaxing the linking constraints to decompose the initial problem into two distinct easier to solve problems. For instance, in Millar and Yang (1994), the authors decompose the CLSP with backlogs into a transportation problem and an uncapacitated multi-item lot-sizing problem.

Approaches using column generation are also used to solve capacitated lot-sizing problems (Jans and Degraeve (2004)). Column generation methods can be embedded in a branch-and-price algorithm to develop an exact approach (Degraeve and Jans (2007)). Branching approaches can also be tightened by the addition of valid inequalities such as the (l, S) inequalities (Barany et al. (1984)) that are used to tighten the LP-relaxation of the CLSP. We can also cite Miller et al. (2000) that define valid inequalities for the CLSP with setup times or Absi and Kedad-Sidhoum (2008) that also take into account lost sales. Tighter LP-relaxation can also be found by considering reformulations of the problems, as in Eppen and Martin (1987) for the CLSP by introducing the shortest path formulation or in Denzel and Süral (2006) for the CLSP with setup times.

Many heuristics combine different approaches (Muller et al. (2012), Fiorotto et al. (2015)). A review of solution approaches for solving dynamic lot-sizing problems can be found in Jans and Degraeve (2007) and for capacitated lot-sizing problems in Maes and Wassenhove (1988) (or more recently in Buschkühl et al. (2010)). An overview of complex characteristics in capacitated lot-sizing problems can be found in Karimi et al. (2003).

1.4 Conclusion

In this section, we have shown that lot-sizing problems are relevant in the industry and are still actively studied in the literature. Capacitated lot-sizing problems especially have many industrial applications, which explains why these problems have been studied during the last decades. In this thesis, we focus on the CLSP with setup times and extend the classical models in order to make them more relevant for industrial applications. We also develop solution approaches to solve the considered problems introduced in the following chapters.

Chapter 2

Capacitated lot-sizing problem with minimum and maximum ending inventories and setup times

In this chapter, we propose a way to mitigate the end-of-horizon effect by adding to the CLSP with setup times and lost sales (Absi and Kedad-Sidhoum (2007, 2008)) both a maximum ending inventory per item and a global minimum ending inventory at the end of the horizon. From the generation scheme of Trigeiro et al. (1989) (still used as a benchmark for capacitated lot-sizing problems, see e.g. Absi and Kedad-Sidhoum (2007) and de Araujo et al. (2015)), we propose a new framework to create instances for this new lot-sizing problem where the end-of-horizon effect is avoided, and that are relevant when solving lot-sizing problems on a rolling horizon where the available information on future demands can be used.

The chapter is organized as follows. Section 2.1 positions our work in the current literature on the subject. Section 2.2 motivates the need to mitigate the end-of-horizon effect by the addition of maximum ending inventory levels as well as a global minimum ending inventory when solving a CLSP with setup times and lost sales. In Section 2.3, an analysis of optimal solutions in a capacitated cyclical configuration is performed in order to evaluate relevant inventory levels on a rolling horizon, in a way similar to the definition of the Time Between Order (Harris (1913)) but taking the capacity into account. Section 2.4 addresses some extensions of the CLSP for which the same analysis can be applied. In Section 2.5, an extension of the CLSP with setup times, lost sales and global minimum ending inventory for all items is introduced. A new generation scheme which extends the one of Trigeiro et al. (1989) to create more relevant instances is then proposed. A computational analysis is carried out in Section 2.6 to compare the solutions obtained by solving the original and the new instances, and the effect of planning on a rolling horizon. Some conclusions and perspectives can be found in Section 2.7.

2.1 Related works

Throughout the years, the focus is increasingly on finding ways to model industrial problems as close to the reality as possible (Jans and Degraeve (2008)). Additional constraints in single-item lot-sizing problems have been extensively considered (Brahimi et al. (2017)). Although some dynamic lot-sizing problems are polynomial, the first one being studied in Wagner and Whitin (1958), they are generally NP-hard, and many heuristics have been proposed in the literature to find good feasible solutions for single-item and multi-item problems. Lagrangian relaxation approaches (see e.g. Brahimi, Dauzère-Pères and Najid (2006) and Süral et al. (2009)) and partial LP-relaxation approaches (see e.g. Absi and Kedad-Sidhoum (2007) and Helber and Sahling (2010)) are two of the most popular methods.

To model the fact that, in many industrial contexts, starting a new product incurs a fixed time to configure the resource, Trigeiro et al. (1989) consider the notion of setup times in the multi-item CLSP. Usually industrial lot-sizing problems are solved on a rolling horizon. In this context, only the decisions for the immediate periods are implemented, after which the horizon is rolled forward and the model is applied once more with updated inventory, demand and capacity parameters. Using this approach enables for each period to be optimized several times and updated according to new information on future demands. However, because each optimization problem only considers a finite time horizon, an end-of-horizon effect can occur. As described in Fisher et al. (2001), most lot-sizing problems have in common that there exists a solution with a zero-ending-inventory policy, meaning that there is no inventory at the end of the time horizon. The fact that the ending inventory is null for an optimal solution raises some issues, and can affect the production plan during the first periods in such a way that the quality of the solution decreases over time. Stadtler (2000), Fisher et al. (2001), van den Heuvel and Wagelmans (2005) propose ways to either define an adequate length of the time horizon or modify the objective function in order to cope with this end-of-horizon effect. However, they only consider single-item uncapacitated lot-sizing problems, and thus deduce inventory valuations based only on the cost, using indicators such as the *Economic Order Quantity* (Harris (1913)). In addition, the proposed approaches do not apply to capacitated lot-sizing problems since capacity is not taken into account when evaluating the ending inventory. These methods do not apply to the CLSP because they assume the Zero Inventory Ordering (ZIO) property (Wagner and Whitin (1958)) and extend the dynamic programming algorithm proposed by Wagner and Whitin (WW). However the CLSP does not have the ZIO property, so an update on the cost coefficients when solving the problem using the WW algorithm as proposed by Stadtler (2000) and van den Heuvel and Wagelmans (2005) cannot be used to determine the ending inventory of a multi-item capacitated problem. In such problems, the obtained solutions will not respect the capacity constraints. We also quote the work of Fisher et al. (2001) that provides a valuation for the ending inventory in the objective function, however this cost is not linear and cannot be solved by a linear solver. Studying the impact of the end-of-horizon effect on a multi-item capacitated lot-sizing problem with setup times cannot be neglected. Yet, in the literature, the ending-horizon effect on the CLSP is very rarely considered. As illustrated in Section 2.2, independently of the number of periods of the planning horizon, the decisions in the first periods might be impacted by the zero-ending-inventory property. To the best of our knowledge, this phenomenon has never been studied in the literature. Clark and Clark (2000) consider a rolling horizon setting for a capacitated lot-sizing problem with multiple machines and setup carry-over and propose a new model that modifies production times according to the average demand in order to get better LP-relaxations. Similarly to our approach, their model also

takes into account a number of setups per period but the setting of this parameter is left to the user and depends neither on the costs nor on the capacity at each period. Moreover, they do not consider additional inventory constraints. [Campbell and Mabert \(1991\)](#) justify the fact that cyclical schedules are often preferred in practice, mainly because they can be efficiently implemented. They also point out that cyclical schedules provide good results on average when solving capacitated lot-sizing problems. They impose a cyclical CLSP where cycle lengths are picked among a set of discrete values based on the Time Between Order (TBO) for each item. This point is also emphasized in [Bahl and Ritzman \(1984\)](#). However in this chapter, cycle lengths are not predefined, setup times are considered in the TBO calculation, and we use a cyclical subproblem to define relevant inventory indicators. The numerical results in [Campbell and Mabert \(1991\)](#) show that cyclical schedules are especially relevant for small demand variability, which is consistent with our computational results. [Campbell and Mabert \(1991\)](#) also point out that tighter capacity constraints provide larger gaps between the costs obtained by solving a cyclical problem and a non-cyclical problem. However, this can be caused by the fact that, when considering the set of cycle lengths, capacity is not taken into account in [Campbell and Mabert \(1991\)](#). In our approach, the theoretical cycle lengths also take capacity into account, adjusting the cycle length for each item accordingly.

In [Chand et al. \(2002\)](#), the authors point out the importance and the impact of the length of the time horizon on the solution quality. As pointed out by [Carlson et al. \(1979\)](#), to reduce the nervousness of a Material Requirement Planning problem (MRP), we ideally want to minimize the changes in the production plan for practical reasons when we add a period on a rolling horizon. Yet, the added information on the new demand might change the optimal order of setups. [Federgruen and Tzur \(1994\)](#) extend the notion of nervousness in MRP ([Carlson et al. \(1979\)](#)) and propose an algorithm to find a minimum *forecast horizon* that is sufficient to not affect the decisions taken over a *planning horizon* for the uncapacitated single-item lot-sizing problem (ULSP). Their numerical results show that the minimum *forecast horizon* varies a lot depending on the parameters, but can be quite large for static costs. Moreover, finding a minimal *forecast horizon* is a problem that is of the same complexity as the ULSP. This importance is even greater when there are setup times and when the capacity is tight. In this case, there is no guarantee that all demands can be satisfied, and lost sales should be allowed and penalized. A common belief among researchers in the field is that extending the planning horizon is enough to ensure that decisions in the first periods are not impacted by the end-of-horizon effect. We show in this chapter that this belief is not true for the the CLSP with setup times, and that allowing zero ending inventories might lead to poor decisions on arbitrarily large planning horizons, in particular when planning on a rolling horizon as it is the case in practice.

2.2 Motivations

Section [2.2.1](#) recalls the multi-item CLSP model with setup times and lost sales. Section [2.2.2](#) illustrates the impact of a global ending inventory, and Section [2.2.3](#) shows how the end-of-horizon effect can affect decisions taken in the first periods when planning on a rolling horizon, and thus the limits of the model and the instances of [Trigeiro et al. \(1989\)](#). Section [2.2.4](#) focuses on the impact of initial inventories.

2.2.1 Problem formulation

Let us consider the capacitated lot-sizing problem with setup times and lost sales, where N items have to be produced over a planning horizon of T periods. The discrete demand of each item i is given by d_{it} at period t . Each unit of item i produced at period t induces a production time b_{it} as well as a fixed setup time s_{it} . We aim at finding an optimal production plan, i.e. a production plan complying with the capacity restriction c_t^{max} for each period t while minimizing the total cost. This cost comprises the fixed and unitary production costs to be incurred each time a production takes place, the inventory holding costs for all the items as well as the lost sales costs penalizing the unsatisfied demand. The cost parameters are the unitary production p_{it} , fixed setup f_{it} and unitary inventory holding h_{it} costs for item i at period t . The lost sales costs penalizing each unit of unsatisfied demand of item i at period t are defined by l_{it} . Let us recall the mathematical formulation of the problem that can be found in [Trigeiro et al. \(1989\)](#) (without lost sales) and [Absi and Kedad-Sidhoum \(2008\)](#).

Let us define the decision variables as follows:

- $X_{it} \geq 0$: Quantity of item i produced at period t ,
- $Y_{it} \in \{0, 1\}$: Setup variable equals to 1 if there is an order for item i at period t , and 0 otherwise,
- $I_{it} \geq 0$: Inventory of item i at the end of period t ,
- $L_{it} \geq 0$: Quantity of lost sales for item i at the end of period t .

We extend the definition of I_{it} with $t = 0$ to model the initial inventory of item i . Moreover, we use $\bar{\cdot}$ to define the average value of a parameter over all items and all periods, e.g. $\bar{f} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T f_{it}$.

The formulation of the CLSP with setup times and lost sales, denoted (P_T) , is given below:

$$\min \sum_{i=1}^N \sum_{t=1}^T (f_{it}Y_{it} + p_{it}X_{it} + h_{it}I_{it} + l_{it}L_{it}) \quad (2.1)$$

$$I_{i,t-1} + X_{it} + L_{it} = d_{it} + I_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (2.2)$$

$$\sum_{i=1}^N (s_{it}Y_{it} + b_{it}X_{it}) \leq c_t^{max}, \quad \forall t \in 1, \dots, T \quad (2.3)$$

$$X_{it} \leq M_{it}Y_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (2.4)$$

$$L_{it} \leq d_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (2.5)$$

$$Y_{it} \in \{0, 1\}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (2.6)$$

$$X_{it}, I_{it}, L_{it} \geq 0, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (2.7)$$

The objective function [\(2.1\)](#) minimizes the total production, setup, inventory and lost sales costs of all items over the planning horizon. Constraints [\(2.2\)](#) are the flow conservation constraints that balance, for each item, the inventory at period $t - 1$ and the production and lost sales quantities at period t with the inventory and the demand at period t . Constraints [\(2.3\)](#) ensure that the capacity consumed by setup and production times does not exceed the maximum production capacity. Constraints [\(2.4\)](#) link the continuous production variables with the binary setup variables, M_{it} being an upper bound on the optimal production quantity (e.g. $M_{it} = \min(\sum_{k=t}^T d_{ik}, c_t^{max} - s_{it})$). Constraints [\(2.5\)](#) state that the quantity

of lost sales cannot exceed the demand. Constraints (2.6) and (2.7) define the domain of the variables.

As in Trigeiro et al. (1989), we only consider the case where there are no production costs, and where the cost parameters are constant over the horizon. Thus, the index t is removed in the cost parameters.

2.2.2 Impact of global ending inventory

In this section, through an illustrative example, we show that the end-of-horizon effect can affect the capacity consumption in the first periods even for large horizons. On a rolling horizon, as discussed in Section 2.2.3, this can lead to significant lost sales. The example shows that the addition of a global ending inventory can mitigate the end-of-horizon effect.

To illustrate the impact on the first periods of a production plan of considering an ending inventory, let us consider the optimal solution of an instance of the problem with 2 items, i.e. $N = 2$, and 20 periods, i.e. $T = 20$. The demand is constant over time and is set to 100, and the holding costs, unitary production times and setup times are set to 1. No setup and production costs are considered and lost sales costs are set to a very high value. The available capacity is $c^{\max} = 201$ in each period. In addition, the initial inventory $I_{1,0}$ is set to 100 for the first item.

Figure 2.1a (resp. 2.1b) shows the optimal solution obtained without (resp. with) a global minimum ending inventory set to 100 (using the method proposed in Section 2.5.2), while Figure 2.1c shows the optimal solution for the first 20 periods when solving the problem with $T = 101$ and Figure 2.1d for $T = 200$.



Figure 2.1: Production quantities in the first 20 periods

The addition of the global ending inventory enables the decisions in the first periods to match the optimal decisions observed over a very long horizon. Even with $T = 20$ and only

two items (Figures 2.1a and 2.1b), not considering inventory constraints at the end of the horizon directly impacts the decisions taken in the first periods. Without a global ending inventory, Figure 2.1a shows that the capacity at the end of the horizon is used to add additional setups at each period. This leads to a poor capacity utilization in the first period, where only a little over half the capacity is consumed. On the opposite, Figure 2.1b shows that capacity is better used with the ending inventory constraints. As shown in Figures 2.1c and 2.1d, the optimal production plan over longer horizon tends to the production plan of Figure 2.1b.

In order to illustrate the impact of the end-of-horizon effect on an instance with more than two items, Figure 2.2 shows the optimal plan for an instance of Trigeiro et al. (1989) with 10 items, where each color corresponds to an item. All the optimal solutions of the instances of Trigeiro et al. (1989) share the same shape, with small production lots in the first periods and an under-utilization of the capacity in the last periods of the horizon, as discussed in more details in Sections 2.2.3 and 2.2.4

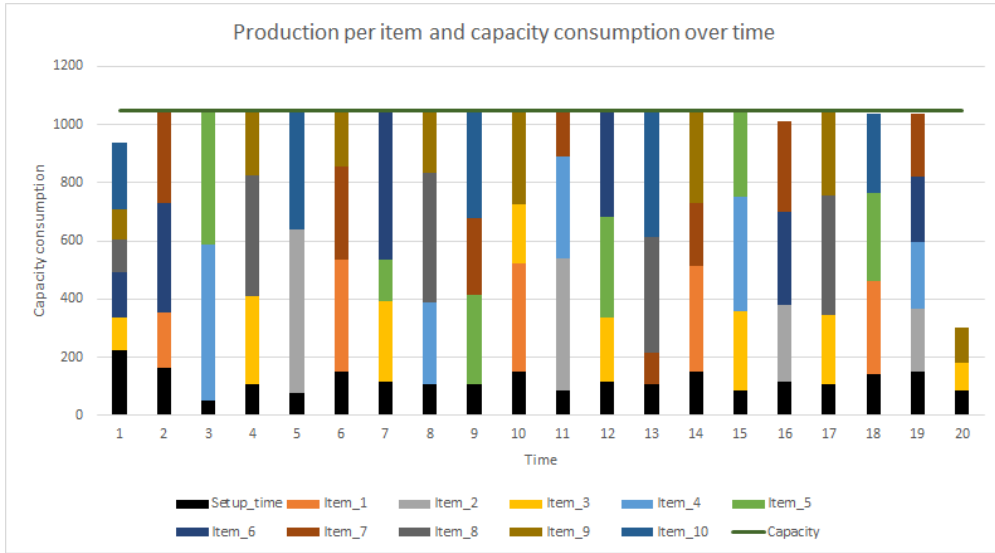


Figure 2.2: Optimal solution for an instance of Trigeiro et al. (1989).

Enough inventories should be available at the end of horizon, in particular on a rolling horizon, to make better use of production capacity in the first periods. The benefits of considering ending inventories on a rolling horizon are discussed in the following section.

It should also be pointed out that the number of periods affected by the edge effect (both the first and last periods of the horizon) can vary depending on the capacity and cost parameters. The impact of this effect is especially hard to evaluate for the CLSP with setup times, where regular indicators such as the TBO or the EOQ, used for instance in Fisher et al. (2001), cannot be applied to capacitated problems. This impact is illustrated in Section 2.2.3, where even large time horizons cannot cope with the end-of-horizon effect.

2.2.3 Planning on a rolling horizon

In this section, we consider the process of planning on a rolling horizon. Let us denote by T the number of periods of each planning horizon and by Ω the number of periods of the total time horizon ($T \leq \Omega$). $\tau \leq T$ is the number of first periods in which decisions are fixed after optimizing the production plan. Let us consider an instance of the problem with

$N = 4$, $T = 10$, $\Omega = 16$ and $\tau = 1$. The demand is constant over time and is set to 100. The holding costs and unitary production times are set to 1. Setup times are fixed to 50. No setup and production costs are considered. The available capacity is $c^{\max} = 450$ for each period of the horizon. An initial inventory of 300 units is considered for each item. Note that, since no setup costs are considered, the best policy, only guided by the holding costs, is to have the lowest possible inventory levels. Ideally, no inventory would be carried and 100 units of both items would be produced at each period. However, this production plan is not feasible because of the limited capacity at each period. In this example, we assume that lost sales are highly penalized.

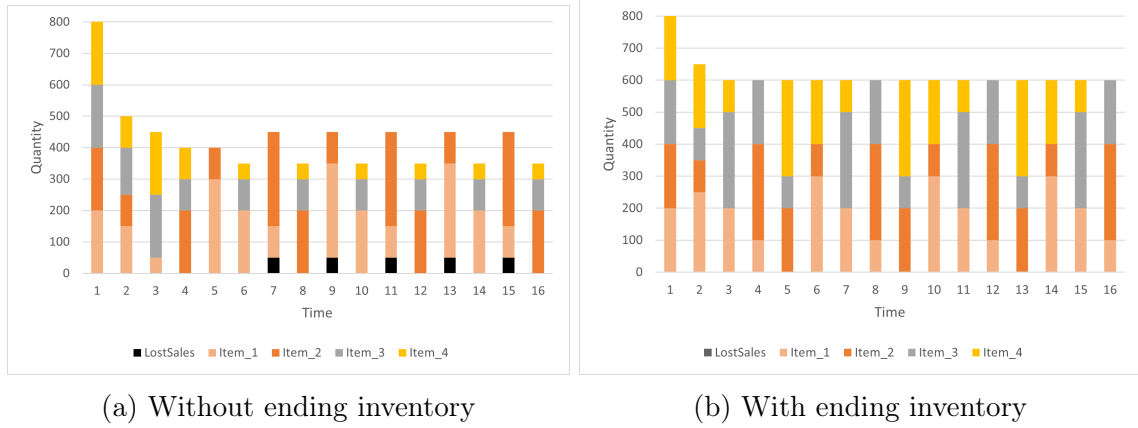


Figure 2.3: Inventory evolution on a rolling horizon

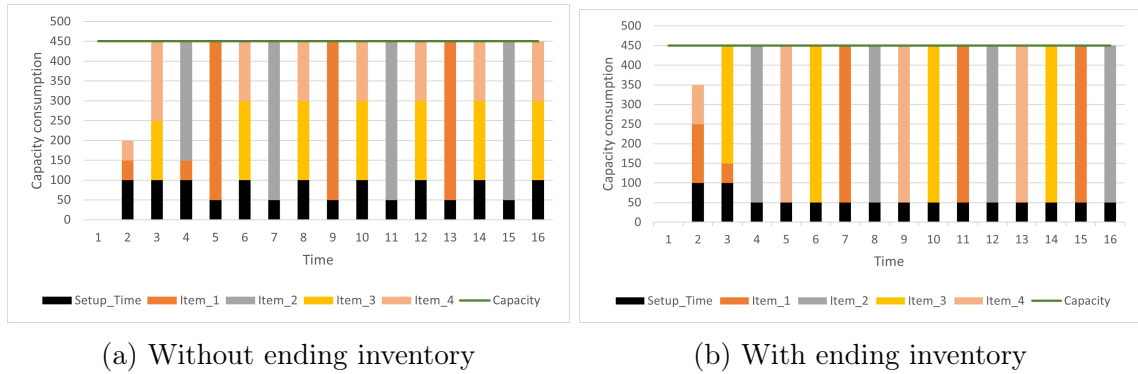


Figure 2.4: Capacity consumption on a rolling horizon

Figure 2.3 (resp. Figure 2.4) shows the inventory levels and lost sales (resp. the capacity consumption) for this instance on a rolling horizon with or without ending inventory constraints as proposed in Section 2.5.2. Let us first analyze the two different cases in Figure 2.3:

- Figure 2.3a shows the inventory levels and lost sales in the first 16 periods when no ending inventory constraints are considered. Without ending inventory constraints, the additional capacity provided at the end of the horizon, because of the ZIO policy, enables the inventory in the first periods of the rolling horizon to be immediately consumed and not kept to satisfy later demands. At each step of the rolling horizon process, the initial inventory decreases until it reaches a point where it is no longer

possible to find a feasible solution without lost sales due to capacity limitations. We then get a cyclic production plan where 50 units are lost every two iterations.

- Figure 2.3b shows the inventory and lost sales evolution for the same instance with the addition of a global minimum ending inventory of 600 units and a maximum ending inventory of 300 units for each item (using the method proposed in Section 2.5.2). The ending inventory constraints force the capacity to be fully used throughout the planning horizon, and the decisions of the first periods are not impacted by the unused capacity at the end of the horizon. We can see that no lost sales are observed in this case.

When lost sales are not allowed, we get an infeasible production plan when there are no ending inventory constraints. When lost sales are penalized and not forbidden, because we have a cyclic production plan after 5 periods (the first 5 periods with a total holding cost of 2550) for the case without ending inventory and 2 periods with ending inventory (the first 2 periods with a total holding cost of 1450), we can calculate the optimal cost over a rolling horizon of $\Omega \geq 5$ periods. Assuming unit lost sales costs of l , we get for the problem without ending inventory (Figure 2.3a) an optimal cost of:

$$C_1 = 2550 + (400 + 50l) \lfloor \frac{\Omega - 5}{2} \rfloor + 350 \lceil \frac{\Omega - 5}{2} \rceil$$

and

$$C_2 = 1450 + (\Omega - 2)600$$

for the problem with ending inventory (Figure 2.3b). In this case, for $\Omega = 100$, we get that $C_1 \geq C_2$ for:

$$38150 + 2350l \geq 60250.$$

Thus, if $l \geq 9.5$, the production plan in Figure 2.3b becomes less expensive in terms of costs than the plan in Figure 2.3a, and the difference increases with l . In many industrial applications, lost sales costs are the primary objective to optimize. Note that, because of the symmetrical structure of the costs, there are multiple optimal production plans when there are no ending inventory constraints.

Let us now analyze the two different cases in Figure 2.4:

- In Figure 2.4a, because there are no ending inventory constraints, more capacity is allocated to setup times since the initial inventory is used in the first periods, and then not enough inventory is kept to fully satisfy part of the demands, leading to the lost sales observed in Figure 2.3a.
- Figure 2.4b shows that ending inventory constraints better allocate setup times throughout the horizon. This explains why there are no lost sales in Figure 2.3b.

One can consider that the production plan patterns in Figures 2.3b and 2.4b could be obtained by significantly increasing the number of periods of the planning horizon. However, this leads to some negative effects. Indeed, the introduction of non accurate demands, because of the lack of information at the end of the horizon, should impact the quality of the obtained solutions. Moreover, increasing the number of periods negatively affects the computational efficiency of solution approaches. As the capacity gets tighter, the number of periods T that needs to be considered on a rolling horizon increases. In contrast, setting a

global ending inventory based on future demand predictions on a rolling horizon allows capacity to be better used. Moreover, contrarily to studies on uncapacitated lot-sizing problems (e.g. [Carlson et al. \(1979\)](#)), there is no theoretical guarantee that, for capacitated lot-sizing problems with setup times, there exists a forecast horizon ensuring that the decisions in the first periods will not be affected by demands outside the planning horizon.

2.2.4 Initial inventories

Figure [2.2](#) shows that there are 5 setups in the first period while, from periods 2 to 18, the number of setups oscillates between 2 and 3. Note also that the fraction of the capacity consumption taken by setup times is larger in the first period than in the following ones. This is because, when there are neither lost sales nor initial inventory, as it is the case for the CLSP, a setup will occur for every item before or in the period corresponding to its first positive demand. The first production periods are not impacted by production and setup costs, which explain the difference in the number of setups. This leads the optimization process to focus on packing the first production quantities to meet the demands of the first periods as well as making full use of the capacity constraints.

The feasibility of the problem highly depends on whether or not the capacity in the first periods is large enough to cover the demands of the first periods. As the capacity is constant over the planning horizon, in order to avoid infeasibility for these instances due to the required capacity for covering the demands during the first periods, 25% of the demands in the first four periods were set to 0 in the instances generated by [Trigeiro et al. \(1989\)](#). This choice was arbitrarily made to guarantee feasibility, and has no practical reality. Optimizing the production quantities in the first production periods, which increases the computational complexity of the optimization problem, also does not make much sense when planning on a rolling horizon.

Initial inventories can also have a significant impact on the feasibility of the solution when planning on a rolling horizon, as the initial inventories are linked to the decisions taken in the first periods. In order to tackle infeasibility issues, some authors have considered initial inventories as decision variables, with high penalty costs ([Vanderbeck \(1998\)](#), [Degraeve and Jans \(2007\)](#), [de Araujo et al. \(2015\)](#))).

2.3 Inventory levels for the capacitated lot-sizing problem with setup times on a rolling horizon

The main goal of this section is to define new indicators to characterize relevant inventory levels for the considered capacitated lot-sizing problem on a rolling horizon. To this end, we use similar arguments that the ones used to define the Time Between Order and the Economic Order Quantity ([Harris \(1913\)](#)). To do that, we define in Section [2.3.1](#) a new problem that enables us to find approximated analytical values whose relevance will be discussed in the numerical analysis of Section [2.6](#). We show in Section [2.3.2](#) that this problem is relevant compared to the CLSP, and more specifically to the problem with static costs and the parameters considered in [Trigeiro et al. \(1989\)](#). This simplified model will be used in Section [2.5.2](#) to update the instance generation scheme proposed in [Trigeiro et al. \(1989\)](#), in order to create instances whose optimal plans will not be affected by the end-of-horizon effect.

2.3.1 Multi-item cyclical production planning with bounded average capacity consumption

Let us consider the multi-item lot-sizing problem with setup times that consists in finding optimal cycle lengths on a rolling horizon. All costs and demands are static. Let us denote by $\phi_i \in \mathbb{N}^*$ the cycle length of an item i , *i.e.* the number of periods between two production periods. The production cycle length $\tilde{\phi} \in \mathbb{N}^*$ is defined as the minimum number of periods, such that each item has an integer number of cycles, *i.e.* the least common multiple of the cycle length of all the items. Additionally, the cycle lengths of each item should be such that the average capacity consumed in each period of a production cycle should not exceed the maximum capacity c^{\max} . Figure 2.5 illustrates the item and the production cycle lengths for an instance with four items and a horizon of ten periods, where $\phi_1 = \phi_2 = 3$, $\phi_3 = \phi_4 = 2$ and $\tilde{\phi} = 6$.

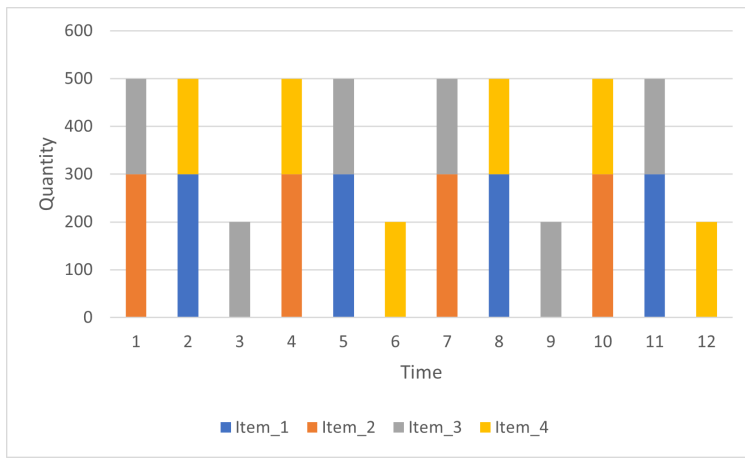


Figure 2.5: Example of cyclical production with 4 items.

The total inventory cost for one cycle of length ϕ_i is $\bar{h}d_i \sum_{t=1}^{\phi_i-1} t = \bar{h}d_i\phi_i \frac{\phi_i-1}{2}$. The total inventory cost in a production cycle of length $\tilde{\phi}$ is then:

$$\sum_{i=1}^N \frac{\tilde{\phi}}{\phi_i} \bar{h}d_i\phi_i \frac{\phi_i-1}{2} = \bar{h}\tilde{\phi} \sum_{i=1}^N d_i \frac{\phi_i-1}{2}.$$

The average inventory at each period \tilde{I}^{inf} , as well as the maximum inventory I_i^{sup} for item i , can be defined as:

$$\tilde{I}^{inf} = \sum_{i=1}^N d_i \frac{\phi_i-1}{2}, \quad (2.8)$$

$$I_i^{sup} = (\phi_i-1)d_i. \quad (2.9)$$

Let us denote $\phi = \{\phi_1, \dots, \phi_N\}$. The optimization problem denoted (P_C) , that minimizes the cost per period in a production cycle ϕ can be defined as follows:

$$\min_{\phi_i \in \mathbb{N}^{*N}} \sum_{i=1}^N \frac{f_i}{\phi_i} + \bar{h} \sum_{i=1}^N d_i \frac{\phi_i-1}{2} \quad (2.10)$$

$$s.t. \quad \sum_{i=1}^N \frac{s_i}{\phi_i} \leq c^{\max} - N\bar{b}\bar{d} \quad (2.11)$$

The objective function (2.10) minimizes the average setup and inventory costs for a period of the cycle length. In a cyclical configuration, the average production time per period is $N\bar{b}\bar{d}$, which means that, on average, the capacity available for setup times is $c^{\max} - N\bar{b}\bar{d}$. Constraint (2.11) thus imposes an upper bound on the average number of setups.

2.3.2 Similarities between (P_C) and the optimization problem of Trigeiro et al. (1989)

To consider a problem with data similar to the ones generated in Trigeiro et al. (1989), average values for costs and demands are set identical for all items. These values correspond to the average values of the parameters defined in Section 2.2.1, except for the production costs that are equal to 0 and the lost sales that are not allowed.

The expected number of setups per period $k_i \in \mathbb{R}_+$ for item i corresponds to the reciprocal of the cycle length ϕ_i :

$$\frac{1}{\phi_i} = k_i.$$

This implies that the average number of setups $k \in \mathbb{R}_+$ in each period of a production cycle is related to the cycle length of each item:

$$k = \sum_{i=1}^N k_i = \sum_{i=1}^N \frac{1}{\phi_i}.$$

The optimization problem denoted (\bar{P}_C) , that minimizes the cost per period in a production cycle ϕ can be derived from problem (P_C) and is defined as follows:

$$\min_{(k, \phi_i) \in \mathbb{R}_+^* \times \mathbb{N}^{*N}} \quad \bar{f}k + \bar{h}\bar{d} \sum_{i=1}^N \frac{\phi_i - 1}{2} \tag{2.12}$$

$$s.t. \quad \sum_{i=1}^N \frac{1}{\phi_i} = k \tag{2.13}$$

$$k \leq \frac{c^{\max} - N\bar{b}\bar{d}}{\bar{s}} \tag{2.14}$$

The objective function (2.12) as well as Constraint (2.14) are similar to Constraints (2.10) and (2.11), where cycle lengths are replaced by the variable corresponding to the average number of setups. Constraint (2.13) links the number of setups per period with the cycle length of each item.

The cyclical approximation is a simple but relevant simplification of the model in order to get a general idea of the behaviour of a solution as it provides insights on the role of the costs and the capacity. In the Trigeiro et al. (1989) generation scheme, all costs are static. The costs can vary between items, but they are all generated by doing slight changes around common average values. Even though the demands are dynamic, they are also generated around the same common value. In (\bar{P}_C) , we only require that the average capacity consumed in each period is lower than c^{\max} . However, compared to the problem addressed in Trigeiro et al. (1989), the upper bound on the capacity consumption is not set in each period. However, with N production cycle lengths, we assume that there is one configuration such that the capacity used in each period is close to the average capacity consumption. Furthermore, in this chapter, we are only interested in approximating the inventories in

each period. The idea is not to consider the dynamics of production. Thus, analyzing the cyclical multi-item problem, where all costs and demands are averaged and where the average capacity consumption is bounded, should provide enough insights on the shape of an optimal solution of the problem addressed in [Trigeiro et al. \(1989\)](#) on a rolling horizon. The relevance of our assumptions is analyzed in the computational experiments of Section [2.6.1](#). The study of the continuous relaxation of (\bar{P}_C) helps to evaluate the average total inventory as well as the maximum inventory per item.

2.3.3 Analysis of the continuous relaxation of (\bar{P}_C)

When the integrality constraints on variables ϕ_i are relaxed in (\bar{P}_C) , the following non-linear problem can be derived:

$$F^* = -\frac{N\bar{h}}{2}\bar{d} + \min_{\{k \in \mathbb{R}_+^* | k \leq \frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}}\}} \left(\bar{f}k + \frac{\bar{h}}{2}\bar{d} \min_{\{\phi_i \in \mathbb{R}_+^* | \sum_{i=1}^N \frac{1}{\phi_i} = k\}} \sum_{i=1}^N \phi_i \right),$$

as the set $\{\phi_i \in \mathbb{R}_+^* | \sum_{i=1}^N \frac{1}{\phi_i} = k\}$ is non-empty for all $k \in \mathbb{R}_+^*$.

Let us show that the continuous relaxation of (\bar{P}_C) can be analytically solved to optimality. Providing an easy-to compute analytical formula might give insights on the links between the costs and the capacity for an instance of the problem.

Theorem 1. *If $\frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}} \geq N\sqrt{\frac{\bar{h}}{2\bar{f}}}$, then the optimal solution is reached for $k^* = N\sqrt{\frac{\bar{h}\bar{d}}{2\bar{f}}}$, and the optimal value is: $-\frac{\bar{h}\bar{d}N}{2} + N\sqrt{2\bar{h}\bar{d}\bar{f}}$. Otherwise, F^* is reached for $k^* = \frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}}$, and is equal to: $\frac{\bar{h}\bar{d}N}{2} \left(\frac{N\bar{s}}{c^{\max} - N\bar{d}\bar{b}} - 1 \right) + \bar{f} \frac{(c^{\max} - N\bar{d}\bar{b})}{\bar{s}}$.*

Proof. Let us first show that:

$$\min_{\{\phi_i \in \mathbb{R}_+^* | \sum_{i=1}^N \frac{1}{\phi_i} = k\}} \sum_{i=1}^N \phi_i = \frac{N^2}{k}$$

Using the Euclidian norm and its corresponding scalar product, the Cauchy-Schwarz inequality states that, for $\phi_i \in \mathbb{R}_+^*$:

$$\left(\sum_{i=1}^N \sqrt{\frac{\phi_i}{\phi_i}} \right)^2 \leq \left(\sum_{i=1}^N \frac{1}{\phi_i} \right) \left(\sum_{i=1}^N \phi_i \right)$$

Thus, by positivity:

$$\sum_{i=1}^N \phi_i \geq \frac{N^2}{\sum_{i=1}^N \frac{1}{\phi_i}}$$

So that:

$$\min_{\{\phi_i \in \mathbb{R}_+^* | \sum_{i=1}^N \frac{1}{\phi_i} = k\}} \sum_{i=1}^N \phi_i \geq \frac{N^2}{k}$$

When $\phi_i = \phi = \frac{N}{k}$, for all i , we have $\sum_{i=1}^N \frac{1}{\phi_i} = k$ and $\sum_{i=1}^N \phi_i = \frac{N^2}{k}$, and then:

$$\min_{\{\phi_i \in \mathbb{R}_+^* | \sum_{i=1}^N \frac{1}{\phi_i} = k\}} \sum_{i=1}^N \phi_i = \frac{N^2}{k}$$

The continuous relaxation of (\bar{P}_C) is equivalent to:

$$F^* = -\frac{\bar{h}\bar{d}N}{2} + \min_{\{k \in \mathbb{R}_+^* | k \leq \frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}}\}} \left(\bar{f}k + \frac{\bar{h}\bar{d}N^2}{2k} \right)$$

Let us consider the function $g(x) = \bar{f}x + \frac{\bar{h}\bar{d}N^2}{2x}$ on \mathbb{R}_+^* . This function of $x \in \mathbb{R}_+^*$ is decreasing until $x^* = N\sqrt{\frac{\bar{h}\bar{d}}{2\bar{f}}}$, and then increasing. Let k^* denotes the optimal average number of setups. Since $k^* \leq \frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}}$, we get:

- If $\frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}} \geq N\sqrt{\frac{\bar{h}\bar{d}}{2\bar{f}}}$, then $k^* = N\sqrt{\frac{\bar{h}\bar{d}}{2\bar{f}}}$ and F^* is equal to: $-\frac{\bar{h}\bar{d}N}{2} + N\sqrt{2\bar{h}\bar{d}\bar{f}}$,
- Otherwise, $k^* = \frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}}$, and F^* is equal to: $\frac{\bar{h}\bar{d}N}{2} \left(\frac{N\bar{s}}{c^{\max} - N\bar{d}\bar{b}} - 1 \right) + \bar{f} \frac{(c^{\max} - N\bar{d}\bar{b})}{\bar{s}}$.

□

Let us introduce $k_{capa} = \frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}}$ and $k_{cost} = N\sqrt{\frac{\bar{h}\bar{d}}{2\bar{f}}}$. In a cyclical configuration, the number of setups k in each period is close to:

$$k = \min(k_{cost}, k_{capa}) \tag{2.15}$$

and the cycle length ϕ_i is close to $\frac{N}{k}$ for all items.

If the capacity constraints are not binding, the production cycles follow the time between order $TBO = \sqrt{2\frac{\bar{f}}{\bar{h}\bar{d}}}$ for each item. On average, $k = k_{cost}$ items will be produced in each period, consuming a capacity of $\frac{N}{k}\bar{s} + N\bar{d}\bar{b}$.

By applying the same analysis to the original [Trigeiro et al. \(1989\)](#) instances, we find that, for some of these instances, $k_{capa} < 1$. This would imply that these instances are only feasible because of the extra capacity freed by the demands randomly set to 0 in the first periods of the planning horizon. For the other instances, we have $k_{capa} > N$, which is not relevant when the capacity is constrained. Moreover, if the costs are defined such that $\frac{N}{k}$ is integer, then the optimal solution of the relaxed problem (\bar{P}_C) is an optimal solution of problem (\tilde{P}_C) . Indeed, because of the relaxation of the integrality property of the ϕ variables we have:

$$\min_{\{\phi \in \mathbb{R}_+^* | \sum_{i=1}^N \frac{1}{\phi_i} = k\}} \sum_{i=1}^N \phi_i = \frac{N^2}{k} \leq \min_{\{\phi \in \mathbb{N}_+^* | \sum_{i=1}^N \frac{1}{\phi_i} = k\}} \sum_{i=1}^N \phi_i.$$

For $\phi_i = \frac{N}{k} \in \mathbb{N}^*$, the value of $\frac{N^2}{k}$ is reached for problem (\tilde{P}_C) .

If the cycle length ϕ_i of item i is integer, then we have already established that the average inventory of item i is equal to $\bar{d}\frac{\phi_i - 1}{2}$. The maximum inventory is $\bar{d}(\phi_i - 1)$. We then apply these formulas to deduce the approximate values of the total average and the maximum inventory for each item. As $\phi_i = \frac{N}{k}$, we get the following values:

$$\tilde{I}_i^{inf} = \bar{d} \sum_{i=1}^N \frac{\frac{N}{k} - 1}{2} = \frac{N(N - k)\bar{d}}{k} \frac{1}{2}, \tag{2.16}$$

$$I_i^{sup} = I^{sup} = \frac{N - k}{k} \bar{d}. \tag{2.17}$$

We can define analytical values for \tilde{I}_i^{inf} and I^{sup} even when $\phi_i = \frac{N}{k}$ is not an integer. We show in Section [2.6.1](#) that this indicator is effective on a rolling horizon.

2.3.4 Lost sales costs

We want to study a special case of the cyclical problem defined in Section 2.3.1, where only a fraction of the demand can be satisfied in each period. Let us denote by $\gamma \in [0, 1]$ the fraction of the average demand that is satisfied in a cyclical production process, and by \bar{l} the average lost sales cost over all items. We want to find sufficient conditions on \bar{l} in order to get $\gamma^* = 1$, and thus an optimal solution where all demands are satisfied. We assume that $\frac{N\bar{h}}{2\bar{l}+\bar{h}} \leq \frac{c^{\max}}{\bar{s}}$ otherwise Formula (2.22) implies that it is optimal to lose all demands ($\gamma^* = 0$).

The mathematical formulation of the problem is given by:

$$\min_{(k, \phi, \gamma) \in \mathbb{R}^* \times \mathbb{N}^{*N} \times [0, 1]} \bar{f}k + \bar{h}\gamma d \sum_{i=1}^N \frac{\phi_i - 1}{2} + N(1 - \gamma)\bar{d}\bar{l} \quad (2.18)$$

$$s.t. \quad \sum_{i=1}^N \frac{1}{\phi_i} = k \quad (2.19)$$

$$k \leq \frac{c^{\max} - N\bar{b}\gamma\bar{d}}{\bar{s}} \quad (2.20)$$

which can be rewritten as:

$$N\bar{d}\bar{l} + \min_{k \in]0, \frac{c^{\max}}{\bar{s}}]} (\bar{f}k + \bar{d} \min_{\{\gamma \in [0, 1] | k \leq \frac{c^{\max} - N\bar{b}\gamma\bar{d}}{\bar{s}}\}} \gamma[\bar{h} \min_{\{\phi \in \mathbb{N}^{*N} | \sum_{i=1}^N \frac{1}{\phi_i} = k\}} (\sum_{i=1}^N \frac{\phi_i}{2}) - N\bar{l} - \frac{\bar{h}N}{2}]) \quad (2.21)$$

The continuous relaxation is equivalent to:

$$N\bar{d}\bar{l} + \min_{k \in]0, \frac{c^{\max}}{\bar{s}}]} (\bar{f}k + \frac{N\bar{d}}{2} \min_{\{\gamma \in [0, 1] | \gamma \leq \frac{c^{\max} - k\bar{s}}{N\bar{b}\bar{d}}\}} \gamma[\bar{h}(\frac{N}{k} - 1) - 2\bar{l}]) \quad (2.22)$$

If $\bar{h}(\frac{N}{k} - 1) - 2\bar{l} \geq 0$, i.e $k \leq \frac{N\bar{h}}{2\bar{l}+\bar{h}}$, then $\gamma^* = 0$. The term $\frac{N\bar{d}}{2}\bar{h}(\frac{N}{k} - 1)$ represents the average inventory costs per period if all demands are satisfied while the term $N\bar{d}\bar{l}$ corresponds to the cost of losing all demands for a period. For a given cycle with k setups per period, when the average inventory costs are higher than the lost sales costs, it will always be better to lose all demands.

When $k \geq \frac{N\bar{h}}{2\bar{l}+\bar{h}}$, the average inventory costs per period are smaller than the lost sales costs, we try to satisfy as much of the demands as possible. Otherwise $\gamma^* = \min(1, \frac{c^{\max} - k\bar{s}}{N\bar{b}\bar{d}})$.

Characterization of optimal solutions

The optimal solution of the problem in Formula (2.22) is denoted by F^{ls^*} . We can divide the problem into three distinct cases depending on the definition domain of the average number of setups k :

Case 1. $K_1 = \{k \in]0, \frac{c^{\max}}{\bar{s}}] | k \leq \frac{N\bar{h}}{2\bar{l}+\bar{h}}\}$:

As shown previously $\gamma^* = 0$ and it is optimal to lose all demands:

$$F^{ls^*} = N\bar{d}\bar{l} + \min_{k \in K_1} \bar{f}k = N\bar{d}\bar{l}$$

Case 2. $K_2 = \{k \in]0, \frac{c^{\max}}{\bar{s}}] | k \geq \frac{N\bar{h}}{2l+h} \text{ and } k \leq \frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}\}$:

We have $\frac{c^{\max}-k\bar{s}}{N\bar{b}\bar{d}} \geq 1$ and thus $\gamma^* = 1$. It is optimal to satisfy all demands:

$$F^{ls*} = N\bar{d}\bar{l} + \min_{k \in K_2} (\bar{f}k + \frac{N\bar{d}}{2}(\bar{h}(\frac{N}{k} - 1) - 2\bar{l}))$$

$$F^{ls*} = -\frac{N\bar{d}\bar{h}}{2} + \min_{k \in K_2} (\bar{f}k + \frac{N^2\bar{d}\bar{h}}{2k}) \quad (2.23)$$

The minimum of F^{ls*} is reached for $k^* = N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ if $N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ is in the definition domain (which would correspond to a solution guided only by the costs), or k^* corresponds to one of the domain bounds.

Case 2.1. $\frac{N\bar{h}}{2l+h} \leq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}} \leq \frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$:

We have $k^* = N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ and then:

$$F^{ls*} = N(\sqrt{2\bar{f}\bar{d}\bar{h}} - \frac{\bar{d}\bar{h}}{2})$$

Case 2.2. $N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}} \leq \frac{N\bar{h}}{2l+h}$:

F^{ls*} is increasing between $\frac{N\bar{h}}{2l+h}$ and $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$, so $k^* = \frac{N\bar{h}}{2l+h}$ and:

$$F^{ls*} = N\bar{d}\bar{h}(\frac{\bar{f}}{\bar{d}(2\bar{l} + \bar{h})} + \frac{\bar{l}}{\bar{h}})$$

It is worth noticing that this case corresponds to a case where the average inventory costs are equal to the lost sales costs, thus all values of γ^* are equivalent. This means that the solution where all demands are satisfied is equivalent to the solution considering all demands as lost sales. We want to avoid this case when defining relevant lost sales costs.

Case 2.3. $N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}} \geq \frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$:

F^{ls*} is decreasing between $\frac{N\bar{h}}{2l+h}$ and $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$, so $k^* = \frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$ and then:

$$F^{ls*} = \frac{N\bar{d}\bar{h}}{2}(\frac{N\bar{s}}{c^{\max} - N\bar{b}\bar{d}} - 1) + \frac{\bar{f}}{\bar{s}}(c^{\max} - N\bar{b}\bar{d})$$

Case 3. $K_3 = \{k \in]0, \frac{c^{\max}}{\bar{s}}] | k \geq \frac{N\bar{h}}{2l+h} \text{ and } k \geq \frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}} \text{ and } k \leq \frac{c^{\max}}{\bar{s}}\}$:

We have $\frac{c^{\max}-k\bar{s}}{N\bar{b}\bar{d}} \leq 1$ and $\gamma^* = \frac{c^{\max}-k\bar{s}}{N\bar{b}\bar{d}}$. It can be optimal to satisfy only a fraction of the demand.

$$F^{ls*} = N\bar{d}\bar{l} + \min_{k \in K_3} (\bar{f}k + \frac{(c^{\max} - k\bar{s})}{2\bar{b}}(\bar{h}(\frac{N}{k} - 1) - 2\bar{l}))$$

Let us define $\tilde{f} = \bar{f} + \frac{\bar{s}(2\bar{l} + \bar{h})}{2\bar{b}}$ and $\tilde{h} = \frac{c^{\max}}{N\bar{b}\bar{d}}\bar{h}$.

$$F^{ls*} = N\bar{d}\bar{l} - \frac{(2\bar{l} + \bar{h})c^{\max} + N\bar{h}\bar{s}}{2\bar{b}} + \min_{k \in K_3} \left(\bar{f}k + \frac{N^2\bar{d}\bar{h}}{2k} \right) \quad (2.24)$$

This corresponds to a new cyclical problem with updated setup and inventory costs. The minimum of F^{ls*} is reached for $k^* = N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ if $N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ is in the definition domain (which would correspond to a solution guided only by the updated costs), or k^* corresponds to one of the domain bounds.

Case 3.1. $\frac{c^{\max}}{\bar{s}} \leq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$:

The minimum is reached for $k^* = \frac{c^{\max}}{\bar{s}}$, which corresponds to $\gamma^* = 0$ and none of the demand is satisfied.

$$F^{ls*} = N\bar{d}\bar{l} + \frac{\bar{f}}{\bar{s}}c^{\max}$$

Case 3.2. $\frac{c^{\max}}{\bar{s}} \geq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ and $\frac{N\bar{h}}{2\bar{l}+\bar{h}} \leq \frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$:

The lower bound of K_3 is $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$ and the minimum of F^{ls*} (see Formula (2.24)) is reached in $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$ or in $N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$.

Case 3.2.1. $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}} \leq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$, we have $k^* = N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ and $\gamma^* \leq 1$. It can be optimal to satisfy only part of the demand, and:

$$F^{ls*} = N\bar{d}\bar{l} - \frac{(2\bar{l} + \bar{h})c^{\max} + N\bar{h}\bar{s}}{2\bar{b}} + N\sqrt{2\bar{f}\bar{h}\bar{d}}$$

Case 3.2.2. $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}} \geq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$:

We have $k^* = \frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}}$ and $\gamma^* = 1$. It is optimal to satisfy all demands, and:

$$F^{ls*} = \frac{N\bar{d}\bar{h}}{2} \left(\frac{N\bar{s}}{c^{\max} - N\bar{b}\bar{d}} - 1 \right) + \frac{\bar{f}}{\bar{s}}(c^{\max} - N\bar{b}\bar{d})$$

Case 3.3. $\frac{c^{\max}}{\bar{s}} \geq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ and $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}} \leq \frac{N\bar{h}}{2\bar{l}+\bar{h}}$:

Case 3.3.1. $\frac{N\bar{h}}{2\bar{l}+\bar{h}} \leq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$:

We have $k^* = N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ and $\gamma^* \leq 1$. It can be optimal to satisfy only part of the demand, and:

$$F^{ls*} = N\bar{d}\bar{l} - \frac{(2\bar{l} + \bar{h})c^{\max} + N\bar{h}\bar{s}}{2\bar{b}} + N\sqrt{2\bar{f}\bar{h}\bar{d}}$$

Case 3.3.2. $\frac{N\bar{h}}{2\bar{l}+\bar{h}} \geq N\sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$:

We have $k^* = \frac{N\bar{h}}{2\bar{l}+\bar{h}}$ and $\gamma^* \leq 1$ because $\frac{c^{\max}-N\bar{b}\bar{d}}{\bar{s}} < \frac{N\bar{h}}{2\bar{l}+\bar{h}}$. It can be optimal to satisfy only part of the demand, and:

$$F^{ls*} = N \left(\bar{d}\bar{l} + \frac{\bar{f}\bar{h}}{2\bar{l} + \bar{h}} \right)$$

Definition of relevant lost sales costs

The goal of the section is to propose sufficient conditions to define lost sales costs that guarantee that the optimal solution does not have a fraction of demand unsatisfied for the problem defined in Formula (2.22). With regards to the different cases introduced previously, we want to consider only the cases where $\gamma^* = 1$ is the only optimal value for γ . This means we want Cases **2.1**, **2.3** and **3.2.2** to dominate the other cases.

Sufficient conditions for not having lost sales in an optimal cycle can be set as follows:

1. In order to prevent Case **3.1**:

$$\frac{c^{\max}}{\bar{s}} \geq N \sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$$

2. In order to prevent Cases **3.1**, **3.3** and **3.2.1**, we want to define \bar{l} such that $\frac{c^{\max} - N\bar{b}\bar{d}}{\bar{s}} \geq N \sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ and $\frac{c^{\max} - N\bar{b}\bar{d}}{\bar{s}} \geq \frac{N\bar{h}}{2l+h}$. Both conditions are respected by setting:

$$\bar{l} \geq \frac{c^{\max}}{(c^{\max} - N\bar{b}\bar{d})} \frac{N\bar{s}\bar{h}}{2(c^{\max} - N\bar{b}\bar{d})}$$

This condition also implies that $\frac{c^{\max}}{\bar{s}} \geq N \sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}}$ and $\frac{c^{\max}}{\bar{s}} \geq \frac{N\bar{h}}{2l+h}$.

3. In order to prevent Case **2.2** and to ensure that Case **2.1** is dominated by Case **1**, we want to define \bar{l} such that $N \sqrt{\frac{\bar{d}\bar{h}}{2\bar{f}}} \geq \frac{N\bar{h}}{2l+h}$ and $N(\sqrt{2\bar{f}\bar{d}\bar{h}} - \frac{\bar{d}\bar{h}}{2}) \leq N\bar{d}\bar{l}$. Both conditions are respected by setting:

$$\bar{l} \geq \sqrt{\frac{2\bar{f}\bar{h}}{\bar{d}}}$$

4. In order to ensure that Cases **2.3** and **3.2.2** are dominated by Case **1**, we want to define \bar{l} such that $\frac{N\bar{d}\bar{h}}{2} (\frac{N\bar{s}}{c^{\max} - N\bar{b}\bar{d}} - 1) + \frac{\bar{f}}{\bar{s}} (c^{\max} - N\bar{b}\bar{d}) \leq N\bar{d}\bar{l}$, which implies:

$$\bar{l} \geq \frac{\bar{h}}{2} (\frac{N\bar{s}}{c^{\max} - N\bar{b}\bar{d}} - 1) + \frac{\bar{f}}{N\bar{d}\bar{s}} (c^{\max} - N\bar{b}\bar{d})$$

It should be pointed out that, because the goal is to find sufficient conditions for \bar{l} that are easy to express, the focus for Points 2. and 3. was not to find the minimum \bar{l} satisfying the required conditions but only a threshold that would guarantee that these conditions are respected.

Consequently, there is an optimal solution without lost sales if we set the lost sales costs for item i as follows:

$$l_i = \max \left(\frac{N\bar{s}\bar{h}c^{\max}}{2(c^{\max} - N\bar{b}\bar{d})^2}, \sqrt{\frac{2\bar{f}\bar{h}}{\bar{d}}}, \frac{\bar{h}}{2} (\frac{N\bar{s}}{c^{\max} - N\bar{b}\bar{d}} - 1) + \frac{\bar{f}}{N\bar{d}\bar{s}} (c^{\max} - N\bar{b}\bar{d}) \right) \quad (2.25)$$

This proposed definition of the lost sales cost better integrates the cost and the capacity parameters of the instances.

2.4 Extensions to more general cases

This section considers some extensions of the CLSP for which the same rationale can be applied and analytical values for the optimal cycle lengths can be deduced from the relaxed problem. The first extension addresses the case where the average demand varies between items, and the second extension the case where, in addition, the setup times and costs are linearly dependent.

2.4.1 Average demand per item

Let us consider the case where the average demand d_i is different for each item i . Let us define an optimal cyclical production plan where the average capacity consumption does not exceed c^{\max} by solving the following problem:

$$\min_{(k, \phi_i) \in \mathbb{R}^* \times \mathbb{N}^{*N}} \bar{f}k + \bar{h} \sum_{i=1}^N d_i \frac{\phi_i - 1}{2} \quad (2.26)$$

$$\text{s.t.} \quad \sum_{i=1}^N \frac{1}{\phi_i} = k \quad (2.27)$$

$$k \leq \frac{c^{\max} - N\bar{b}\bar{d}}{\bar{s}} \quad (2.28)$$

Constraint (2.28) still applies to this model to limit the average capacity consumption, since the average capacity consumed in each period by the production setup time is $\sum_{i=1}^N d_i \bar{b} = N\bar{b}\bar{d}$. Hence, the capacity available for setup times is still $c^{\max} - N\bar{b}\bar{d}$. The continuous relaxation of the problem can be written as:

$$-\bar{h} \sum_{i=1}^N \frac{d_i}{2} + \min_{\{k \in \mathbb{R}_+^* \mid k \leq \frac{c^{\max} - N\bar{b}\bar{d}}{\bar{s}}\}} (\bar{f}k + \frac{\bar{h}}{2} \min_{\{\phi_i \in \mathbb{R}^{*N} \mid \sum_{i=1}^N \frac{1}{\phi_i} = k\}} \sum_{i=1}^N d_i \phi_i)$$

Theorem 2. *If $\frac{c^{\max} - N\bar{b}\bar{d}}{\bar{s}} \geq \sqrt{\frac{\bar{h}}{2\bar{f}}} \sum_{i=1}^N \sqrt{d_i}$, then the optimal solution is reached for $k^* = \sqrt{\frac{\bar{h}}{2\bar{f}}} \sum_{i=1}^N \sqrt{d_i}$, and the optimal value is: $\sqrt{2\bar{f}\bar{h}} \sum_{i=1}^N \sqrt{d_i} - \bar{h} \sum_{i=1}^N \frac{d_i}{2}$. Otherwise, the optimal solution is reached for $k^* = \frac{c^{\max} - N\bar{b}\bar{d}}{\bar{s}}$ and the optimal value is: $\frac{\bar{f}}{\bar{s}}(c^{\max} - N\bar{b}\bar{d}) + \frac{\bar{s}\bar{h}}{2} \frac{(\sum_{i=1}^N \sqrt{d_i})^2}{c^{\max} - N\bar{b}\bar{d}} - \bar{h} \sum_{i=1}^N \frac{d_i}{2}$.*

The proof of Theorem 2 follows similar arguments than the proof of Theorem 1.

2.4.2 Average demand per item and correlated setup costs and times

Let us now assume that the setup times and costs are correlated, i.e. there exists $\lambda \in \mathbb{R}^*$ such that $f_i = \lambda s_i$, $\forall i \in \{1, \dots, N\}$, and also that the average demand d_i is not the same for all items. We set $\tilde{k} = \sum_{i=1}^N \frac{s_i}{\phi_i}$, which corresponds to the average setup time per period. The average setup cost per period is then $\sum_{i=1}^N \frac{f_i}{\phi_i} = \lambda \sum_{i=1}^N \frac{s_i}{\phi_i} = \lambda \tilde{k}$.

Similarly, the following cyclical production planning problem is solved:

$$\min_{(\tilde{k}, \phi_i) \in \mathbb{R}^* \times \mathbb{N}^{*N}} \lambda \tilde{k} + \bar{h} \sum_{i=1}^N d_i \frac{\phi_i - 1}{2} \quad (2.29)$$

$$s.t. \quad \sum_{i=1}^N \frac{s_i}{\phi_i} = \tilde{k} \quad (2.30)$$

$$\tilde{k} \leq c^{\max} - N\bar{b}\bar{d} \quad (2.31)$$

The continuous relaxation of the problem can be written as:

$$-\bar{h} \sum_{i=1}^N \frac{d_i}{2} + \min_{\{\tilde{k} \in \mathbb{R}_+^* | \tilde{k} \leq c^{\max} - N\bar{b}\bar{d}\}} (\lambda \tilde{k} + \frac{\bar{h}}{2} \min_{\{\phi_i \in \mathbb{R}^{*N} | \sum_{i=1}^N \frac{s_i}{\phi_i} = \tilde{k}\}} \sum_{i=1}^N d_i \phi_i)$$

Theorem 3. *If $c^{\max} - N\bar{b}\bar{d} \geq \sqrt{\frac{\bar{h}}{2\lambda}} \sum_{i=1}^N \sqrt{d_i s_i}$, then the optimal solution is reached for $\tilde{k}^* = \sqrt{\frac{\bar{h}}{2\lambda}} \sum_{i=1}^N \sqrt{d_i s_i}$, and the optimal value is: $\sqrt{2\lambda\bar{h}} \sum_{i=1}^N \sqrt{d_i s_i} - \bar{h} \sum_{i=1}^N \frac{d_i}{2}$. Otherwise, the optimal solution is reached for $\tilde{k}^* = c^{\max} - N\bar{b}\bar{d}$, and the optimal value is: $\lambda(c^{\max} - N\bar{b}\bar{d}) + \frac{\bar{h}}{2} \frac{(\sum_{i=1}^N \sqrt{d_i s_i})^2}{c^{\max} - N\bar{b}\bar{d}} - \bar{h} \sum_{i=1}^N \frac{d_i}{2}$.*

The proof of Theorem 3 follows similar arguments than the proof of Theorem 1.

2.5 New instance generation scheme

The new generation scheme proposed in this chapter is based on the one proposed in Trigeiro et al. (1989) with additional enhancements and parameters. The original instance generation scheme is recalled in Section 2.5.1, and the new generation scheme is outlined in Section 2.5.2. The parameters are described and analyzed in details in Section 2.5.3.

2.5.1 Original generation scheme

In the CLSP instances of Trigeiro et al. (1989), the cost and capacity parameters are constant over time. The number of items varies from 10 to 30, and the production costs are equal to 0. The instances were built as follows:

- **Demand range.** Demands are dynamic with an average value $\bar{d} = 100$. Half of the instances have demands following a uniform probability distribution in the range $[75, 125]$, the other half in the range $[0, 200]$. In addition, 25% of the demands in the first four periods are set to 0.
- **Time Between Order (TBO).** The time between order, defined as $TBO = \sqrt{2 \frac{\bar{f}}{\bar{h}}}$ (Harris (1913)), is in $\{1, 2, 4\}$. In all the original instances, $\bar{h} = 1$.
- **Production and Setup times.** Half of the instances have an average setup time of $\bar{s} = 11$, and of $\bar{s} = 43$ for the other half. All unitary production times are set to $\bar{b} = 1$.

- **Capacity tightness.** For each instance, an average capacity use per period is computed following the EOQ of [Harris \(1913\)](#). This capacity consumption is divided by a factor $\rho \in \{0.75, 0.85, 0.95\}$ to define the instance capacity per period: $c^{\max} = \frac{N}{\rho}(\frac{\bar{s}}{TBO} + \bar{b}\bar{d})$.
- **Variability between items.** Setup times as well as inventory and setup costs for each item are generated based on their average values multiplied by coefficients taking values uniformly in the range $[0.5, 1.5]$.

2.5.2 New generation scheme

In this section, we propose a new generation scheme integrating the features discussed in Section [2.2](#). Since we consider an extension of the CLSP with lost sales, new related parameters will be defined. Lost sales are allowed but at a very high cost.

In order to obtain optimal solutions with limited end-of-horizon effect at the beginning or at the end of the planning horizon, we solve a new mixed integer linear problem that we denote (\overline{P}_T) based on the CLSP formulation (Constraints [\(2.1\)](#)-[\(2.7\)](#)) with the following additional parameters and constraints:

- A global ending inventory I^{inf} , so that the inventory level is not equal to 0 at the end of the horizon, subject to the following constraint:

$$\sum_{i=1}^N I_{iT} \geq I^{inf} \quad (2.32)$$

- An upper bound on the final inventory I^{sup} of each item in order to have enough item diversity in the ending stock, subject to the following constraint:

$$I_{iT} \leq I^{sup}, \quad \forall i \in \{1, \dots, N\} \quad (2.33)$$

- An initial inventory per item I_{i0} to have enough stock to satisfy the first demands.

The tricky point is the set up of the new parameters I^{inf} , I^{sup} and I_{i0} , so that they are in line with the practical considerations discussed in Section [2.2](#). The parameters that do not appear in the outline of the generation scheme below, are generated according to the original scheme described in Section [2.5.1](#). For the new parameters, a reference to the section with the detailed analysis is provided.

- **Demand range.** Demands are dynamic with an average value $\bar{d} = 100$. Half of the instances have demands following a uniform probability distribution in the range $[75, 125]$, the other half in the range $[0, 200]$. Contrary to [Trigeiro et al. \(1989\)](#), no demands in the first four periods are set to 0.
- **TBO, Setup times.** As defined in Section [2.5.1](#).
- **Lost sales cost for item i .** As defined in Section [2.3.4](#).
- **Maximum inventory per item.** $I^{sup} = \frac{N-k}{k}\bar{d}$ (see [\(2.17\)](#) derived in Section [2.3.3](#)), where $k = \min(k_{cost}, k_{capa})$ (see [\(2.15\)](#) derived in Section [2.3.3](#)), with $k_{cost} = N\sqrt{\frac{\bar{d}h}{2f}}$ and $k_{capa} = \frac{c^{\max} - N\bar{d}\bar{b}}{\bar{s}}$.

- **Global ending inventory.** $I^{inf} = \sum_{i=1}^N I_{iT}^*$, where the values of I_{iT}^* are obtained by solving the Mixed Integer Linear Program (MILP) (P_f) in Section 2.5.3.
- **Initial inventory per item.** $I_{i0} = I_{i0}^*$, where the values of I_{i0}^* are obtained by solving (P_f) in Section 2.5.3.
- **Capacity tightness.** $c^{\max} = k\bar{s} + N\bar{d}\bar{b} + c_o^*$, where the value of c_o^* is obtained by solving (P_f) in Section 2.5.3.

Section 2.3 showed how the last four parameters of the new generation scheme were derived. These parameters are fitted in Section 2.5.3. In Section 2.3.3, we defined an approximate global ending inventory \tilde{I}^{inf} , and a maximum inventory per item, denoted by I^{sup} , based on the value of a time between order deduced from the average value of the demands, the average holding and setup costs and the maximum capacity per period. From these parameters, we then define in Section 2.5.3 a new MILP which, given a global inventory \tilde{I}^{inf} as well as a maximum inventory per item I^{sup} , determines feasible initial inventories for each item as well as a capacity limit. The value of I^{inf} is also fitted in order to follow the dynamic nature of the demand.

2.5.3 Fitting I^{inf} and setting initial inventories

In Section 2.3, a global ending inventory and a maximum inventory based on a static cyclical model were proposed. In order to find fitted values for the initial inventories of the dynamic CLSP with lost sales, we solve a MILP where all the constraints of the original model (2.2)-(2.7), as well as additional global ending inventory and maximum inventory constraints, are considered. The initial inventory of each item must be set so that the total initial inventory should be close to \tilde{I}^{inf} , yet individually each initial inventory should be lower than I^{sup} . This comes from the fact that, ideally, the total inventory is constant throughout the time horizon, and individual inventories should not exceed the value of I^{sup} deduced in Section 2.3.3.

We want an inventory configuration at the end of the horizon that is similar to the one at the beginning of the horizon. Therefore, the goal is to find a feasible solution that minimizes the absolute value of the difference between the total initial inventory and the total inventory at the end of the planning horizon. Let us denote by (P_f) the following MILP:

$$\min Kc_o + \delta \quad (2.34)$$

$$I_{i,t-1} + X_{it} = d_{it} + I_{it}, \quad \forall i \in 1, \dots, N, \quad \forall t \in 1, \dots, T \quad (2.35)$$

$$\sum_{i=1}^N (s_{it}Y_{it} + b_{it}X_{it}) \leq \tilde{c}^{\max} + c_o, \quad \forall t \in 1, \dots, T \quad (2.36)$$

$$0 \leq I_{i0} \leq I^{sup}, \quad \forall i \in 1, \dots, N \quad (2.37)$$

$$0 \leq I_{iT} \leq I^{sup}, \quad \forall i \in 1, \dots, N \quad (2.38)$$

$$\gamma \tilde{I}^{inf} \leq \sum_{i=1}^N I_{i0} \leq \tilde{I}^{inf}, \quad (2.39)$$

$$\delta \geq \sum_{i=1}^N (I_{i0} - I_{iT}), \quad (2.40)$$

$$\delta \geq \sum_{i=1}^N (I_{iT} - I_{i0}), \quad (2.41)$$

$$X_{it} \leq M_{it}Y_{it}, \quad \forall i \in 1, \dots, N, \quad \forall t \in 1, \dots, T \quad (2.42)$$

$$Y_{it} \in \{0, 1\}, \quad \forall i \in 1, \dots, N, \quad \forall t \in 1, \dots, T \quad (2.43)$$

$$X_{it}, I_{it} \geq 0, \quad \forall i \in 1, \dots, N, \quad \forall t \in 1, \dots, T \quad (2.44)$$

$$c_o, \delta \geq 0 \quad (2.45)$$

where \tilde{c}^{max} is defined using average values on the costs and capacity, $\tilde{c}^{max} = k\bar{s} + N\bar{d}\bar{b}$ with k defined as shown in Section 2.3.3.

The capacity limit is the sum of the fixed capacity \tilde{c}^{max} and the overtime c_o . The fixed parameter is the estimated capacity consumption in the cyclical configuration. Overtime is added to guarantee the feasibility of the problem with dynamic demands, but is highly penalized in the objective function (2.34) by parameter K . The fixed capacity is defined according to the EOQ provided either by the average costs or by the capacity c^{max} . Variable δ is the gap between the initial inventory and the ending inventory. This gap has to be minimized as well. Constraints (2.36) are the capacity constraints. Constraints (2.37) and (2.38) set bounds on the initial inventories and the ending inventories. Constraint (2.39) ensures that the total initial inventory is close to the global ending inventory, and parameter γ defines the tightness in Constraint (2.39), where $0 \leq \gamma \leq 1$. Constraints (2.40) and (2.41) link δ with the inventory gap, while Constraints (2.42) connect the production and setup variables. Finally, the domains of the variables are given by Constraints (2.43)-(2.45).

In the proposed generation scheme, we define $c^{max} = \tilde{c}^{max} + c_o^*$, where c_o^* is the optimal value obtained by solving (P_f) . In addition, the global ending inventory is fitted to guarantee that a feasible solution without lost sales can be found, $I^{inf} = \sum_{i=1}^N I_{iT}^*$, where I_{iT}^* is the optimal ending inventory obtained by solving (P_f) . In our case, we want to avoid adding overtime to the analytical capacity except to avoid infeasibility without lost sales. We set K the penalization per unit of overtime to an order of magnitude higher than the unit penalization of the gap between the initial and the ending inventories. Regarding γ , ideally this parameter should be close to 1 to keep the same global inventory at the beginning and at the end of the horizon. However, in order to allow some slack due to the variability of the costs and the parameters, we can set a slightly lower value. In our computational experiments, we set $K = 100$ and $\gamma = 0.95$.

The initial inventory values I_{i0} are considered as decision variables because we want to create instances that are relevant on a rolling horizon. However a possible extension could be to assume that the initial inventories are known parameters and to define an adequate minimum ending inventory level I^{inf} that is close to \tilde{I}^{inf} but takes into account the potential lack of inventory during the first periods and the impact of the capacity. Ideally, the ending inventory should be close to \tilde{I}^{inf} without lost sales or extra capacity required. The analytical minimum inventory target \tilde{I}^{inf} might be reached after a few iterations over the rolling horizon even without initial inventories. If we assume I_{i0} are known parameters, we can remove Constraints (2.37) and (2.39) and modify Constraints (2.40) and (2.41) from (P_f) as follows:

$$\delta \geq \tilde{I}^{inf} - \sum_{i=1}^N I_{iT}, \quad (2.46)$$

$$\delta \geq \sum_{i=1}^N I_{iT} - \tilde{I}^{inf}. \quad (2.47)$$

Constraints (2.46) and (2.47) define δ as the absolute value of the difference between the ending inventory and the analytical ending inventory \tilde{I}^{inf} .

2.6 Computational experiments

The original instances of [Trigeiro et al. \(1989\)](#) are compared with our new instances in Section [2.6.1](#) while, in Section [2.6.2](#), the relevance of considering ending inventories on a rolling horizon is shown.

2.6.1 Comparison of original and new instances

As a benchmark for the creation of the new proposed instances, we use the set of 180 instances of the library LOTSIZELIB proposed by [Trigeiro et al. \(1989\)](#) with $N = 10$ items and $T = 20$ periods, and the 180 instances with $N = 30$ items. All the instances have an average demand of 100 units per period and per item and, for a given instance, all items have the same cost and demand pattern. Half of the instances have low demand variability (demand between 75 and 125), and the other half have high demand variability (demand between 0 and 200). We first modified the demands that were originally set to 0 in the first periods to ensure the feasibility of the instances, by assigning them a random value generated as the strictly positive demand (see Section [2.5.1](#)). The original instances will be referred as “Orig.” and correspond to problem (P_T) (see Section [2.2.1](#)), and the instances created by applying the new generation scheme summarized in Section 5.2 as “New” and correspond to problem (\overline{P}_T) (see Section [2.5.2](#)).

The mathematical models are solved using IBM ILOG CPLEX 12.7 on a computer with 2.6 GHz PC, 64 GB of RAM and 2 processors, with a maximum running time of 600 seconds for each instance, except for Table [2.3](#) where the time limit is set to 100 seconds.

Adding the initial inventory and the ending inventory constraints is supposed to mitigate the end-of-horizon effect, whose potential main impacts are a drop of production in the last periods and a large number of setups in the first periods to satisfy the initial demands. A way to measure this edge effect is to analyze the variation of the number of setups and of the production between periods. Indeed, for the original instances, this variation is high because of the end-of-horizon effect, as observed in Figure [2.2](#). The initial and ending inventories are established using a cyclical submodel, where we assume a constant production and number of setups over time. Even if the CLSP model with minimum and maximum ending inventories is not cyclical, we expect to find an optimal production plan with low variability between periods, which should lead to a reduction of the impact of the end-of-horizon effect (see Figure [2.6](#)). In order to show that the constraints on the minimum and maximum ending inventories do not make the model easier to solve, the average relative optimality gaps, denoted Gap , were computed for both the original model and the new model. The relative optimality gap corresponds to the relative difference between the best upper bound UB and the best lower bound LB obtained by IBM ILOG CPLEX: $Gap = \frac{UB-LB}{UB}$.

Let us define a variability coefficient as the ratio between the standard deviation and the average value over all the periods. For the original and new instances, Table [2.1](#) provides the variability coefficients for the number of setups per period, the total quantity produced per period and the total inventory per period. The results are classified according to different parameters: Number of items (N), Time Between Orders (TBO), demand range, average setup time and capacity tightness. Except for the capacity tightness, the classification parameters are not affected by the modification of the original instances using the new generation scheme. It is worth noticing that the capacity tightness relies on the computation of EOQ, and hence on the average demand. The following observation is interesting to note about the original instances. Because the capacity is defined by dividing the capacity required

for an EOQ production with a coefficient that is smaller than 1, in the problem studied in Section 2.3.1, the shape of the relaxed solution is only guided by the costs, that is $k = k_{cost}$. By following the new generation scheme in Section 2.5, it should also be the case for the new instances. However, in all the original instances of Trigeiro et al. (1989), the capacity was deduced by taking an average demand of 90 to compute the EOQ. This should not be the case as Trigeiro et al. (1989) state that the average demand is equal to 100. However, unlike described in their generation scheme, not 25% but 50% of the demands in the first four periods were set to 0 when generating the original instances, and the demands that were removed were not balanced among the demands at other periods. When recomputing k_{capa} with the same capacity but with an average demand of 100 for the new instances, there are cases where $k_{capa} < k_{cost}$, thus an optimal solution guided by the capacity.

		Variability coefficient					
		Setup		Production		Inventory	
		Orig.	New	Orig.	New	Orig.	New
N	10	0.29	0.15	0.26	0.04	0.70	0.40
	30	0.26	0.11	0.25	0.03	0.59	0.27
TBO	1	0.22	0.09	0.23	0.06	1.17	0.75
	2	0.24	0.13	0.25	0.03	0.42	0.18
	4	0.37	0.16	0.28	0.01	0.35	0.08
Demand range	[75;125]	0.27	0.09	0.25	0.02	0.76	0.32
	[0;200]	0.28	0.16	0.25	0.05	0.53	0.35
Average setup time	11	0.28	0.14	0.25	0.02	0.68	0.34
	43	0.27	0.11	0.26	0.04	0.61	0.33
Capacity tightness ^(*)	EOQ/0.75	0.31	0.12	0.31	0.04	0.66	0.37
	EOQ/0.85	0.27	0.12	0.26	0.04	0.76	0.46
	EOQ/0.95	0.25	0.14	0.18	0.02	0.52	0.17

Table 2.1: Comparison of the variability for the original and new instances.

^(*)This classification only applies on the original instances.

Table 2.1 shows that the expected behavior is observed, i.e. the variability is greatly reduced when adding the initial inventories and the global ending inventory. The variability of the inventory is larger than the variability of the setup and production, but always significantly lower for the new instances than for the original ones. On average, the variability coefficients for the quantity produced are between 5 to 28 times smaller when the ending inventory constraints are added (from 0.28 to 0.01 for the instances with a TBO of 4). This implies that the deviation from the average quantity produced at each period is much smaller for the new instances. As illustrated in Figure 2.2, this variability in the produced quantities was mostly caused by the end-of-horizon effect. Note that the setup range, defined by the difference between the maximum and the minimum number of setups in a period, is also lower for the new instances.

Let us now analyze the capacity utilization in more detail. Table 2.2 shows the mean capacity utilization over all periods and the standard deviation between the periods. The mean capacity utilization is defined as the average value of the ratio between the consumed capacity and the available capacity over the time horizon. The capacity parameters of the original instances were generated by taking an average value of the capacity required to have a production based on the EOQ of each item and by dividing it by a coefficient of 0.75, 0.85

		Capacity Utilization			
		Mean (%)		Variability coefficient	
		Orig.	New	Orig.	New
N	10	84.0	98.2	0.25	0.03
	30	84.1	98.4	0.24	0.02
TBO	1	83.0	95.8	0.22	0.05
	2	84.2	99.3	0.24	0.02
	4	85.0	99.8	0.28	0.00
Demand range	[75;125]	84.6	99.4	0.25	0.01
	[0;200]	83.5	97.2	0.25	0.04
Average setup time	11	84.5	98.9	0.24	0.02
	43	83.6	97.7	0.25	0.03
Capacity tightness ^(*)	EOQ/0.75	74.4	97.6	0.31	0.03
	EOQ/0.85	84.5	98.1	0.26	0.03
	EOQ/0.95	93.2	99.3	0.17	0.01

Table 2.2: Capacity utilization comparison.

(^{*})This classification only applies on the original instances.

and 0.95 (Trigeiro et al. (1989)). It is clear that the mean capacity utilization is much larger for the new instances, always larger than 95% and most often close or larger than 98%. On the opposite, the mean capacity utilization for the original instances is nearly always smaller than 85%, and is even equal to 74.4% when the capacity tightness is equal to EOQ/0.75. The results are even more impressive when considering the variability, which is never larger than 0.05 in the new instances, whereas it is always larger than 0.17 in the original instances with a peak at 0.31, again when the capacity tightness is equal to EOQ/0.75.

Tables 2.3 and 2.4 compare the average optimality gaps and the average computational times of the original and new instances as well as the maximum optimality gap observed for each set of instances. With a maximum computational time of 100 seconds, the average optimality gap for the new instances is as large as 2.3% when TBO is equal to 4, whereas it is never larger than 0.9% for the original instances. When the maximum computational time is increased to 600 seconds, the differences between the optimality gaps remain large, up to 1.7% when TBO is equal to 4. Average computational times are also much larger when solving the new instances.

Finally, Tables 2.3 and 2.4 show that the new instances are harder to solve than the original ones, i.e. adding initial inventories and a global ending inventory does not make the problem easier to solve and raises issues as how to solve the new instances efficiently.

Figure 2.6 displays the optimal plans for an original instance and its associated new instance. Note that the production is relatively constant over time, as is the number of setups in each period. However, in the first and last periods, there is both a decrease in the number of setups and an increase of the capacity utilization, in line with the other periods, in Figure 2.6b compared to Figure 2.6a, i.e. in the new instance compared to the original one. By adding initial inventories and a global ending inventory, the capacity is fully consumed in all the periods of the horizon in the optimal production plan of the new instance. Moreover, except for slight variations caused by differences in the costs between items, the fraction of capacity used for setup times is rather stable throughout the planning horizon.

Initial and global ending inventories were deduced from the study of a relaxed version of the problem defined in Section 2.3.1. In Sections 2.3.3 and 2.4, analytical values for objective

		Gap (%)		MaxGap (%)		T (s)	
		Orig.	New	Orig.	New	Orig.	New
N	10	0.4	1.7	4.9	6.6	34.5	71.7
	30	0.2	0.8	1.4	8.1	38.7	78.2
TBO	1	0.1	0.3	1.9	4.3	25.6	34.1
	2	0.2	1.0	2.0	5.8	31.7	90.7
	4	0.5	2.3	4.9	8.1	52.6	100.2
Demand range	[75;125]	0.4	1.5	4.9	8.1	41.9	78.2
	[0;200]	0.2	0.9	3.4	4.6	31.3	71.8
Average setup time	11	0.3	1.5	4.9	8.1	39.0	81.6
	43	0.3	0.9	2.7	4.6	34.2	68.3
Capacity tightness ^(*)	EOQ/0.75	0.0	1.0	0.0	5.8	1.2	68.0
	EOQ/0.85	0.0	0.9	0.7	6.4	19.2	67.2
	EOQ/0.95	0.9	1.7	4.9	8.1	89.4	89.7

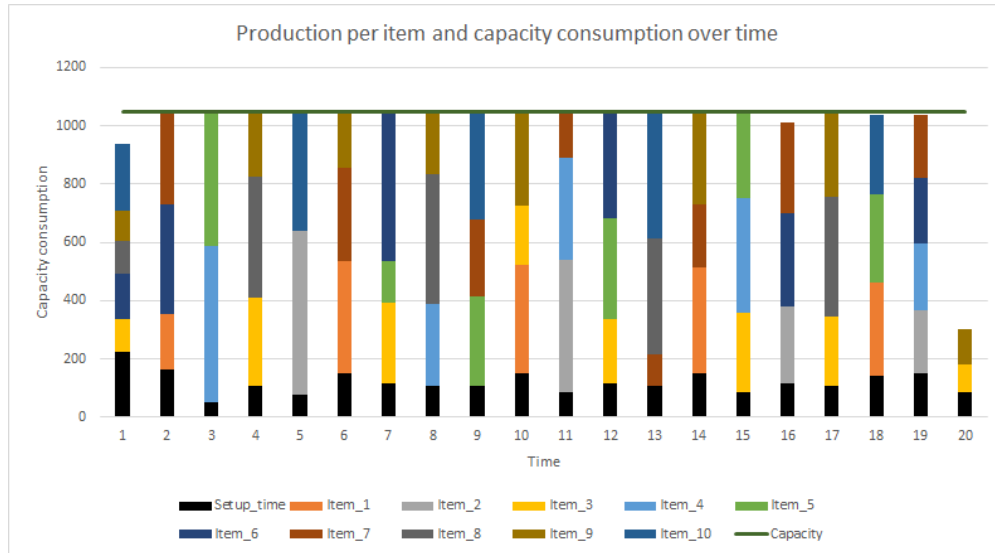
Table 2.3: Average optimality gaps and computational times for the original and new instances with $T_{lim} = 100$ sec.

^(*)This classification only applies on the original instances.

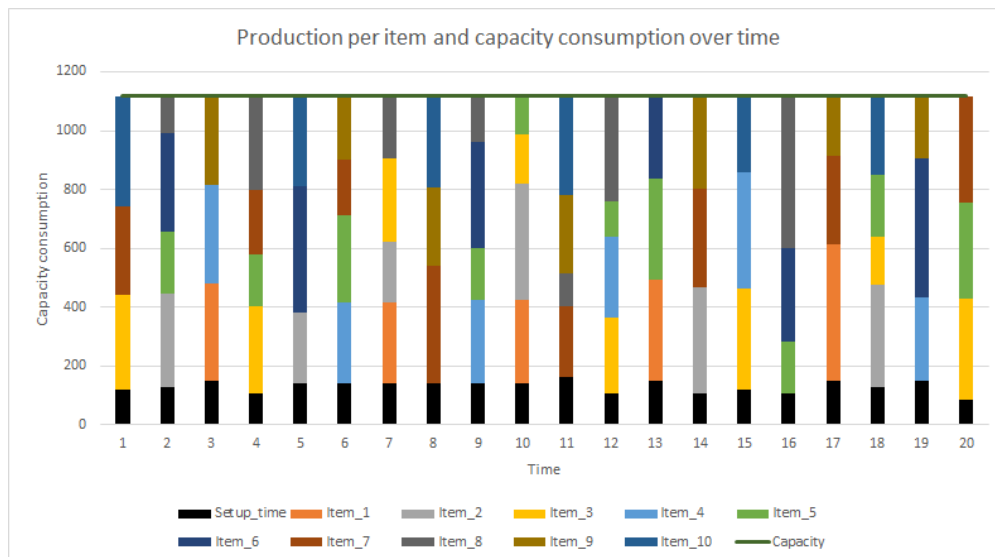
		Gap (%)		MaxGap (%)		T (s)	
		Orig.	New	Orig.	New	Orig.	New
N	10	0.3	1.3	4.5	5.8	171.7	385.4
	30	0.1	0.5	1.0	2.7	212.5	453.8
TBO	1	0.1	0.2	1.1	2.3	138.7	165.5
	2	0.2	0.8	1.8	5.4	180.1	494.6
	4	0.4	1.7	4.5	5.8	257.5	598.6
Demand range	[75;125]	0.3	1.2	4.5	5.8	225.4	453.6
	[0;200]	0.2	0.6	3.2	3.1	158.7	385.5
Average setup time	11	0.2	1.1	4.5	5.8	206.2	467.8
	43	0.2	0.7	2.4	3.6	178.0	371.4
Capacity tightness ^(*)	EOQ/0.75	0.0	0.8	0.0	5.5	5.6	367.9
	EOQ/0.85	0.0	0.7	0.0	5.5	62.9	361.3
	EOQ/0.95	0.7	1.3	4.5	5.8	507.7	529.5

Table 2.4: Average optimality gaps and computational times for the original and new instances with $T_{lim} = 600$ sec.

^(*)This classification only applies on the original instances.



(a) Solution for original instance.



(b) Solution for new instance.

Figure 2.6: Comparing the optimal plans of two related original and new instances.

functions are provided. To validate the study of this simplified problem to deduce values for the CLSP with setup times, Table 2.5 displays the gaps between the best upper bound obtained by solving the MILP model and the analytical optimal values for each instance.

		Gap (%)	
		Orig.	New
N	10	9.7	4.3
	30	10.6	5.9
TBO	1	13.3	5.1
	2	10.4	5.2
	4	6.7	5.1
Demand range	[75;125]	6.2	2.4
	[0;200]	14.1	7.8
Average setup time	11	10.3	5.0
	43	10.0	5.3
Capacity tightness^(*)	EOQ/0.75	11.4	5.1
	EOQ/0.85	10.5	5.2
	EOQ/0.95	8.5	5.1

Table 2.5: Gap with predicted objective value.

(*)This classification only applies on the original instances.

Table 2.5 shows that, for the original and new instances, the gap between the predicted and the optimized objective values is on average equal to 5.1%, which implies that the problem defined in Section 2.3.1 is a good relaxation of the CLSP with setup times. When the demand range is small, the approximation is even better. The gap is equal to 2.4% for a small demand range, whereas it increases to 7.8% when the demand has a larger range. That makes sense because the smaller the demand range, the closer each demand is to its average value. For the original instances, the average gap is equal to 10.2%, so the approximation is less precise. That can in part be explained by the fact that the original instances have neither initial nor ending inventories, hence approximating the original problem by a problem on a rolling horizon might be too constraining. Note also that, in Table 2.5, the quality of the approximation does not seem to depend on the TBO for the new instances. The TBO is theoretically linked to $k = k_{capa} = k_{cost}$ by the formula $k = \frac{N}{TBO}$ in the new generated instances. This is an interesting point as the average number of setups per period k is the main factor shaping the production plan. Consequently, the optimal production plan varies greatly depending on the TBO but the approximation remains of the same quality. In the original instances of Trigeiro et al. (1989), the analytical optimal value better approximates the best upper bound for the problem as the TBO increases (13.3% of average gap for a TBO of 1 to 6.7% for a TBO of 4). The poor evaluation of the optimal objective value for smaller TBO can be partly explained by the fact that, in the original instances of Trigeiro et al. (1989), 50% of the demands in the first 4 periods are set to 0. For the instances with a TBO close to 1, when the capacity is not constraining, the number of setups during the first periods can be reduced compared to the analytical average number of setups. This leads to an overevaluation of the optimal objective value. For the instances with tight capacity, some of the later demands need to be satisfied during the first periods where demands were removed. This leads to additional inventory costs and an underevaluation of the optimal objective value.

2.6.2 Analysis on a rolling horizon

To test the impact of the global ending inventory on a rolling horizon, we extend the instances of [Trigeiro et al. \(1989\)](#) with $N = 30$ by using the same generation scheme to create instances with $\Omega = 100$. The global minimum ending inventory and maximum ending inventory per item are generated by solving the continuous relaxation of Problem (P_C) defined in Section [2.3.1](#). The continuous relaxation of (P_C) is:

$$\min_{(k_i, \phi_i) \in \mathbb{R}_+^* \times \mathbb{R}_+^*} \sum_{i=1}^N f_i k_i + \bar{h} \sum_{i=1}^N d_i \frac{\phi_i - 1}{2} \quad (2.48)$$

$$s.t. \quad \sum_{i=1}^N s_i k_i \leq c^{\max} - N\bar{b}\bar{d} \quad (2.49)$$

$$k_i \phi_i = 1, \forall i \in 1, \dots, N \quad (2.50)$$

This problem is a *Quadratic Constraint Problem* that can be solved to optimality using IBM ILOG CPLEX with the barrier algorithm. The obtained minimum and maximum inventory levels remain constant through the rolling horizon. Three settings for the time horizon T are considered ($T \in \{5, 10, 20\}$) in order to fix the decisions period by period. Compared to the original instances, we modified the instances and the capacity by taking $k_{capa} \in \{3.75, 7.5, 15\}$ and set $c_t^{\max} = N\bar{b}\bar{d} + k\bar{s}$, where $k = \max(k_{capa}, N\sqrt{\frac{d\bar{h}}{2\bar{f}}})$. The initial inventory for each item has been set as half the maximum ending inventory for each item ($I_{i0} = \frac{I_i^{sup}}{2}$), which corresponds to the average inventory level based on the cycle length determined in Section [2.4](#).

When optimizing on a rolling horizon, Constraint [\(2.32\)](#) was slightly modified to allow lost sales on the global ending inventory:

$$\sum_{i=1}^N I_{iT} \geq I^{inf} + l_{T+1}.$$

The new parameter l_{T+1} , which is the unit cost of lost sales for the ending inventory, is defined so that it is less costly to have lost sales in the last period of the horizon than in previous periods.

		Fraction of lost sales (%)					
		T=5		T=10		T=20	
		Global Ending Inventory		Global Ending Inventory		Global Ending Inventory	
		w/o	with	w/o	with	w/o	with
		Avg Max	Avg Max	Avg Max	Avg Max	Avg Max	Avg Max
Demand range	[75;125]	6.74 9.68	0.61 1.87	4.14 5.85	0.09 0.40	2.13 3.02	0.13 0.65
	[0;200]	6.20 10.65	0.93 3.45	3.87 6.45	0.52 2.29	1.99 3.28	0.38 2.36
Average	11	4.38 6.24	0.25 1.36	2.93 4.46	0.22 1.29	1.60 3.06	0.19 2.36
setup time	43	8.56 10.65	1.29 3.45	5.08 6.45	0.40 2.29	2.52 3.28	0.33 1.19

Table 2.6: Fraction of lost sales (%) for $N = 30$ and $k_{capa} = 3.75$

		Fraction of lost sales (%)					
		T=5		T=10		T=20	
		Global Ending Inventory		Global Ending Inventory		Global Ending Inventory	
		w/o	with	w/o	with	w/o	with
		Avg Max	Avg Max	Avg Max	Avg Max	Avg Max	Avg Max
Demand range	[75;125]	5.06 7.16	0.24 0.72	2.78 3.58	0.17 0.70	1.23 1.61	0.19 0.47
	[0;200]	4.56 7.71	0.79 1.89	2.53 4.15	0.58 1.63	1.04 1.77	0.32 1.12
Average setup time	11	3.56 5.13	0.47 1.89	2.22 3.57	0.39 1.63	1.06 1.77	0.22 1.12
	43	6.05 7.71	0.56 1.88	3.09 4.15	0.36 1.37	1.21 1.71	0.28 0.66

Table 2.7: Fraction of lost sales (%) for $N = 30$ and $k_{capa} = 7.5$

		Fraction of lost sales (%)					
		T=5		T=10		T=20	
		Global Ending Inventory		Global Ending Inventory		Global Ending Inventory	
		w/o	with	w/o	with	w/o	with
		Avg Max	Avg Max	Avg Max	Avg Max	Avg Max	Avg Max
Demand range	[75;125]	1.71 2.87	0.15 0.47	0.75 1.54	0.14 0.40	0.21 0.65	0.07 0.20
	[0;200]	1.39 3.06	0.47 2.15	0.57 1.85	0.28 1.43	0.16 0.89	0.11 0.58
Average setup time	11	1.48 3.06	0.44 2.15	0.79 1.85	0.29 1.43	0.29 0.89	0.11 0.58
	43	1.62 3.01	0.18 0.74	0.53 1.05	0.13 0.40	0.08 0.24	0.07 0.20

Table 2.8: Fraction of lost sales (%) for $N = 30$ and $k_{capa} = 15$

Tables 2.6, 2.7 and 2.8 compare the fraction of lost sales on a rolling horizon on instances classified according to their demand range and average setup time. Each table presents the results for a specific value of k_{capa} without or with the global minimum ending inventory and for different planning horizons (T). The influence of each parameter is similar in each table, even if it can be noticed that the average fraction of lost sales seems to increase when the capacity becomes tightened. With $k_{capa} = 3.75$, the average fraction of lost sales is equal to 6.74% for $T = 5$ and no ending inventory constraint while, for the same instances and $k_{capa} = 15$, the average fraction of lost sales decreases to 1.96%.

For a specific value of k_{capa} , several remarks can be raised. In terms of lost sales, the results are much better with ending inventory constraints. In Table 2.6 with $T = 5$, the average lost sales of 6.47% without ending inventory constraints drops to 0.77%. In Table 2.7, the average lost sales decrease from 4.81% to 0.52% and, in Table 2.8, from 1.55% to 0.31%. The average lost sales without ending inventory constraints are almost always larger than 1%, except when both the planning horizon and the capacity are large. With ending inventory constraints, even with a small planning horizon and a tight capacity, the average lost sales are almost always smaller than 1%. With ending inventories, there are on average both less lost sales and less setup times. In Table 2.6 for $T = 10$, the average lost sales are equal to 2.93% with $\bar{s} = 11$ and 5.08% with $\bar{s} = 43$. This makes sense because, when setup times are small compared to the available capacity, the impact they can have on the feasibility of the problem is less relevant. When setup times increase, up to a certain point, it becomes more difficult to recover from a lack of production in a previous period, leading to an increase of the lost sales. However, this effect disappears when setup times reach a threshold, because then the number of setups per period is fixed and the decisions in the first periods will not lead to more sales later on.

To illustrate this point, let us consider the case with $N = 3$, $T = 5$, $d = 100$, with no

setup and production costs. The holding costs and unitary production times are set to 1. Let us consider different values for the setup times, $\bar{s} \in \{0, 10, 100\}$. Let us set $k_{capa} = 1$ and define $c^{max} = 300 + k_{capa}\bar{s}$. To model the lack of production in a previous period, we consider the case where the initial inventory is set to 0 for all items. Lost sales are highly penalized. The optimal production plans can be found in Figure 2.7.

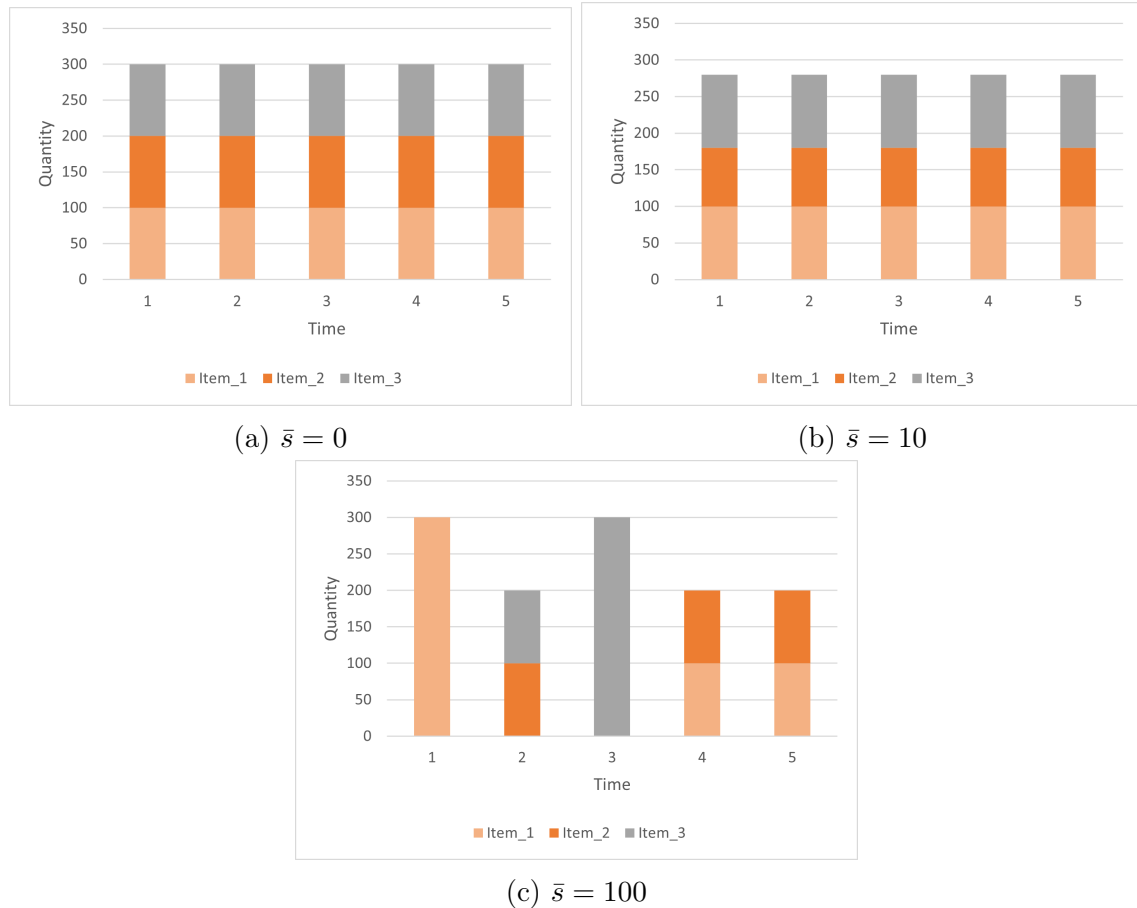


Figure 2.7: Production quantities for different values of \bar{s}

- For $\bar{s} = 0$: There are no lost sales in the optimal solution. This can be explained by the fact that a lack of production from a previous period can easily be recovered because all demands at a period can be satisfied by a production at the same period, which is not the case for $\bar{s} > 0$.
- For $\bar{s} = 10$: There are 80 units of lost sales (20 units of item 2 at each period) in the optimal solution. The production plan is similar to the one obtained for $\bar{s} = 0$. However, when there are 3 setups in a period, only 280 units can be produced, leading to a deficit of 20 units per period.
- For $\bar{s} = 100$: There are 300 units of lost sales (100 units of item 2 at $t = 1$ and at $t = 3$, 100 units of item 3 at $t = 1$) in the optimal solution. Because each setup takes $\frac{1}{4}$ of the capacity, it is suboptimal to have 3 setups in the first period. However, we get a production plan that is optimal on a rolling horizon if one period is fixed in each optimization run, leading to less lost sales compared to Figure 2.7b.

- For $\bar{s} \geq 300$: Each setup takes more than half the capacity available during a period, so the optimal production plan consists in exactly one setup in each period. There are also 300 units of lost sales in the optimal solution. Because the production capacity forces to have at most one setup per period, the resulting production plan is the optimal production plan on a rolling horizon.

Note that, without constraints on inventory levels at the end of the planning horizon, the results depend a lot on the length of the planning horizon. The fluctuation of the fraction of lost sales when the length of the planning horizon increases is larger without ending inventories. Table 2.7 shows that, for $\bar{s} = 43$, the fraction of lost sales decreases from 6.05% to 3.09% when T increases from 5 to 10. With ending inventories, the best results are obtained when there is a small variability in demands. For instance, for $T = 10$ in Table 2.7, the fraction of lost sales drops from 0.79% for instances with a large demand range to 0.24% for instances with a small demand variability. This also makes sense because the computation of the ending inventories are based on models using averages. When looking more specifically at the evolution of the inventories in each period, the ending inventories enable for inventories to be kept throughout the horizon rather than being depleted in the first periods.

2.7 Conclusions

In this chapter, we first highlight the issues associated with classical instances of the literature for the capacitated lot sizing problem with setup times. Our analysis shows that the lack of ending inventories creates inconsistencies with the industrial reality where planning is performed on a rolling horizon. In particular, even when using a long planning horizon, decisions in the first periods might be negatively impacted. Building on this analysis and to mitigate the end-of-horizon effect, we propose to modify the problem by considering a global minimum ending inventory and maximum ending inventories for items. The values of these parameters are deduced from an analysis of cyclical lot-sizing problems with setup times. A new scheme, extending the one of Trigeiro et al. (1989), is proposed to generate instances with initial and ending inventories. Numerical results on the new instances show the practical relevance of the new problem.

As future research perspectives, we believe it is interesting to study how constraints on ending inventories impact various extensions of the CLSP with setup times, such as demand and production time windows or inventory capacity constraints. Another interesting research topic would be to pursue the work started at the end of Section 2.5.3, namely investigating approaches to determine relevant ending inventories from known initial inventories and probably also demand forecasts.

Chapter 3

Capacitated lot-sizing problem with inventory constraints within periods

In Chapter 2, we have extended the CLSP with setup times by adding ending inventory constraints to mitigate the end-of-horizon effect. These ending inventory constraints enabled for a better inventory management on a rolling horizon. In this chapter, still to improve inventory management, we consider the evolution of inventory within periods. In case the inventory is bounded and under specific assumptions on how demand and production evolve over time, we propose new constraints to bound the inventory levels reached within each period. Our contributions in this chapter are twofold. First, we propose different CLSP models to capture the inventory evolution within periods. Second, we present and discuss numerical experiments to assess the practical relevance of the proposed models.

The chapter is organized as follows. Section 3.1 positions our work in the current literature on the subject. Section 3.2 motivates the need to consider the dynamic evolution of the inventory within each period when solving a CLSP. Section 3.3 proposes a first model with a uniform production rate and a bounded demand rate. Section 3.4 proposes a second model where the production occurs at maximum rate and the demand is instantaneous. Computational experiments are presented in Section 3.5. Potential extensions are introduced in Section 3.6 and some conclusions and perspectives can be found in Section 3.7.

3.1 Related works

There are some works in the literature dealing with inventory bounds that limit the products that can be kept in the inventory (Love (1973), Gutiérrez et al. (2003), Atamtürk and Küçükyavuz (2008), Hwang and van den Heuvel (2012), Liu (2008), Önal et al. (2012)). These bounds are practically relevant in various industrial applications. A survey on lot-sizing problems with inventory bounds can be found in Minner (2009). We can also quote Absi and Kedad-Sidhoum (2009) that consider penalty costs if the products kept in inventory do not meet a safety stock level. Moreover, the increasing research considering stochastic demands emphasizes the need to have a minimum stock level on the inventory at each period. Some works consider production and demand rates to model the evolution of the production and the inventory in production planning (Dauzère-Pérès et al. (2000)). Grigoriev et al. (2020) consider, among other models, a lot-sizing problem with production and demand rates (under strong assumptions on these rates). It is also considered in the Economic Lot-Scheduling literature (Elmaghraby (1978)). However, in the classical lot-sizing literature, inventory bounds are always set on discrete periods, not taking into account the dynamic nature of the inventory evolution within each period. Most of the time, inventory bounds are considered at the beginning of the period but, in some cases, the inventory can also take into account a production occurring within the period (Gutiérrez et al. (2013)). Inventory constraints have also been combined with other constraints such as lost sales (Hwang et al. (2013)) or backlogs (Hwang and van den Heuvel (2012)). In the recent literature, inventory bounds are still modeled assuming that quantities are produced and demands taken from the inventory only at the end of each period. The focus in the past years has mostly been on proposing more realistic problem formulations for discrete problems and not on modeling the evolution of the inventory within periods (Brahimi et al. (2017), Zhong et al. (2016), Jing and Mu (2020)).

3.2 Motivations

Section 3.2.1 recalls the CLSP with setup times, lost sales and inventory bounds, while Section 3.2.2 motivates why considering the inventory evolution within periods can be important.

We consider the CLSP with minimum and maximum ending inventories defined in Chapter 2. We add for each item i at period t a lower (resp. upper) bound denoted by I_{it} (resp. \overline{I}_{it}). The maximum ending inventory per item I_i^{sup} for $i \in \llbracket 1, N \rrbracket$ introduced in Chapter 2 is replaced by \overline{I}_{iT} . Let us first recall the mathematical formulation of the problem before addressing inventory constraints within periods.

3.2.1 Original problem formulation

The quantity of item i produced at period t is given by variable X_{it} . The binary variable Y_{it} indicates whether a setup for item i occurs at period t or not. Variable I_{it} is the inventory variable for item i at the end of period t . Finally, variable L_{it} defines the quantity of lost sales for item i at the end of period t . We extend the definition of I_{it} with $t = 0$ to describe the initial inventory of item i . The formulation of the CLSP with setup times, lost sales and minimum ending inventory follows.

$$\min \sum_{i=1}^N \sum_{t=1}^T (f_{it}Y_{it} + p_{it}X_{it} + h_{it}I_{it} + l_{it}L_{it}) \quad (3.1)$$

$$\begin{aligned} & \text{(2.2)} - \text{(2.7)} \\ & \underline{I}_{it} \leq I_{it} \leq \overline{I}_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \end{aligned} \quad (3.2)$$

$$\sum_{i=1}^N I_{iT} \geq I^{inf} \quad (3.3)$$

The objective function (3.1) minimizes the total production, setup, inventory and lost sales costs for all items over the planning horizon. Constraints (3.2) ensure the lower and upper bounds for the inventory. Constraint (3.3) sets a lower bound for the ending inventory. It should be noted that the global minimum ending inventory constraint only applies at the end of the planning horizon, once all satisfied demands are consumed and all production processes are finished. Consequently, contrary to the other inventory constraints, the satisfaction of Constraint (3.3) does not depend on the demand and production rates within each period.

3.2.2 Inventory evolution within periods

Even if it is common to consider inventory bounds in an industrial manufacturing problem, Constraints (3.2) only apply at the end of each period, and do not capture the dynamic nature of the inventory evolution. To illustrate this point, let us consider the specific case of a single item to be produced during a single period with no initial inventory. The demand is $d = 500$. The length of the period is $c^{max} = 1000$ and the demand occurs instantaneously at $t^d = 250$. The item is produced at a unit production rate of $b = 1$ and there are no setup times. For that specific case, it can be noticed that it is not possible to satisfy the entire demand when considering the dynamic evolution of the inventory since no more than 250 units can be produced at t^d (Figure 3.1). Yet, in a regular lot-sizing model using Constraints (3.2), only c^{max} and b are considered, leaving enough capacity to satisfy the demand entirely.

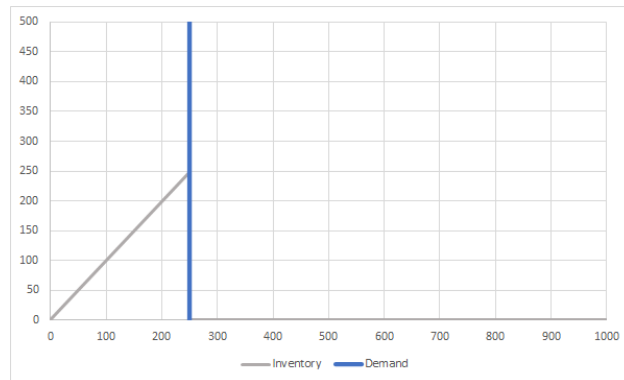


Figure 3.1: Dynamic evolution of the inventory

This consideration is relevant in at least two cases:

1. The production of one item in a workshop determines the production in another workshop. There is a need for synchronization between these two levels. Hence, even in short periods, the exact time at which the demand occurs is well known. In that specific case, it is important to have a precise description of the shape of the demand, possibly at the expense of strong assumptions on the shape of the production. This case corresponds to the industrial application introduced in Section 1.2 of Chapter 1.

2. Periods are long (weeks/months) and the interval during which the demand is consumed is much smaller than the period (hours/days). In this case, because the time at which the demand occurs is small in comparison to the period length, it can be assumed that the demand is instantaneous. Not considering a production rate when optimizing the production decisions might lead to underestimating the inventory bound constraints. To illustrate a lot-sizing problem where decisions are made on a monthly basis, we can take the example of a clothing manufacturer making both production and routing decisions between the production sites and the warehouses in order to satisfy the monthly demand from the retailers.

Let us consider in the following two models to handle a finer modeling of the inventory constraints within periods. The model proposed in Section 3.3 corresponds to a model compatible with the first case, while the model proposed in Section 3.4 corresponds to the second case.

3.3 Uniform production and demand within a cone of uncertainty

We consider in this section that the production occurs at a uniform rate right after the setup time. We propose in Section 3.3.1 an approximation of the demand within a cone of uncertainty and in Section 3.3.2 new linear constraints to guarantee that, with these assumptions, the inventory remains within its bounds in each period.

3.3.1 Demand approximation within a cone of uncertainty

Because it covers a large set of possible scenarios of demand evolution within a period, we propose to approximate each demand by two slopes and two offsets. The first slope corresponds to the demand at its earliest time, the second to the demand at its latest time. The offsets allow for these slopes to be adjusted as tightly as necessary. We define four new parameters for each period t : o_t^b (resp. o_t^e) is the offset before the start of the early (resp. late) demand, α_t (resp. β_t) is the early (resp. late) demand rate.

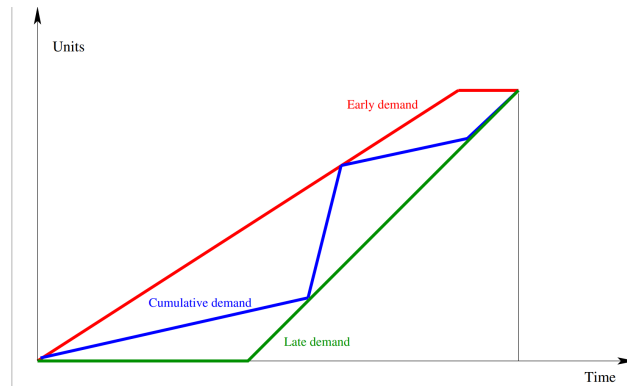


Figure 3.2: Convex demand

Figure 3.2 illustrates the approximation of a demand that is consumed within a given time interval. This approximation can still be used even in the case of uncertainties on the exact distribution of the demand. These two slopes frame not only a linear distribution but also approximate any given distribution that is inside this *cone of uncertainty*.

The proposed demand approximation allows the following specific cases of demand shapes to be considered:

1. If the demand has a convex shape, as illustrated in Figure 3.3, then the line that corresponds to the early demand rate and the one corresponding to a uniform demand are identical. To create the line for the late demand rate, β_t is set equal to the slope at the end of the demand consumption.

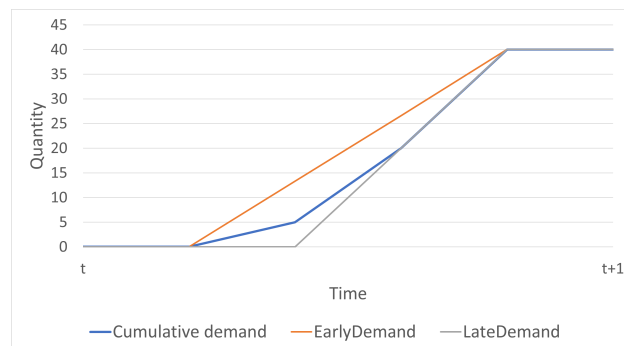


Figure 3.3: Convex demand

2. If the demand has a concave shape, as illustrated in Figure 3.4, then the line that corresponds to the late demand rate and the one corresponding to a uniform demand are identical. To create the line for the early demand rate, α_t is set equal to the slope at the beginning of the demand consumption.

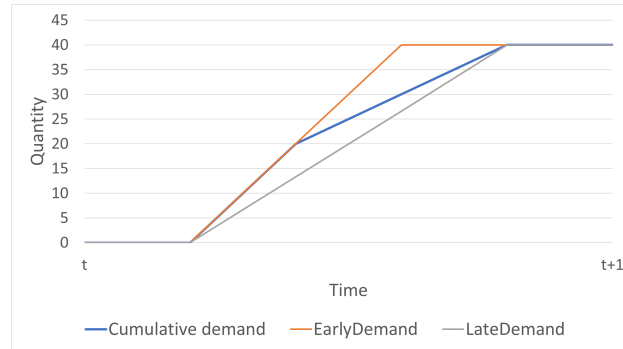


Figure 3.4: Concave demand

3. If the demand is neither convex nor concave, as in Figure 3.5, then the slope for the early rate α_t corresponds to the maximum value when considering all slopes from the beginning of the demand consumption to one of the local extrema. The slope for the late demand rate β_t corresponds to the maximum value when considering all slopes from one of the local extrema to the end of the demand consumption.

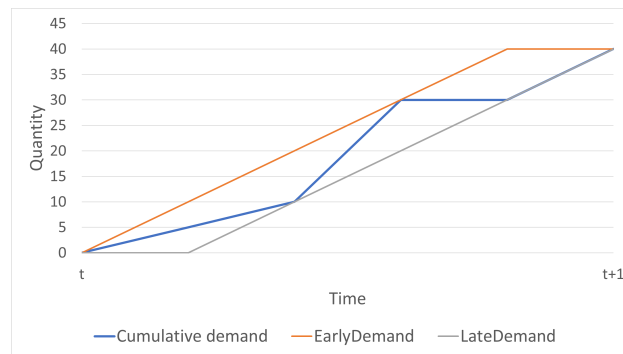


Figure 3.5: Generic demand

3.3.2 Modeling inventory constraints

Let us consider the single-item capacitated lot-sizing problem with setup times where each demand is approximated as in Section 3.3.1. With the assumptions on the production and demand rates, new constraints for inventory levels within periods can be linearly modeled. Several cases associated to minimum inventory constraints and maximum inventory constraints are considered in the following. We assume that lost sales are allowed under the specific condition that the time taken to consume the demand is constant and that only the demand rates and the quantity of satisfied demand are affected. Thus the fact that the production and demand lines cross does not depend on the fraction of the demand that is satisfied.

Minimum inventory constraints

To consider minimum inventory levels within periods, we only need to consider the slope for the demand approximation of the earliest demand. Three cases need to be considered, leading to the addition of one or two new constraints in the original lot-sizing model.

Case 1. The production starts before the demand ($s_t \leq o_t^b$), see example of Figure [3.6](#).

In that case, the minimum inventory level is reached either at the beginning of the period or when the satisfied demand $d_t - L_t$ is fully consumed. In the latest case, the inventory is $I_{t-1} - d_t + L_t + \frac{X_t}{c_t^{max} - s_t} (o_t^b + \frac{d_t}{\alpha_t} - s_t)$. Thus, to respect the minimum inventory constraint, one new constraint needs to be added to the model to guarantee that this inventory level is above the inventory lower bound:

$$I_{t-1} - d_t + L_t + \frac{X_t}{c_t^{max} - s_t} (o_t^b + \frac{d_t}{\alpha_t} - s_t) \geq \underline{I}_t \quad (3.4)$$

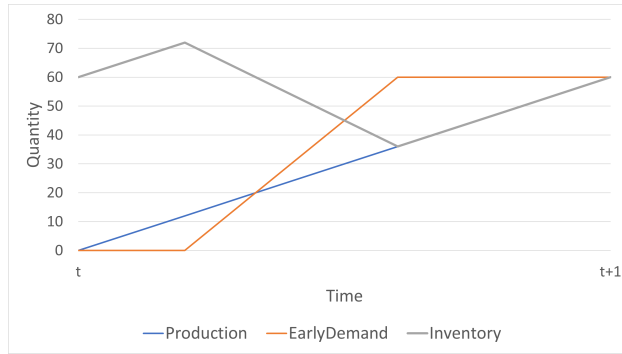


Figure 3.6: Case 1

Case 2. The production starts after the demand but before the demand is fully consumed ($o_t^b \leq s_t \leq o_t^b + \frac{d_t}{\alpha_t}$), see example of Figure [3.7](#).

In that case, the minimum inventory level is reached either (1) when the production starts, with an inventory of $I_{t-1} - \alpha_t(1 - \frac{L_t}{d_t})(s_t - o_t^b)$ or (2) when the demand is fully consumed, with an inventory of $I_{t-1} - d_t + L_t + \frac{X_t}{c_t^{max} - s_t} (o_t^b + \frac{d_t}{\alpha_t} - s_t)$. Thus to respect the minimum inventory constraint, two new constraints need to be added to the model:

$$I_{t-1} - \alpha_t(1 - \frac{L_t}{d_t})(s_t - o_t^b) \geq \underline{I}_t \quad (3.5)$$

$$I_{t-1} - d_t + L_t + \frac{X_t}{c_t^{max} - s_t} (o_t^b + \frac{d_t}{\alpha_t} - s_t) \geq \underline{I}_t \quad (3.6)$$

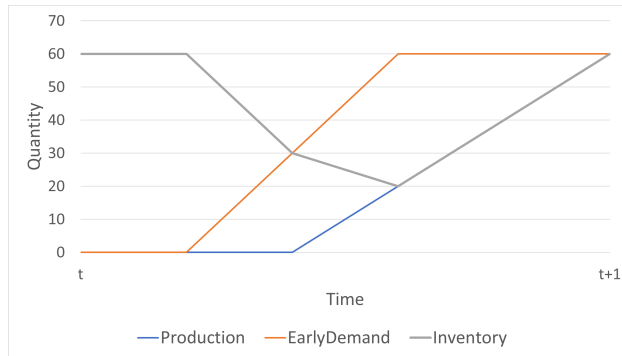


Figure 3.7: Case 2

Case 3. The production starts after the demand is fully consumed ($s_t \geq o_t^b + \frac{d_t}{\alpha_t}$), see example of Figure 3.8.

In that case, the minimum inventory level is reached when the demand is fully consumed, with an inventory of $I_{t-1} - d_t + L_t$. Thus, to respect the minimum inventory constraint, the following constraint needs to be added to the model:

$$I_{t-1} - d_t + L_t \geq \underline{I}_t \quad (3.7)$$

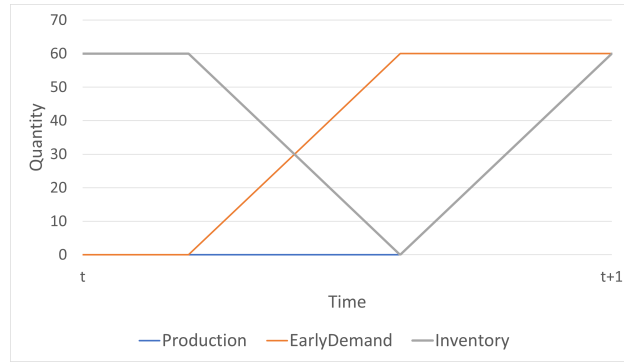


Figure 3.8: Case 3

Maximum inventory constraints

To consider the maximum inventory levels within periods, we only need to consider the slope for the demand approximation of latest demand. There is only one case where the maximum inventory level is not reached at the beginning or at the end of the period, and it leads to the addition of one new constraint to the model.

Case 1. The production starts before the demand ($s_t \leq o_t^e$), see example of Figure 3.9

In that case, the maximum inventory level is reached either right before the demand starts to be consumed with an inventory of $I_{t-1} + \frac{X_t}{c_t^{max} - s_t}(o_t^e - s_t)$, or at the end of the period. Thus, to respect the maximum inventory constraint, the following constraint needs to be added to the model:

$$I_{t-1} + \frac{X_t}{c_t^{max} - s_t}(o_t^e - s_t) \leq \bar{I}_t \quad (3.8)$$

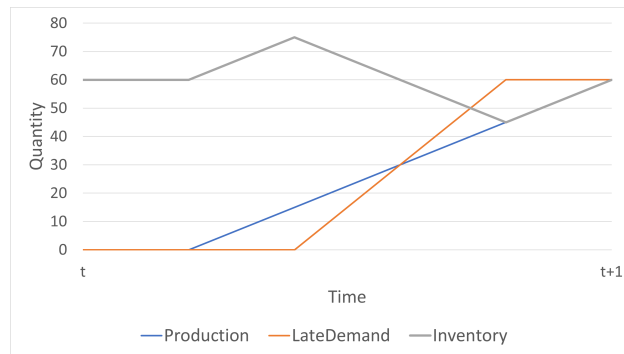


Figure 3.9: Case 1

3.3.3 Worst case errors

Under the assumptions made on the shapes of the demand and production rates, let us now consider the multi-item version of the problem. The capacity constraint at period t (Constraint (2.3)) is expressed as such:

$$\sum_{i=1}^N (s_{it}Y_{it} + b_{it}X_{it}) \leq c_t^{max}.$$

The unit production time b_{it} can be defined as the inverse of the production rate. The capacity constraint at period t implies that, after optimization, we can find a production schedule with a production at maximum rate that respects the length of the period. With the assumptions made in this section, the optimal production schedule found once the quantities to produce at each period are defined might be infeasible with regards to the minimum and maximum inventory bounds. However, we show in this section that adding these constraints enables to lower the worst case inventory error even when the shapes of the demand and production rates differ from the assumptions. Because the analysis is identical for each item, the index i has been removed in the rest of the section. In this analysis, we assume that the inventory at the beginning of period t respects the inventory bounds at period t : $\underline{I}_t \leq I_{t-1} \leq \bar{I}_t$.

Worst case for the minimum inventory error

Let us consider the maximum difference between the minimum inventory level \underline{I}_t and the minimum inventory level I_t^{min} that can be reached with a production at maximum capacity at time t . Ideally we want $I_t^{min} \geq \underline{I}_t$.

– **No inventory constraints within periods**

Without the inventory constraints, the inventory gap between \underline{I}_t and I_t^{min} is at most:

$$\underline{I}_t - I_t^{min} = \min(d_t, X_t).$$

Indeed, the gap can never exceed the value of the demand at period t , else the inventory constraint would not have been respected at the beginning of the period. The gap also cannot exceed the quantity produced at t , otherwise the inventory constraint would not be satisfied at the end of the period. The maximum gap occurs when the demand is fully consumed before the production starts in case $I_{t-1} = \underline{I}_t$. Considering the maximum quantity that can be produced $\frac{(c_t^{max} - s_t)}{b_t}$, we get:

$$\underline{I}_t - I_t^{min} = \min(d_t, \frac{(c_t^{max} - s_t)}{b_t}).$$

– **Inventory constraints within periods**

If we add the inventory constraints for the uniform production rate with setup times then, at period t , the maximum difference in minimum inventory occurs when the production at maximum rate occurs as late as possible. If quantity X_t is produced, then the time required to produce the quantity at maximum production rate is $b_t X_t$.

The worst case occurs when the production at maximum rate starts at $c_t^{max} - b_t X_t$. From s_t to $c_t^{max} - b_t X_t$, at a uniform production rate $\frac{X_t}{c_t^{max} - s_t}$ the quantity $\frac{X_t}{c_t^{max} - s_t} (c_t^{max} -$

$b_t X_t - s_t) = X_t(1 - \frac{b_t X_t}{(c_t^{max} - s_t)})$ is produced. Thus, we know that there is at most a difference of $X_t(1 - \frac{b_t X_t}{(c_t^{max} - s_t)})$ for the produced quantity between considering a production at maximum rate and a uniform production. The inventory constraints within periods guarantee that, for a uniform production rate the inventory cannot go below \underline{I}_t . The gap still cannot exceed the demand for the period, thus the previous result implies that the maximum inventory gap is:

$$\underline{I}_t - I_t^{min} = \min(d_t, X_t(1 - \frac{b_t X_t}{(c_t^{max} - s_t)})).$$

The function $f(x) = x(1 - \frac{b_t x}{(c_t^{max} - s_t)})$ for $x \in [0; \frac{(c_t^{max} - s_t)}{b_t}]$ reaches its maximum for $x^* = \frac{(c_t^{max} - s_t)}{2b_t}$, so that:

$$\underline{I}_t - I_t^{min} = \min(d_t, \frac{(c_t^{max} - s_t)}{4b_t}).$$

It is interesting to note that, by considering or not the inventory constraints within periods, the potential extreme points for the inventory gaps differ.

Worst case for the maximum inventory error

Let us consider the maximum difference between the maximum inventory level \bar{I}_t and the inventory level I_t^{max} that can be reached with a production at maximum capacity at time t . Ideally we want $I_t^{max} \leq \bar{I}_t$.

– No inventory constraints within periods

Without the inventory constraints, the maximum inventory gap between \bar{I}_t and I_t^{max} is:

$$I_t^{max} - \bar{I}_t = \min(d_t, X_t).$$

Indeed, the gap can never exceed the value of the demand at period t , else the inventory constraint would not be satisfied at the end of the period. The gap also cannot exceed the quantity produced at t , otherwise the inventory constraint would not have been respected at the beginning of the period. The maximum gap occurs when the production is fully completed before the demand starts to be consumed, in case $I_{t-1} = \bar{I}_t$. Considering the maximum quantity that can be produced $\frac{(c_t^{max} - s_t)}{b_t}$, we get:

$$I_t^{max} - \bar{I}_t = \min(d_t, \frac{(c_t^{max} - s_t)}{b_t}).$$

– Inventory constraints within periods

If we add the inventory constraints for the a uniform production rate and a production with setup times, then at period t the maximum difference in maximum inventory occurs when the production at maximum rate occurs right after the setup time. If quantity X_t is produced, then the time required to produce the quantity at maximum production rate is $b_t X_t$

The worst case occurs when the production at maximum rate starts at s_t . From s_t to the end of production $s_t + b_t X_t$, at a uniform production rate $\frac{X_t}{c_t^{max} - s_t}$ the quantity $\frac{X_t}{c_t^{max} - s_t} b_t X_t = \frac{b_t X_t^2}{X_t(c_t^{max} - s_t)}$ is produced. Thus we know that there is at most a difference of $X_t(1 - \frac{b_t X_t}{(c_t^{max} - s_t)})$ for the produced quantity between considering a production at maximum rate and a uniform production. The inventory constraints within periods guarantee that, for a uniform production rate, the inventory cannot go over \bar{I}_t . The gap still cannot exceed the demand for the period, thus the previous result implies that the maximum inventory gap is:

$$I_t^{max} - \bar{I}_t = \min(d_t, X_t(1 - \frac{b_t X_t}{(c_t^{max} - s_t)})).$$

For the same reasons as for the minimum inventory level, we have in the worst case that:

$$I_t^{max} - \bar{I}_t = \min(d_t, \frac{(c_t^{max} - s_t)}{4b_t}).$$

These results are especially relevant for small-bucket problems with sparse demands. If the production capacity at t is lower than the demand at the same period, it is possible to have a deficit or an excess of inventory which is equal to the entirety of the production during the period. We avoid this worst case with the additional inventory constraints, reducing the worst possible inventory deficit or excess by a factor of 4.

3.4 Production at maximum rate and instantaneous demand

In this section, we define a lot-sizing model based on assumptions on the production and demand rates that differ from the ones in Section 3.3. We introduce for each demand at period t a new parameter $t_{it}^d \leq c_t^{max}$ that corresponds to the time at which the demand occurs within the period. We also assume that the demand for item i at period t is instantaneous. This assumption can be reasonable for long periods. According to the model described in Section 3.2.1, each unit of item i at period t takes b_{it} to be produced. Consequently, in this section, we assume that the production of item i at period t occurs at a production rate of $\frac{1}{b_{it}}$. Under these assumptions, additional constraints to bound the inventory within periods are introduced in Section 3.4.1. In Section 3.4.2, we show that the problem introduced in the section can be polynomially solved once the capacity constraints are relaxed.

3.4.1 Modeling inventory constraints

We split each production decision variable X_{it} into two decision variables X_{it}^b and X_{it}^a that respectively represents the production before and after the demand occurs. We also assume that each production either ends strictly before or starts strictly after t_{it}^d . To that intent, we split each setup decision variable into two setup decision variables Y_{it}^b and Y_{it}^a .

$$Y_{it} = Y_{it}^b + Y_{it}^a, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.9)$$

$$X_{it} = X_{it}^b + X_{it}^a, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.10)$$

$$X_{it}^b \leq M_{it} Y_{it}^b, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.11)$$

$$X_{it}^a \leq M_{it} Y_{it}^a, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.12)$$

$$\sum_{j, t_{jt}^d \leq t_{it}^d} (b_{jt} X_{jt}^b + s_{jt} Y_{jt}^b) \leq t_{it}^d, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.13)$$

$$\sum_{j, t_{jt}^d \geq t_{it}^d} (b_{jt} X_{jt}^a + s_{jt} Y_{jt}^a) \leq c_t^{max} - t_{it}^d, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.14)$$

$$I_{i,t-1} + X_{it}^b \leq \bar{I}_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.15)$$

$$I_{i,t-1} + X_{it}^b - d_{it} \geq \underline{I}_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (3.16)$$

Constraints (3.9) imply that a setup induces a production either before or after the demand occurs. Constraints (3.10) link the production variables before and after the demand to the production quantity in the period. Constraints (3.11) and (3.12) link the production and the setup variables. Constraints (3.13) (resp. (3.14)) state that, for a given demand for item i at period t , the sum of all setup and production times for all demands occurring before (resp. after) t_{it}^d must be lower than t_{it}^d (resp. the remaining time $c_t^{max} - t_{it}^d$).

The advantage of this formulation is that it takes into account all previous production quantities when deciding on the quantity to produce for each item. A good scheduling algorithm minimizing the excess of inventories should find better solutions with these additional constraints without affecting the quality of the solution. This model is less constraining than the one assuming that the production is uniform throughout the period, which can restrict the feasibility domain, especially for longer periods.

3.4.2 The uncapacitated lot-sizing problem with constant inventory bounds and inventory constraints within periods

Let us consider the problem defined in Section 3.4.1 where all capacity constraints (Constraints (3.13), Constraints (3.14) and Constraints (2.3)) are removed. Because the capacity constraints are relaxed, the initial problem can be decomposed into independent uncapacitated single-item problems. Lost sales are also not allowed and we assume that the inventory bounds are constant. We show in this section that this problem can be polynomially solved. This result could provide insights in order to develop a heuristic based on the relaxation of the capacity constraints of the problem. Let us denote by \underline{I} and \bar{I} the minimum and maximum inventory bounds.

Let us consider 4 inventory values at the end of each period t :

- (1) $I_t^1 = \bar{I}$,
- (2) $I_t^2 = \underline{I}$,
- (3) $I_t^3 = \bar{I} - d_t$,
- (4) $I_t^4 = \underline{I} + d_{t+1}$.

We refer to these values by the term inventory points of type $i \in \{1, 2, 3, 4\}$ in the rest of the section, and assume that the initial inventory is an inventory point. We refer to a period with an inventory point as an inventory period.

Theorem 4. *There is an optimal solution such that, between two successive production periods j and k with $j < k$, there is at least one inventory period l with $j \leq l < k$.*

Proof. Let us define $c_j = p_j + \sum_{m=j}^{k-1} h_m$ and $c_k = p_k$. We need to analyze four different cases. We assume that for all l such that $j \leq l < k$, I_l is not an inventory point.

Case 1. $X_j = X_j^a$ and $X_k = X_k^b$.

We have $I_j < \bar{I}$ and $I_{k-1} > \underline{I}$.

- If $c_j < c_k$, then $\delta = \min(X_k, \min_{j \leq m < k}(\bar{I} - I_m)) = \min(X_k, \bar{I} - I_j)$ units can be moved from period k to period j without affecting the feasibility of the solution. This new solution, with either a new inventory at period j of $I'_j = \bar{I}$ or a new production quantity at period k of $X'_k = 0$, improves the objective function by at least $\delta(c_k - c_j) \geq 0$.
- If $c_j \geq c_k$, then $\delta = \min(X_j, \min_{j \leq m < k}(I_m - \underline{I})) = \min(X_j, I_{k-1} - \underline{I})$ units can be moved from period j to period k without affecting the feasibility of the solution. This new solution, with either a new inventory at period $k - 1$ of $I'_{k-1} = \underline{I}$ or a new production quantity at period j of $X'_j = 0$, improves the objective function by at least $\delta(c_j - c_k) \geq 0$.

Case 2. $X_j = X_j^b$ and $X_k = X_k^b$.

In that case, the inventory right before time t_j^d is $I_j + d_j < \bar{I}$. We have $I_j < \bar{I} - d_j$ and $I_{k-1} > \underline{I}$.

- If $c_j < c_k$, then $\delta = \min(X_k, \bar{I} - d_j - I_j, \min_{j \leq m < k}(\bar{I} - I_m)) = \min(X_k, \bar{I} - d_j - I_j)$ units can be moved from period k to period j without affecting the feasibility of the solution. This new solution, with either a new inventory at period j of $I'_j = \bar{I} - d_j$ or a new production quantity at period k of $X'_k = 0$, improves the objective function by at least $\delta(c_k - c_j) \geq 0$.
- If $c_j \geq c_k$, then $\delta = \min(X_j, \min_{j \leq m < k}(I_m - \underline{I})) = \min(X_j, I_{k-1} - \underline{I})$ units can be moved from period j to period k without affecting the feasibility of the solution. This new solution, with either a new inventory at period $k - 1$ of $I'_{k-1} = \underline{I}$ or a new production quantity at period j of $X'_j = 0$, improves the objective function by at least $\delta(c_j - c_k) \geq 0$.

Case 3. $X_j = X_j^b$ and $X_k = X_k^a$.

In that case, the inventory right after t_k^d is $I_{k-1} - d_k > \underline{I}$. We have $I_j < \bar{I} - d_j$ and $I_{k-1} > \underline{I} + d_k$.

- If $c_j < c_k$, then $\delta = \min(X_k, \bar{I} - d_j - I_j, \min_{j \leq m < k}(\bar{I} - I_m)) = \min(X_k, \bar{I} - d_j - I_j)$ units can be moved from period k to period j without affecting the feasibility of the solution. This new solution, with either a new inventory at period j of $I'_j = \bar{I} - d_j$ or a new production quantity at period k of $X'_k = 0$, improves the objective function by at least $\delta(c_k - c_j) \geq 0$.

- If $c_j \geq c_k$, then $\delta = \min(X_j, I_{k-1} - d_k - \underline{I}, \min_{j \leq m < k} (I_m - \underline{I})) = \min(X_j, I_{k-1} - d_k - \underline{I})$ units can be moved from period j to period k without affecting the feasibility of the solution. This new solution, with either a new inventory at period $k - 1$ of $I'_{k-1} = \underline{I} + d_k$ or a new production quantity at period j of $X'_j = 0$, improves the objective function by at least $\delta(c_j - c_k) \geq 0$.

Case 4. $X_j = X_j^a$ and $X_k = X_k^a$.

In that case, the inventory right after t_k^d is $I_{k-1} - d_k > \underline{I}$. We have $I_j < \bar{I}$ and $I_{k-1} > \underline{I} + d_k$.

- If $c_j < c_k$, then $\delta = \min(X_k, \min_{j \leq m < k} (\bar{I} - I_m)) = \min(X_k, \bar{I} - I_j)$ units can be moved from period k to period j without affecting the feasibility of the solution. This new solution, with either a new inventory at period j of $I'_j = \bar{I}$ or a new production quantity at period k of $X'_k = 0$, improves the objective function by at least $\delta(c_k - c_j) \geq 0$.
- If $c_j \geq c_k$, then $\delta = \min(X_j, I_{k-1} - d_k - \underline{I}, \min_{j \leq m < k} (I_m - \underline{I})) = \min(X_j, I_{k-1} - d_k - \underline{I})$ units can be moved from period j to period k without affecting the feasibility of the solution. This new solution, with either a new inventory at period $k - 1$ of $I'_{k-1} = \underline{I} + d_k$ or a new production quantity at period j of $X'_j = 0$, improves the objective function by at least $\delta(c_j - c_k) \geq 0$.

The proof also implies that we can find an optimal solution where either period j or period $k - 1$ is an inventory period. \square

Corollary 1. *There is an optimal solution such that, between two successive inventory periods j and k with $j < k$, there is at most one production period l such that $j < l \leq k$.*

Proof. If we have two successive production periods l_1 and l_2 with $j < l_1 < l_2 \leq k$ then, according to Theorem 4 we can find an optimal solution where either l_1 or $l_2 - 1$ is an inventory period. Both cases contradict the fact that periods j and k are successive inventory periods. \square

From Corollary 1 we can derive a polynomial dynamic programming algorithm to solve the relaxed problem.

Dynamic programming algorithm

Let us denote $C(j, i)$ the minimum cost from period $j + 1$ to period T where j is an inventory period of type $i \in \{1, 2, 3, 4\}$.

There is an optimal solution where the inventory at T is an inventory point of type 2. Indeed, if there is a solution with an ending inventory level $I_T > \underline{I}$, the production quantity in the last production period $k \leq T$ can be decreased such that $\min(X_k, I_T - \underline{I}) = 0$.

We have $C(T, 2) = 0$.

Corollary 1 provides the following recursion:

$$C(j, i) = \min_{\{(k, i') | k > j; i' \in \{1, 2, 3, 4\}\}} G(j, k, i, i') + C(k, i'),$$

where $G(j, k, i, i')$ corresponds to the minimum cost to go from inventory period j of type i to inventory period k of type i' with at most one production period $j < l \leq k$, i.e.

$$G(j, k, i, i') = \min_{j < l \leq k} [f_l + p_l(I_k^{i'} - I_j^i + D_{j+1,k}) + \sum_{m=j+1}^{l-1} h_m(I_j^i - D_{j+1,m}) + \sum_{m=l}^k h_m(I_k^{i'} + D_{m+1,k})],$$

where $D_{jk} = \sum_{l=j}^k d_l$, $D_{jk} = 0$ if $j > k$.

If l is a production period, then either $X_l = X_l^a$ or $X_l = X_l^b$. This implies that, in a feasible solution, either $(I_j^i - D_{j+1,l}) \geq \underline{I}$ or $(I_k^{i'} + D_{lk}) \leq \bar{I}$.

Let us only consider feasible production plans from j to k with production at period l . Hence, if $(I_k^{i'} - I_j^i + D_{j+1,k}) < 0$, $(I_j^i - D_{j+1,l-1}) < \underline{I}$ or $(I_k^{i'} + D_{l+1,k}) > \bar{I}$, we set a value of ∞ for this value of l during the computation of $G(j, k, i, i')$. We also set a value of ∞ if both $(I_j^i - D_{j+1,l}) < \underline{I}$ and $(I_k^{i'} + D_{lk}) > \bar{I}$.

If we assume that all sums are precomputed (which can be done in $O(T^2)$), then each computation of $G(j, k, *, *)$, for fixed j and k , can be done in $O(T)$. Consequently, if we compute $C(j, *)$ in decreasing order of j , each $C(j, *)$ can be computed in at most $O(T^2)$. The optimal value of the problem considered in this section corresponds to $C(0, i)$, where i is the type of the initial inventory. $C(j, *)$ needs to be computed for $j \in \{T-1, \dots, 0\}$, which implies that the complexity of the algorithm is $O(T^3)$.

3.5 Computational experiments

Computational experiments were performed on 180 instances based on the instances of [Trigeiro et al. \(1989\)](#) with $N = 30$ items and $T = 20$ periods modified according to the generation scheme of Section 2.5.2 of Chapter 2. We set a lower bound for the inventory of each item i at each period t of $\underline{I}_{it} = 0$, and an upper bound equal to the minimum between the maximum ending inventory and twice the average demand per period for each item. All mixed-integer linear programs were solved using IBM ILOG CPLEX 12.10 with the solver default settings. A time limit for each run is set to 600 seconds. We denote by deficit (resp. surplus) the units below (resp. above) the minimum (resp. maximum) inventory value at each period for each item. After each run, we have computed the deficit, surplus and lost sales obtained after reconstructing a production schedule. The results are reported in Tables 3.1, 3.2 and 3.3 that show the ratio of the obtained values and the total demand. We sort the instances according to their TBO defined by $\sqrt{2 \frac{\bar{f}}{\bar{h}d}}$. We also report the average relative optimality gap, denoted Gap , for each set of instances, that corresponds to the relative difference between the best upper bound UB and the best lower bound LB obtained by CPLEX: $Gap = \frac{UB-LB}{UB}$.

3.5.1 Planning reconstruction

Using the information provided by solving the lot-sizing problem on quantities to produce at each period, we can reconstruct a production plan that respects as much as possible the inventory constraints by considering that the production for item i at period t occurs at rate $\frac{1}{b_{it}}$ after a setup time of s_{it} . This assumption makes sense with regards to the capacity constraints of the CLSP, which allow a feasible production plan to be determined.

3.5.2 Analysis

Let us first consider two settings in which we assume that we know exactly when the demands occur within the periods.

1. In the first setting, all demands are instantaneous and randomly generated within the period.
2. In the second setting, all demands are instantaneous and randomly generated within the first half of the period.

Let us denote by $M0$ the original model in Section 3.2.1, and $M1$ (resp. $M2$) the model with the additional constraints introduced in Section 3.3 (resp. Section 3.4).

Model		Deficit (%)			Surplus (%)			Lost Sales (%)			Gap (%)		
		M0	M1	M2	M0	M1	M2	M0	M1	M2	M0	M1	M2
TBO	1	11.30	2.06	0.91	0.65	2.68	0.06	0.00	2.39	0.88	0.88	0.58	1.65
	2	7.88	3.00	0.32	3.04	3.07	0.01	0.02	1.50	1.94	2.76	1.05	9.94
	4	3.94	2.62	0.01	3.82	2.92	0.00	0.16	0.58	1.72	2.34	3.30	18.56
Average		7.71	2.56	0.41	2.50	2.89	0.02	0.06	1.49	1.51	1.99	1.65	10.05

Table 3.1: Comparison between models for randomly generated demands (maximum CPU time of 600 seconds)

Table 3.1 shows that, for the original model $M0$, the lost sales are very small (always less than 1%), the inventory surplus or deficit can be significant, especially for small values of TBO (11.30% of deficit on average for TBO=1). Model $M1$ has larger lost sales (2.39% for instances with TBO=1), but the fraction of deficit and surplus is greatly reduced compared to $M0$. The optimality gaps reported for $M0$ and $M1$ are equivalent. Model $M2$, being more precise, performs much better than $M0$ with regards to the inventory deficit and surplus (from 3.94% to 0.01% for the inventory deficit for the instances with TBO=4). The lost sales are greater for TBO=4 compared to $M1$, but the total loss (deficit, surplus and lost sales) is much smaller. These inventory excesses can be considered as lost sales for strict inventory constraints. However, because $M2$ is more precise and $2N$ new variables have been added for each period, the optimality gaps are larger within the same CPU time limit of 600 seconds. By setting a larger CPU time limit, a lost sales decrease can be reached.

Model		Deficit (%)			Surplus (%)			Lost Sales (%)			Gap (%)		
		M0	M1	M2	M0	M1	M2	M0	M1	M2	M0	M1	M2
TBO	1	43.99	10.48	3.65	0.25	0.93	0.00	0.00	3.15	2.56	0.78	0.67	1.97
	2	24.46	10.12	1.59	0.95	1.25	0.00	0.02	2.06	2.85	2.67	0.98	8.24
	4	9.86	6.47	0.09	0.81	1.14	0.00	0.16	0.68	1.83	2.39	2.92	21.86
Average		26.11	9.03	1.77	0.67	1.11	0.00	0.06	1.96	2.42	1.95	1.52	10.69

Table 3.2: Comparison between models for randomly generated demands during the first half of the period (maximum CPU time of 600 seconds)

In the specific case where demands occur at the beginning of the period, not considering the dynamic evolution of the inventory may lead to inventory management issues because

decision making does not take into account the inventory drop at the beginning of the period. Table 3.2 shows the results obtained when all demands occur during the first half of each period. In this case, we notice a great increase in deficit for $M0$. Indeed, if we take the case of $TBO=1$ (43.99%), without the additional inventory constraints, the initial inventory level at each period will most likely be close to 0, meaning that if the production occurs after the demand, then there will be a negative inventory. The deficit increases for $M1$ and $M2$, but they still comparatively perform much better than the original model. $M2$ still has the best performances.

Model		Deficit (%)			Surplus (%)			Lost Sales (%)			Gap (%)		
		M0	M1	M2	M0	M1	M2	M0	M1	M2	M0	M1	M2
TBO	1	87.53	0.00	0.00	0.00	0.00	0.00	0.00	4.97	5.06	0.70	0.68	0.81
	2	49.47	0.00	0.00	0.00	0.00	0.00	0.02	4.67	5.30	2.29	1.06	1.61
	4	20.85	0.00	0.00	0.00	0.00	0.00	0.16	2.18	3.08	2.43	2.59	3.94
Average		52.61	0.00	0.00	0.00	0.00	0.00	0.06	3.94	4.48	1.81	1.44	2.12

Table 3.3: Comparison between models for instantaneous demand at the beginning of the period (maximum CPU time of 600 seconds)

Model		Deficit (%)			Surplus (%)			Lost Sales (%)			Gap (%)		
		M0	M1	M2	M0	M1	M2	M0	M1	M2	M0	M1	M2
TBO	1	0.00	0.00	0.00	3.03	0.00	0.00	0.00	0.35	0.44	0.77	0.83	3.47
	2	0.00	0.00	0.00	13.79	0.00	0.00	0.02	1.87	2.60	2.35	1.89	3.04
	4	0.00	0.00	0.00	20.19	0.00	0.00	0.16	2.12	3.07	2.44	7.93	6.45
Average		0.00	0.00	0.00	12.34	0.00	0.00	0.06	1.45	2.04	1.85	3.55	4.32

Table 3.4: Comparison between models for instantaneous demand at the end of the period (maximum CPU time of 600 seconds)

Table 3.3 (resp. Table 3.4) reports the fraction of deficit, surplus and lost sales in the extreme case where all demands are instantaneous and occur at the beginning (resp. the end) of the period. When demands occur at the beginning of the period, if the inventory at the beginning of the period is smaller than the demand at this period, the difference will be lost. When demands occur at the end of the period, if the inventory at the beginning of the period plus the production during the period exceed the maximum inventory, the difference will be lost. Because $M0$ does not consider the inventory evolution within the periods, this model performs the worse. For a demand at the beginning of the period, the worst cases occur for instances with a small TBO where the initial inventory at each period is close to 0 (87.53% of deficit for the instances with $TBO=1$). For a demand at the end of the period, the worst cases occur for instances with a high TBO where the production quantities per item at each period are larger (20.19% of surplus for the instances with $TBO=4$). Both $M1$ and $M2$ handle the case where all demands occur at the beginning (resp. the end) of each period. For these two models, we can find production plans without inventory deficit or surplus. With a maximum CPU time of 600 seconds, the average optimality gaps of all models are close (1.44% for $M1$ to 2.12% for $M2$) when demands occur at the beginning of the period. In that case, $M1$ and $M2$ provide very similar results, with solutions for $M2$

with slightly higher lost sales (2.18% for $M1$ to 3.08% for $M2$ of lost sales for the instances with $TBO=4$).

3.6 Additional studies

Several extensions of the modeling of production and demand within periods can be considered. In this section, we propose additional models that could be used within this framework for further research. The models introduced in Sections [3.6.2](#) and [3.6.3](#) were not implemented and tested with computational experiments.

3.6.1 Backlogs

Since we consider when demand occurs within each period, adding backlogs to the problem is a natural extension to our models. Instead of being considered as lost sales, a demand not satisfied at the end of a period can be satisfied with production quantities in following periods assuming a penalty cost ([Zangwill \(1966\)](#)). To model backlogs in our problem, we introduce the variable $B_{it} \geq 0$ which is the quantity of demand of item i that is backlogged at the end of period t . Each unit of backlogs of item i at period t is penalized by a coefficient h_{it}^- . Each unit of backlogs at period T is considered as a lost sale and is highly penalized in the objective function. The extension of model $M0$ with backlogs is defined as follows:

$$\min \sum_{i=1}^N \sum_{t=1}^T (f_{it}Y_{it} + p_{it}X_{it} + h_{it}I_{it} + h_{it}^-B_{it}) \quad (3.17)$$

$$I_{i,t-1} + X_{it} + B_{it} = d_{it} + I_{it} + B_{i,t-1}, \quad \forall i, t \quad (3.18)$$

$$B_{it} - B_{i,t-1} \leq d_{it}, \quad \forall i, t \quad (3.19)$$

$$X_{it}, I_{it}, B_{it} \geq 0, \quad \forall i, t \quad (3.20)$$

$$\text{\a href="#">(2.3), \a href="#">(2.4), \a href="#">(2.6), \a href="#">(3.2), \a href="#">(3.3)}$$

Backlog costs have been added in the objective function ([3.17](#)). In the flow conservation constraints ([3.18](#)), the lost sales variables L_{it} of Constraints ([2.2](#)) are replaced with the difference $B_{it} - B_{i,t-1}$, which corresponds to the new backlogs added at the end of period t for item i . Constraints ([3.19](#)) state that we cannot add to the backlogs a quantity larger than the demand for item i at period t . We can extend models $M1$ and $M2$ by considering backlogs instead of lost sales under the assumption that, similarly to the lost sales, backlogs do not change the time interval during which each satisfied demand is consumed. Under this assumption, we simply need to replace L_t by $B_{it} - B_{i,t-1}$ in all the inventory constraints introduced in Sections [3.3](#) and [3.4](#).

Computational experiments

Table [3.5](#) reports the fraction of deficit, surplus, lost sales and backlogs when solving the instances introduced in Section [3.5](#) with a time limit of 600 seconds and with backlog costs $h_{it}^- = 5h_{it}$.

We first notice, by comparing Table [3.5](#) and Table [3.1](#) that, for all models we get similar results with regards to the respect of the inventory constraints within periods whether backlogs are allowed or not. However some of the lost sales become backlogs. For model $M1$,

Model		Deficit (%)			Surplus (%)			Lost Sales (%)			Backlogs (%)			Gap (%)		
		M0	M1	M2	M0	M1	M2	M0	M1	M2	M0	M1	M2	M0	M1	M2
TBO	1	11.18	2.29	0.89	0.67	3.47	0.02	0.00	0.03	0.09	0.05	4.88	2.57	0.75	1.60	3.60
	2	8.24	3.59	0.19	3.52	4.70	0.01	0.00	0.06	0.71	0.65	5.23	11.31	1.56	3.52	20.35
	4	4.71	3.97	0.00	4.68	4.18	0.00	0.01	0.10	1.42	2.18	4.66	8.86	2.84	16.34	19.34
Average		8.04	3.28	0.36	2.96	4.11	0.01	0.00	0.06	0.74	0.96	4.93	7.58	1.71	7.15	14.43

Table 3.5: Comparison between models for randomly generated demands and backlogs (maximum CPU time of 600 seconds)

the average lost sales drop from 1.49% to 0.06% when backlogs are allowed. The problem with backlogs seems to be computationally harder to solve, especially for $M1$, where the average optimality gap increases from 1.44% to 7.15%. $M2$ has the most backlogged demand (7.58% on average), but it still outperforms $M0$ and $M1$ when considering the inventory management within periods, with no inventory excess and limited lost sales (1.42%) for the instances with $TBO=4$.

3.6.2 Production and demand at maximum rates

Let us consider a single-item lot-sizing problem with inventory bounds, and consider the case where production (resp. demand) occurs with a production (resp. demand) time per unit of b_t (resp. b_t^d). We know the time t^d at which the demand starts being consumed. We can model the fact that we want to avoid the worst case for the respect of both the inventory lower bound I_t and upper bound \bar{I}_t within period t .

This model guarantees that the inventory constraints will be respected independently of when production and demand occur within the period. This model can be useful for problems where robustness is required, e.g. when there is uncertainty in the value of t^d or if being able to find a production schedule respecting the inventory constraints is mandatory. These constraints are less constraining than considering that all production quantities has to be completed before the demand occurs or that the demand is fully consumed before the production starts, which is what is usually done to guarantee the respect of the inventory constraints. The additional constraints can be found below. Because these constraints are not linear, we propose a way to linearize them.

Minimum inventory constraints

The worst case occurs when production ends right at the end of the period. In that case, production starts at $c_t^{max} - b_t X_t$. We need to consider three cases:

Case 1. Demand ends before production starts, i.e. $t^d + b_t^d d_t \leq c_t^{max} - b_t X_t$. The extremum is reached at time $t^d + b_t^d d_t$, when d_t units of demand have been consumed and nothing has been produced. The following constraint needs to be added to the model:

$$I_{t-1} - d_t \geq \underline{I}_t$$

Case 2. Demand starts before production starts ($t^d \leq c_t^{max} - b_t X_t$) and ends after production starts ($t^d + b_t^d d_t \geq c_t^{max} - b_t X_t$). Two cases need to be considered.

(a) The extremum is reached at time $c_t^{max} - b_t X_t$ right before production starts:

$$I_{t-1} - \frac{(c_t^{max} - b_t X_t - t^d)}{b_t^d} \geq \underline{I}_t \quad (3.21)$$

(b) The extremum is reached at time $t^d + b_t^d d_t$, when demand is fully consumed:

$$I_{t-1} + \left(\frac{b_t^d}{b_t} - 1\right)d_t - \frac{(c_t^{max} - b_t X_t - t^d)}{b_t} \geq \underline{I}_t \quad (3.22)$$

Case 3. Demand starts after production starts, i.e. $t^d \geq c_t^{max} - b_t X_t$. The extremum is reached at time $t^d + b_t^d d_t$ when demand is fully consumed

$$I_{t-1} - d_t + \frac{b_t^d}{b_t} d_t - \frac{(c_t^{max} - b_t X_t - t^d)}{b_t} \geq \underline{I}_t \quad (3.23)$$

If we define $\underline{\gamma}_t = \min(d_t, \frac{(c_t^{max} - b_t X_t - t^d)}{b_t^d})$, Constraints (3.21)-(3.23) can be reduced into these two constraints:

$$I_{t-1} - \underline{\gamma}_t \geq \underline{I}_t \quad (3.24)$$

$$I_{t-1} - \left(1 - \frac{b_t^d}{b_t}\right)d_t - \frac{b_t^d}{b_t} \underline{\gamma}_t \geq \underline{I}_t \quad (3.25)$$

Maximum inventory constraints

The worst case occurs when the production starts at the beginning of the period. We have three cases:

Case 1. Production ends before demands starts, i.e. $s_t + b_t X_t \leq t^d$. The extremum is reached at time $s_t + b_t X_t$, at the end of the production when X_t units have been produced:

$$I_{t-1} + X_t \leq \bar{I}_t \quad (3.26)$$

Case 2. Production ends after demand starts, i.e. $s_t + b_t X_t \geq t^d$, and production starts before demand starts, i.e. $s_t \leq t^d$:

(a) The extremum is reached at time t^d , right before the demand starts being consumed:

$$I_{t-1} + \frac{(t^d - s_t)}{b_t} \leq \bar{I}_t \quad (3.27)$$

(b) The extremum is reached at time $s_t + b_t X_t$, at the end of the production:

$$I_{t-1} + X_t - \frac{b_t}{b_t^d} X_t - \frac{(s_t - t^d)}{b_t^d} \leq \bar{I}_t \quad (3.28)$$

Case 3. Production starts after demand starts, i.e. $s_t \geq t^d$. The extremum is reached at the end of the production:

$$I_{t-1} + X_t - \frac{b_t}{b_t^d} X_t - \frac{(s_t - t^d)}{b_t^d} \leq \bar{I}_t \quad (3.29)$$

If we define $\bar{\gamma}_t = \min(X_t, \frac{(t^d - s_t)}{b_t})$, Constraints (3.26)-(3.29) can be reduced into these two constraints:

$$I_{t-1} + \bar{\gamma}_t \leq \bar{I}_t \quad (3.30)$$

$$I_{t-1} + \left(1 - \frac{b_t}{b_t^d}\right)X_t + \frac{b_t}{b_t^d} \bar{\gamma}_t \leq \bar{I}_t \quad (3.31)$$

Linearization

For each item, we only need to add two constraints per period and per item to model the worst case for the minimum (resp. maximum) inventory. Constraints (3.24), (3.25), (3.30) and (3.31) are not linear because $\underline{\gamma}_t$ (resp. $\overline{\gamma}_t$) is not linear. However, we can use additional boolean variables to linearize a minimum between two variables. Let us define $\gamma_t^b \in \{0, 1\}$ such that, for the minimum inventory constraints, we add the following constraints to the model:

$$\underline{\gamma}_t \leq d_t \quad (3.32)$$

$$\underline{\gamma}_t \leq \frac{(c_t^{max} - b_t X_t - t^d)}{b_t^d} \quad (3.33)$$

$$\underline{\gamma}_t \geq d_t - \frac{c_t^{max}}{b_t^d} \gamma_t^b \quad (3.34)$$

$$\underline{\gamma}_t \geq \frac{(c_t^{max} - b_t X_t - t^d)}{b_t^d} - \frac{c_t^{max}}{b_t^d} (1 - \gamma_t^b) \quad (3.35)$$

and remove the constraint $\underline{\gamma}_t = \min(d_t, \frac{(c_t^{max} - b_t X_t - t^d)}{b_t^d})$.

The same process can be applied to $\overline{\gamma}_t$ by adding the variables $\overline{\gamma}_t^b \in \{0, 1\}$.

3.6.3 Joint storage capacity

Another interesting extension is to consider the case where the inventory bounds are on the total joint inventory of all items combined (Akbalik et al. (2015), Melo and Ribeiro (2017)). It makes sense when considering upper bounds on the inventories when the storage location has a limited capacity. Because, in both models considered in this chapter, the inventory evolution for each item is piecewise linear, the extremum of the total inventory corresponds to one of the extrema of the individual inventory. We have already defined these extreme points in the previous sections and can evaluate the inventory at each of these points. In this section, we consider the extension of the model defined in Section 3.3 where the joint inventory of all items is bounded. Let us consider a specific item i , in the following we provide linear constraints for each period to bound the minimum and maximum joint inventories.

Minimum joint inventory constraints

Case 1. Production starts before demand, i.e. $s_{it} \leq o_{it}^b$. The extremum is reached at time $o_{it}^b + \frac{d_{it}}{\alpha_{it}}$. For $j \neq i$, the inventory at time $o_{it}^b + \frac{d_{it}}{\alpha_{it}}$ that we denote I_{jt}^i is:

$$\underline{I}_{jt}^i = I_{j,t-1} + \delta_{\{o_{it}^b + \frac{d_{it}}{\alpha_{it}} \geq s_{jt}\}} \frac{X_{jt}}{c_t^{max} - s_{jt}} (s_{it} - s_{jt}) - \delta_{\{o_{it}^b + \frac{d_{it}}{\alpha_{it}} \geq o_{jt}^b\}} (d_{jt} - L_{jt}) \min(1, \frac{\alpha_{jt}}{d_{jt}} (s_{it} - o_{jt}^b))$$

The additional constraint to be added is:

$$I_{i,t-1} - d_{it} + L_{it} + \frac{X_{it}}{c_t^{max} - s_{it}} (o_{it}^b + \frac{d_{it}}{\alpha_{it}} - s_{it}) + \sum_{j \neq i} \underline{I}_{jt}^i \geq \underline{I}_t \quad (3.36)$$

Case 2. Production starts before demand is fully consumed, i.e. $o_{it}^b \leq s_{it} \leq o_{it}^b + \frac{d_{it}}{\alpha_{it}}$.

(a) The extremum is reached at time s_{it}^b . For $j \neq i$, the inventory at time s_{it}^b is:

$$\underline{I}_{jt}^i = I_{j,t-1} + \delta_{\{s_{it} \geq s_{jt}\}} \frac{X_{jt}}{c_t^{max} - s_{jt}} (s_{it} - s_{jt}) - \delta_{\{s_{it} \geq o_{jt}^b\}} \min(d_{jt} - L_{jt}, \alpha_{jt} (1 - \frac{L_{jt}}{d_{jt}})) (s_{it} - o_{jt}^b)$$

The additional constraint to be added is:

$$I_{i,t-1} - \alpha_{it} (1 - \frac{L_{it}}{d_{it}}) (s_{it} - o_{it}^b) + \sum_{j \neq i} \underline{I}_{jt}^i \geq \underline{I}_t \quad (3.37)$$

(b) The extremum is reached at time $o_{it}^b + \frac{d_{it}}{\alpha_{it}}$. For $j \neq i$, the inventory at time $o_{it}^b + \frac{d_{it}}{\alpha_{it}}$ is:

$$\underline{I}_{jt}^i = I_{j,t-1} + \delta_{\{o_{it}^b + \frac{d_{it}}{\alpha_{it}} \geq s_{jt}\}} \frac{X_{jt}}{c_t^{max} - s_{jt}} (s_{it} - s_{jt}) - \delta_{\{o_{it}^b + \frac{d_{it}}{\alpha_{it}} \geq o_{jt}^b\}} (d_{jt} - L_{jt}) \min(1, \frac{\alpha_{jt}}{d_{jt}} (s_{it} - o_{jt}^b))$$

The additional constraint to be added is:

$$I_{i,t-1} - d_{it} + L_{it} + \frac{X_{it}}{c_t^{max} - s_{it}} (o_{it}^b + \frac{d_{it}}{\alpha_{it}} - s_{it}) + \sum_{j \neq i} \underline{I}_{jt}^i \geq \underline{I}_t \quad (3.38)$$

Case 3. Production starts after demand is fully consumed, i.e. $s_{it} \geq o_{it}^b + \frac{d_{it}}{\alpha_{it}}$. The extremum is reached at time $o_{it}^b + \frac{d_{it}}{\alpha_{it}}$. For $j \neq i$, the inventory at time $o_{it}^b + \frac{d_{it}}{\alpha_{it}}$ is:

$$\underline{I}_{jt}^i = I_{j,t-1} + \delta_{\{o_{it}^b + \frac{d_{it}}{\alpha_{it}} \geq s_{jt}\}} \frac{X_{jt}}{c_t^{max} - s_{jt}} (s_{it} - s_{jt}) - \delta_{\{o_{it}^b + \frac{d_{it}}{\alpha_{it}} \geq o_{jt}^b\}} (d_{jt} - L_{jt}) \min(1, \frac{\alpha_{jt}}{d_{jt}} (s_{it} - o_{jt}^b))$$

The additional constraint to be added is:

$$I_{i,t-1} - d_{it} + L_{it} + \sum_{j \neq i} \underline{I}_{jt}^i \geq \underline{I}_t \quad (3.39)$$

Maximum joint inventory constraints

Case 1. Production starts before demand, i.e. $s_{it} \leq o_{it}^e$. The extremum is reached at o_{it}^e . For $j \neq i$, the inventory at o_{it}^e is:

$$\overline{I}_{jt}^i = I_{j,t-1} + \delta_{\{s_{jt} \leq o_{it}^e\}} \frac{X_{jt}}{c_t^{max} - s_{jt}} (o_{it}^e - s_{jt}) - \delta_{\{o_{jt}^e \leq o_{it}^e\}} \beta_{jt} (1 - \frac{L_{jt}}{d_{jt}}) (o_{it}^e - o_{jt}^e)$$

The additional constraint that needs to be added is:

$$I_{i,t-1} + \frac{X_{it}}{c_t^{max} - s_{it}} (o_{it}^e - s_{it}) + \sum_{j \neq i} \overline{I}_{jt}^i \leq \overline{I}_t \quad (3.40)$$

3.7 Conclusions

In this chapter, we have first shown that there are incentives to consider the evolution of inventory within each period when solving lot-sizing problems, even when these problems are based on a time discretization of the planning horizon. Assuming that the production and the demand are both instantaneous and simultaneous is a strong assumption that is implicitly made when modeling inventory bounds in lot-sizing models. This assumption can lead to operational production schedules that do not respect the inventory bounds. We have proposed two ways to model and bound the maximum (resp. minimum) inventory level reached for each item at each period under specific hypotheses on the shapes of the production and demand rates. These models guarantee that a production plan that respects inventory bounds can be determined. We have also shown that the hypotheses for each of the two models are relevant in some cases. Computational experiments showed that these models help to determine production schedules that better respect inventory constraints, even when the shapes of the production and demand rates differ from the ones in the hypotheses. In particular, we have shown that, in the extreme case where the demand is instantaneous at either the beginning or the end of the period, both models ensure a production schedule in which the inventory levels always fall within their bounds.

A perspective is to focus on solution methods that are specific to the proposed models as well as on a detailed complexity analysis for the new identified problems. Based on the numerical results, it could also be interesting to only generate a subset of the new inventory constraints within periods and analyze whether they provide more relevant production plans without affecting running times.

Chapter 4

Parallelized decomposition approaches for capacitated lot-sizing problems

In Chapter 2, we have extended the CLSP with setup times to better take into account the impact of the end-of-horizon effect on optimized production plans. The addition of a minimum global ending inventory was proposed as well as a maximum ending inventory per item. This allowed for the capacity in the last periods to not be underused and we showed that, on a rolling horizon, the production decisions are less affected by the length of the planning horizon. In Chapter 3, we have added constraints to bound the inventory levels in each period. We have shown that these new problems are computationally harder to solve. We now want to efficiently solve large instances of these problems by using two classical decomposition heuristics of the literature: Relax-and-fix and Lagrangian relaxation. These heuristics are adapted by splitting the problem into independent subproblems that can be solved in parallel.

The chapter is organized as follows. Section 4.1 positions our work in the literature. In Section 4.2, the lot-sizing problem with setup times, lost sales and global ending inventory is recalled, and the choice to use parallelized decomposition approaches is motivated. In Section 4.3, a parallelized version of the relax-and-fix algorithm is proposed, combined with a reconstruction technique at each iteration. In Section 4.4, we propose a Lagrangian relaxation based heuristic that uses a reconstruction method with an additional vertical pass compared to the literature. In Section 4.5, we show that, once the capacity constraints in the problem without lost sales are relaxed, each subproblem can be polynomially solved. Computational experiments are carried out in Section 4.6 where our methods are compared to the standard solver IBM ILOG CPLEX 12.10.

4.1 Related works

The relax-and-fix heuristic relaxes the integrality of some of the binary variables to decompose the problem into a series of subproblems that are solved in an iterative process (Absi and Kedad-Sidhoum (2007)). A recent study on the performances of the relax-and-fix heuristic can be found in Absi and van den Heuvel (2019). Our parallelization approach can also apply to the fix-and-optimize algorithm that fixes the binary decision variables to a specific value (Helber and Sahling (2010)). The Lagrangian relaxation based heuristic consists in combining a relaxation of the capacity constraints with a reconstruction heuristic to find a feasible plan (see e.g. Trigeiro et al. (1989), Brahimi, Dauzère-Pérès and Najid (2006) and Süral et al. (2009)). A major novelty in this thesis is that we parallelize these approaches not only to accelerate the solution times, but also to propose new relax-and-fix strategies. To solve the relaxed problem in the Lagrangian relaxation heuristic, a polynomial dynamic programming algorithm is also introduced for the uncapacitated single-item lot-sizing problem with ending inventory constraints and without lost sales. Even though there are some research works focusing on the use of a parallel framework to solve production planning problems (Roux et al. (1999)), there are few references on this subject (Haase and Kohlmorgen (1996)).

Decomposing a problem into subproblems is a well-known way to handle complex problems and large problem instances. In practice, an industrial problem can be split into subproblems that are solved in parallel, for instance by decomposing the problem by item or by location. This reduces the search tree when solving a problem using a Branch-And-Bound approach. Nowadays, it is easy to use threads in multicore platforms to perform many tasks in parallel. Almost every commercial solver uses threads in their exploration methods, exploiting the capacity of the machine in order to speed optimization processes. Once the problem is decomposed, the resulting subproblems can also be easier to solve than the original problem with regards to the time and space complexity of the problem. A problem that might be difficult to solve can become a series of polynomial subproblems for which there exist effective solution approaches in the literature. When solving the CLSP with setup times as defined in Trigeiro et al. (1989), methods based on the relaxation of the capacity constraints transform an NP-hard problem into a number of polynomial subproblems that can be solved in parallel using dynamic programming. The main drawback of using a decomposition method is that the reconstructed solution, while being fast to solve, might be of poor quality with regards to the initial problem. A compromise needs to be found between computational times and solution quality.

4.2 Problem formulations

4.2.1 General problem formulation

Let us consider the capacitated lot-sizing problem (CLSP) with minimum and maximum ending inventories, setup times and lost sales. N items have to be produced over a planning horizon of T periods. The discrete demand of each item i is given by d_{it} at period t . Each unit of item i produced at period t induces a production time b_{it} as well as a fixed setup time s_{it} . At the end of the horizon, there is a global minimum ending inventory I^{inf} as well as a maximum ending inventory per item I_i^{sup} .

The goal is to find an optimal production plan respecting the capacity restriction c_t^{\max} for

each period t while minimizing the total cost. This total cost comprises the fixed and unitary production costs to be incurred each time a production takes place, the inventory holding costs for all the items as well as the lost sales costs penalizing the unsatisfied demand. The cost parameters are the unitary production p_{it} , fixed setup f_{it} and unitary inventory holding h_{it} costs for item i at period t . The lost sales costs penalizing each unit of unsatisfied demand of item i at period t are defined by l_{it} .

Let us recall the mathematical formulation of the problem that is defined in Chapter 2 and that can be found in Charles et al. (2021), and first define the decision variables as follows:

- $X_{it} \geq 0$: Quantity of item i produced at period t ,
- $Y_{it} \in \{0, 1\}$: Setup variable which is equal to 1 if there is an order for item i at period t , and 0 otherwise,
- $I_{it} \geq 0$: Inventory of item i at the end of period t ,
- $L_{it} \geq 0$: Quantity of lost sales for item i at the end of period t .

We extend the definition of I_{it} with $t = 0$ to describe the initial inventory of item i . Moreover, $\bar{\cdot}$ is used to define the average value of a parameter over all items and all periods, e.g. $\bar{f} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T f_{it}$.

Let us define problem (P) as follows:

$$\min \sum_{i=1}^N \sum_{t=1}^T (f_{it}Y_{it} + p_{it}X_{it} + h_{it}I_{it} + l_{it}L_{it}) \quad (4.1)$$

$$I_{i,t-1} + X_{it} + L_{it} = d_{it} + I_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.2)$$

$$\sum_{i=1}^N (s_{it}Y_{it} + b_{it}X_{it}) \leq c_t^{max}, \quad \forall t \in 1, \dots, T \quad (4.3)$$

$$X_{it} \leq M_{it}Y_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.4)$$

$$L_{it} \leq d_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.5)$$

$$\sum_{i=1}^N I_{iT} \geq I^{inf} \quad (4.6)$$

$$I_{iT} \leq I_i^{sup}, \quad \forall i \in 1, \dots, N \quad (4.7)$$

$$Y_{it} \in \{0, 1\}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.8)$$

$$X_{it}, I_{it}, L_{it} \geq 0, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.9)$$

The objective function (4.1) minimizes the total production, setup, inventory and lost sales costs of all items over the planning horizon. Constraints (4.2) are the flow conservation constraints that balance, for each item, the inventory at period $t - 1$ and the production and lost sales quantities at period t with the inventory and the demand at period t . Constraints (4.3) ensure that the capacity consumed by the setup and production times does not exceed the maximum production capacity. Constraints (4.4) link the continuous production variables with the binary setup variables, M_{it} being an upper bound on the optimal production quantity (e.g. $M_{it} = \min(\sum_{k=t}^T d_{ik} + I_i^{sup}, c_t^{max} - s_{it})$). Constraint (4.6) states that the global ending inventory is bounded by I^{inf} . Constraints (4.7) add an upper bound for the ending inventory of each item. Constraints (4.5) state that the lost sales cannot exceed the demand. Constraints (4.8) and (4.9) define the domain of the variables.

4.2.2 Disaggregated formulation

In this chapter, we are also considering a disaggregated formulation for the CLSP with minimum and maximum ending inventories, setup times and lost sales. This formulation consists in disaggregating the production variables at period t and to consider not only the production period but also the demand it satisfies. Disaggregated formulations are often used to solve the LP-relaxation of the CLSP because they usually provide better lower bounds than the classical formulation (Eppen and Martin (1987), Denizel and Süral (2006)). The reformulation considered in this section was introduced in Krarup and Bilde (1977) and can be found in Degraeve and Jans (2007) and de Araujo et al. (2015) for the CLSP with setup times. Let us denote by X_{ikt} the quantity produced at period k to satisfy the demand for item i at period t ($k \leq t$). We extend this definition to $t = T + 1$ to model the quantity produced to satisfy the ending inventory constraint and to $k = 0$ to model the distribution of the initial inventory. With these new variables, the inventory variables can be removed, and by defining the new parameters

$$\tilde{p}_{ikt} = p_{ik} + \sum_{l=k}^{t-1} h_{il}$$

the following model can be written:

$$\min \sum_{i=1}^N \sum_{t=1}^T (f_{it} Y_{it} + l_{it} L_{it}) + \sum_{i=1}^N \sum_{k=0}^T \sum_{t=k}^{T+1} \tilde{p}_{ikt} X_{ikt} \quad (4.10)$$

$$L_{it} + \sum_{k=0}^t X_{ikt} = d_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.11)$$

$$\sum_{i=1}^N (s_{ik} Y_{ik} + \sum_{t=k}^{T+1} b_{ik} X_{ikt}) \leq c_k^{max}, \quad \forall k \in 1, \dots, T \quad (4.12)$$

$$X_{ikt} \leq d_{it} Y_{it}, \quad \forall i \in 1, \dots, N, \forall k \in 0, \dots, T, \forall t \in 1, \dots, T + 1 \quad (4.13)$$

$$X_{ik,T+1} \leq I_i^{sup} Y_{ik}, \quad \forall i \in 1, \dots, N, \forall k \in 0, \dots, T, \forall t \in 1, \dots, T + 1 \quad (4.14)$$

$$\sum_{t=1}^{T+1} X_{i0t} = I_{i0}, \quad \forall i \in 1, \dots, N \quad (4.15)$$

$$\sum_{i=1}^N \sum_{k=0}^T X_{ik,T+1} \geq I^{inf} \quad (4.16)$$

$$\sum_{k=0}^T X_{ik,T+1} \leq I_i^{sup}, \quad \forall i \in 1, \dots, N \quad (4.17)$$

$$Y_{it} \in \{0, 1\}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.18)$$

$$X_{ikt} \geq 0, \quad \forall i \in 1, \dots, N, \forall k \in 0, \dots, T, \forall t \in 1, \dots, T + 1 \quad (4.19)$$

$$L_{it} \geq 0, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.20)$$

In the new objective function (4.10), the inventory variables are removed and replaced by

the updated production cost parameters. Constraints (4.11) link the production variables to satisfy the demand of item i at period t to the corresponding demand and lost sales variables. Constraints (4.12) are equivalent to Constraints (4.3), where the X_{it} variables are replaced by $\sum_{k=t}^{T+1} X_{itk}$. Constraints (4.13)-(4.14) link the production variables with the binary setup variables. Constraints (4.16)-(4.17) model the minimum and maximum ending inventory constraints.

4.3 Parallelized relax-and-fix

The aim of the parallelized relax-and-fix approach is to give more flexibility to the production plan than in the classical relax-and-fix approach described in Section 4.3.1 by not necessarily fixing production decisions chronologically. The main issue of a relax-and-fix heuristic is that it might underestimate the impact of the limited capacity in the future periods when fixing some of the binary setup variables after optimization. This aspect might be even more relevant when optimizing on a rolling horizon, because a minimum level of inventory is required at the end of the horizon, and the relax-and-fix heuristic tends to underestimate the impact of the capacity in the last periods.

To illustrate the impact of the order in which the subproblems are optimized, let us consider an instance of the CLSP with two items, i.e. $N = 2$, and $T = 5$. Except for the first period where $d_{11} = d_{21} = 0$, demand is constant over time and is set to 100 for both items. Holding costs and unitary production times are set to 1 and setup times are set to 100. No setup and production costs are considered. The available capacity is $c^{\max} = 300$ in each period. The optimal objective function for this instance is 400 and can be found for instance by setting $Y_{12} = Y_{14} = Y_{21} = Y_{23} = Y_{25} = 1$ and all other setup variables to 0. An optimal production plan is shown in Figure 4.1. Let us consider a framework that optimizes production decisions one period at a time, fixing decisions that have already been taken and relaxing the integrality of all other binary variables. If we solve each subproblem in chronological order, after the first optimization we fix $Y_{11} = Y_{12} = 0$, which cannot lead to a feasible production plan without lost sales. We get the production plan shown in Figure 4.2. There are however several optimization orders leading to a feasible production plan without lost sales. For instance, by optimizing subproblems in the reverse chronological order, we get the production plan of Figure 4.1.

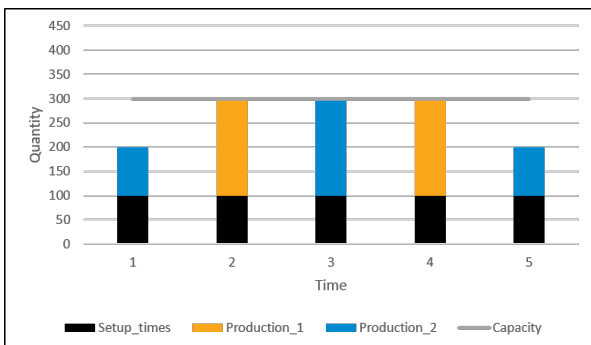


Figure 4.1: Optimal production plan

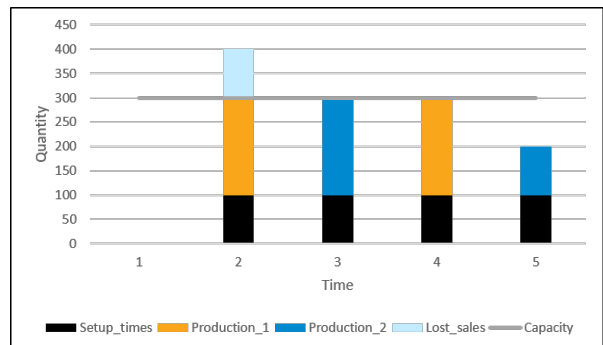


Figure 4.2: Chronological optimization

4.3.1 Principle

The relax-and-fix heuristic (Absi and Kedad-Sidhoum (2007), Helber and Sahling (2010)) is an iterative method. Let us define two parameters $\delta \in \mathbb{N}^*$ and $\gamma \in \mathbb{N}^*$ such that $\delta \leq T$ and $\gamma \leq \delta$. These parameters are respectively the number of periods where setup variables are optimized and the additional number of periods in which binary setup variables are fixed between two successive iterations. Let us denote by $K = \lceil \frac{T}{\gamma} \rceil$ the number of iterations necessary to fix all binary decision variables. More precisely, the planning horizon is divided into three sets at each iteration $k \in \llbracket 1, K \rrbracket$.

$\mathcal{S}_k^{fix} = \{1, \dots, (k-1)\gamma\}$ for $k = 2, \dots, K$ ($\mathcal{S}_1^{fix} = \emptyset$) is the set of periods where all binary variables are fixed.

$\mathcal{S}_k^{opt} = \{(k-1)\gamma+1, \dots, \min((k-1)\gamma+\delta, T)\}$ is the set of periods where all the constraints hold.

$\mathcal{S}_k^{rel} = \{(k-1)\gamma + \delta + 1, \dots, T\}$ is the set of periods where all the integrality constraints are relaxed.

At each iteration k , we solve subproblem (P_k) derived from problem (P) where, for each $t \in \llbracket 1, T \rrbracket$:

- If $t \in \mathcal{S}_k^{fix}$: variables Y_{it} are fixed to the values that were obtained up to iteration k , denoted $\overline{Y_{itk}}$. (P_k) thus has the following additional constraints:

$$Y_{it} = \overline{Y_{itk}} \quad (4.21)$$

- If $t \in \mathcal{S}_k^{rel}$: variables Y_{it} are considered as continuous variables. Constraints (4.8) are replaced in (P_k) by:

$$Y_{it} \in [0, 1]. \quad (4.22)$$

Optimal values Y_{it}^* obtained at iteration k after solving (P_k) are used to define $\overline{Y_{it,k+1}} = Y_{it}^*$ for $t \in \llbracket (k-1)\gamma, k\gamma \rrbracket$. The algorithm is formalized in Algorithm 4.1.

Algorithm 4.1 Relax-and-fix algorithm

```

1 ≤ δ ≤ T, 1 ≤ γ ≤ δ
for k ∈ ⌊1, K⌋ do
    Skfix ← {1, ..., (k-1)γ}
    Skopt ← {(k-1)γ+1, ..., (k-1)γ+δ}
    Skrel ← {(k-1)γ+δ+1, ..., T}
    Solve (Pk) where:
    - Yit ←  $\overline{Y_{itk}}$  for (i, t) ∈ ⌊1, N⌋ × Skfix
    - Yit ∈ [0, 1] for (i, t) ∈ ⌊1, N⌋ × Skrel
    Y* is an optimal solution of (Pk)
     $\overline{Y_{it,k+1}}$  ← Yit* for (i, t) ∈ ⌊1, N⌋ × ⌊(k-1)γ+1, kγ⌋
end for

```

In the fix-and-optimize algorithm (Helber and Sahling (2010)), all binary variables Y_{it} for $t \in \mathcal{S}_k^{rel}$ are fixed to $\overline{Y_{itk}}$ whose values come from an initial solution. At the end of iteration k , all $\overline{Y_{itk}}$ are updated for $t \in \mathcal{S}_k^{opt}$.

4.3.2 Parallelized relax-and-fix

The parallelized version of the relax-and-fix heuristic introduces a more generic definition of the sets defined in Section 4.3.1. At iteration k , at most $J = \lceil \frac{T}{\gamma} \rceil$ subproblems (P_{jk}) for $j \in \mathcal{J}_k \subset \{1, \dots, J\}$ are solved. We assume that γ divides T . These subproblems can be solved in parallel. The decomposition of the time horizon is provided by the triplet $(\mathcal{S}_{jk}^{fix}, \mathcal{S}_{jk}^{opt}, \mathcal{S}_{jk}^{rel})$ for each subproblem (P_{jk}) , where \mathcal{S}_{jk}^{fix} is the same for all $j \in \mathcal{J}_k$. More precisely, at iteration k and for each subproblem $j \in \mathcal{J}_k$, we have:

- \mathcal{S}_{jk}^{opt} , \mathcal{S}_{jk}^{rel} and \mathcal{S}_{jk}^{fix} form a partition of $\{1, \dots, T\}$,
- $\mathcal{S}_{jk}^{opt} = \{(j-1)\gamma + 1, \dots, \min((j-1)\gamma + \delta, T)\}$,
- $\mathcal{S}_{j_0}^{fix} = \emptyset$,
- $\mathcal{S}_{j,k+1}^{fix} = \mathcal{S}_{jk}^{fix} \cup \{t \in \llbracket (j_k - 1)\gamma + 1, j_k\gamma \rrbracket \mid t \notin \mathcal{S}_{jk}^{opt}\}$.

The sets \mathcal{J}_k are defined as follows. $\mathcal{J}_{k+1} = \mathcal{J}_k \setminus \{j_k\}$ for $k = 1, \dots, J-1$, where j_k corresponds to a selected subproblem and $\mathcal{J}_1 = \{1, \dots, J\}$. The selection is based on a score denoted v_{jks} that depends on the selected strategy $s \in \llbracket 1, C \rrbracket$ where C is the number of strategies. The subproblem with the lowest score is selected. All the binary variables Y_{it} for $t \in \llbracket (j_k - 1)\gamma + 1, j_k\gamma \rrbracket$ are fixed for all iterations $k' > k$. Algorithm 4.2 formalizes the parallelized relax-and-fix heuristic.

Algorithm 4.2 Parallelized relax-and-fix algorithm

```

1 ≤ δ ≤ T, 1 ≤ γ ≤ δ, Sfix ← ∅, s ∈ [1, C]
J1 = {1, ⋯, T/γ}
for k ∈ [1, K] do
  for j ∈ Jk do
    Sjkopt ← {γ(j-1) + 1, ⋯, γ(j-1) + δ}, Sjkfix ← Sfix \ Sjkopt, Sjkrel ← {1, ⋯, T} \ (Sjkfix ∪ Sjkopt)
    Solve (Pjk) where:
    - Yit ← Ykit* for (i, t) ∈ [1, N] × Sjkfix
    - Yit ∈ [0, 1] for (i, t) ∈ [1, N] × Sjkrel
    Yj* is an optimal solution of (Pjk)
    Set vjks as defined in Section 4.3.2
  end for
  jk ← argminj ∈ Jk vjks
  Yk+1,it ← Yjkit* for (i, t) ∈ [1, N] × [(jk - 1)γ + 1, jkγ]
  Sfix ← Sfix ∪ Sjkkopt
  Jk+1 ← Jk \ {jk}
end for

```

The iterative method is illustrated in Figure 4.3 for $J = 3$. At each iteration different subproblems are optimized in parallel, the most promising partial solution is kept and all binary variables of the corresponding set of periods are fixed in the next iterations. We get a complete production plan after J iterations.

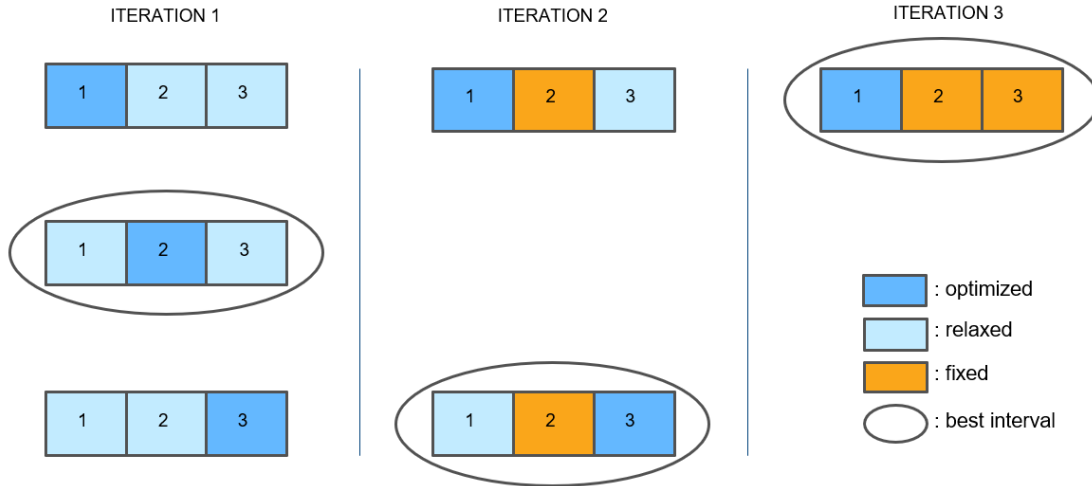


Figure 4.3: Illustration of the parallelized relax-and-fix heuristic

Finding the best interval

At each iteration, different strategies can be considered to select the partial solution that leads to the interval to fix in the next iterations. In the present work, the number of strategies $C = 4$. The four strategies are detailed below.

1. The first strategy ($s = 1$) consists in randomly selecting an interval among the remaining subproblems to be fixed in the next iterations, i.e v_{jk1} is randomly fixed in $[0, 1]$. This strategy, which should be the worst, is used as a baseline for comparison with the three other strategies.
2. The second strategy ($s = 2$) considers the intervals in chronological order, which corresponds to the classical relax-and-fix heuristic of Section 4.3.1. At iteration k , $j_k = k$ and $v_{jk2} = j$.
3. The third strategy ($s = 3$) consists in selecting at each iteration the interval associated to the subproblem with the lowest objective value. The idea is to fix at each iteration the interval that degrades the least the objective function. Let us denote by $v(P_{jk})$ the objective value of an optimal solution for (P_{jk}) , then $v_{jk3} = v(P_{jk})$.
4. The fourth strategy ($s = 4$) consists in selecting at each iteration the interval associated to the subproblem with the largest number of integer setup variables. Even if the objective function at the iteration is not the lowest, it might be the solution that will be the least affected during the following iterations, hence $v_{jk4} = \sum_{i=1}^N \sum_{t=1}^T (1 - Y_{jit}^*) Y_{jit}^*$.

The different strategies are illustrated in Figure 4.4 that shows that different strategies can lead to different selected intervals. Figure 4.4 corresponds to a possible outcome after the first iteration of the parallelized relax-and-fix heuristic. A comparative analysis is provided in Section 4.6.

4.3.3 Reconstruction method

After solving subproblem (P_{jk}) , we obtain a partial solution where setup variables have either binary or fractional values. Let us denote $0 \leq Y_{jit}^* \leq 1$ the values obtained for the

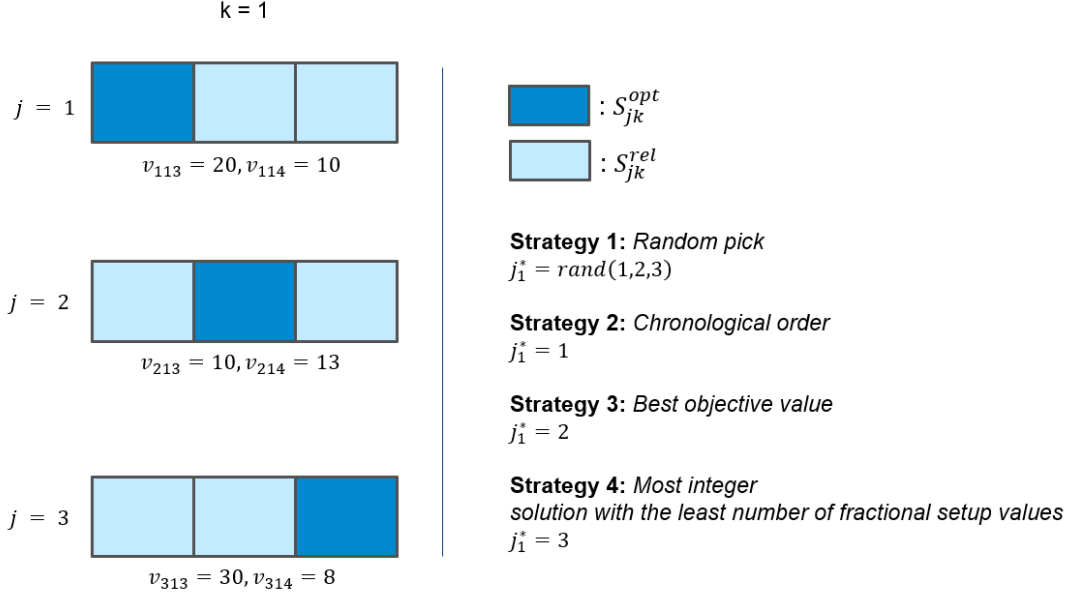


Figure 4.4: Strategies for selecting the best interval for $T = 3$.

binary variables after optimization of subproblem j . The information provided by the partial production plan can be used to build a feasible solution to the initial problem.

Let us define the following sets:

- $Y_{jk}^{frac} = \{(i, t) \in \llbracket 1, N \rrbracket \times \llbracket 1, T \rrbracket, 0 < Y_{jit}^* < 1\}$,
- $Y_{jk}^{prod} = \{(i, t) \in \llbracket 1, N \rrbracket \times \llbracket 1, T \rrbracket, Y_{jit}^* = 1\}$,
- $Y_{jk}^{off} = \{(i, t) \in \llbracket 1, N \rrbracket \times \llbracket 1, T \rrbracket, Y_{jit}^* = 0\}$.

We reconstruct a feasible solution based on each partial solution obtained using the relax-and-fix heuristic as defined in Section 4.3.2. At iteration l of the relax-and-fix heuristic, all binary variables in set S_l^{opt} are optimized. A new problem (P_{jk}^l) derived from problem (P_{jk}) is optimized where, for all binary variables whose integrality constraints were initially relaxed:

- If $(i, t) \in Y_{jk}^{off} \setminus S_l^{opt}$, the binary variables are fixed to 0:

$$Y_{it} = 0$$

- If $(i, t) \in Y_{jk}^{prod} \setminus S_l^{opt}$, the binary variables are fixed to 1:

$$Y_{it} = 1$$

- If $(i, t) \in Y_{jk}^{frac} \setminus S_l^{opt}$, the binary variables have their integrality relaxed:

$$0 \leq Y_{it} \leq 1$$

At the end of iteration l , we update the sets Y_{jk}^{frac} , Y_{jk}^{prod} and Y_{jk}^{off} using the new setup values. Fixing all binary variables in Y_{jk}^{off} and Y_{jk}^{prod} simplifies the problem by reducing the number of binary variables that are optimized, hence fastening the optimization process. Performing the reconstruction heuristic during the first iterations of the algorithm provides more flexibility as less binary variables are fixed. In the last iterations, the reconstruction method is faster but less flexible.

4.4 A Lagrangian relaxation based heuristic

Lagrangian relaxation consists in approximating an optimization problem by relaxing some of the problem constraints. Lagrangian multipliers are used to penalize the violation of the relaxed constraints in the objective function. By fitting the Lagrangian multipliers, the convex hull of the optimization problem can be found in some cases (Süral et al. (2009)). Indeed, the Lagrangian dual bound is usually better than the linear relaxation bound, and the usually infeasible solutions of the relaxed problems found in the iterative algorithm to determine the Lagrangian dual, such as the subgradient algorithm, can be used to derive feasible solutions by means of reconstruction heuristics.

Heuristics based on Lagrangian relaxation have been widely used to solve capacitated lot-sizing problems since the early work of Trigeiro et al. (1989). Our Lagrangian relaxation based heuristic is similar to the ones proposed in Brahimi, Dauzère-Pérès and Najid (2006) or Absi et al. (2013) and is divided into two phases. After relaxing the capacity constraints, we solve the Lagrangian relaxed problem for given Lagrangian multipliers using the algorithms proposed in Section 4.5. Starting from the obtained solution, we restore the feasibility with respect to the capacity constraints by smoothing the produced quantities using a forward and a backward passes known in the literature (Trigeiro et al. (1989)) and recalled in Section 4.4.3. In addition, the global minimum ending inventory constraint adds a linking constraint between all items. Hence, one of the new features of our approach compared to the literature is a vertical pass, presented in Section 4.4.4, which aims at satisfying the global minimum ending inventory by moving inventories between items. In the original problem (P), lost sales are allowed. However, contrary to Absi et al. (2013), lost sales are not allowed in the Lagrangian subproblem (\bar{P}) that we define and study in Section 4.4.2. We propose, in Section 4.4.5, a way to efficiently transfer any quantity of overtime, that corresponds to the excess beyond the capacity limit, to lost sales in order to find a feasible solution to the original problem (P). This approach is mainly relevant to solve problems where lost sales are allowed but highly penalized (for instance multi-objective problems where the first and most important objective consists in minimizing the total lost sales).

The resulting uncapacitated lot-sizing problem without lost sales is analyzed in Section 4.5. We show in Section 4.5.1 that the general problem is NP-hard. A heuristic is thus proposed in Section 4.5.2. Using this heuristic to solve the relaxed problem implies that the Lagrangian relaxation based heuristic does not provide a valid lower bound for given Lagrangian multipliers. We show in Section 4.5.3 that, when all maximum ending inventories are identical, the uncapacitated problem without lost sales becomes polynomial and an exact algorithm is proposed. Hence, in this special case, a lower bound is obtained at each iteration of the Lagrangian relaxation based heuristic when there are no lost sales in an optimal solution of the original problem.

4.4.1 Relaxation of the capacity constraints

Let us consider Problem (P) modeled in Section 4.2.1 where the capacity Constraints (4.3) are relaxed. The related Lagrangian multipliers are denoted $\lambda_t \geq 0$ for each period t , leading to the following problem (\bar{P}):

$$\min \sum_{i=1}^N \sum_{t=1}^T (f_{it}Y_{it} + p_{it}X_{it} + h_{it}I_{it} + l_{it}L_{it} + \lambda_t(s_{it}Y_{it} + b_{it}X_{it} - \frac{c_t^{max}}{N})) \quad (4.23)$$

(4.2), (4.4) – (4.9)

The objective function can also be written as follows:

$$\min \sum_{i=1}^N \sum_{t=1}^T ((f_{it} + \lambda_t s_{it})Y_{it} + (p_{it} + \lambda_t b_{it})X_{it} + h_{it}I_{it} + l_{it}L_{it}) - \sum_{t=1}^T \lambda_t c_t^{max} \quad (4.24)$$

The only linking constraint in Problem (\tilde{P}) is the global minimum ending inventory constraint (Constraint (4.6)).

4.4.2 Lagrangian heuristic

The iterative algorithm described in this section is designed to determine a feasible production plan for the CLSP with setup times and ending inventory constraints.

At each iteration of the algorithm several steps are performed:

1. We first remove the lost sales from Problem (\tilde{P}) in the objective function (4.23) and the material flow constraints (4.2), and obtain Problem (\bar{P}).
2. Problem (\bar{P}) is solved.
3. Based on the solution that is obtained, any excess in capacity (overtime) becomes lost sales.

The method to transfer overtime into lost sales is formalized in Section 4.4.5. We assume that the initial inventory for each item i is not greater than the total demand for this item over the time horizon:

$$I_{i0} \leq \sum_{t=1}^T d_{it}.$$

Because lost sales are not allowed, the initial inventory can be spread on the demands in the first periods. By updating these demands, we can reduce (\tilde{P}) to a problem with no initial inventory and without lost sales. Let us denote by $f'_{it} = f_{it} + \lambda_t s_{it}$ and $p'_{it} = p_{it} + \lambda_t b_{it}$ the updated setup and production costs using the Lagrangian multipliers defined in Section 4.4.1. The resulting problem (\bar{P}) is written below:

$$\min \sum_{i=1}^N \sum_{t=1}^T (f'_{it}Y_{it} + p'_{it}X_{it} + h_{it}I_{it}) \quad (4.25)$$

$$\sum_{i=1}^N I_{iT} = I^{inf} \quad (4.26)$$

$$I_{i,t-1} + X_{it} = d_{it} + I_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (4.27)$$

(4.7) – (4.8)

Constraints (4.6) are replaced by an equality as there is no need to produce more than required if the costs are positive. At each iteration k , we obtain, after solving Problem (\tilde{P}) a production plan that can be described by the decision variables Y_{kit}^* , X_{kit}^* and I_{kit}^* that are the setup, production and inventory decisions for each item i at each period t . Using this production plan, we can define:

$$o_k = \sum_{t=1}^T \max(0, \sum_{i=1}^N (s_{it}Y_{kit}^* + b_{it}X_{kit}^*) - c_t^{max}) \quad (4.28)$$

the overtime that would be required to satisfy production plan (Y^*, X^*, I^*) and L_k the corresponding quantity of lost sales. We update the Lagrangian multipliers using the formula below:

$$\lambda_{k+1,t} = \max(0, \lambda_{kt} - \psi_k(c_t^{max} - \sum_{i=1}^N (s_{it}Y_{kit}^* + b_{it}X_{kit}^*))) \quad (4.29)$$

where $\psi_k > 0$ is a coefficient that is set to:

$$\psi_k = \phi_k \frac{o_k}{\sqrt{\sum_{t=1}^T (c_t^{max} - \sum_{i=1}^N (s_{it}Y_{kit}^* + b_{it}X_{kit}^*))^2 + \epsilon}} \quad (4.30)$$

where $\epsilon > 0$ is a parameter set to a very small value in order to avoid the denominator to be equal to 0. At iteration k , ϕ_k is updated according to the learning based expander-contractor defined in [Zamani and Lau \(2010\)](#), where the overtime o_k is compared to the minimum overtime found \underline{o} . If $o_k \leq \underline{o}$, we set $\phi_{k+1} = \mu_1 \phi_k$, otherwise we set $\phi_{k+1} = \mu_2 \phi_k$, with $\phi_0 > 0$, $\mu_1 > 1$ and $0 < \mu_2 < 1$. This implies that, if the minimum overtime is improved, the step size increases while the step size decreases when there is no improvement.

If the capacity constraint is not respected ($c_t^{max} < \sum_{i=1}^N (s_{it}Y_{kit}^* + b_{it}X_{kit}^*)$), the Lagrangian coefficient λ_{kt} increases. The overtime is reduced by performing runs of vertical, forward and backward passes as described in [Section 4.4.3](#) and [Section 4.4.4](#) until no further improvement can be made. Let us denote by \underline{L} the smallest number of lost sales found during the run of the algorithm. The algorithm is formalized in [Algorithm 4.3](#).

Algorithm 4.3 Lagrangian relaxation based heuristic

```

 $k \leftarrow 0, k_{max}, \underline{L} \leftarrow \infty, \underline{Q} \leftarrow \infty,$ 
 $\phi_0 > 0, \mu_1 > 1, 0 < \mu_2 < 1$ 
 $\lambda_{0t} \leftarrow 0 \forall t \in \{1, \dots, T\}$ 
while  $k \leq k_{max}$  do
    Solve  $(\bar{P})$  as defined in Section 4.4.2 using Algorithm 4.4
     $(Y_k^*, X_k^*)$  is the obtained solution
    Compute  $o_k$  using formula (4.28)
    if  $o_k < \underline{Q}$  then
         $\underline{Q} \leftarrow o_k$ 
         $\phi_{k+1} \leftarrow \mu_1 \phi_k$ 
    else
         $\phi_{k+1} \leftarrow \mu_2 \phi_k$ 
    end if
    Compute  $\psi_k$  using formula (4.30)
     $\lambda_{k+1,t} \leftarrow \max(0, \lambda_{kt} - \psi_k(c_t^{max} - \sum_{i=1}^N (s_{it} Y_{kit}^* + b_{it} X_{kit}^*)))$  for  $t \in \llbracket 1, T \rrbracket$ 
    while stopping criteria do
        Perform forward, backward and vertical passes as defined in Section 4.4.3 and Section 4.4.4
    end while
    Transfer excess production to lost sales as defined in Section 4.4.5
    Compute  $L_k$ 
    if  $L_k < \underline{L}$  then
         $\underline{L} \leftarrow L_k$ 
    end if
     $k \leftarrow k + 1$ 
end while
    
```

In Sections 4.4.3 and 4.4.4, index k is removed to simplify the notations.

4.4.3 Forward and backward passes

Although these passes are known in the literature (Trigeiro et al. (1989), Brahimi, Dauzère-Pérès and Najid (2006)), they are recalled in this section for the sake of completeness. At each iteration k , once the uncapacitated problem (\bar{P}) is solved, we get a production plan that requires overtime in some periods to satisfy all demands over the horizon. This initial production plan is usually smoothed out by performing forward and backward passes. At iteration k , a forward pass moves the production quantity of an item to a later period while a backward pass moves the production quantity of an item to a previous period. When performing a forward pass from period t' to a later period t , enough inventory should be kept to satisfy all demands before t . Hence, we cannot move a quantity greater than the inventory at any given period between t' and $t-1$. Let us denote by c_t^a the available capacity at period t . The quantity of i that is moved forward cannot consume more time than the available capacity at t , including a setup time if i was initially not produced at t . With a forward pass, the maximum quantity $Q_{it't}^f$ of item i that can be moved from period t' to period t ($t' < t$) without affecting the feasibility of the production plan is:

$$Q_{it't}^f = \min\left(\frac{c_t^a - (1 - Y_{it}^*)s_{it}}{b_{it}}, X_{it'}^*, \min_{t' \leq l < t} I_{il}^*\right) \quad (4.31)$$

With a backward pass, the quantity moved from period t' to period t similarly cannot consumed more time than the available capacity c_t^a , including a setup time if there was no previous production of i at t . The maximum quantity $Q_{itt'}^b$ of item i that can be moved from period t' to an earlier period t ($t < t'$) without affecting the feasibility of the production plan is:

$$Q_{itt'}^b = \min\left(\frac{c_t^a - (1 - Y_{it}^*)s_{it}}{b_{it}}, X_{it'}^*\right) \quad (4.32)$$

For both passes, at each iteration of a pass, the best move from t' to t is chosen according to the following criterion:

$$\min_{(t,t') \in T^2} \frac{\lambda_t}{\max(Q_{itt'}, \epsilon)\lambda_{t'}}, \quad (4.33)$$

where $Q_{itt'} = Q_{itt'}^f$ or $Q_{itt'} = Q_{itt'}^b$ depending on whether a forward or a backward pass is performed and $\epsilon > 0$. Production is less penalized for periods with smaller Lagrangian multipliers, thus this criterion moves production quantities from a period with a low incentive to produce to a period with a high incentive to produce. As in [Trigeiro et al. \(1989\)](#), forward and backward passes are performed three times each. Each pass stops when no further move is possible.

4.4.4 Vertical pass

As the problem with ending inventory bounds has an additional constraint linking the ending inventory of all items, at each iteration k an additional pass that enables units of ending inventories to be moved from one item to another is added to the forward and backward passes. This pass does not apply for the CLSP with setup times with the linking global ending inventory constraint because demands are independent for each item. We show in Section [4.5](#) that, when solving the uncapacitated version of the problem with ending inventory constraints, all ending inventories can only take a restricted set of values. This property does not apply to the capacitated problem, justifying the need for this additional pass. With a vertical pass, the maximum quantity of ending inventory of item i' produced at period t' that can be replaced by ending inventory of i produced at period t (denoted $Q_{i'ikt}^v$) without affecting the feasibility of the production plan is:

$$Q_{i'ikt}^v = \min\left(\frac{c_t^a - (1 - Y_{it}^*)s_{it}}{b_{it}}, I_i^{sup} - I_{iT}^*, X_{i't'}^*, \min_{t' \leq l \leq T} I_{i'l}^*\right) \quad (4.34)$$

Indeed, the quantity of item i produced at period t cannot consume more time than the available capacity c_t^a . Moreover, we cannot remove a quantity of i' at t' that is greater than the available inventory at any given period between t' and T . Finally, the ending inventory of i cannot exceed I_i^{sup} . At each iteration of the pass, we find the best move from item i' produced at period t' to item i produced at period t according to the criterion:

$$\min_{(i,i',t,t') \in N^2 \times T^2} \frac{\lambda_t}{\max(Q_{i'ikt}^v, \epsilon)\lambda_{t'}} \quad (4.35)$$

Similarly to the forward and backward passes, we move production quantities from a period with a low incentive to produce to a period with a high incentive to produce.

4.4.5 Deriving a feasible solution for (P)

At the end of the passes performed at iteration k , the resulting production plan has no lost sales but may require overtime o_t at period t in order to satisfy all demands.

We compute for each item i with $Y_{it}^* = 1$ the ratio between the setup times and the production times:

$$r_{it} = \begin{cases} \frac{s_{it}+1}{b_{it}X_{it}^*} & \text{if } b_{it}X_{it}^* \leq o_t \\ \frac{1}{b_{it}X_{it}^*} & \end{cases} \quad (4.36)$$

By not producing an item, we also delete the setup time for this item. Thus, it makes sense to first remove items for which the setup times are the largest compared to the processing times. This is why, at each period t , we remove units of item i in decreasing order of r_{it} until all overtime has been removed. Runs of the passes described in Section 4.4.3 are then performed.

4.5 Uncapacitated lot-sizing problem without lost sales and with ending inventory constraints

We consider in this section the uncapacitated lot-sizing problem without lost sales and with ending inventory constraints (\bar{P}) . We prove in Section 4.5.1 that (\bar{P}) is NP-hard. Then, a heuristic to find a feasible solution in polynomial time is proposed in Section 4.5.2. Finally, we show in Section 4.5.3 that, when the ending inventory upper bound is the same for all items, the problem can be solved polynomially.

4.5.1 Problem analysis

We first prove a preliminary result that is used in the proof of Theorem 6, the main theorem of this section.

Theorem 5. *There exists an optimal solution for (\bar{P}) where at most one item $i \in \llbracket 1, N \rrbracket$ has an ending inventory $0 < I_{iT} < I_i^{sup}$. All other ending inventories take values in $\{0, I_i^{sup}\}$.*

Proof. The proof is done by contradiction. Let us assume there exists an optimal solution where the ending inventories of two items i and i' are such that:

$$0 < I_{iT} < I_i^{sup},$$

$$0 < I_{i'T} < I_{i'}^{sup}.$$

We can assume that this solution is such that all demands as well as the ending inventory are produced during the latest production period. Otherwise, because the problem is uncapacitated and because inventory costs are non-negative, we can find a feasible solution as least as good that satisfies this requirement. Let us denote by t_i^e and $t_{i'}^e$ the last production periods for items i and i' in this solution. Such periods exist because ending inventories for items i and i' are strictly greater than 0 and the problem does not have initial inventory. Let us consider the costs:

$$c_i = p_{it_i^e} + \sum_{t=t_i^e}^T h_{it}$$

and

$$c'_i = p_{i't'_i} + \sum_{t=t'_i}^T h_{i't}.$$

Let us consider the case where $c_i \leq c_{i'}$. Then $\delta = \min(I_{i'T}, I_i^{sup} - I_{iT})$ units of item i' can be replaced by units of item i without affecting the feasibility of the solution. The new solution is at least as good as the previous one with either $I_{iT} = I_i^{sup}$ or $I_{i'T} = 0$. By symmetry, this proposition holds for $c_i > c_{i'}$. \square

The theorem below shows that the problem is NP-hard.

Theorem 6. *The uncapacitated lot-sizing problem without lost sales and with ending inventory constraints is NP-hard.*

Proof. Let us consider a single period case where $f_i = I_i^{sup}$ for $i \in \llbracket 1, N \rrbracket$. Demands as well as production and holding costs are equal to 0. Let us denote this problem by (\overline{P}_1) and its optimal value by F_1^* . (\overline{P}_1) is given by:

$$\min \sum_{i=1}^N I_i^{sup} Y_i \tag{4.37}$$

$$s.t. \quad \sum_{i=1}^N X_i \geq I^{inf} \tag{4.38}$$

$$X_i \leq I_i^{sup} Y_i, \quad \forall i \in \llbracket 1, N \rrbracket \tag{4.39}$$

$$Y_{it} \in \{0, 1\}, \quad \forall i \in \llbracket 1, N \rrbracket \tag{4.40}$$

$$X_{it} \geq 0, \quad \forall i \in \llbracket 1, N \rrbracket \tag{4.41}$$

We show that (\overline{P}_1) is equivalent to (\overline{P}_2) , whose optimal value is denoted F_2^* and which is given by:

$$\min \sum_{i=1}^N I_i^{sup} Z_i \tag{4.42}$$

$$s.t. \quad \sum_{i=1}^N I_i^{sup} Z_i \geq I^{inf} \tag{4.43}$$

$$Z_{it} \in \{0, 1\}, \quad \forall i \in \llbracket 1, N \rrbracket \tag{4.44}$$

Let us first show that $F_2^* = F_1^*$:

- Let us denote (Z^*) an optimal solution for (\overline{P}_2) . The solution consisting in setting $Y_i = Z_i^*$ and $X_i = I_i^{sup} Z_i^*$ is a feasible solution for (\overline{P}_1) of objective value F_2^* , hence $F_2^* \geq F_1^*$.
- $F_2^* \leq F_1^*$: Let us denote (Y_i^*, X_i^*) an optimal solution for (\overline{P}_1) . The solution consisting in setting $Z_i = Y_i^*$ is a feasible solution for (\overline{P}_2) of objective value F_1^* . Indeed we have $\sum_{i=1}^N I_i^{sup} Y_i^* \geq \sum_{i=1}^N X_i^* \geq I^{inf}$, hence $F_2^* \leq F_1^*$.

Let us now consider the subset sum problem, which consists in deciding whether a sum of any subset of a set of positive integers is equal to a specific target value. We show in the following that, for all instances of the subset sum problem, we can define an instance of (\overline{P}_2) such that by finding an optimal solution for this instance, we can solve the corresponding subset sum instance.

Let us denote an instance of the subset problem $S = \{S_1, \dots, S_N\}$ of N positive integers and a target sum TS . Let us also define a corresponding instance of (\overline{P}_2) with $I_i^{sup} = S_i$ for all $i \in \llbracket 1, N \rrbracket$ and $I_i^{inf} = TS$ of optimal solution (Z^*) and optimal value $F_2^* \geq TS$.

- If $F_2^* = TS$, then the subset of $i \in \llbracket 1, N \rrbracket$ such that $Z_i^* = 1$ provides a solution to the instance of the subset sum problem.
- If $F_2^* > TS$, then no subset of the set of integers sums to TS , otherwise a feasible solution for the instance (\overline{P}_2) would exist with a value of TS .

This result shows that (\overline{P}_2) is at least as hard as the subset sum problem. The subset sum problem is known to be NP-complete (Karp (1972)). Thus (\overline{P}_2) is NP-hard and consequently (\overline{P}_1) is NP-hard as well. \square

4.5.2 Heuristic approach

For each item i , let us consider the single-item uncapacitated lot-sizing problem (\overline{P}_i) with ending inventory $I_i^e \geq 0$ below:

$$\min \sum_{t=1}^T (f'_{it} Y_{it} + p'_{it} X_{it} + h_{it} I_{it}) \quad (4.45)$$

$$s.t. \quad I_{i,t-1} + X_{it} = d_{it} + I_{it}, \quad \forall t \in \llbracket 1, T \rrbracket \quad (4.46)$$

$$I_{iT} = I_i^e \quad (4.47)$$

$$X_{it} \leq M_{it} Y_{it}, \quad \forall t \in \llbracket 1, T \rrbracket \quad (4.48)$$

$$Y_{it} \in \{0, 1\}, \quad \forall t \in \llbracket 1, T \rrbracket \quad (4.49)$$

$$X_{it} \geq 0, \quad \forall t \in \llbracket 1, T \rrbracket \quad (4.50)$$

$$I_{it} \geq 0, \quad \forall t \in \llbracket 1, T \rrbracket \quad (4.51)$$

(\overline{P}_i) is equivalent to an uncapacitated single-item lot-sizing problem with $T + 1$ periods, where the last demand is equal to I_i^e and the setup and production costs at the last period are large enough (for instance $f'_{i,T+1} + p'_{i,T+1} > f'_{iT} + p'_{iT} + h_{iT}$), so that, in an optimal solution, no production occurs at period $T + 1$. This last problem can be solved in at most $O(T^2)$ (Wagner and Whitin (1958)) (or in $O(T \log T)$, Wagelmans et al. (1992)). Consequently, (\overline{P}_i) can also be polynomially solved. Let us denote by $F_i^*(I)$ the optimal value of an instance of problem (\overline{P}_i) with $I_i^e = I$.

Theorem 5 shows that there is an optimal solution for (\overline{P}) such that the ending inventory for each item $i \in \llbracket 1, N \rrbracket$ is in $\{0, I_i^{sup}\}$ except for at most one item. Algorithm 4.4 consists in solving one instance of (\overline{P}_i) for each item i with ending inventory I_i^e , with at most one item i such that $I_i^e \notin \{0, I_i^{sup}\}$.

Let us sort the items in a set $s = \{i_1, \dots, i_N\}$ and iterate over this set. At each iteration j , we keep a parameter I_j^{left} indicating how much ending inventory is still left to be satisfied. Initially, $I_0^{left} = I^{inf}$.

At iteration j , when considering item i_j :

- If $I_j^{left} > I_{i_j}^{sup}$, we set $I_{i_j}^e = I_{i_j}^{sup}$.
- Else if $I_j^{left} > 0$, we set $I_{i_j}^e = I_j^{left}$.
- Otherwise, we set $I_{i_j}^e = 0$.

At the end of each iteration, we set $I_{j+1}^{left} = I_j^{left} - I_{i_j}^e$. At most one item i will have an ending inventory not in $\{0, I_i^{sup}\}$. Assume the items are sorted according to their maximum ending inventories. In Algorithm 4.4, let us denote by $s = \{i_1, \dots, i_N\}$ the set of items sorted by decreasing order of maximum ending inventory. The heuristic solving (\bar{P}) when the maximum ending inventory varies between items is formalized in Algorithm 4.4:

Algorithm 4.4 Heuristic to solve (\bar{P})

```

 $F^* \leftarrow 0, I_0^{left} \leftarrow I^{inf}$ 
 $s \leftarrow \{i_1, \dots, i_N\}$ 
for  $j \in \{1, \dots, N\}$  do
  if  $I_j^{left} \geq I_{i_j}^{sup}$  then
     $F^* \leftarrow F^* + F_{i_j}^*(I_{i_j}^{sup})$ 
     $I_{j+1}^{left} \leftarrow I_j^{left} - I_{i_j}^{sup}$ 
  else if  $I_j^{left} > 0$  then
     $F^* \leftarrow F^* + F_{i_j}^*(I_{i_j}^{sup} - I_j^{left})$ 
     $I_{j+1}^{left} \leftarrow 0$ 
  else
     $F^* \leftarrow F^* + F_{i_j}^*(0)$ 
  end if
end for

```

The heuristic runs in $O(N(\log N + T \log T))$, with a complexity of $O(N \log N)$ to sort the set of maximum ending inventories and a complexity of $O(NT \log T)$ to solve N single-item uncapacitated lot-sizing problems.

4.5.3 Polynomial case

In this section, we consider the case where $I_i^{sup} = I^{sup}, \forall i \in \{1, \dots, N\}$, i.e. the maximum ending inventory is the same for all items, and show that (\bar{P}) can be polynomially solved. Let us define $k \in \mathbb{N}$ such that $I^{inf} = kI^{sup} + I^R$ with $0 \leq I^R < I^{sup}$. k is the quotient of the euclidean division of I^{inf} by I^{sup} . This means that, according to Theorem 5, we can find an optimal solution where exactly k items have an ending inventory equal to I^{sup} .

Let us show that an optimal solution of (\bar{P}) can be determined by solving a polynomial number of instances of subproblems (\bar{P}_i) . Let us denote by $G \subset \mathbb{R}^N$ the set of possible ending inventory combinations $g = (g_1, \dots, g_N)$ such that each element $g_i \in g$ is in $\{0, I^{sup}, I^R\}$, $\sum_{i=1}^N g_i = I^{inf}$, and at most one g_i of g is equal to I^R .

Let us denote by (\bar{P}_g) for $g \in G$ the uncapacitated multi-item lot-sizing problem where $I_i^e = g_i$ for $i \in \{1, \dots, N\}$.

Theorem 7. *Finding an optimal solution of (\bar{P}) is equivalent to finding an optimal solution among all (\bar{P}_g) , for $g \in G$.*

Proof. Theorem 5 implies that there is an optimal solution (Y^*, X^*, I^*) of (\bar{P}) such that the combination of ending inventories $\{I_{1T}^*, \dots, I_{NT}^*\}$ corresponds to an element of G .

Reciprocally, every element of $g = \{g_1, \dots, g_N\} \in G$ corresponds to a feasible solution for (\bar{P}) with ending inventory g_i for item i . \square

It should be pointed out that the size of G is exponential. If $I^R = 0$, then there are exactly k items with ending inventory equal to I^{sup} and the rest equal to 0, so $|G| = \binom{N}{k}$. If $I^R \neq 0$, then there are exactly k items with ending inventory equal to I^{sup} , one item with an ending inventory equal to I^R and the rest equal to 0, so $|G| = (N - k) \binom{N}{k}$. However, we will show that finding the optimal solution among all (\bar{P}_g) is polynomial, and that it is not necessary to explore all the possible values for g .

Theorem 8. *For any item $i \in \{1, \dots, N\}$, let us denote $K_i = \{g \in G \mid g_i = I^R\}$. An optimal solution g^* of (\bar{P}_g) restricted to $g \in K_i$ can be determined in polynomial time.*

Proof. Let us denote $F^*(g) = \sum_{j=1}^N F_j^*(g_j)$ for $g = \{g_1, \dots, g_N\}$.

For each $g \in K_i$, the indices of g can be divided in three disjoint sets S_g^{sup} , S_g^0 and S_g^R such that $S_g^R = \{i\}$ and $|S_g^{sup}| = k$ by defining:

- (1) $S_g^R = \{i\}$,
- (2) $S_g^{sup} = \{j \in \{1, \dots, N\} \setminus \{i\} \mid g_j = I^{sup}\}$,
- (3) $S_g^0 = \{j \in \{1, \dots, N\} \setminus \{i\} \mid g_j = 0\}$.

Reciprocally, each partition of $\{1, \dots, N\}$ can be divided into three disjoint sets S^R , S^{sup} and S^0 such that $S^R = \{i\}$ and $|S^{sup}| = k$ defines an element $g = \{g_1, \dots, g_N\}$ of K_i where:

- (1) $g_i = I^R$,
- (2) $g_j = I^{sup}$ for $j \in S^{sup}$,
- (3) $g_j = 0$, for $j \in \llbracket 1, N \rrbracket \setminus (\{i\} \cup S^{sup})$.

Let us denote by S^{part} the set of partitions $\{1, \dots, N\}$ into three disjoint sets S^R , S^{sup} and S^0 such that $S^R = \{i\}$ and $|S^{sup}| = k$. We show in the following that there is an optimal partition to solve (\bar{P}_g) restricted to $g \in K_i$.

$$\min_{g \in K_i} F^*(g) = \min_{\{S^R, S^{sup}, S^0\} \in S^{part}} (F_i^*(I^R) + \sum_{j \in S^{sup}} F_j^*(I^{sup}) + \sum_{j \in S^0} F_j^*(0)) \quad (4.52)$$

Let us denote $\Delta_j = F_j^*(I^{sup}) - F_j^*(0)$. We have:

$$\min_{g \in K_i} F^*(g) = F_i^*(I^R) + \sum_{j \in \{1, \dots, N\} \setminus \{i\}} F_j^*(0) + \min_{\{S^R, S^{sup}, S^0\} \in S^{part}} \left(\sum_{j \in S^{sup}} \Delta_j \right). \quad (4.53)$$

Finding an optimal solution $S^* = \{S^{R*}, S^{sup*}, S^{0*}\} \in S^{part}$ can be done by sorting each Δ_j for $j \in \llbracket 1, N \rrbracket \setminus \{i\}$ and by taking the first k indices in increasing order. As problem (\bar{P}_g) is polynomial, computing and sorting all Δ_j is polynomial.

Because we have shown that each element of S^{part} corresponds to an element of K_i , we can define an optimal solution $g^* = \{g_1^*, \dots, g_N^*\} \in K_i$ for (\bar{P}_g) restricted to $g \in K_i$. \square

The problem complexity of solving (\bar{P}) is defined in the theorem below.

Theorem 9. (\bar{P}) can be polynomially solved.

Proof. We want to find:

$$\min_{i \in [1, N]} (\min_{g \in K_i} F^*(g)) \quad (4.54)$$

Because finding an optimal solution to $\min_{g \in K_i} F^*(g)$ can be done in polynomial time according to Theorem 8 for each $i \in \{1, \dots, N\}$, finding an optimal solution for each sub-problem for $i \in \{1, \dots, N\}$ and taking the solution with the lowest value provides an optimal solution for (\bar{P}) . □

Algorithm 4.5 solving to optimality the uncapacitated lot-sizing problem with ending inventory constraints when the maximum ending inventory is identical for all items is formalized as follows:

Algorithm 4.5 Exact method to solve (\bar{P}) when the maximum ending inventory is the same for all items

```

 $F^* \leftarrow \infty$ 
for  $i \in \{1, \dots, N\}$  do
    Compute  $F_i^*(I^{sup})$ ,  $F_i^*(0)$  and  $F_i^*(I^R)$ 
     $\Delta_i \leftarrow F_i^*(I^{sup}) - F_i^*(0)$ 
end for
 $s \leftarrow \{i_1, \dots, i_N\}$  sorted in increasing order of  $\Delta = \{\Delta_1, \dots, \Delta_N\}$ 
for  $i \in \{1, \dots, N\}$  do
     $s_i \leftarrow \{i_{\phi(1)}, \dots, i_{\phi(N-1)}\} = s \setminus \{i\}$ 
     $F_{|K_i}^* \leftarrow F_i^*(I^R)$ 
    for  $j \in \{1, \dots, N-1\}$  do
        if  $j \leq k$  then
             $F_{|K_i}^* \leftarrow F_{|K_i}^* + F_{i_{\phi(j)}}^*(I^{sup})$ 
        else
             $F_{|K_i}^* \leftarrow F_{|K_i}^* + F_{i_{\phi(j)}}^*(0)$ 
        end if
    end for
    if  $F_{|K_i}^* < F^*$  then
         $F^* \leftarrow F_{|K_i}^*$ 
    end if
end for

```

Similarly to the heuristic formalized in Algorithm 4.4, Algorithm 4.5 runs in $O(N(N + T \log T))$.

4.6 Computational experiments

4.6.1 Instance generation scheme

We generate instances according to the generation scheme introduced in Charles et al. (2021) and presented in Section 2.5.2 of Chapter 2. We added slight changes in order to make these

instances more realistic. The number of items N takes values in $\{50; 80; 100\}$ with $T = 24$. The average demand is set to $\bar{d} = 100$, and this average demand takes into account the fact that some items are high runners while other items are low runners. To do so, we define $\theta \leq 1$ the fraction of high-running items. These items correspond to 80% of the total demand and have a demand generated at each period. Low-running items only have a 20% chance to have a positive demand at each period. In our computational experiments, $\theta = 0.3$. We have also defined the average setup time \bar{s} according to the average fraction of capacity consumed by setups at each period $\{0.05; 0.2; 0.4\}$. The average number of setups based on the costs ($k_{cost} = \sqrt{\frac{h}{2f}} \sum_{i=1}^N \sqrt{d_i}$), as well as the average number of setups based on the capacity ($k_{capa} = \frac{c^{max} - N\bar{b}\bar{d}}{\bar{s}}$), as defined in Chapter 2, take values in $\{\frac{N}{5}; \frac{N}{10}; \frac{N}{20}\}$. The *Time Between Order* (TBO) that extends the definition of Harris (1913) is defined as $TBO = \frac{N}{k_{cost}}$. For all these instances, there exists a feasible production plan without lost sales. In order for this production plan to be optimal, lost sales are highly penalized in the objective function.

The mathematical models are solved using IBM ILOG CPLEX 12.10 with a maximum running time of 50 seconds and the optimization is stopped if the optimality gap is lower than 1%.

4.6.2 Analysis

We compute the fraction of lost sales over the total demand obtained after using the parallelized relax-and-fix (c.f. Algorithm 4.2) for the different strategies introduced in Section 4.3.2 and the results are compared to the ones obtained using the Lagrangian relaxation based heuristic using Algorithm 4.3 with a maximum number of iterations $k^{max} = 2000$ and to the ones obtained using Algorithm 4.1. We set the following parameters for the Lagrangian relaxation based heuristic: $\phi_0 = 1$, $\mu_2 = 0.96$ and $\mu_1 = \frac{3-\mu_2}{2} = 1.02$.

The choice was made to show the fraction of lost sales instead of a computed optimality gap because of the high lost sales costs that lead to high optimality gaps when evaluating the quality of the solution. For the relax-and-fix heuristic, in all the tables except for Tables 4.3 and 4.4, we set $\gamma = \delta = 2$. In Tables 4.3 and 4.4, we set $\gamma = 2$ and $\delta = 4$. This makes more sense with regards to the classical use of the relax-and-fix heuristic where we usually allow for the binary variables to be revised several times. In Tables 4.1 and 4.3, the instances are sorted according to the TBO while, in Tables 4.2 and 4.4, the instances are sorted according to the average fraction of capacity consumed by setups. These two parameters are the ones leading to the most variations between instances. The parallelized relax-and-fix heuristic is referred to by the strategy used (classical -chronological order-, random, best first or most integer). The Lagrangian heuristic is denoted LR. We refer to the fraction of lost sales over the total demand by LS.

Table 4.1 shows that, except for the Lagrangian relaxation based heuristic, instances with a smaller TBO are the hardest to solve and lead to the largest lost sales (from 5.10% to 2.36% for the relax-and-fix with a selection based on the most integer partial solution). This can be explained by the fact that the capacity in the relaxed periods might be underestimated for smaller production quantities. For the Lagrangian relaxation based heuristic, we get the opposite effect because the larger the number of production periods, the more we have flexibility when performing the passes. For all methods, note in Table 4.2 that the instances become harder to solve as the average capacity consumed by setup times at each period increases. Smaller setup times provide more flexibility to add additional production periods to compensate for a prior underestimation of the impact of the capacity on the relaxed periods

(from 2.77% to 11.13% for the relax-and-fix with a selection based on the chronological order). In Tables 4.1 and 4.2, note that the chronological order provides in our case the worst results with regards to the fraction of lost sales (11.74% for the instances with a TBO of 5 compared to 5.10% by picking the interval with the most integer partial solution). The random and best first strategies seem comparable (they both have an average value of around 5% of lost sales), both when sorting the instances according to the TBO or the fraction of capacity consumed by the setups. With comparable running times, the Lagrangian relaxation always provides better results (always less than 1%).

TBO	Parallelized relax-and-fix								LR	
	Classical		Random		Best first		Most integer			
	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)
5	11.74	11.48	8.30	32.93	7.66	28.07	5.10	14.19	0.24	18.43
10	7.98	12.11	4.82	33.74	5.20	35.00	3.17	29.33	0.21	12.61
20	5.87	13.37	4.62	29.56	5.30	56.41	2.36	12.59	0.51	15.68

Table 4.1: Numerical results with respect to TBO

Capa. for setups (%)	Parallelized relax-and-fix								LR	
	Classical		Random		Best first		Most integer			
	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)
5	2.77	11.19	2.95	30.41	3.32	29.15	2.04	23.74	0.24	14.82
20	6.94	11.96	5.40	29.63	5.93	37.41	3.19	14.48	0.34	11.98
40	11.13	13.81	6.80	36.19	6.86	52.93	3.87	17.89	0.39	19.92

Table 4.2: Numerical results with respect to capacity settings

In Tables 4.3 and 4.4, only half of the binary variables are fixed when selecting an interval. Because the set of binary variables that are optimized at each iteration is twice as large ($\delta = 4$ while in the previous tables $\delta = 2$), the running times are larger (165 seconds on average for the classical order selection compared to less than 10 seconds without reoptimization). However, the fraction of lost sales decreases compared to the previous tables (from 11.13% to 5.16% with the chronological order). The most integer strategy still outperforms the other ones, and we still get better results with the Lagrangian relaxation heuristic, this time with much better running times.

TBO	Parallelized relax-and-fix							
	Classical		Random		Best first		Most integer	
	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)
5	4.35	162.52	1.35	192.52	2.44	119.19	0.88	181.48
10	3.44	169.96	1.14	205.04	1.67	137.56	0.54	222.81
20	3.67	162.96	1.60	231.48	2.54	246.59	0.65	215.89

Table 4.3: Numerical results with respect to TBO with reoptimization

In Table 4.5, the different strategies are compared using the disaggregated formulation introduced in Section 4.2.2. The instances are sorted according to the average fraction of the capacity that is consumed by setup times. When comparing the results with Table 4.2

Capa. for setups (%)	Parallelized relax-and-fix							
	Classical		Random		Best first		Most integer	
	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)
5	1.09	118.00	0.97	187.19	1.11	135.85	0.57	203.85
20	3.52	159.04	1.18	190.89	1.99	154.59	0.63	192.22
40	5.16	218.41	1.80	250.96	2.94	212.89	0.73	224.11

Table 4.4: Numerical results with respect to capacity settings with reoptimization

that uses the aggregated formulation to the results of Table 4.5, we can point out that the running times are always larger (from around 12 seconds for the relax-and-fix heuristic using the aggregated formulation to 100 seconds using the disaggregated formulation). This can be explained by the fact that the disaggregated formulation has more variables ($O(NT^2)$) than the aggregated formulation ($O(NT)$). However, the disaggregated formulation always provides better reconstructed solutions (always less than 0.5% of lost sales using the disaggregated formulation). With this formulation, the chronological relax-and-fix outperforms the other strategies (0.21% of lost sales for the instances with 20% of capacity being consumed by setup times to 0.39% of lost sales with the strategy based on the most integer partial solution). All strategies are very similar and provide reconstructed solutions of similar qualities. The quality of the reconstructed solution however seems to depend on the setup times. When a large fraction of the capacity is consumed by setup times, it becomes harder to recover from decisions taken in the early iterations of the algorithm. On average, when 5% of the capacity is consumed by setup times, about 0.09% of demands are lost while, when 40% of the capacity is consumed by setup times, the fraction of lost sales increases to 0.33%.

Capa. for setups (%)	Parallelized relax-and-fix							
	Classical		Random		Best first		Most integer	
	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)	LS (%)	T (s)
5	0.06	94.78	0.11	92.44	0.09	106.22	0.10	106.59
20	0.12	96.26	0.22	103.19	0.23	104.11	0.23	113.78
40	0.21	99.70	0.32	119.26	0.38	119.85	0.39	117.22

Table 4.5: Numerical results with respect to capacity settings using disaggregated formulation

Using the disaggregated formulation, Table 4.6 shows the results of the parallelized relax-and-fix using the reconstruction method of Section 4.3.3, and compares it to the results of the chronological relax-and-fix heuristic with reoptimization of half of the binary variables that were fixed during the previous iteration. The number of lost sales for 1000 units of demand is displayed. The fraction of lost sales is very small, however note that the reconstruction heuristic outperforms the chronological relax-and-fix both in terms of running times (78.33 seconds on average compared to 135.63 seconds for the instances where 40% of the capacity is consumed by setup times) and in terms of lost sales (4.2 ‰ to 8.12 ‰).

Capa. for setups (%)	Parallelized relax-and-fix			
	Classical		Reconstruction	
	LS (‰)	T (s)	LS (‰)	T (s)
5	3.19	101.48	2.68	47.63
20	5.28	90.67	3.91	60.26
40	8.12	135.63	4.25	78.33

Table 4.6: Numerical results with respect to capacity settings using disaggregated formulation and with reconstruction

4.7 Conclusions

The main focus of this chapter was the use of parallelization to enhance well-known heuristics for capacitated lot-sizing problems. We considered the relax-and-fix heuristic as well as a Lagrangian relaxation based heuristic to solve the CLSP with the addition of the minimum and maximum ending inventory constraints introduced in Chapter 2. We have extended the relax-and-fix heuristic and proposed a parallelized version that optimizes several intervals in parallel and selects the best interval to fix for the next iterations based on strategies to evaluate the partial solutions. We have shown that, under specific assumptions on the maximum ending inventory constraints and when relaxing the capacity constraints, the problem can be decomposed into independent uncapacitated single-item lot-sizing problems. We have proposed a new version of the Lagrangian relaxation based heuristic introduced by Trigeiro et al. (1989) which includes an additional vertical pass that moves the ending inventory between items. We have considered both the aggregated formulation of the problem as well as the facility location (a disaggregated) formulation. Computational experiments show that, with the aggregate formulation, the relax-and-fix heuristic can provide better reconstructed solutions when the selection strategy introduced in Section 4.3.2 differs from the chronological order. We have also shown that the Lagrangian relaxation based heuristic outperforms the relax-and-fix heuristic with similar running times. When using the facility location formulation, all selection strategies of the parallelized relax-and-fix heuristic give results of similar qualities, with the chronological order providing slightly better reconstructed solutions. We have finally shown that, using a reconstruction strategy to determine feasible solutions based on the partial solutions at each iteration, solutions with very few lost sales can be reached.

As a perspective, it would be interesting to use the solution of the Lagrangian relaxation based heuristic defined in Section 4.4 as an initial solution for the parallelized relax-and-fix heuristic. We have seen that the Lagrangian relaxation based heuristic provided solutions with few lost sales, and the relax-and-fix heuristic could improve even more this solution. We have considered the CLSP with additional ending inventory constraints, however it has been proven that many uncapacitated lot-sizing problems can be solved polynomially using dynamic programming algorithms (minimum or maximum production quantities, inventory bounds...). Hence it would be interesting to adapt the Lagrangian relaxation based heuristic to solve these problems. In Chapter 5, we extend the parallelized relax-and-fix to tackle more generic lot-sizing problems, such as the ones that are faced at DecisionBrain. Additional perspectives are discussed in Chapter 6.

Chapter 5

Industrial application

5.1 Introduction

The previous chapters focused on single-level single-machine multi-item capacitated lot-sizing problems. However, the problems faced at DecisionBrain often have multiple levels and multiple resources. Manufacturing problems modeled as lot-sizing problems can in reality be a combination of lot-sizing, scheduling and routing problems. In order to efficiently solve a large variety of lot-sizing problems, DecisionBrain designed the Planning Engine, an optimization tool using IBM ILOG CPLEX that can model and solve multi-level lot-sizing problems with parallel machines and shared resources. The differences between the models studied so far in this thesis and the lot-sizing problem modeled in the Planning Engine affect the feasibility and the performances of the heuristics proposed in Chapter 4 when solving real instances. One of the main differences is that production decision variables are linked to processes and not items. The available capacity on the resources is consumed by the execution of these processes and each execution of a process corresponds to the production of some items. Another difference is the fact that setup carry-over is allowed on the machines. This means that, if the last item produced on a machine at the end of a period is the same than the first item at the beginning of the next period, then no setup cost and time are required in the next period, which is defined as setup carry-over. Considering setup carry-over affects the parallelized relax-and-fix heuristic because we have implicitly assumed in Chapter 4 that fixing decision variables at one period does not affect the feasibility of the problem. This assumption does not apply when setup carry-over is considered because setup variables are linked between the periods.

In this chapter, we extend the parallelized relax-and-fix heuristic to take into account some of the constraints modeled in the Planning Engine, particularly setup carry-overs. These constraints and how they affect the heuristics are detailed in Section 5.2. We specifically consider the industrial application introduced in Section 1.2 of Chapter 1. Section 5.3 describes the updated version of the parallelized relax-and-fix heuristic. Computational experiments are conducted and discussed in Section 5.4.

5.2 Complex lot-sizing constraints and industrial application

The lot-sizing formulations introduced in the previous chapters mostly model big time bucket problems, where setups are not carried over from one period to the next. However, some of the problems solved by DecisionBrain, including the industrial problem described in Section 1.2, allow setups to be carried from one period to the next. Setup carry-over in lot-sizing problems has been extensively studied (Karmarkar and Schrage (1985), Dillenberger et al. (1993), Dillenberger et al. (1994), Suerie and Stadtler (2003)). Using small time buckets with only one or two setups per period, modeling setup carry-over becomes necessary as the production of an item might last several periods (Sahling et al. (2009)). This section describes some of the complex lot-sizing constraints that apply to the industrial application described in Section 1.2 and that are modeled in the Planning Engine. We explain how each of these constraints can affect the parallelized relax-and-fix heuristics introduced in Chapter 4. We also provide a mathematical formulation of the problem.

5.2.1 Minimum production quantities over several periods

One of the complex constraints of the problem introduced in Section 1.2 consists in considering a minimum production quantity \underline{X}_i for item i that exceeds the available capacity during a single period. We define by δ_{it} the minimum number of periods required to produce the minimum production quantity starting at the beginning of period t . If the available capacity between t and T is not enough to produce quantity \underline{X}_i , then $\delta_{it} = T - t + 1$ and the minimum ordering quantity corresponds to the available capacity between t and T (assuming setup carry-over). The minimum production quantity constraints are modeled as:

$$\sum_{k=0}^{\delta_{it}-1} X_{i,t+k} \geq \min\left(\sum_{k=0}^{\delta_{it}-1} \frac{c_{i,t+k}^{max} - s_{i,t+k}}{b_{i,t+k}}, \underline{X}_i\right)(Y_{it} - Y_{i,t-1}), \quad \forall i \in 1, \dots, N, \quad \forall t \in 1, \dots, T \quad (5.1)$$

$$X_{it} \geq \frac{c_t^{max} - s_{it}}{b_{it}}(Y_{i,t-1} + Y_{i,t+1} - 1), \quad \forall i \in 1, \dots, N, \quad \forall t \in 2, \dots, T - 1 \quad (5.2)$$

Constraints (5.1) model the fact that, if a production starts at t for item i , then at least $\min(\sum_{k=0}^{\delta_{it}-1} \frac{c_{i,t+k}^{max} - s_{i,t+k}}{b_{i,t+k}}, \underline{X}_i)$ units of item i need to be produced in the next δ_i periods. Constraints (5.2) ensure that, if item i is produced for three consecutive periods $t - 1$, t and $t + 1$, then period t needs to produce only item i at full capacity.

This can raise issues because a decision at period t should be taken assuming that there will be enough capacity available to produce \underline{X}_i in the next δ_{it} periods, even if some of these periods were relaxed.

5.2.2 Shared resources and setup states

The Planning Engine developed by DecisionBrain not only considers fixed setup times, through binary variables Y_{it} that model when item i is produced in period t , but each setup requires a specific state u among U possible states for a given set of R resources. A resource set to a specific state can only produce one category of items. Changing the state of a

resource r in period t requires an additional setup time to set the resource in the correct configuration. In the industrial application we are considering, an additional initial state binary variable denoted Z_{rut}^0 is added for each resource r and each period t . In that case, there is no need to configure the machine in state u because the state is the same as the one observed at the end of the previous period. Setup times only occur when there is a change of states for a given resource r at t , that is modeled using the binary variables denoted Z_{rut}^+ .

In the Planning Engine, for each resource $r \in \{1, \dots, R\}$, the following binary decision variables are used:

- $Z_{rut}^0 \in \{0, 1\}$: This variable is equal to 1 if state u is the initial state of resource $r \in \{1, \dots, R\}$ at the beginning of time bucket t , and is equal to 0 otherwise,
- $Z_{rut}^+ \in \{0, 1\}$: This variable is equal to 1 if there is a change to state u on resource $r \in \{1, \dots, R\}$ at time bucket t , and is equal to 0 otherwise,
- $Z_{rut} \in \{0, 1\}$: This variable is equal to 1 if there is a production using state u on resource $r \in \{1, \dots, R\}$ at time bucket t , and is equal to 0 otherwise.

Setup state constraints

Setup times and setup costs at each period t for each resource r only occur if there is a change of states in the period. This implies that the initial setup state does not induce a setup time. The following set of constraints link the Z_{rut}^0 , Z_{rut}^+ and Z_{rut} variables:

$$Z_{rut}^+ \leq Z_{rut}, \quad \forall r \in 1, \dots, R, \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.3)$$

$$Z_{rut} \leq Z_{rut}^0 + Z_{rut}^+, \quad \forall r \in 1, \dots, R, \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.4)$$

$$Z_{rut}^0 + Z_{rut}^+ \leq 1, \quad \forall r \in 1, \dots, R, \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.5)$$

Constraints (5.3) ensure that if the state of resource r changes to u in time bucket t , then u is one of the states of resource r in t . Constraints (5.4) state that if resource r is producing in period t using state u , then it either means that there was a change or that it was the initial state of the resource at the beginning of t . Constraints (5.5) model the fact that state u on resource r cannot be both an initial state and a state that has changed at t .

The continuity between a previous state and the initial state of a resource is expressed as follows:

$$Z_{ru_1, t+1}^0 + Z_{ru_1 t}^0 + Z_{ru_2 t} \leq 2, \quad \forall r \in 1, \dots, R, \forall (u_1, u_2) \in \{1, \dots, U\}^2, \forall t \in 1, \dots, T \quad (5.6)$$

$$Z_{ru, t+1}^0 \leq Z_{rut}^+ + Z_{rut}^0, \quad \forall r \in 1, \dots, R, \forall u \in 1, \dots, U, \forall t \in 2, \dots, T \quad (5.7)$$

Constraints (5.6) model the fact that, if the initial state of resource r remains the same in two consecutive periods t and $t + 1$, then there are no other states in t . Constraints (5.7) ensure that the initial state of resource r at each period is either a state that has changed or the initial state of the previous period.

Additionally, the following constraints imply that there is exactly one initial state for each resource at each period:

$$\sum_{u=1}^U Z_{rut}^0 = 1, \quad \forall r \in 1, \dots, R, \forall t \in 1, \dots, T \quad (5.8)$$

Linking resources to production

We need to link the setup state at the resource level to the setup variables at the production level. A setup for item i can only occur if each used resource r_i is in the correct state u_i .

$$Y_{it} \leq Z_{r_i u_i t}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (5.9)$$

$$\sum_{i \in \{1, \dots, N\}, (r_i, u_i) = (r, u)} Y_{it} \geq Z_{rut}, \quad \forall r \in 1, \dots, R, \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.10)$$

Constraints (5.9) imply that item i can only be produced if the required resource r_i is in the required state u_i . Constraints (5.10) state that, at period t , if a resource r is in a specific state u , at least one item i requiring resource r to be in setup state u is such that $Y_{it} = 1$.

5.2.3 Problem formulation for a single resource

For sake of simplicity, let us consider a single resource problem (P_{PE}). Let us denote by s_{ut}^r (resp. f_{ut}^r) the setup time (resp. setup cost) induced if there is a change to state u at period t .

The formulation of (P_{PE}) is given below:

$$\min \sum_{t=1}^T \left(\sum_{i=1}^N (f_{it} Y_{it} + p_{it} X_{it} + h_{it} I_{it} + l_{it} L_{it}) \right) + \sum_{s=1}^S f_{ut}^r Z_{ut}^+ \quad (5.11)$$

$$I_{i,t-1} + X_{it} + L_{it} = d_{it} + I_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (5.12)$$

$$X_{it} \leq M_{it} Y_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (5.13)$$

$$L_{it} \leq d_{it}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (5.14)$$

$$\sum_{i=1}^N (s_{it} Y_{it} + b_{it} X_{it}) + \sum_{s=1}^S s_{ut}^r Z_{ut}^+ \leq c_t^{max}, \quad \forall t \in 1, \dots, T \quad (5.15)$$

$$\sum_{i \text{ req. } s} Y_{it} \leq N Z_{ut}, \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.16)$$

$$\sum_{i \text{ req. } s} Y_{it} \geq Z_{ut}, \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.17)$$

$$Z_{ut}^+ \leq Z_{ut}, \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.18)$$

$$Z_{ut} \leq Z_{ut}^0 + Z_{ut}^+, \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.19)$$

$$Z_{ut}^0 + Z_{ut}^+ \leq 1, \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.20)$$

$$Z_{u_1, t+1}^0 + Z_{u_1 t}^0 + Z_{u_2 t} \leq 2, \quad \forall (u_1, u_2) \in \{1, \dots, U\}^2, u_1 \neq u_2, \forall t \in 1, \dots, T \quad (5.21)$$

$$Z_{u, t+1}^0 \leq Z_{ut}^+ + Z_{ut}^0, \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.22)$$

$$\sum_{s=1}^S Z_{ut}^0 = 1, \quad \forall t \in 1, \dots, T \quad (5.23)$$

$$Y_{it} \in \{0, 1\}, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (5.24)$$

$$Z_{ut}, Z_{ut}^0 \in \{0, 1\}, \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.25)$$

$$Z_{ut}^+ \in [0, 1], \quad \forall u \in 1, \dots, U, \forall t \in 1, \dots, T \quad (5.26)$$

$$X_{it}, I_{it}, L_{it} \geq 0, \quad \forall i \in 1, \dots, N, \forall t \in 1, \dots, T \quad (5.27)$$

The objective function (5.11) includes the costs of a change in the setup configuration at each period for each state. Only the state change variables Z_{ut}^+ imply a setup cost. In the capacity constraints (5.15), the setup times corresponding to a state change have been added. Constraints (5.16) and (5.17) ensure that the machine state is linked to the production of one of the item requiring this state. Constraints (5.18) to (5.22) link the setup state variables to the initial and change setup state variables. Constraints (5.23) state that the machine is in exactly one configuration at the beginning of each period. Constraints (5.24) to (5.26) define the domains of the variables.

5.2.4 Multi-objective optimization

The Planning Engine supports lexicographic multi-objective optimization. The problem to solve is divided into a set of criteria, that are associated to a lexicographic order. The objective functions are optimized in this order, meaning that at step i all criteria of order i are optimized. Lexicographic optimization allows the different costs to be strongly prioritized. For instance, lost sales might be the main objective to optimize, even though setup and holding costs should be minimized as well. After a given criterion has been optimized, upper bounds are set on the criteria already considered when optimizing the next criterion in the lexicographical order.

The problem considered in this chapter has demands of different importance. We first want to satisfy the most important demands, and then use the remaining capacity on the resources to satisfy the remaining demands. This first criterion to minimize is the number of lost sales for the most important demands. The second criterion to optimize is the number of lost sales for the less important demands.

If we consider a bi-level optimization problem where all lost sales would first be minimized, and then the production, inventory and setup costs would be minimized, the Lagrangian relaxation based heuristic proposed in Section 4.4.2 would be effective because it focuses on finding a solution minimizing the lost sales, but also takes the different costs into consideration when deciding on the quantities to produce in each period.

5.3 Extending the relax-and-fix heuristic

In Section 5.3.1, we show that Algorithm 4.2 cannot be used to solve lot-sizing problems with minimum production quantities and with setup carry-over. In Section 5.3.2, we discuss how the relax-and-fix heuristic can be extended to solve multi-objective lot-sizing problems. In Section 5.3.3, we discuss how some of the infeasibility issues can be dealt with by considering other decomposition of the binary variables. In Section 5.3.4, we propose an updated version of Algorithm 4.2 to solve lot-sizing problems with minimum production quantities and setup carry-over.

5.3.1 Infeasibility issues

Minimum production quantities

Applying the relax-and-fix heuristic described in Section 4.3.2 will not work if the minimum production quantity constraints (Constraints (5.1)) are added, because Algorithm 4.2 might fix setup variables to 1 that will make the problem infeasible even if there is a feasible solution with lost sales.

Let us consider a simple example with $N = 2$, $T = 2$, $\gamma = \delta = 1$, a demand per period of $d_{1t} = 400$, $d_{21} = 0$ and $d_{22} = 200$, a setup time of $s_{it} = 200$, holding costs $h_{it} = 1$, lost sales costs $l_{it} = 100$ and a maximum capacity $c_t^{max} = 1000$ for $i \in \{1, \dots, N\}$ and $t \in \{1, \dots, T\}$. All the other costs are set to 0. We are not considering the setup states on the resources, although each production incurs a setup time (similar to the models described in the previous chapters). In the Planning Engine, M_{it} (the upper bound on the production linking the setup and production variables) is set to $c_t^{max} - s_{it} = 800$. Let us consider the strategy that consists in selecting the subproblem with the lowest objective value and set a minimum production quantity $\underline{X}_1 = 1600$ for item 1.

If we optimize the binary variables for the first period, because of Constraints (5.1), the demand for item 2 cannot be satisfied, leading to an objective value of 200,000. If the binary variables for the second period are optimized, we can satisfy the demands for both item 1 and item 2 and get a solution with an objective value of 0 if we set $(Y_{11}^*, X_{11}^*) = (0.5, 400)$, $(Y_{12}^*, X_{12}^*) = (1, 400)$, $(Y_{21}^*, X_{21}^*) = (0, 0)$ and $(Y_{22}^*, X_{22}^*) = (1, 200)$. Because of Constraints (5.1), if $Y_{11}^* = 0.5$, the minimum production quantity of item 1 between period 1 and period 2, if production starts at period 1, is 800 units instead of 1600 units.

If we fix $Y_{12} = Y_{22} = 1$ for the second iteration, it would not be possible to find a feasible solution satisfying the minimum production quantity because the production for item 1 cannot fully consume the capacity in the second period.

A first solution would be to relax these constraints and only penalize a production quantity that does not satisfy the minimum production quantity for each item. This can raise issues because, even after several iterations of the algorithm, it might not be possible to find a solution satisfying these constraints. Another possible solution is to only fix the non-production decision variables, i.e. $Y_{it}^* = 0$, and allow the binary variables that were set to 1, i.e. $Y_{it}^* = 1$, to be reoptimized. In that case, if a production decision was taken in the previous iterations but the capacity is not sufficient to satisfy the capacity constraints and the minimum production quantities, then previous decisions can be revised.

Setup carry-overs

Let us give an example to show that Algorithm 4.2 of Chapter 4 cannot be used to solve (P_{PE}) . Let us consider an instance with $N = 1$, $T = 3$, $S = 2$, $d_t = 100$, $h_t = 1$, $s_{u_1 t}^r = 800$, $c_t^{max} = 500$ for $t \in \{1, \dots, 3\}$, all other costs or parameters are set to 0. The item that needs to be produced requires setup state u_1 . However, in this instance, the initial setup state at the beginning of the planning horizon is u_0 .

The only feasible solution for this instance is to consider that all demands are lost. However, if we optimize first the binary variables of the third period and relax the binary variables for the other periods, we find the following solution with objective value 0:

- $(X_1^*, Y_1^*, Z_{u_1 1}^{0*}, Z_{u_1 1}^{+*}) = (100, 0.5, 0, 0.5)$
- $(X_2^*, Y_2^*, Z_{u_1 2}^{0*}, Z_{u_1 2}^{+*}) = (100, 0.5, 0.5, 0.5)$
- $(X_3^*, Y_3^*, Z_{u_1 3}^{0*}, Z_{u_1 3}^{+*}) = (100, 1, 1, 0)$

Fixing $Z_{u_1 3}^0 = 1$ leads to an infeasible instance because the setup time required to change the setup state to state u_1 exceeds the production capacity.

5.3.2 Extending the parallelized relax-and-fix to multi-objective optimization

The parallelized relax-and-fix heuristic proposed in Chapter 4 only considers a single objective function. In order to extend the algorithm to a multi-objective optimization setting, different options can be considered.

At each iteration and for each subproblem, we could find a partial solution based on the optimization of all the objective functions. The main issue with this method is that, at each iteration, decisions with regards to the first objective to optimize can be affected by optimizing objectives that are less important in previous iterations. Moreover, some of the score definitions would need to be updated to take into account the fact that multiple objectives were minimized.

Alternatively, we could perform runs of Algorithm 4.2 for each objective function iteratively. In this way, the first feasible solution found after a first run of the algorithm would only have minimized the first objective function without considering the others. We then set an upper bound on this objective and, based on this solution, perform another run of Algorithm 4.2 where only the second objective function is minimized.

Both of these methods still need to be compared.

5.3.3 Other types of decompositions

An option to cope with the feasibility issues that come with splitting the binary decision variables according to the periods would be to consider other types of decompositions.

Another decomposition that has already been studied for the relax-and-fix heuristic to solve a multi-level capacitated lot-sizing problem (Helber and Sahling (2010)) consists in splitting the binary variables by item. With a decomposition per item, Constraints (5.1) are compatible with the relax-and-fix heuristic. However, this raises the concern as to how to consider the setup states of the resource. Indeed, different items can share the same required setup state. Depending on the number of setup states, it would be possible to consider a decomposition per item sharing the same setup state.

As pointed out in Helber and Sahling (2010), the order in which the subset of items are considered in the relax-and-fix heuristic can affect the solution obtained by the heuristic. Further work should be done to find the best way to split the items. In Helber and Sahling (2010), the items are sorted according to their cost in an optimal solution of the LP-relaxation of the problem. Their numerical experiments, especially the comparison with the work of Tempelmeier and Derstroff (1996) and Stadtler (2003), show that large gains can be obtained by considering a decomposition per item.

5.3.4 Updated parallelized relax-and-fix

Based on the discussions in the previous sections, we propose an updated version of Algorithm 4.2 of Chapter 4. The binary variable decomposition is generalized, so that the decomposition can be done by period, item, or a combination of both. It would be an interesting topic to analyze what would be the best way to decompose the binary variables depending on the structure of the problem, especially for large instances. Once the binary variables are split, K subproblems are defined that can be solved in parallel. Then, once the set of binary variables to fix has been picked, for each binary variable of this set:

- If the optimized value of the binary variable is 0, we fix this variable to 0 for the next iterations.
- Otherwise, the variable will be optimized as a binary variable in the next iterations without being fixed as in Algorithm 4.2.

When the problem to solve has both setup carry-over and minimum production quantities, the initial setup state variables are not relaxed. Indeed, if at an iteration, all changes of states variables are set to 0 between period t and period $t' > t$, then the initial setup state binary variables between period $t + 1$ and t' need to be changed.

Algorithm 5.1 First run of the updated parallelized relax-and-fix algorithm

```

 $K, s \in \llbracket 1, C \rrbracket$ 
A partition of  $\llbracket 1, N \rrbracket \times \llbracket 1, T \rrbracket$  into sets  $\mathcal{S}_j^{opt,I}$  for  $j \in \{1, \dots, K\}$ 
A partition of  $\llbracket 1, U \rrbracket \times \llbracket 1, T \rrbracket$  into sets  $\mathcal{S}_j^{opt,U}$  for  $j \in \{1, \dots, K\}$ 
 $\mathcal{S}^{fix,I} \leftarrow \emptyset, \mathcal{S}^{fix,U} \leftarrow \emptyset$ 
 $\mathcal{J}_1 = \{1, \dots, K\}$ 
for  $k \in \llbracket 1, K \rrbracket$  do
  for  $j \in \mathcal{J}_k$  do
     $\mathcal{S}_{jk}^{fix,I} \leftarrow \mathcal{S}^{fix,I} \setminus \mathcal{S}_j^{opt,I}, \mathcal{S}_{jk}^{rel,I} \leftarrow \llbracket 1, N \rrbracket \times \llbracket 1, T \rrbracket \setminus (\mathcal{S}_{jk}^{fix,I} \cup \mathcal{S}_j^{opt,I})$ 
     $\mathcal{S}_{jk}^{fix,U} \leftarrow \mathcal{S}^{fix,U} \setminus \mathcal{S}_j^{opt,U}, \mathcal{S}_{jk}^{rel,U} \leftarrow \llbracket 1, U \rrbracket \times \llbracket 1, T \rrbracket \setminus (\mathcal{S}_{jk}^{fix,U} \cup \mathcal{S}_j^{opt,U})$ 
    Solve  $(P_{PE})_{jk}$  where:
    -  $Y_{it} \leftarrow 0$  if  $\overline{Y_{kit}} = 0$  for  $(i, t) \in \mathcal{S}_{jk}^{fix,I}$ 
    -  $Z_{ut} \leftarrow 0$  if  $\overline{Z_{kut}} = 0$  for  $(u, t) \in \mathcal{S}_{jk}^{fix,U}$ 
    -  $Y_{it} \in [0, 1]$  for  $(i, t) \in \mathcal{S}_{jk}^{rel,I}$ 
    -  $Z_{ut} \in [0, 1]$  for  $(u, t) \in \mathcal{S}_{jk}^{rel,U}$ 
     $(Y_j^*, Z_j^*)$  an optimal solution
    Set  $v_{jks}$  as defined in Section 4.3.2
  end for
   $j_k \leftarrow \operatorname{argmin}_{j \in \mathcal{J}_k} v_{jks}$ 
  for  $(i, t) \in \mathcal{S}_{j_k}^{opt,I}$  do
     $\overline{Y_{k+1,it}} \leftarrow Y_{j_k it}^*$ 
  end for
  for  $(u, t) \in \mathcal{S}_{j_k}^{opt,U}$  do
     $\overline{Z_{k+1,ut}} \leftarrow Z_{j_k ut}^*$ 
  end for
   $\mathcal{S}^{fix,I} \leftarrow \mathcal{S}^{fix,I} \cup \mathcal{S}_{j_k}^{opt,I}$ 
   $\mathcal{S}^{fix,U} \leftarrow \mathcal{S}^{fix,U} \cup \mathcal{S}_{j_k}^{opt,U}$ 
   $\mathcal{J}_{k+1} \leftarrow \mathcal{J}_k \setminus \{j_k\}$ 
end for

```

We can define a number of runs per objective in a multi-objective problem. During the first run(s), only the first objective is optimized. At the end of the first run, a first feasible solution is obtained. During the following runs, the other objectives are optimized in lexicographic order and an upper bound is set on the first objective, similarly to what is currently being done in the Planning Engine.

5.4 Computational experiments

Computational experiments were performed on a set of 6 industrial instances denoted II_i for $i \in \{1, \dots, 6\}$ of the problem introduced in Section 1.2. Although the mathematical model introduced in Section 5.2.3 has been written for a single resource, there are multiple resources in the industrial instances, and each item requires a set of resources to be in a specific setup state. Each demand has a high or low priority. Two criteria are minimized in a lexicographic order: The total lost sales of demands with a high priority are first minimized, and then the total lost sales of demands with a low priority.

Let us compare the results obtained using Algorithm 5.1 and the standard solver IBM ILOG CPLEX 12.10. The maximum optimality gap is set to 0.1%, and a time limit of 30 seconds is set to solve each subproblem. To solve each instance, Algorithm 5.1 runs four times: (1) Two times where only the first objective function is optimized and (2) Two times where both objective functions are optimized in lexicographic order. The binary decision variables are split in $\{4, 6, T\}$ sets using a maximum number of 6 threads. The splitting of the binary decision variables is alternatively time-based or item-based, similarly to Helber and Sahling (2010).

Instances	T	N	R	U
II_1	14	110	79	10
II_2	13	110	79	10
II_3	15	97	67	8
II_4	14	105	77	10
II_5	14	99	77	9
II_6	14	99	77	9

Table 5.1: Instance characteristics

In the considered industrial problem, the resources are either the machines that mold the items, or the molds that can be shared between different machines. Each change of state on a machine induces a setup time that lasts for half of a period, implying that at most two different items can be produced on each machine at each period. Table 5.1 shows, for each instance, the number of periods T , the number of items N , the number of resources R and the average number of states per resource U . The six instances have similar sizes. The number of periods varies between 13 for instance II_2 to 15 for instance II_3 , and the number of resources varies between 97 items for instance II_3 and 110 for instances II_1 and II_2 . On average, around 75 resources are available, each with about 9 possible setup states.

Instances	CPLEX				
	Obj. 1	Obj. 2	LB 1	LB 2	T (s)
II_1	1,691	0	1,691	0	0
II_2	0	0	0	0	1
II_3	871	0	871	0	14
II_4	51,500	152,580	51,500	145,592	300
II_5	18,285	0	17,540	0	300
II_6	6,500	135,096	6,500	124,862	300

Table 5.2: Numerical results using IBM ILOG CPLEX (maximum CPU time of 300 seconds per objective)

Table 5.2 shows the objective values and the lower bounds given by IBM ILOG CPLEX after a time limit of 300 seconds. Note that Instances II_1 , II_2 and II_3 are easy to solve and an optimal solution is found within less than 14 seconds. Instances II_4 and II_6 are the only instances that have demands of both high and low priorities, hence two objective functions to minimize in a lexicographic order. An optimal solution for the first objective function is found within a few seconds. This is not the case for the second objective function which is not solved to optimality for both instances (for Instance II_6 , the best feasible solution has a second objective value of 135,096 while the lower bound is 124,862). Instance II_5 has no demand of low priority, but an optimal solution for the demands of high priority is not found within the time limit.

Instances	Chronological		CPLEX		T (s)	Most integer		CPLEX		T (s)
	Obj. 1	Obj. 2	Obj. 1	Obj. 2		Obj. 1	Obj. 2	Obj. 1	Obj. 2	
II_1	1,691	0	1,691	0	11	1,691	0	1,691	0	28
II_2	0	0	0	0	10	0	0	0	0	24
II_3	871	0	871	0	47	873	0	871	0	79
II_4	51,500	153,970	51,500	154,847	19	51,500	156,381	51,500	158,107	49
II_5	18,285	0	18,285	0	90	18,485	0	18,285	0	95
II_6	6,500	134,957	6,500	271,980	31	6,500	140,943	6,500	147,980	55

Table 5.3: Numerical results with 4 intervals

Instances	Chronological		CPLEX		T (s)	Most integer		CPLEX		T (s)
	Obj. 1	Obj. 2	Obj. 1	Obj. 2		Obj. 1	Obj. 2	Obj. 1	Obj. 2	
II_1	1,691	0	1,691	0	23	1,691	0	1,691	0	44
II_2	0	0	0	0	21	0	0	0	0	41
II_3	871	0	871	0	78	871	0	871	0	126
II_4	51,500	155,767	51,500	154,270	33	51,500	155,270	51,500	154,037	90
II_5	18,485	0	18,285	0	116	18,285	0	18,285	0	179
II_6	6,500	137,400	6,500	286,580	48	6,500	141,167	6,500	235,337	130

Table 5.4: Numerical results with 6 intervals

Instances	Chronological		CPLEX		T (s)	Most integer		CPLEX		T (s)
	Obj. 1	Obj. 2	Obj. 1	Obj. 2		Obj. 1	Obj. 2	Obj. 1	Obj. 2	
II_1	1,691	0	1,691	0	65	1,691	0	1,691	0	100
II_2	0	0	0	0	48	0	0	0	0	86
II_3	871	0	871	0	147	925	0	871	0	231
II_4	51,500	155,163	51,500	160,514	58	51,500	156,146	51,500	153,597	182
II_5	18,485	0	18,285	0	268	18,685	0	18,285	0	301
II_6	6,500	139,790	6,500	204,910	99	6,500	137,033	6,500	141,023	223

Table 5.5: Numerical results with T intervals

Tables 5.3, 5.4 and 5.5 show the results obtained for the two objective functions after using four iterations of Algorithm 5.1, by dividing the binary variables in respectively 4, 6 and T sets using either the *Chronological* strategy or the *Most integer* strategy, as defined in Chapter 4, to select the best set of binary variables to fix at each iteration. At each

iteration of Algorithm 5.1, the binary variables are divided by sets of periods or by sets of items alternatively. The corresponding objective values of the best solution obtained by IBM ILOG CPLEX with a time limit that is equal to the running time $T(s)$ of the heuristic are also displayed.

For Instances II_1 and II_2 , an optimal solution is obtained by both the heuristics and IBM ILOG CPLEX. However, IBM ILOG CPLEX solves these instances in less than 1 second while the heuristic takes at least 10 seconds (when the binary variables are divided into 4 sets). When selecting the binary variables based on the chronological order, Instance II_3 is always solved to optimality by the updated relax-and-fix heuristic. This is not the case when the selection is based on the most integer relaxed solution (for instance an objective value of 925 is obtained when the binary variables are split in T sets, while the optimal value is 871). For Instance II_6 , the hardest instance to solve by IBM ILOG CPLEX, the heuristic always provides better results than IBM ILOG CPLEX within the same time limit (134,957 for a selection of binary variables based on the chronological order and 4 sets, while IBM ILOG CPLEX finds a best upper bound for the second objective function of 271,980). When the selection of binary variables is based on the chronological order, the running times as well as the objective values are better as the number of sets dividing the binary variables decreases (for Instance II_6 , the second objective value is 134,957 and the running time is 31 seconds when 4 sets are considered, while the second objective value is 139,790 and the running time is 99 seconds when T sets are considered). When the selection is based on the most integer relaxed solution, this observation does not seem to apply. As an example, the best solution for Instance II_6 is obtained when the binary variables are divided in T sets with a second objective value of 137,033.

5.5 Conclusion and perspectives

In this section, we have adapted the parallelized relax-and-fix heuristic introduced in Chapter 4 to solve an industrial lot-sizing problem with setup carry-over and minimum production quantities that span over several periods. We have shown that the relax-and-fix algorithms (Algorithms 4.2 and 5.1) can lead to infeasibility issues when linking constraints between periods are added to the model. As an alternative, we have proposed an adaptation of Algorithm 5.1 where not all binary variables are fixed. The heuristic was implemented on the Planning Engine of DecisionBrain and tested on a set of real instances. For these instances, the heuristic returned feasible solutions that are close to the ones returned by IBM ILOG CPLEX within the same time limit.

These results are just preliminary results, and more tests of the updated parallelized relax-and-fix heuristic should be performed on instances that are hard to solve by IBM ILOG CPLEX. Future works also have to be done on the way binary variables are split and fixed. Finally, it would be interesting to test the proposed heuristic on another production planning problem, with additional constraints or levels.

Chapter 6

Conclusions and perspectives

6.1 Conclusions

In the first part of this manuscript, we have considered how we could improve the modeling of the CLSP with setup times and lost sales to better take into account inventory management. We first focused in Chapter 2 on the end-of-horizon effect caused by the ZIO property. We have shown that, for capacitated lot-sizing problems, simply extending the planning horizon is not sufficient to compensate the end-of-horizon effect. This is why we proposed to add both a minimum ending inventory constraint and a maximum ending inventory constraint per item to model the fact that, at the end of the planning horizon, there should be enough stock to satisfy some of the future demands. To define these ending inventory levels, we have studied a simpler cyclical lot-sizing problem. From this problem, we derived the cycle length for each item. We have shown that these additional constraints indeed mitigated the end-of-horizon effect, and that this was especially useful on a rolling horizon. We then focused in Chapter 3 on the impact of discretizing the time when solving lot-sizing problems with inventory bounds. We have shown that, when the inventory is constrained, not considering the evolution of demand and production within periods can lead to infeasibility. To consider the inventory evolution, we have proposed two models with specific assumptions on the production and demand rates of items, and we have shown that these models help to better respect the inventory constraints when optimizing production plans.

The numerical results in Chapters 2 and 3 showed that the new models we have proposed are numerically harder to solve. Thus the second part of the manuscript was devoted to finding generic ways to solve lot-sizing problems, using parallelization to fasten the optimization process. We have proposed in Chapter 4 two adaptations of heuristics used to solve lot-sizing problems: A parallelized version of the relax-and-fix algorithm and a heuristic based on Lagrangian relaxation. We have shown that, for the classical formulation of the problem, these heuristics could provide better results than when using the original version of the relax-and-fix algorithm. Finally, in Chapter 5, we have extended the relax-and-fix algorithm to tackle more generic lot-sizing problems, such as the ones that are solved at DecisionBrain. We studied a specific problem with setup carry-over and proposed a new version of the parallelized relax-and-fix algorithm.

6.2 Perspectives

This section proposes various extensions of the studies presented in the manuscript. These perspectives are classified based on the time we believe they require to be investigated.

6.2.1 Short-term perspectives

Lot-sizing problems with seasonal demands

We could extend Chapter 2 by considering variations of the demand outside the planning horizon, e.g. a seasonal demand. The analysis of the minimum and maximum ending inventories was based on the average costs and demands. We could define these parameters by weighing part of the costs and demands outside of the planning horizon. For instance, the average demand \bar{d} used in Section 2.3 could be defined as:

$$\bar{d} = \alpha \bar{d}^{in} + (1 - \alpha) \bar{d}^{out},$$

with $\alpha \in [0, 1]$. \bar{d}^{in} corresponds to the average demand in the planning horizon and \bar{d}^{out} a forecasted demand after the planning horizon. By performing numerical analysis varying the values of α , a relevant weight could be determined for the demand after the planning horizon.

Studies on cyclical lot-sizing problems

Another perspective is to use cyclical lot-sizing problems to perform cost analysis. We have shown in Chapter 2 that we can analytically find the optimal solution for the relaxed problem proposed in Section 2.2 and get an average number of setups per period in an optimal solution based on the capacity and the costs. We can use these formulas to study the relevance of the instances, especially the consistency between the capacity and the costs. Some parameters, such as the lost sales costs, are in practice hard to define because attributing a monetary value for these parameters is not straightforward. Features such as lost sales are usually defined in order to guarantee feasibility of the problem during the optimization process, but should ideally be avoided. The analysis of the relaxed problem with lost sales in Section 2.3.4 provided a value for lost sales costs based on the other parameters that would imply an optimal solution without lost sales for the relaxed cyclical problem. It would be interesting to study if this analysis can be applied for other costs or with additional constraints, such as a limited storage or production capacity.

Models for other production and demand approximations

Several extensions could be considered with regards to the modeling of inventory evolution within periods of Chapter 3. Indeed, there are only few references on using production and demand rates for lot-sizing problems and on the links between lot-sizing and scheduling problems. As shown in Section 3.6, other shapes of production and demand rates could be considered and modeled. Further experiments could be performed to test these new models. Tests on industrial instances, especially when inventory management is important, should be performed as well.

Replace lost sales with overtimes

We could also consider overtimes instead of lost sales to deal with infeasibility issues in all our models. This could provide more flexibility on the user's side in case extra capacity can be provided and will reduce the number of variables in our models. We simply need to add new variables o_t to model the overtime in period t and replace the capacity constraints (2.3) of the model in Section 2.2.1 by:

$$\sum_{i=1}^N (s_{it}Y_{it} + b_{it}X_{it}) \leq c_t^{max} + o_t, \quad \forall t \in 1, \dots, T \quad (6.1)$$

$$o_t \geq 0, \quad \forall t \in 1, \dots, T \quad (6.2)$$

and add overtime penalties on the objective function. It would be interesting to analyze if the parallelized relax-and-fix heuristic proposed in Chapter 4 performs differently. For the Lagrangian relaxation based heuristic proposed in Section 4.4, we could remove the last step that consists in converting the overtime at each period into lost sales. The method proposed to convert overtimes to lost sales could be used to derive good feasible solutions to the problem with lost sales using feasible solutions to the problem with overtime.

6.2.2 Mid-term perspectives

Modeling other constraints to mitigate the end-of-horizon effect

In Chapter 2, we have shown that, for a capacitated lot-sizing problem, extending the length of the planning horizon was not a guarantee that the optimal solution would not be affected by the end-of-horizon effect. This is a phenomenon that has not been studied in the literature for capacitated lot-sizing problems, even though it can significantly impact production plans on a rolling horizon. The minimum ending inventory constraints we proposed were a way to partially cope with this effect. However, other ways to enforce a minimum production quantity in the last periods could have been implemented. It would be interesting to try to mitigate the end-of-horizon effect using another method. We could, as in Fisher et al. (2001) for an uncapacitated lot-sizing problem, add a negative cost to the ending inventory in order to avoid the ZIO property. However, the penalization used in Fisher et al. (2001) leads to a non-linear objective function and only applies on uncapacitated lot-sizing problems. It would be a challenge to find relevant costs for the CLSP. We could alternatively set lower bounds on the capacity utilization or set a minimal total production quantity in the last periods.

Solution approaches for lot-sizing problems with inventory constraints within periods

We have seen in Chapter 3 that the new models are harder to solve than the CLSP because of the additional inventory constraints. It could be interesting to:

1. Develop a Lagrangian relaxation based heuristic for these new models to figure out if an adequate penalization of the variables in the objective function can provide good solutions with regards to the satisfaction of the inventory constraints within each period. We have shown in Section 3.4.2 that the model with instantaneous demand can be polynomially solved when the capacity constraints are relaxed.

2. Analyze if only a subset of these constraints are sufficient to guarantee the feasibility of the problems under the production and demand assumptions. It would be interesting to implement a method based on constraints generation that would sequentially generate a subset of the inventory constraints.

Adapt our heuristics to solve complex lot-sizing problems

With regards to the industrial applications, the methods developed in Chapter 4 could be generalized as a framework to solve complex lot-sizing problems of any sort using parallelization to fasten the optimization process. This framework could be implemented in the Planning Engine of DecisionBrain. The dynamic programming method used to solve the uncapacitated problem in the Lagrangian relaxation based heuristic can be extended to consider the additional constraints that were not implemented in the model defined in Section 4.2.1. The generalized version of the parallelized relax-and-fix algorithm has the advantage to cope with a wide variety of lot-sizing problems, and can easily be adapted by changing the selection strategy or the decomposition method used to split the boolean variables.

We have only tested our model on one industrial application and it would be relevant to extend these tests to other industrial applications. As observed in Chapter 5, constraints that were not considered when developing a heuristic can affect the performances of the method. In order to develop a heuristic that solves generic lot-sizing problems, it is necessary to perform tests on several types of industrial lot-sizing problems.

6.2.3 Long-term perspectives

Links between lot-sizing and scheduling

In Chapter 3, we have modeled the evolution of inventory based on production and demand rates, which are parameters that are usually associated to scheduling problems. The correlation between lot-sizing and scheduling problems is a topic that, even though is known in literature (Dauzère-Pérès and Lasserre (1994), Dauzère-Pérès and Lasserre (2002), Gomez Urrutia et al. (2014), Wolosewicz et al. (2015)) with various integrated models, would be interesting to further explore. Decisions made at a tactical level when solving lot-sizing problems are used as inputs of scheduling problems. Being able to provide more information on the lot-sizing problem without significantly increasing the complexity of the lot-sizing or scheduling problems should help to determine better schedules and is an important industrial topic.

General decomposition approach using parallelization

Following Chapter 4, the use of parallelization to fasten and improve the performances of heuristics is also an interesting topic to explore, within and outside the field of manufacturing. For instance, the basic idea behind the parallelized relax-and-fix algorithm could be adapted to a wide range of optimization problems with binary variables. By splitting the binary variables into different sets, we could potentially use or adapt the algorithm proposed in Section 5.3.4.

Appendix A

Extended summary in French

Chapitre 1 : Contexte industriel et état de l'art

Résoudre un problème de planification de la production (*lot-sizing*) consiste à déterminer un plan de production qui permet de satisfaire un ensemble de demandes connues dans un horizon temporel discrétisé en périodes de temps (typiquement un jour ou une semaine). Ce plan de production doit entre autres minimiser les coûts de production, de stock et de setup. Les problèmes de *lot-sizing* sont très courants en logistique en raison de l'intérêt économique lié au fait de prendre des décisions de production rationnelles. Durant cette thèse nous nous sommes intéressés au problème de *lot-sizing* avec capacité limitée (CLSP), temps de setup et ventes perdues. La littérature en *lot-sizing* est très riche et a été l'objet de nombreuses recherches depuis le premier article de [Harris](#) (1913). Plus particulièrement, les problèmes avec plusieurs produits et avec une capacité de production limitée par période sont particulièrement intéressants pour modéliser des problèmes industriels. Certains problèmes de *lot-sizing* peuvent être résolus optimalement de manière polynomiale à l'aide d'algorithmes basés sur la programmation dynamique. Cependant, il est souvent nécessaire d'utiliser des heuristiques pour résoudre les problèmes les plus complexes, en particulier les problèmes avec contraintes de capacité de production. Parmi les heuristiques possibles, nous nous sommes concentrés durant cette thèse sur les heuristiques basées sur la décomposition d'un problème en un ensemble de sous-problèmes plus simples (*relaxation lagrangienne*, *relax-and-fix*, ...). Cette thèse s'inscrit dans une collaboration CIFRE entre Mines Saint-Etienne et DecisionBrain, une entreprise spécialisée dans les solutions en optimisation, et plus particulièrement dans le domaine de la planification de la production. Les problèmes que doivent résoudre DecisionBrain incluent souvent des contraintes particulièrement complexes intégrant à la fois des concepts de planification de la production et d'ordonnancement. L'outil de planification optimisée développé par DecisionBrain (*Planning Engine*) permet de modéliser et de résoudre des problèmes génériques de planification de la production. L'objectif principal de cette thèse a été de prendre en compte des aspects industriels souvent négligés dans la littérature, et d'enrichir les modèles existants par de nouvelles contraintes permettant d'obtenir des plans de production de meilleure qualité. Nous avons donc étendu dans les Chapitres 2 et 3 le CLSP avec temps de setup et ventes perdues afin de prendre en compte des aspects importants dans le milieu industriel, notamment en ce qui concerne une meilleure gestion des stocks. Nous avons ensuite proposé dans le Chapitre 4 deux approches de décomposition pouvant être accélérées à l'aide de techniques de parallélisation. Enfin, nous avons étendu dans le Chapitre 5 l'une de ces heuristiques afin de résoudre un problème industriel avec contraintes de continuité dans les setups.

Chapitre 2 : Problème de lot-sizing avec capacité, temps de setup et stock final minimum et maximum

La prise en compte des stocks et leurs évolutions ont été au coeur des enjeux de modélisation de cette thèse. Dans le Chapitre 2, nous analysons le fait que le CLSP considère habituellement un stock nul en fin d’horizon, car il est sous-optimal de produire pour satisfaire des demandes au-delà de l’horizon considéré. Avoir un stock nul en fin d’horizon peut poser problème car la nature des solutions obtenues n’est pas compatible avec une optimisation en horizon glissant, alors que ce type d’optimisation est classique lors de la résolution de problèmes industriels. Les effets de fin d’horizon peuvent influencer les décisions prises durant les premières périodes de l’horizon, et nous avons montré que cela était le cas même pour des horizons temporels très longs en raison d’une mauvaise gestion de la capacité.

Pour compenser les effets de fin d’horizon, nous proposons d’ajouter une nouvelle contrainte de stock final minimal global ainsi que de nouvelles contraintes de stock maximal par produit. Les valeurs des bornes du stock ont été déduites d’une analyse d’un problème cyclique similaire au CLSP avec temps de setup dont la relaxation linéaire peut être résolue de manière analytique. En plus de nous fournir des informations sur une longueur de cycle optimale par produit, cette analyse nous a permis de définir et d’introduire deux nouveaux indicateurs basés sur le nombre moyen de setups par période dans une solution optimale théorique. Le premier indicateur (k_{cost}) se base sur les coûts moyens, tandis que le second indicateur (k_{capa}) se base sur la capacité moyenne consommée par période. Ces indicateurs étendent la notion de Time Between Order introduite par Harris (1913), permettant de déterminer le nombre optimal de périodes entre deux productions successives.

Nous avons utilisé ces indicateurs pour définir un stock minimal global I^{inf} ainsi qu’un stock maximal I_i^{sup} pour chaque produit $i \in \llbracket 1, N \rrbracket$, où N est le nombre de produits. Ces bornes sur les stocks ont été utilisées pour proposer un nouveau schéma de génération d’instances, adapté du schéma de Trigeiro et al. (1989), qui permet de créer des instances plus pertinentes d’un point de vue industriel et moins soumises aux effets de fin d’horizon. L’étude d’autres modèles de lot-sizing cycliques a permis d’étendre les analyses sur les longueurs optimales de cycles à des problèmes plus génériques, notamment dans le cas où la demande moyenne varie selon les produits. L’un de ces modèles permet aussi de proposer une façon de définir des coûts de ventes perdues de manière plus pertinente.

Les résultats numériques montrent que l’ajout de bornes sur les stocks finaux permet de limiter les effets de fin d’horizon, en particulier dans une résolution en horizon glissant. Les résultats numériques montrent aussi que les nouvelles instances générées sont moins soumises aux effets de bord que les instances créées par Trigeiro et al. (1989). Une extension de ce chapitre serait d’utiliser les nouveaux indicateurs k_{capa} et k_{cost} pour définir des coûts ainsi que des capacités adéquates se basant sur les analyses théoriques du problème relâché, notamment pour certains coûts étant en pratique particulièrement difficiles à évaluer. Ces indicateurs pourraient aussi permettre d’analyser si la capacité est bien dimensionnée par rapport au plan de production optimal induit par les coûts. Il serait aussi intéressant de définir des niveaux de stock en fin d’horizon dépendant de projections futures sur les demandes ainsi que sur la capacité moyenne disponible à chaque période. Ainsi, ce stock final pourrait permettre de s’adapter à des futures modifications dans les ressources ou dans les demandes.

Chapitre 3 : Problème de lot-sizing avec capacité et contraintes de stock à l'intérieur des périodes

Toujours dans le but d'avoir une meilleure gestion opérationnelle des stocks, dans le Chapitre 3 nous nous intéressons à l'évolution des stocks à l'intérieur des périodes. Cet aspect est particulièrement important lorsque les valeurs des stocks sont bornées, car les contraintes sur les stocks peuvent être particulièrement importantes pour certaines applications. Avoir un stock de sécurité permet de se protéger en cas de rupture dans la production, tandis qu'un stock maximal peut être nécessaire pour prendre en compte des considérations physiques, telles que la taille des entrepôts. Deux cas concrets rencontrés à DecisionBrain illustrent l'importance de considérer les évolutions de stock à l'intérieur des périodes.

- Lorsque les périodes sont très longues, de l'ordre de la semaine ou du mois : négliger l'évolution du stock à l'intérieur de la période peut être problématique car les variations de stock peuvent être particulièrement importantes. Il est par exemple possible, si les périodes sont suffisamment longues, que toute la demande d'un produit soit consommée avant que sa production ne démarre à l'intérieur d'une période donnée. Dans ce cas précis, tenir compte de la production dans le calcul du stock minimum atteint peut autoriser des solutions entraînant en pratique des déficits de stock à l'intérieur d'une période.
- Lorsque la production doit être synchronisée sur plusieurs niveaux : la fabrication d'un produit s'effectue souvent sur plusieurs niveaux, et la production d'un atelier à un niveau sert à alimenter les demandes d'un atelier à un second niveau. Dans ce cas, il est très important d'être sûr qu'il y ait suffisamment de stock quand les produits changent d'ateliers, ce qui peut arriver au milieu d'une période.

Les évolutions dynamiques de stock et les modélisations par taux de production et de demande sont très peu analysées dans les problèmes de lot-sizing alors qu'elles sont particulièrement importantes dans les problèmes d'ordonnancement. Sous des hypothèses spécifiques de modélisation d'évolution de la demande et de la production à l'intérieur des périodes, nous avons proposé de nouvelles contraintes permettant de limiter les excès et les déficits de stock lors de l'ordonnancement du plan de production obtenu après résolution du problème de lot-sizing.

Dans un premier modèle, la demande est approchée par deux droites. La première droite représente une demande uniforme au plus tôt, et la seconde droite représente une demande uniforme au plus tard. Un des avantages de cette approximation est qu'elle permet d'approcher de manière assez précise et simple des demandes ayant des formes très génériques. Elle apporte aussi de la robustesse dans le cas où la demande n'est pas exactement connue, sans excessivement contraindre le problème. On suppose dans cette modélisation que la production a lieu de manière uniforme au cours de la période, une fois que le temps de setup a été effectué. Cette approximation est très simple et permet de modéliser linéairement les stocks minimal et maximal atteints à chaque période.

Dans un second modèle, la demande est considérée comme instantanée et la production a lieu à chaque période avec un taux de production maximal. Alors que le premier modèle est plus pertinent pour résoudre des problèmes de lot-sizing où les mailles temporelles sont petites, ce second modèle est plus pertinent lorsque les mailles temporelles sont grandes et que la production ne peut être considérée comme uniforme à l'intérieur de chaque période. Dans ce second modèle, chaque variable de décision associée à un setup ou à une production

est décomposée en deux variables. Ces nouvelles variables correspondent à une production avant et après le moment où la demande est satisfaite, pour chaque produit et à chaque période. Sous ces hypothèses sur la production et la demande, le stock maximal (resp. minimal) est atteint juste après (resp. avant) le moment où une production a lieu durant la période.

Les résultats numériques montrent que ces nouvelles contraintes peuvent permettre de construire des plans de production respectant davantage les bornes sur les stocks. Ces plans seront donc moins susceptibles à des modifications ultérieures au niveau opérationnel. Pour comparer les différents modèles avec ajout de contraintes sur les stocks, chaque solution est utilisée pour créer un ordonnancement des lots de production minimisant les excès et les déficits de stock à chaque période. Le second modèle permet de créer des plans de production respectant davantage les bornes sur les stocks à l'intérieur de chaque période, car il apporte plus d'informations par l'ajout de nouvelles variables.

Une question théorique ouverte soulevée dans ce chapitre concerne la complexité du modèle où la production est uniforme dans la période lorsque les contraintes de capacité sont relâchées. Il serait aussi intéressant de trouver des modélisations linéaires du CLSP prenant en compte d'autres hypothèses sur les évolutions de la production et de la demande à l'intérieur des périodes. Nous avons proposé d'autres modélisations de l'évolution des stocks à l'intérieur des périodes lorsque les stocks sont bornés, notamment avec une modélisation où la demande peut être satisfaite avec du retard, mais aussi si les contraintes sur les stocks concernent le stock total et non le stock par produit.

Les nouveaux modèles étant plus difficiles à résoudre, la seconde partie de la thèse s'est concentrée sur l'implémentation de méthodes de résolution génériques, et plus particulièrement de méthodes décomposant un problème complexe en un sous-ensemble de problèmes plus simples pouvant être résolus en parallèle. Le but est de réduire la taille des problèmes à résoudre et d'utiliser les outils disponibles à DecisionBrain. Comme le chapitre sur l'état de l'art l'a montré, les heuristiques basées sur des relaxations partielles des variables booléennes (*relax-and-fix*, *fix-and-optimize*), ainsi que les heuristiques basées sur la relaxation lagrangienne font parties des heuristiques les plus utilisées pour résoudre des problèmes de lot-sizing.

Chapitre 4 : Approches de décompositions parallèles pour le problème de lot-sizing avec capacité

Dans le Chapitre 4, nous avons adapté et étendu des heuristiques de décomposition en les parallélisant pour résoudre le CLSP avec temps de setup, ventes perdues et contraintes de stock final.

Une première heuristique étend l'heuristique de relax-and-fix en proposant une version qui résout à chaque itération plusieurs sous-problèmes en parallèle. L'heuristique de relax-and-fix est une heuristique itérative consistant à décomposer l'horizon temporel en un ensemble d'intervalles où chaque intervalle contient une partie des variables binaires du modèle (e.g. les variables de setup). A chaque itération, les variables binaires d'un des intervalles sont optimisées, tandis que les autres variables binaires sont soit relâchées, soit fixées à des valeurs déterminées à des itérations précédentes. Les solutions partiellement relâchées obtenues à chaque itération sont utilisées pour fixer une partie des variables binaires dans les itérations futures. Le relax-and-fix parallélisé que nous proposons diffère du relax-and-fix classique par dans lequel les intervalles temporels sont fixés. Au lieu d'optimiser les intervalles de façon chronologique, comme c'est le cas pour l'heuristique classique, divers intervalles temporels sont optimisés en parallèle à chaque itération afin de sélectionner l'intervalle le plus pertinent

à fixer pour les itérations futures. Différentes stratégies de sélection du meilleur intervalle à fixer à chaque itération ont été comparées en définissant un score pour chaque solution partielle en se basant sur:

1. Une sélection chronologique, qui correspond à la version classique de l'heuristique de relax-and-fix,
2. Un score aléatoire, afin de servir de point de comparaison avec les autres stratégies de sélection,
3. La fonction objectif de la solution au sous-problème, dans l'idée que la meilleure solution à chaque itération fournira la meilleure solution finale,
4. Le nombre de variables binaires relâchées ayant une valeur optimale entière, dans l'idée que la solution la plus entière sera la moins affectée dans les itérations futures.

Une stratégie de sélection additionnelle consiste à reconstruire à chaque itération une solution réalisable pour le problème à partir de chaque solution partielle. L'heuristique de reconstruction est une heuristique de relax-and-fix où une partie des variables binaires est fixée à des valeurs calculées durant les itérations précédentes. Cette heuristique peut permettre de trouver de bonnes solutions réalisables dès les premières itérations de l'algorithme.

La seconde heuristique de décomposition considérée est basée sur de la relaxation lagrangienne, comme cela est classiquement le cas pour les problèmes de lot-sizing avec capacité de production. Dans cette heuristique, les contraintes de capacité de production à chaque période sont relâchées. Contrairement au CLSP classique, le problème sans contraintes de capacité contient toujours la contrainte de stock minimal global liant les stocks finaux de chaque produit. Nous avons montré que le problème de lot-sizing sans capacité avec contraintes de stock final est NP-difficile dans le cas général, mais peut être résolu en temps polynomial si les stocks minimaux par produit sont identiques. L'algorithme sous-jacent décompose le problème initial en un ensemble de sous-problèmes à un seul produit, avec un stock final imposé et sans contraintes de capacité de production. Chacun de ces sous-problèmes peut être résolu par un algorithme de programmation dynamique. Tous ces sous-problèmes sont indépendants et leur résolution peut donc être parallélisée. Dans le cas où le stock maximal varie selon les produits, nous avons proposé une heuristique pouvant aussi être accélérée à l'aide de la parallélisation. Une fois le problème relâché résolu, l'heuristique lagrangienne proposée reconstruit une solution réalisable pour le problème avec contraintes de capacité à partir de la solution obtenue lorsque les contraintes de capacité sont relâchées. Les heuristiques classiques de reconstruction dans une heuristique lagrangienne pour résoudre des problèmes de lot-sizing utilisent en général deux phases de lissage :

- Une première passe en amont, qui consiste à retarder des quantités de production dans les périodes futures,
- Une seconde passe en aval, qui consiste à avancer des quantités de production dans les périodes passées.

Ces phases de lissage ont pour but de libérer de la capacité durant les périodes contenant de l'excès de capacité, que l'on définit comme l'excès de capacité nécessaire afin de satisfaire le plan de production obtenu après avoir résolu le problème sans contraintes de capacité. Comme notre problème contient aussi une contrainte de stock final minimal global liant les produits entre eux, nous avons ajouté une troisième phase, dite phase *verticale*, qui

consiste à déplacer du stock final d'un produit vers un autre. Ce déplacement doit se faire en étant sûr que les contraintes de stock final maximal soient bien respectées. Après avoir résolu le problème sans contraintes de capacité et effectué les différentes phases de lissage, la solution obtenue peut ne toujours pas respecter les contraintes de capacité de production à chaque période. Pour reconstruire une solution réalisable de notre problème initial, nous avons ajouté à l'heuristique une dernière phase qui transforme l'excès de capacité en ventes perdues.

Ces heuristiques ont permis de fournir des solutions avec peu de ventes perdues, ce qui était le but recherché. En utilisant la formulation agrégée du problème étudié, nous avons aussi montré que l'heuristique de relax-and-fix parallèle pouvait, selon le choix de stratégie pour la sélection des intervalles à chaque itération, fournir de meilleurs résultats que l'heuristique de relax-and-fix dans sa version classique. Cette heuristique a aussi l'avantage de proposer un ensemble de solutions partielles à chaque itération. La reconstruction des solutions partielles a permis d'obtenir dès les premières itérations de bonnes solutions réalisables pour le problème. Les résultats numériques montrent aussi que l'heuristique basée sur la relaxation lagrangienne est particulièrement intéressante pour trouver rapidement des solutions avec peu de ventes perdues. Il serait intéressant d'étendre l'heuristique lagrangienne proposée à des problèmes avec d'autres contraintes industrielles (quantités minimales à produire par période, bornes sur les stocks), dont on sait que les versions sans contraintes de capacité peuvent être résolues en un temps polynomial à l'aide de la programmation dynamique. Il ne s'agirait dans ce cas là que de changer la méthode de résolution des sous-problèmes et d'adapter les mouvements locaux présentés dans les Sections [4.4.3](#) et [4.4.4](#), afin que ces mouvements locaux n'affectent pas la réalisabilité de la solution. L'avantage de l'heuristique de relax-and-fix est qu'elle est plus facilement modulable pour résoudre d'autres modèles plus complexes de lot-sizing.

Chapitre 5 : Application industrielle

L'objectif final de cette thèse a été l'implémentation des heuristiques du Chapitre [4](#) dans l'outil d'optimisation développé par DecisionBrain, ainsi que des tests de performance sur des instances industrielles. Dans le Chapitre [5](#) nous proposons une modification de l'algorithme du relax-and-fix parallélisé pour le rendre plus générique et capable de résoudre des problèmes avec des contraintes liant les productions entre les périodes, telle que la continuité de l'état d'une ressource entre deux périodes successives. En effet, nous avons montré que si le problème contenait certaines contraintes liant les périodes entre elles, l'heuristique de relax-and-fix, et en particulier le fait de fixer des décisions de production, peut aboutir à des solutions non réalisables. Nous avons considéré un cas industriel de peinture de pièces automobiles dont la modélisation contient des contraintes de continuité de la production entre les périodes, ainsi que des contraintes de quantités minimales de production sur plusieurs périodes consécutives. Contrairement à ce qui a été considéré dans les autres chapitres, pour le problème étudié dans ce chapitre, les setups correspondent à des états sur les ressources parmi un ensemble d'états possibles requis pour produire un certain type de produit. Dans la nouvelle version du relax-and-fix proposée, la décomposition des variables binaires se fait successivement selon les périodes ou selon les produits. Afin de pouvoir remettre en question les décisions de production si ces dernières peuvent entraîner l'infaisabilité du modèle, seules certaines variables binaires sont fixées à chaque itération. De plus, cette heuristique permet de prendre en compte une optimisation multiobjectif, où les différents objectifs sont optimisés dans un ordre lexicographique. L'heuristique proposée dans ce chapitre a été implémentée

dans l'outil de modélisation et de résolution de problème de planification de la production de DecisionBrain (*Planning Engine*) et testée sur un ensemble d'instances industrielles. Pour ces instances, les résultats obtenus par l'heuristique sont similaires à ceux obtenus en utilisant IBM ILOG CPLEX pour des temps comparables.

Certaines contraintes industrielles n'ont pas été prises en compte et il serait intéressant de voir comment les heuristiques pourraient être modifiées pour les prendre en compte. Nous avons seulement considéré une décomposition chronologique et une décomposition par produit des variables binaires, mais l'heuristique présentée dans l'Algorithme 5.1 pourrait s'appliquer à toute décomposition des variables binaires. Une extension possible serait donc de considérer d'autres décompositions, notamment en explorant davantage la décomposition par produit. Des résultats numériques doivent aussi être obtenus pour des instances plus difficiles à résoudre.

Chapitre 6 : Perspectives

Différentes perspectives à court, moyen et long terme peuvent être considérées pour prolonger les travaux entrepris durant cette thèse.

En ce qui concerne les perspectives à court terme, une première perspective concernant la considération des effets de fin d'horizon introduite dans le Chapitre 2 serait de prendre en compte des variations dans la demande au-delà de l'horizon temporel. Il serait alors possible d'étendre un problème statique à un problème avec entre autres des variations saisonnières. Les valeurs de stock minimal et maximal pourraient se déduire par pondération des coûts et des demandes au-delà de l'horizon à optimiser. Dans le Chapitre 2, nous avons défini un problème de lot-sizing cyclique pour établir des niveaux pertinents de stock en fin d'horizon. Ce modèle pourrait aussi être utilisé pour faire de l'analyse des coûts. En effet, il serait par exemple possible d'analyser les instances testées, notamment en ce qui concerne la corrélation entre les coûts et la capacité de production. Certains paramètres, tels que les paramètres de ventes perdues, sont difficiles à définir en pratique car la définition de tels coûts n'est pas directe. L'analyse du problème de lot-sizing cyclique a pu permettre de définir des coûts de ventes perdues permettant l'existence théorique d'une solution optimale sans ventes perdues, ce qui en pratique peut représenter une définition pertinente des coûts de ventes perdues. Le sujet de la prise en compte de l'évolution des stocks à l'intérieur des périodes du Chapitre 3 n'a été que peu étudié dans la littérature et de nombreuses extensions peuvent être envisagées. Différents modèles peuvent être proposés en supposant d'autres hypothèses sur les évolutions de la production et de la demande à l'intérieur de chaque période. Davantage d'expérimentations doivent être effectuées afin de valider l'utilité de ces nouveaux modèles.

A moyen terme, une perspective possible serait de s'intéresser à d'autres moyens d'atténuer les effets de fin d'horizon. Nous avons montré que ces effets de bords pouvaient grandement affecter la qualité des plans de production obtenus en horizon glissant. Les travaux effectués ont consisté à rajouter des contraintes sur les stocks en fin d'horizon, mais d'autres approches pourraient être considérées. Il serait possible de modifier les coûts de stockage durant la dernière période afin d'éviter d'avoir un stock nul en fin d'horizon. Il serait aussi possible de rajouter des contraintes de production ou de consommation de capacité minimale durant les dernières périodes de l'horizon temporel. Le Chapitre 3 a montré que les nouveaux modèles proposés étaient plus difficiles à résoudre que le CLSP. Il serait donc intéressant de proposer des heuristiques adaptées à la résolution de ces nouveaux modèles. Il pourrait par exemple s'agir d'une heuristique basée sur une relaxation des nouvelles contraintes de stock. Une approche basée sur de la génération de contraintes, qui consisterait à n'ajouter qu'un

sous-ensemble des contraintes de stock à l'intérieur des périodes, pourrait aussi donner de bons résultats numériques. Les méthodes développées dans le Chapitre 4 pourraient être généralisées afin de mettre en place une heuristique permettant de résoudre des problèmes de lot-sizing plus génériques. L'objectif serait à terme d'implémenter cette méthode dans l'outil de planification optimisée de DecisionBrain. L'heuristique de relax-and-fix parallélisé n'a été testé que sur une seule application industrielle, et il serait nécessaire d'étendre ces tests à d'autres types problèmes industriels. En effet, comme cela a été observé dans le Chapitre 5, l'ajout de nouvelles contraintes peut facilement affecter les performances d'une heuristique.

L'une des principales pistes de perspectives à long terme serait d'étudier les liens entre lot-sizing et ordonnancement. Dans les problèmes intégrés, les décisions prises au niveau tactique lors de la résolution du problème de lot-sizing impactent directement les décisions prises au niveau opérationnelle lors de l'ordonnancement des lots à produire. Au cours de cette thèse, nous avons considéré l'évolution de la production et de la demande à l'intérieur de périodes, ce qui est normalement du ressort de l'ordonnancement. Être en mesure de fournir des informations au modèle de lot-sizing sans augmenter de manière significative sa complexité pourrait permettre de trouver des plans de production optimaux plus réalistes, ce qui aurait une portée industrielle importante. L'utilisation de la parallélisation pour améliorer les performances d'une heuristique, comme nous avons pu le faire pour l'heuristique de relax-and-fix dans le Chapitre 4, est aussi un sujet intéressant à explorer. Par exemple, la manière dont l'heuristique de relax-and-fix a été parallélisée peut facilement s'adapter à n'importe quel problème contenant des variables binaires. Une réflexion sur la meilleure manière de séparer les variables binaires dépendant du problème à résoudre serait aussi un sujet intéressant à étudier, surtout si la réalisabilité de l'heuristique dépend de la décomposition choisie.

Introduction générale

Chapitre 1 : Contexte industriel et état de l'art

1.1 Planification de la production

1.2 Exemple industriel

1.3 Etat de l'art

1.4 Conclusions

Chapitre 2 : Problème de lot-sizing avec capacité, temps de setup et stock final minimum et maximum

2.1 Etat de l'art

2.1 Motivations

2.2 Niveaux de stocks pour le problème de lot-sizing avec capacité et temps de setup en horizon glissant

2.3 Généralisations

2.4 Schéma de génération de nouvelles instances

2.5 Résultats numériques

2.6 Conclusions

Chapitre 3 : Problème de lot-sizing avec capacité et contraintes de stock à l'intérieur des périodes

3.1 Etat de l'art

3.2 Motivations

3.3 Production uniforme et approximation de l'évolution de la demande

3.4 Production à taux maximal et demande instantanée

3.5 Résultats numériques

3.6 Extensions

3.7 Conclusions

Chapitre 4 : Approches de décompositions parallèles pour le problème de lot-sizing avec capacité

4.1 Etat de l'art

4.2 Formations du problème

4.3 Relax-and-fix parallèle

4.4 Heuristique lagrangienne

4.5 Problème de lot-sizing sans ventes perdues avec contraintes de stock final

4.6 Résultats numériques

4.7 Conclusions

Chapitre 5 : Application industrielle

5.1 Introduction

5.2 Contraintes additionnelles et application industrielle

5.3 Adaptation du l'algorithme du relax-and-fix

5.4 Résultats numériques

5.5 Conclusions

Chapitre 6 : Conclusion et perspectives

6.1 Conclusion

6.2 Perspectives

List of Figures

1.1	Painting spare parts	8
2.1	Production quantities in the first 20 periods	21
2.2	Optimal solution for an instance of [Trigeiro et al. (1989)].	22
2.3	Inventory evolution on a rolling horizon	23
2.4	Capacity consumption on a rolling horizon	23
2.5	cyclical production	26
2.6	Comparing the optimal plans of two related original and new instances.	43
2.7	Production quantities for different values of \bar{s}	47
3.1	Dynamic evolution of the inventory	53
3.2	Convex demand	55
3.3	Convex demand	55
3.4	Concave demand	56
3.5	Generic demand	56
3.6	Minimum inventory	57
3.7	Minimum inventory	57
3.8	Minimum inventory	58
3.9	Maximum inventory	58
4.1	Optimal production plan	81
4.2	Chronological optimization	81
4.3	Illustration of the parallelized relax-and-fix heuristic	84
4.4	Selecting the best interval	85

List of Tables

2.1 Comparison of the variability for the original and new instances. (*)This classification only applies on the original instances.	40
2.2 Results	41
2.3 Average optimality gaps and computational times for the original and new instances with $T_{lim} = 100$ sec. (*)This classification only applies on the original instances.	42
2.4 Average optimality gaps and computational times for the original and new instances with $T_{lim} = 600$ sec. (*)This classification only applies on the original instances.	42
2.5 Gap with predicted objective value. (*)This classification only applies on the original instances.	44
2.6 Fraction of lost sales (%) for $N = 30$ and $k_{capa} = 3.75$	45
2.7 Fraction of lost sales (%) for $N = 30$ and $k_{capa} = 7.5$	46
2.8 Fraction of lost sales (%) for $N = 30$ and $k_{capa} = 15$	46
3.1 Comparison between models for randomly generated demands (maximum CPU time of 600 seconds)	66
3.2 Comparison between models for randomly generated demands during the first half of the period (maximum CPU time of 600 seconds)	66
3.3 Comparison between models for instantaneous demand at the beginning of the period (maximum CPU time of 600 seconds)	67
3.4 Comparison between models for instantaneous demand at the end of the period (maximum CPU time of 600 seconds)	67
3.5 Comparison between models for randomly generated demands and backlogs (maximum CPU time of 600 seconds)	69
4.1 Numerical results with respect to TBO	98
4.2 Numerical results with respect to capacity settings	98
4.3 Numerical results with respect to TBO with reoptimization	98
4.4 Numerical results with respect to capacity settings with reoptimization	99
4.5 Numerical results with respect to capacity settings using disaggregated formulation	99
4.6 Numerical results with respect to capacity settings using disaggregated formulation and with reconstruction	100
5.1 Instance characteristics	111
5.2 Numerical results using IBM ILOG CPLEX (maximum CPU time of 300 seconds per objective)	111

5.3 Numerical results with 4 intervals	112
5.4 Numerical results with 6 intervals	112
5.5 Numerical results with T intervals	112

List of Algorithms

4.1	Relax-and-fix algorithm	82
4.2	Parallelized relax-and-fix algorithm	83
4.3	Lagrangian relaxation based heuristic	89
4.4	Heuristic to solve (\bar{P})	94
4.5	Exact method to solve (P) when the maximum ending inventory is the same for all items	96
5.1	First run of the updated parallelized relax-and-fix algorithm	110

Bibliography

- Absi, N., Detienne, B. and Dauzère-Pérès, S. (2013). Heuristics for the multi-item capacitated lot-sizing problem with lost sales, *Computers & Operations Research* **40**(1): 264–272.
- Absi, N. and Kedad-Sidhoum, S. (2007). MIP-based heuristics for multi-item capacitated lot-sizing problem with setup times and shortage costs, *RAIRO - Operations Research* **41**(2): 171–192.
- Absi, N. and Kedad-Sidhoum, S. (2008). The multi-item capacitated lot-sizing problem with setup times and shortage costs, *European Journal of Operational Research* **185**(3): 1351–1374.
- Absi, N. and Kedad-Sidhoum, S. (2009). The multi-item capacitated lot-sizing problem with safety stocks and demand shortage costs, *Computers & Operations Research* **36**(11): 2926–2936.
- Absi, N., Kedad-Sidhoum, S. and Dauzère-Pérès, S. (2011). Uncapacitated lot-sizing problem with production time windows, early productions, backlogs and lost sales, *International Journal of Production Research* **49**(9): 2551–2566.
- Absi, N. and van den Heuvel, W. (2019). Worst case analysis of Relax and Fix heuristics for lot-sizing problems, *European Journal of Operational Research* **279**(2): 449–458.
- Aggarwal, A. and Park, J. K. (1993). Improved algorithms for economic lot size problems, *Operations research* **41**(3): 549–571.
- Akbalik, A., Penz, B. and Rapine, C. (2015). Capacitated lot sizing problems with inventory bounds, *Annals of Operations Research* **229**(1): 1–18.
- Aksen, D., Altinkemer, K. and Chand, S. (2003). The single-item lot-sizing problem with immediate lost sales, *European Journal of Operational Research* **147**(3): 558–566.
- Almeder, C., Klabjan, D., Traxler, R. and Almada-Lobo, B. (2015). Lead time considerations for the multi-level capacitated lot-sizing problem, *European Journal of Operational Research* **241**(3): 727–738.
- Anderson, E. J. and Cheah, B. S. (1993). Capacitated lot-sizing with minimum batch sizes and setup times, *International Journal of Production Economics* **30-31**: 137–152.
- Atamtürk, A. and Küçükyavuz, S. (2008). An algorithm for lot sizing with inventory bounds and fixed costs, *Operations Research Letters* **36**(3): 297–299.
- Bahl, H. C. and Ritzman, L. P. (1984). A cyclical scheduling heuristic for lot sizing with capacity constraints, *International Journal of Production Research* **22**(5): 791–800.

- Barany, I., Van Roy, T. J. and Wolsey, L. A. (1984). Strong Formulations for Multi-Item Capacitated Lot Sizing, *Management Science* **30**(10): 1255–1261.
- Berretta, R., França, P. M. and Armentano, V. A. (2005). Metaheuristic approaches for the multilevel resource-constrained lot-sizing problem with setup and lead times, *Asia-Pacific Journal of Operational Research* **22**(2): 261–286.
- Billington, P. J., McClain, J. O. and Thomas, L. J. (1986). Heuristics for multilevel lot-sizing with a bottleneck, *Management Science* **32**(8): 989–1006.
- Bitran, G. R. and Matsuo, H. (1986). The Multi-Item Capacitated Lot Size Problem: Error Bounds of Manne’s Formulations, *Management Science* **32**(3): 350–359.
- Bitran, G. R. and Yanasse, H. H. (1982). Computational complexity of the capacitated lot size problem, *Management Science* **28**(10): 1174–1186.
- Brahimi, N., Absi, N., Dauzère-Pérès, S. and Nordli, A. (2017). Single-item dynamic lot-sizing problems: An updated survey, *European Journal of Operational Research* **263**(3): 838–863.
- Brahimi, N., Dauzère-Pérès, S. and Najid, N. M. (2006). Capacitated Multi-Item Lot-Sizing Problems with Time Windows, *Operations Research* **54**(5): 951–967.
- Brahimi, N., Dauzère-Pérès, S., Najid, N. M. and Nordli, A. (2006). Single item lot sizing problems, *European Journal of Operational Research* **168**(1): 1–16.
- Buschkühl, L., Sahling, F., Helber, S. and Tempelmeier, H. (2010). Dynamic capacitated lot-sizing problems: a classification and review of solution approaches, *Or Spectrum* **32**(2): 231–261.
- Campbell, G. M. and Mabert, V. A. (1991). Cyclical Schedules for Capacitated Lot Sizing with Dynamic Demands, *Management Science* **37**(4): 409–427.
- Carlson, R. C., Jucker, J. V. and Kropp, D. H. (1979). Less Nervous MRP Systems: A Dynamic Economic Lot-Sizing Approach, *Management Science* **25**(8): 754–761.
- Chand, S., Hsu, V. N. and Sethi, S. (2002). Forecast, Solution, and Rolling Horizons in Operations Management Problems: A Classified Bibliography, *Manufacturing & Service Operations Management* **4**(1): 25–43.
- Charles, M., Dauzère-Pérès, S., Kedad-Sidhoum, S. and Mazhoud, I. (2021). Minimum and maximum ending inventories for the capacitated lot-sizing problem with setup times, *Technical Report EMSE CMP-SFL 2021/1*, Mines Saint-Etienne, Gardanne, France.
- Chen, W. H. and Thizy, J. M. (1990). Analysis of relaxations for the multi-item capacitated lot-sizing problem, *Annals of Operations Research* **26**(1): 29–72.
- Clark, A. R. and Clark, S. J. (2000). Rolling-horizon lot-sizing when set-up times are sequence-dependent, *International Journal of Production Research* **38**(10): 2287–2307.
- Copil, K., Wörbelauer, M., Meyr, H. and Tempelmeier, H. (2017). Simultaneous lotsizing and scheduling problems: a classification and review of models, *OR Spectrum* **39**(1): 1–64.

- Dauzère-Pérès, S., Brahimi, N., Najid, N. M. and Nordli, A. (2002). Uncapacitated lot-sizing problems with time windows, *Research report 02/4/AUTO*, Ecole des Mines de Nantes, France.
- Dauzère-Pérès, S., Gershwin, S. B. and Sevaux, M. (2000). Models and solving procedures for continuous-time production planning, *IIE Transactions* **32**(2): 93–103.
- Dauzère-Pérès, S. and Lasserre, J.-B. (1994). Integration of lotsizing and scheduling decisions in a job-shop, *European Journal of Operational Research* **75**(2): 413–426.
- Dauzère-Pérès, S. and Lasserre, J.-B. (2002). On the importance of sequencing decisions in production planning and scheduling, *International Transactions in Operational Research* **9**(6): 779–793.
- de Araujo, S. A., De Reyck, B., Degraeve, Z., Fragkos, I. and Jans, R. (2015). Period Decompositions for the Capacitated Lot Sizing Problem with Setup Times, *INFORMS Journal on Computing* **27**(3): 431–448.
- Degraeve, Z. and Jans, R. (2007). A New Dantzig-Wolfe Reformulation and Branch-and-Price Algorithm for the Capacitated Lot-Sizing Problem with Setup Times, *Operations Research* **55**(5): 909–920.
- Denizel, M. and Süral, H. (2006). On alternative mixed integer programming formulations and LP-based heuristics for lot-sizing with setup times, *Journal of the Operational Research Society* **57**(4): 389–399.
- Diaby, M., Bahl, H. C., Karwan, M. H. and Zionts, S. (1992). A Lagrangean Relaxation Approach for Very-Large-Scale Capacitated Lot-Sizing, *Management Science* **38**(9): 1329–1340.
- Díaz-Madroño, M., Mula, J. and Peidro, D. (2014). A review of discrete-time optimization models for tactical production planning, *International Journal of Production Research* **52**(17): 5171–5205.
- Dillenberger, C., Escudero, L. F., Wollensak, A. and Zhang, W. (1993). On solving a large-scale resource allocation problem in production planning, *Operations research in production planning and control*, Springer, pp. 105–119.
- Dillenberger, C., Escudero, L. F., Wollensak, A. and Zhang, W. (1994). On practical resource allocation for production planning and scheduling with period overlapping setups, *European Journal of Operational Research* **75**(2): 275–286.
- Dixon, P. S. and Silver, E. A. (1981). A heuristic solution procedure for the multi-item, single-level, limited capacity, lot-sizing problem, *Journal of operations management* **2**(1): 23–39.
- Dobson, G. (1987). The Economic Lot-Scheduling Problem: Achieving Feasibility Using Time-Varying Lot Sizes, *Operations Research* **35**(5): 764–771.
- Drexl, A. and Kimms, A. (1997). Lot sizing and scheduling—survey and extensions, *European Journal of operational research* **99**(2): 221–235.
- Elmaghraby, S. E. (1978). The Economic Lot Scheduling Problem (ELSP): Review and Extensions, *Management Science* **24**(6): 587–598.

- Eppen, G. D. and Martin, R. K. (1987). Solving Multi-Item Capacitated Lot-Sizing Problems Using Variable Redefinition, *Operations Research* **35**(6): 832–848.
- Federgruen, A., Meissner, J. and Tzur, M. (2007). Progressive Interval Heuristics for Multi-Item Capacitated Lot-Sizing Problems, *Operations Research* **55**(3): 490–502.
- Federgruen, A. and Tzur, M. (1991). A simple forward algorithm to solve general dynamic lot sizing models with n periods in $O(n \log n)$ or $O(n)$ time, *Management Science* **37**(8): 909–925.
- Federgruen, A. and Tzur, M. (1994). Minimal Forecast Horizons and a New Planning Procedure for the General Dynamic Lot Sizing Model: Nervousness Revisited, *Operations Research* **42**(3): 456–468.
- Fiorotto, D. J., de Araujo, S. A. and Jans, R. (2015). Hybrid methods for lot sizing on parallel machines, *Computers & Operations Research* **63**: 136–148.
- Fisher, M., Ramdas, K. and Zheng, Y.-S. (2001). Ending Inventory Valuation in Multiperiod Production Scheduling, *Management Science* **47**(5): 679–692.
- Fleischmann, B. (1990). The discrete lot-sizing and scheduling problem, *European Journal of Operational Research* **44**(3): 337–348.
- Fleischmann, B. and Meyr, H. (1997). The general lotsizing and scheduling problem, *Operations-Research-Spektrum* **19**(1): 11–21.
- Florian, M. and Klein, M. (1971). Deterministic Production Planning with Concave Costs and Capacity Constraints, *Management Science* **18**(1): 12–20.
- Florian, M., Lenstra, J. K. and Rinnooy Kan, A. H. G. (1980). Deterministic Production Planning: Algorithms and Complexity, *Management Science* **26**(7): 669–679.
- Friedman, Y. and Hoch, Y. (1978). A dynamic lot-size model with inventory deterioration, *INFOR: Information Systems and Operational Research* **16**(2): 183–188.
- Gomez Urrutia, E. D., Aggoune, R. and Dauzère-Pérès, S. (2014). Solving the integrated lot-sizing and job-shop scheduling problem, *International Journal of Production Research* **52**(17): 5236–5254.
- Grigoriev, A., Kreuzen, V. J. and Oosterwijk, T. (2020). Cyclic lot-sizing problems with sequencing costs, *Journal of Scheduling* **24**(2): 123–135.
- Gutiérrez, J., Colebrook, M., Abdul-Jalbar, B. and Sicilia, J. (2013). Effective replenishment policies for the multi-item dynamic lot-sizing problem with storage capacities, *Computers & Operations Research* **40**(12): 2844–2851.
- Gutiérrez, J., Sedeño-Noda, A., Colebrook, M. and Sicilia, J. (2003). A new characterization for the dynamic lot size problem with bounded inventory, *Computers & Operations Research* **30**(3): 383–395.
- Haase, K. and Kohlmorgen, U. (1996). Parallel Genetic Algorithm for the Capacitated Lot-Sizing Problem, *Operations Research Proceedings 1995*, Vol. 1995, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 370–375.

- Harris, F. W. (1913). How Many Parts to Make at Once, *Factory, The Magazine of Management* **10**(2): 135–136.
- Helber, S. and Sahling, F. (2010). A fix-and-optimize approach for the multi-level capacitated lot sizing problem, *International Journal of Production Economics* **123**(2): 247–256.
- Hellion, B., Mangione, F. and Penz, B. (2012). A polynomial time algorithm to solve the single-item capacitated lot sizing problem with minimum order quantities and concave costs, *European Journal of Operational Research* **222**(1): 10–16.
- Holmbom, M. and Segerstedt, A. (2014). Economic order quantities in production: From Harris to economic lot scheduling problems, *International Journal of Production Economics* **155**: 82–90.
- Hwang, H.-C. and van den Heuvel, W. (2012). Improved algorithms for a lot-sizing problem with inventory bounds and backlogging: Improved Algorithms for a Lot-Sizing Problem with Inventory Bounds, *Naval Research Logistics (NRL)* **59**(3-4): 244–253.
- Hwang, H.-C., Van Den Heuvel, W. and Wagelmans, A. P. (2013). The economic lot-sizing problem with lost sales and bounded inventory, *IIE Transactions* **45**(8): 912–924.
- Jans, R. and Degraeve, Z. (2004). Improved lower bounds for the capacitated lot sizing problem with setup times, *Operations Research Letters* **32**(2): 185–195.
- Jans, R. and Degraeve, Z. (2007). Meta-heuristics for dynamic lot sizing: A review and comparison of solution approaches, *European journal of operational research* **177**(3): 1855–1875.
- Jans, R. and Degraeve, Z. (2008). Modeling industrial lot sizing problems: a review, *International Journal of Production Research* **46**(6): 1619–1643.
- Jing, F. and Mu, Y. (2020). Dynamic lot-sizing model under perishability, substitution, and limited storage capacity, *Computers & Operations Research* **122**: 104978.
- Karimi, B., Fatemi Ghomi, S. and Wilson, J. (2003). The capacitated lot sizing problem: a review of models and algorithms, *Omega* **31**(5): 365–378.
- Karmarkar, U. S., Kekre, S. and Kekre, S. (1987). The dynamic lot-sizing problem with startup and reservation costs, *Operations Research* **35**(3): 389–398.
- Karmarkar, U. S. and Schrage, L. (1985). The Deterministic Dynamic Product Cycling Problem, *Operations Research* **33**(2): 326–345.
- Karp, R. M. (1972). Reducibility among combinatorial problems, *Complexity of computer computations*, Springer, pp. 85–103.
- Krarup, J. and Bilde, O. (1977). Plant location, set covering and economic lot size: An 0 (mn)-algorithm for structured problems, *Numerische methoden bei optimierungsaufgaben band 3*, Springer, pp. 155–180.
- Lee, C.-Y., Çetinkaya, S. and Wagelmans, A. P. (2001). A dynamic lot-sizing model with demand time windows, *Management Science* **47**(10): 1384–1395.

- Liu, T. (2008). Economic lot sizing problem with inventory bounds, *European Journal of Operational Research* **185**(1): 204–215.
- Liu, X. and Tu, Y. (2008). Production planning with limited inventory capacity and allowed stockout, *International Journal of Production Economics* **111**(1): 180–191.
- Love, S. F. (1973). Bounded Production and Inventory Models with Piecewise Concave Costs, *Management Science* **20**(3): 313–318.
- Maes, J., McClain, J. O. and Van Wassenhove, L. N. (1991). Multilevel capacitated lotsizing complexity and LP-based heuristics, *European Journal of Operational Research* **53**(2): 131–148.
- Maes, J. and Wassenhove, L. V. (1988). Multi-Item Single-Level Capacitated Dynamic Lot-Sizing Heuristics: A General Review, *The Journal of the Operational Research Society* **39**(11): 991.
- Manne, A. S. (1958). Programming of Economic Lot Sizes, *Management Science* **4**(2): 115–135.
- Melo, R. A. and Ribeiro, C. C. (2017). Formulations and heuristics for the multi-item uncapacitated lot-sizing problem with inventory bounds, *International Journal of Production Research* **55**(2): 576–592.
- Millar, H. H. and Yang, M. (1994). Lagrangian heuristics for the capacitated multi-item lot-sizing problem with backordering, *International Journal of Production Economics* **34**(1): 1–15.
- Miller, A., Nemhauser, G. and Savelsbergh, M. (2000). Solving multi-item capacitated lot-sizing problems with setup times by branch-and-cut, *Technical report*, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).
- Minner, S. (2009). A comparison of simple heuristics for multi-product dynamic demand lot-sizing with limited warehouse capacity, *International Journal of Production Economics* **118**(1): 305–310.
- Muller, L. F., Spoorendonk, S. and Pisinger, D. (2012). A hybrid adaptive large neighborhood search heuristic for lot-sizing with setup times, *European Journal of Operational Research* **218**(3): 614–623.
- Okhrin, I. and Richter, K. (2011). An $O(T^3)$ algorithm for the capacitated lot sizing problem with minimum order quantities, *European Journal of Operational Research* **211**(3): 507–514.
- Park, Y. W. and Klabjan, D. (2015). Lot sizing with minimum order quantity, *Discrete Applied Mathematics* **181**: 235–254.
- Pochet, Y. and Van Vyve, M. (2004). A general heuristic for production planning problems, *INFORMS Journal on Computing* **16**(3): 316–327.
- Pochet, Y. and Wolsey, L. A. (2006). *Production planning by mixed integer programming*, Springer series in operations research and financial engineering, Springer, New York ; Berlin.

- Roux, W., Dauzère-Pérès, S. and Lasserre, J. B. (1999). Planning and scheduling in a multi-site environment, *Production Planning & Control* **10**(1): 19–28.
- Sahling, F., Buschkühl, L., Tempelmeier, H. and Helber, S. (2009). Solving a multi-level capacitated lot sizing problem with multi-period setup carry-over via a fix-and-optimize heuristic, *Computers & Operations Research* **36**(9): 2546–2553.
- Sandbothe, R. A. and Thompson, G. L. (1990). A forward algorithm for the capacitated lot size model with stockouts, *Operations Research* **38**(3): 474–486.
- Seeanner, F. and Meyr, H. (2013). Multi-stage simultaneous lot-sizing and scheduling for flow line production, *OR Spectrum* **35**(1): 33–73.
- Silver, E. A. and Meal, H. C. (1973). A heuristic for selecting lot size quantities for the case of a deterministic time-varying demand rate and discrete opportunities for replenishment, *Prod. Inventory Manage.* **14**: 64–77.
- Stadtler, H. (2000). Improved Rolling Schedules for the Dynamic Single-Level Lot-Sizing Problem, *Management Science* **46**(2): 318–326.
- Stadtler, H. (2003). Multilevel Lot Sizing with Setup Times and Multiple Constrained Resources: Internally Rolling Schedules with Lot-Sizing Windows, *Operations Research* **51**(3): 487–502.
- Suerie, C. and Stadtler, H. (2003). The Capacitated Lot-Sizing Problem with Linked Lot Sizes, *Management Science* **49**(8): 1039–1054.
- Süral, H., Denizel, M. and Van Wassenhove, L. N. (2009). Lagrangean relaxation based heuristics for lot sizing with setup times, *European Journal of Operational Research* **194**(1): 51–63.
- Tempelmeier, H. and Derstroff, M. (1996). A lagrangean-based heuristic for dynamic multi-level multiitem constrained lotsizing with setup times, *Management Science* **42**(5): 738–757.
- Thizy, J. and Van Wassenhove, L. N. (1985). Lagrangean relaxation for the multi-item capacitated lot-sizing problem: A heuristic implementation, *IIE transactions* **17**(4): 308–313.
- Trigeiro, W. W., Thomas, L. J. and McClain, J. O. (1989). Capacitated Lot Sizing with Setup Times, *Management Science* **35**(3): 353–366.
- van den Heuvel, W. and Wagelmans, A. P. (2005). A comparison of methods for lot-sizing in a rolling horizon environment, *Operations Research Letters* **33**(5): 486–496.
- Vanderbeck, F. (1998). Lot-sizing with start-up times, *Management Science* **44**(10): 1409–1425.
- Vofß, S. and Woodruff, D. L. (2006). *Introduction to computational optimization models for production planning in a supply chain*, Vol. 240, Springer Science & Business Media.
- Wagelmans, A., van Hoesel, S. and Kolen, A. (1992). Economic Lot Sizing: An $O(n \log n)$ Algorithm That Runs in Linear Time in the Wagner-Whitin Case, *Operations Research* **40**(1-supplement-1): S145–S156.

- Wagner, H. M. and Whitin, T. M. (1958). Dynamic Version of the Economic Lot Size Model, *Management Science* **5**(1): 89–96.
- Wolosewicz, C., Dauzère-Pérès, S. and Aggoune, R. (2015). A lagrangian heuristic for an integrated lot-sizing and fixed scheduling problem, *European Journal of Operational Research* **244**(1): 3–12.
- Zamani, R. and Lau, S. K. (2010). Embedding learning capability in lagrangean relaxation: An application to the travelling salesman problem, *European Journal of Operational Research* **201**(1): 82–88.
- Zangwill, W. I. (1966). A deterministic multi-period production scheduling model with backlogging, *Management Science* **13**(1): 105–119.
- Zangwill, W. I. (1969). A Backlogging Model and a Multi-Echelon Model of a Dynamic Economic Lot Size Production System—A Network Approach, *Management Science* **15**(9): 506–527.
- Zhong, J., Chu, F., Chu, C. and Yang, S. (2016). Polynomial dynamic programming algorithms for lot sizing models with bounded inventory and stockout and/or backlogging, *Journal of Systems Science and Systems Engineering* **25**(3): 370–397.
- Önal, M., van den Heuvel, W. and Liu, T. (2012). A note on “The economic lot sizing problem with inventory bounds”, *European Journal of Operational Research* **223**(1): 290–294.

École Nationale Supérieure des Mines
de Saint-Étienne

NNT : 2021LYSEM039

Mehdi Charles

Modeling and solving multi-item lot-sizing problems with inventory constraints

Speciality: Industrial Engineering

Keywords: Capacitated lot-sizing problems, lost sales, inventory constraints, heuristics

Abstract:

In this thesis, we considered the capacitated multi-item lot-sizing problem with setup times and lost sales. We extended this problem to take into account important industrial aspects, especially with regards to inventory management. We first studied the end-of-horizon effects on optimal solutions of lot-sizing problems that, even on a rolling horizon, can lead to important additional costs. To reduce these effects, we have added a global minimum ending inventory constraint as well as a maximum ending inventory constraint for each item. These values were deduced from the analysis of a cyclical capacitated lot-sizing problem with setup times, whose linear relaxation can be analytically solved. Then, we modeled the inventory evolution within each period. This point is especially relevant when the storage capacity is limited. We added new inventory constraints to better respect inventory bounds when scheduling productions within each period. The constraints differ based on hypotheses on the shapes of evolution of production and demand. Numerical experiments showed that these new constraints enable to schedule production plans with a better inventory management. Decomposition approaches (*Lagrangian relaxation*, *relax-and-fix*) were developed in order to propose generic approaches to solve capacitated lot-sizing problems with setup times. An original use of parallelization was proposed in order to reduce the size of the subproblems to solve and to use Decisionbrain's tools. Finally, the parallelized relax-and-fix was implemented into DecisionBrain's optimization tool and tests were performed on industrial instances.

École Nationale Supérieure des Mines
de Saint-Étienne

NNT : 2021LYSEM039

Mehdi Charles

Modélisation et résolution de problèmes complexes de lot-sizing multi-produits avec contraintes sur les stocks

Spécialité: Génie Industriel

Mots clefs :

Lot-sizing avec capacité, ventes perdues, gestion de stock, heuristiques

Résumé:

Nous nous sommes intéressés au problème de lot-sizing multi-produits avec capacité, temps de lancement et ventes perdues. Nous avons étendu ce problème afin de prendre en compte des aspects industriels importants, en particulier des contraintes sur les stocks. Nous avons d'abord étudié les effets de fin d'horizon des solutions aux problèmes de lot-sizing, qui peuvent entraîner des coûts importants même pour des horizons temporels très longs. Pour compenser ces effets, nous avons proposé de rajouter une contrainte de stock final minimal ainsi que des contraintes de stock maximal par produit. Ces valeurs ont été déduites d'une analyse d'un problème de lot-sizing cyclique avec temps de lancement et capacité dont la relaxation linéaire peut être résolue de manière analytique. Par la suite, nous nous sommes intéressés à la modélisation de l'évolution des stocks intra-périodes. Cet aspect est particulièrement important lorsque les capacités de stockage sont limitées. Nous avons proposé des nouvelles contraintes qui diffèrent en fonction des hypothèses sur la production et la demande. L'objectif est de limiter les excès et les déficits de stock lors de l'ordonnancement détaillé du plan de production à chaque période. Nos résultats numériques ont montré que ces nouvelles contraintes permettent de construire des plans de production respectant davantage les contraintes sur les stocks. Des méthodes de résolution génériques et plus particulièrement des méthodes de décomposition (*relaxation Lagrangienne, relax-and-fix*) ont été développées. Une approche originale de parallélisation a été proposée, dont l'objectif est de réduire la taille des sous-problèmes à résoudre et d'utiliser les outils disponibles à DecisionBrain. L'objectif final de cette thèse a été l'implémentation des heuristiques proposées dans l'outil d'optimisation développé par DecisionBrain ainsi que des tests de performance sur des instances industrielles.